RISK- AND RELIABILITY ANALYSIS WITH APPLICATIONS

by

ARNE BANG HUSEBY and KRISTINA ROGNLIEN DAHL

Department of Mathematics
Faculty of Mathematics and Natural Sciences
University of Oslo

Abstract

This is a compendium indended for use in the course STK3405 and STK4405 at the Department of Mathematics, University of Oslo, from fall 2017. The compendium provides the foundation of probability theory needed to compute the reliability of a system, i.e., the probability that a system is functioning, when the reliabilities of the components which the system is composed of are known. Examples of systems are energy systems and networks. In addition, various examples of the use of risk analysis for industrial applications are studied. The material is illustrated by different simulation techniques. This material can serve as an introduction to more advanced studies, but is also suitable as support literature to studies in other fields and as further education for realists and engeneers.

Contents

List of Figures

Chapter 1

Introduction

The purpose of these notes is to provide the reader with a basic knowledge on the theoretical foundations of risk- and reliability analysis. In addition, the compendium covers a wide array of current applications of these fields as well as simulation techniques and an introduction to software suitable for applying the theory to practical problems.

Today, various aspects of risk- and reliability theory are used in many different sectors. Examples include transport, project planning, time scheduling, energy networks and aerospace. At the core of all these applications is a need to quantify and control risk, determine how reliable a system is and how to monitor and maintain systems is the best way possible. Clever use of the risk- and reliability theory can provide safer and more efficient systems which are supervised in an optimal way. This leads to financial savings, well-functioning systems and prevention of accidents.

These notes are suitable for a bachelor or master level course in modern risk- and reliability theory, but also as further education for practitioners (engineers, risk analysts etc.).

The structure of the notes is as follows: In Chapter 2 we introduce the theoretical foundation of system analysis. We define a system of components and study various kinds of systems. The structure/reliability function of the system is defined based on component reliabilities. The chapter also includes important defining properties of systems such as monotonicity, coherency and duality. In Chapter 3 we introduce the basic tools for calculating reliability, including pivotal decompositions, path and cut sets as well as modular decompositions. Then, in Chapter 4, we consider different methods for exact computation of the reliability of systems. In Chapter 5 we introduce various measures for computing the structural and the reliability importance of the system components. In some cases, computing the exact reliability is too time-consuming, so in Chapter 6 we give upper and lower bounds for the system reliabilities. In this chapter, we also define a specific kind of dependence between the components of a system

called association, and study how this property can be used to derive bounds for the system reliability without requiring independence of components.

In Chapter 8 we introduce discrete event simulation and show how this is connected to the system analysis of Chapter 2.

Finally, in Chapter 9 we present various applications and examples of practical use of risk- and reliability theory. Some examples include project management and time scheduling, analyzing a network for the transmission of electronic pulses and a conditional Monte Carlo approach to system reliability evaluation. The simulation software Riscue[1] is used to analyze these systems.

Throughout the notes, there are various exercises intended to provide the reader with a deeper understanding of the material. Some of the exercises are more theoretical, some are oriented towards analyzing smaller systems by hand and other rely on simulation techniques and use of Riscue.

The notation used in the compendium is summarized in Chapter 10.

Parts of these notes are based on the compendium by Natvig [35] (in Norwegian). In particular, this is the case for Chapter 2. However, compared to Natvig [35] there are several new examples, exercises, some new topics are included and some others are excluded. The presentation has also been reorganized and reworked.

[1] A free version of Riscue can be downloaded from www.riscue.org

Chapter 2

System analysis

In this chapter, we will introduce the basic concepts of system analysis and reliability theory. In contrast to several other books on this topic, for instance Barlow and Proschan [6] and Natvig [35], we will begin by introducing *stochastic systems* (i.e., when the component states are random). The more classical approach is to first consider *deterministic system* analysis (i.e., where the component states are deterministic), and then move on to the stochastic case. However, as the two are very similar, we go straight to the general stochastic case, and highlight important aspects of the deterministic case whenever necessary. This hopefully saves the reader some work and also makes the presentation more unified. At first, we analyse systems at a particular time. We then move on to *dynamic system analysis* where we consider the development of the systems over time.

2.1 Binary monotone systems

When we use the term *system*, we think of some technological unit consisting of a finite set of *components* which are operating together. A *binary* system has only two possible states: *functioning* or *failed*. Moreover, each component is either functioning or failed as well[1]. We consider a binary system of n components, $1, \ldots, n$, and introduce the component state variable X_i denoting the *state* of component i, $i = 1, \ldots, n$, defined as:

$$X_i := \begin{cases} 1 & \text{if the } i\text{th component is functioning} \\ 0 & \text{otherwise.} \end{cases} \tag{2.1}$$

[1] In so-called *multistate reliability theory*, models where the system and its components have more than two potential states are considered. Much of the theory for binary systems can be extended to multistate systems. Still such systems typically are more computationally challenging. See Natvig [36].

Furtermore, we introduce the variable ϕ representing the *state* of the system, defined as:

$$\phi := \begin{cases} 1 & \text{if the system is functioning} \\ 0 & \text{otherwise.} \end{cases} \tag{2.2}$$

The variables X_i, $i = 1, \ldots, n$ and ϕ are said to be *binary*, since they only have two possible values, 0 and 1. We use upper case letters for the component state variables indicating that these variables are stochastic. Specific values of random variables will be denoted by lower case letters.

The state of the system is assumed to be uniquely determined by the state of the components. Thus, if $\mathbf{X} := (X_1, X_2, \ldots, X_n)$ is the *component state vector*, we may express the system state ϕ as:

$$\phi = \phi(\mathbf{X}).$$

The function $\phi : \{0,1\}^n \to \{0,1\}$ is called the *structure function* of the system.

Intuitively, repairing a component should not make the system worse, and breaking a component should not make system better. Mathematically this property implies that the structure function should be non-decreasing in each argument. This property is an essential part of the definition of a *binary monotone system* which can be stated as follows:

Definition 2.1.1 *A binary monotone system is an ordered pair (C, ϕ), where $C = \{1, \ldots, n\}$ is the component set and ϕ is the structure function. The function ϕ is assumed to be non-decreasing in each argument. The number of components, n, is referred to as the* order *of the system.*

Note that according to this definition the class of binary monotone systems includes systems where the structure function is constant, i.e., systems where either $\phi(\mathbf{x}) = 1$ for all $\mathbf{x} \in \{0,1\}^n$ or $\phi(\mathbf{x}) = 0$ for all $\mathbf{x} \in \{0,1\}^n$. We will refer to such systems as *trivial* systems[2].

We let $\mathbf{0}$ denote the vector where all entries are 0, and similarly, let $\mathbf{1}$ be the vector where all entries are 1. The following result provides a useful property of non-trivial systems.

Theorem 2.1.2 *Let (C, ϕ) be a non-trivial binary monotone system. Then $\phi(\boldsymbol{0}) = 0$ and $\phi(\boldsymbol{1}) = 1$.*

Proof: Since (C, ϕ) is assumed to be non-trivial, there exists at least one vector $\mathbf{x}_0$ such that $\phi(\mathbf{x}_0) = 0$ and another vector $\mathbf{x}_1$ such that $\phi(\mathbf{x}_1) = 1$. Since (C, ϕ) is assumed to be monotone it follows that $0 \leq \phi(\mathbf{0}) \leq \phi(\mathbf{x}_0) = 0$ and $1 = \phi(\mathbf{x}_1) \leq \phi(\mathbf{1}) \leq 1$. Hence, we conclude that $\phi(\mathbf{0}) = 0$ and $\phi(\mathbf{1}) = 1$ □

In order to describe the logical relation between the components and the system, we often use a *reliability block diagram*. In such diagrams components

[2]Some textbooks exclude trivial systems from the class of monotone systems. We have chosen not to do so in order to have a class which is closed with respect to conditioning.

are represented as circles and connected by lines. The system is functioning if and only it is possible to find a way from the left-hand side to the right-hand side of the diagram passing only functioning components. Note that in order to represent arbitrarily complex binary monotone systems, it is sometimes necessary to allow components to occur multiple places in the block diagram. In such cases the reliability block diagram should not be interpreted as a picture of a physical system.

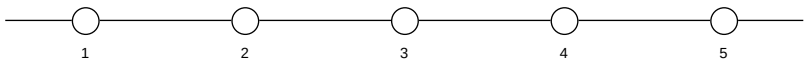

Figure 2.1: A reliability block diagram of a series system of order 5 with component set $C = \{1, \ldots, 5\}$

Example 2.1.3 *Figure 2.1 shows a reliability block diagram of a* series system *of order* 5. *From the diagram we see that the series system is functioning if and only if all of its components are functioning, i.e., if and only if* $X_1 = \cdots = X_5 = 1$. *Hence, the state of the system can be expressed as a function of the component state variables as follows:*

$$\phi = X_1 \cdot X_2 \cdot X_3 \cdot X_4 \cdot X_5 = \prod_{i=1}^{5} X_i.$$

Alternatively, the structure function of the series system can be written as:

$$\phi(\boldsymbol{X}) = \min\{X_1, \ldots, X_5\}$$

since the minimum of $X_1, \ldots, X_5$ *is* 1 *if and only if* $X_1 = \cdots = X_5 = 1$.

It may be useful to think of the product operator $(\cdot)$ of binary variables as representing the logical *and*-operation. Thus, for the series system to function, components 1 *and* 2 *and* $\cdots$ *and* 5 all need to function.

Both the *product*-expression and the *minimum*-expression can be extended to series systems of arbitrary orders. Thus, the structure function of a series system of order n can be written as:

$$\phi(\mathbf{X}) = \prod_{i=1}^{n} X_i = \min_{1 \le i \le n} X_i.$$

In order to derive convenient representations of structure functions of other types of binary monotone system, we also need another operator, called the *coproduct operator*. This operator is denoted by the symbol $\amalg$ and defined as follows:

Notation 2.1.4 *For any $a_1, a_2, \ldots, a_n$ where $a_i \in [0,1]$, $i = 1, 2, \ldots, n$, we define:*

$$a_1 \amalg a_2 := 1 - (1 - a_1)(1 - a_2),$$
$$\coprod_{i=1}^{n} a_i := 1 - \prod_{i=1}^{n}(1 - a_i).$$

It is easy to verify that $1 \amalg 1 = 1 \amalg 0 = 0 \amalg 1 = 1$, while $0 \amalg 0 = 0$. More generally, if $a_1, \ldots, a_n$ are *binary numbers* (i.e., either 0 or 1), their coproduct is 1 if $a_i = 1$ for at least one i, and 0 if $a_i = 0$ for all i.

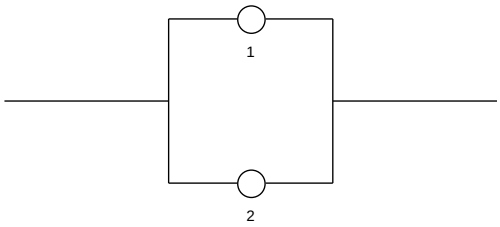

Figure 2.2: A parallel system of order 2

Example 2.1.5 *Figure 2.2 shows a reliability block diagram of a* parallel system *of order* 2. *We see that the parallel system functions if and only if at least one of its components is functioning. The structure function of this system can be written as:*

$$\phi(\boldsymbol{X}) = X_1 \amalg X_2 = 1 - (1 - X_1) \cdot (1 - X_2). \tag{2.3}$$

To see this, note that the only way the system is not functioning is that X_1 and X_2 are both zero. In this case, the right hand side of (2.3) *is clearly* 0. *For any other combination of values for X_1 and X_2, the right hand side of* (2.3) *is* 1.

Alternatively, the structure function of the parallel system can be written as:

$$\phi(\boldsymbol{X}) = \max\{X_1, X_2\}$$

since the maximum of X_1 and X_2 is 0 *if and only if $X_1 = X_2 = 0$.*

From Example 2.1.5 we see that the coproduct-operator, $\amalg$, may be though of as representing the logical *or*-operation. Thus, for the above parallel system to function components 1 *or* 2 must function.

Both the *coproduct*-expression and the *maximum*-expression can be extended to parallel systems of arbitrary orders. Thus, the structure function of a parallel system of order n can be written as:

$$\phi(\mathbf{X}) = \coprod_{i=1}^{n} X_i = \max_{1 \le i \le n} X_i.$$

2.2 Coherent systems

Every component of a binary monotone system should have some impact on the system state. More precisely, if i is a component in a system, there should ideally exist at least some state of the rest of the system where the state of component i is crucial for the system state. Before we formalize this concept further, we introduce some notation:

$$\begin{aligned}(1_i, \mathbf{x}) &:= (x_1, \ldots, x_{i-1}, 1, x_{i+1}, \ldots, x_n),\\ (0_i, \mathbf{x}) &:= (x_1, \ldots, x_{i-1}, 0, x_{i+1}, \ldots, x_n),\\ (\cdot_i, \mathbf{x}) &:= (x_1, \ldots, x_{i-1}, \cdot, x_{i+1}, \ldots, x_n),\end{aligned}$$

Definition 2.2.1 *Let (C, ϕ) be a binary monotone system, and let $i \in C$. The component i is said to be* relevant *for the system (C, ϕ) if:*

$$0 = \phi(0_i, \boldsymbol{x}) < \phi(1_i, \boldsymbol{x}) = 1 \text{ for some } (\cdot_i, \boldsymbol{x}).$$

If this is not the case, component i is said to be irrelevant *for the system.*

Relevance of components plays a very important role in reliability theory. Based on this concept we also introduce the following crucial concept.

Definition 2.2.2 *A binary monotone system (C, ϕ) is* coherent *if all its components are relevant.*

Note that a coherent system is obviously non-trivial as well, since in a trivial system *all* components are irrelevant.

One might think that all binary monotone systems we encounter should be coherent. However, coherency should be viewed a dynamic property. As parts of a system breaks, some of the remaining components may become irrelevant. Similarly, if we, e.g., as part of a reliability calculation, condition on the state of a component, the resulting system may not be coherent.

Finally, depending on how we model a system, we may encounter components which improve the performance of the system even though they are not relevant according to our definition.

Example 2.2.3 *Figure 2.3 shows a part of a large machine consisting of an electrical unit and a condensator. The condensator's job is to prevent the electrical unit from being broken due to high voltage. Viewed as a component, the condensator is irrelevant. However, this does not mean that it is not important! The condensator can prolong the lifetime of the electrical unit, and hence also the lifetime of the whole machine.*

Note that for this example, the problem can be avoided by considering the electrical unit and the condensator as one component instead of two.

We now prove the intuitively reasonable result that within the class of non-trivial monotone systems, the parallel structure is "the best" while the series system is "the worst".

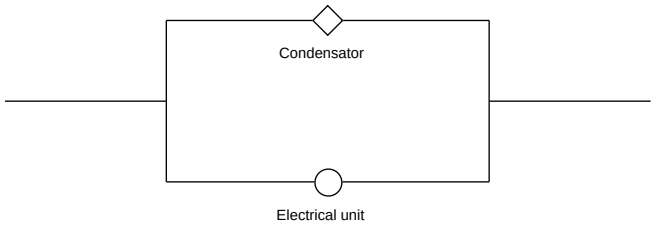

Figure 2.3: A system with an irrelevant component.

Theorem 2.2.4 *Let (C, ϕ) be a non-trivial binary monotone system of order n. Then for all $\boldsymbol{x} \in \{0,1\}^n$ we have:*

$$\coprod_{i=1}^{n} x_i \leq \phi(\boldsymbol{x}) \leq \coprod_{i=1}^{n} x_i. \tag{2.4}$$

Proof: We focus on the left-hand inequality. If $\prod_{i=1}^{n} x_i = 0$, this inequality is trivial since $\phi(\mathbf{x}) \in \{0,1\}$ for all $\mathbf{x} \in \{0,1\}^n$. If on the other hand $\prod_{i=1}^{n} x_i = 1$, we must have $\mathbf{x} = \mathbf{1}$. Since (C, ϕ) is assumed to be non-trivial, it follows by Theorem 2.1.2 that $\phi(\mathbf{1}) = 1$. Thus, the inequality is valid in this case as well. This completes the proof of the left-hand inequality. The right-hand inequality is proved in a similar fashion and is left as an exercise. □

By introducing redundancy into a system one could improve the reliability. However, there are many different ways of doing this. In particular, we may introduce redundancy at the component level or at the system level. The next result shows that redundancy at the component level is better than redundancy at the system level. Moreover, the opposite is true for series connections. In order to prove this result, we introduce the product and coproduct operators for vectors:

$$\begin{aligned} \mathbf{x} \cdot \mathbf{y} &:= (x_1 \cdot y_1, x_2 \cdot y_2, \ldots, x_n \cdot y_n), \\ \mathbf{x} \amalg \mathbf{y} &:= (x_1 \amalg y_1, x_2 \amalg y_2, \ldots, x_n \amalg y_n). \end{aligned}$$

Theorem 2.2.5 *Let (C, ϕ) be a binary monotone system of order n. Then for all $\boldsymbol{x}, \boldsymbol{y} \in \{0,1\}^n$ we have:*

(i) $\phi(\boldsymbol{x} \amalg \boldsymbol{y}) \geq \phi(\boldsymbol{x}) \amalg \phi(\boldsymbol{y})$,

(ii) $\phi(\boldsymbol{x} \cdot \boldsymbol{y}) \leq \phi(\boldsymbol{x}) \cdot \phi(\boldsymbol{y})$.

Moreover, assume that (C, ϕ) is coherent. Then equality holds in (i) for all $\boldsymbol{x}, \boldsymbol{y} \in \{0,1\}^n$ if and only if (C, ϕ) is a parallel structure. Similarly, equality holds in (ii) for all $\boldsymbol{x}, \boldsymbol{y} \in \{0,1\}^n$ if and only if (C, ϕ) is a series structure.

Proof: Since ϕ is non-decreasing and $x_i \amalg y_i \geq x_i$ for $i = 1, \ldots, n$, it follows that:

$$\phi(\mathbf{x} \amalg \mathbf{y}) \geq \phi(\mathbf{x}).$$

Similarly, we see that

$$\phi(\mathbf{x} \amalg \mathbf{y}) \geq \phi(\mathbf{y}).$$

Hence,

$$\phi(\mathbf{x} \amalg \mathbf{y}) \geq \max\{\phi(\mathbf{x}), \phi(\mathbf{y})\} = \phi(\mathbf{x}) \amalg \phi(\mathbf{y}).$$

This proves (i). The proof of (ii) is similar and is left as an exercise.

We then assume that (C, ϕ) is coherent. We shall prove the first equivalence, i.e., that equality holds in (i) for all $\mathbf{x}, \mathbf{y} \in \{0,1\}^n$ if and only if (C, ϕ) is a parallel structure.

If (C, ϕ) is a parallel system, it follows that:

$$\begin{aligned}\phi(\mathbf{x} \amalg \mathbf{y}) &= \amalg_{i=1}^{n}(x_i \amalg y_i) = \max_{1 \leq i \leq n} \big\{ \max\{x_i, y_i\} \big\} \\ &= \max\big\{ \max_{1 \leq i \leq n} x_i, \max_{1 \leq i \leq n} y_i \big\} = \phi(\mathbf{x}) \amalg \phi(\mathbf{y}),\end{aligned}$$

which proves the "if"-part of the equivalence.

Assume conversely that $\phi(\mathbf{x} \amalg \mathbf{y}) = \phi(\mathbf{x}) \amalg \phi(\mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in \{0,1\}^n$. Since (C, ϕ) is coherent it follows that for any $i \in C$ there exists a vector $(\cdot_i, \mathbf{x})$ such that:

$$\phi(1_i, \mathbf{x}) = 1 \text{ and } \phi(0_i, \mathbf{x}) = 0.$$

For this particular vector $(\cdot_i, \mathbf{x})$ we have:

$$\begin{aligned}1 = \phi(1_i, \mathbf{x}) &= \phi\big((1_i, \mathbf{0}) \amalg (0_i, \mathbf{x})\big) \\ &= \phi(1_i, \mathbf{0}) \amalg \phi(0_i, \mathbf{x}) = \phi(1_i, \mathbf{0}) \amalg 0 \\ &= \phi(1_i, \mathbf{0}).\end{aligned}$$

Since $\phi(\mathbf{0}) = 0$, it follows that:

$$\phi((x_i)_i, \mathbf{0}) = x_i \qquad x_i = 0, 1; \quad i = 1, \ldots, n.$$

Hence, for any $\mathbf{x}$, we have:

$$\begin{aligned}\phi(\mathbf{x}) &= \phi\big(((x_1)_1, \mathbf{0}) \amalg ((x_2)_2, \mathbf{0}) \amalg \ldots \amalg ((x_n)_n, \mathbf{0})\big) \\ &= \phi((x_1)_1, \mathbf{0}) \amalg \phi((x_2)_2, \mathbf{0}) \amalg \ldots \amalg \phi((x_n)_n, \mathbf{0}) \\ &= x_1 \amalg x_2 \amalg \ldots \amalg x_n = \coprod_{i=1}^{n} x_i.\end{aligned}$$

Thus, we have shown that (C, ϕ) is a parallel system which proves the "only if"-part of the equivalence. The other equivalence is proved similarly and is left as an exercise. □

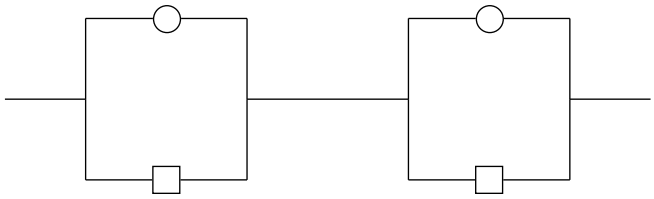

Figure 2.4: A componentwise parallel structure: "better" than systemwise parallel.

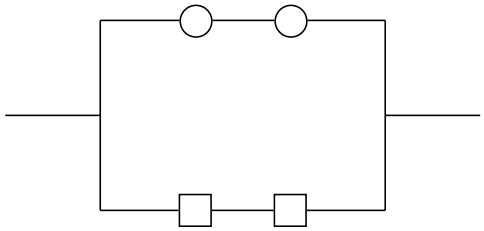

Figure 2.5: A systemwise parallel structure: "worse" than componentwise parallel.

Example 2.2.6 *To illustrate Theorem 2.2.5, let ϕ be a series system of order 2. Then the statement of the theorem is that componentwise connection of the nodes provides a better result than systemwise connection, i.e.,*

$$(x_1 \amalg y_1) \cdot (x_2 \amalg y_2) \geq (x_1 \cdot x_2) \amalg (y_1 \cdot y_2).$$

See Figures 2.4 and 2.5 for an illustration of this result. Here x_1 and x_2 represent the states of the circular nodes, while y_1 and y_2 represent the states of the square nodes.

Assume e.g., that $x_1 = y_2 = 1$ while $x_2 = y_1 = 0$. Then the componentwise parallel structure is functioning, while the systemwise parallel structure is failed. By Theorem 2.2.5 the converse will never hold true. If the systemwise parallel structure is functioning, the componentwise parallel structure must be functioning as well. Similarly, if the componentwise parallel structure is failed, then the systemwise parallel structure must be failed as well.

2.3 Dual systems

In this section we will introduce a concept of *duality*:

Definition 2.3.1 *Let ϕ be a structure function of a binary monotone system of order n. We then define the* dual structure function*, ϕ^D for all $\boldsymbol{y} \in \{0,1\}^n$ as*[3]*:*

$$\phi^D(\boldsymbol{y}) := 1 - \phi(\boldsymbol{1} - \boldsymbol{y}).$$

Furthermore, if $\boldsymbol{X}$ is the component state vector of a binary monotone system, we define the dual component state vector $\boldsymbol{X}^D$ as:

$$\boldsymbol{X}^D := (X_1^D, \ldots, X_n^D) = (1 - X_1, \ldots, 1 - X_n) = \boldsymbol{1} - \boldsymbol{X}$$

Note that the relation between ϕ and ϕ^D should be interpreted as a relation between two *functions*, while the relation between $\mathbf{X}$ and $\mathbf{X}^D$ is a relation between two stochastic vectors. Corresponding to the dual component state vector $\mathbf{X}^D$ we introduce the dual component set $C^D = \{1^D, \ldots, n^D\}$, where it is understood that the dual component i^D is functioning if the component i is failed, while i^D is failed if the component i is functioning.

By combining duality on the component level and on the system level, we see that we have the following relation between the two stochastic variables $\phi(\mathbf{X})$ and $\phi^D(\mathbf{X}^D)$:

$$\phi^D(\mathbf{X}^D) = 1 - \phi(\mathbf{1} - \mathbf{X}^D) = 1 - \phi(\mathbf{X}).$$

Hence, the dual system is functioning if and only if the original system is failed and vice versa.

Example 2.3.2 *Let ϕ be the structure function of a system of order 3 where:*

$$\phi(\boldsymbol{y}) = y_1 \amalg (y_2 \cdot y_3),$$

The dual structure function is then given by:

$$\begin{aligned}
\phi^D(\boldsymbol{y}) &= 1 - \phi(\boldsymbol{1} - \boldsymbol{y}) \\
&= 1 - (1 - y_1) \amalg ((1 - y_2) \cdot (1 - y_3)) \\
&= 1 - [1 - (1 - (1 - y_1))(1 - (1 - y_2) \cdot (1 - y_3))] \\
&= 1 - [1 - y_1 \cdot (1 - (1 - y_2) \cdot (1 - y_3))] \\
&= y_1 \cdot (y_2 \amalg y_3)
\end{aligned}$$

Example 2.3.3 *Let ϕ be the structure of a series system of order n. That is, ϕ is given by:*

$$\phi(\boldsymbol{y}) = \prod_{i=1}^{n} y_i.$$

[3]Note that we use $\mathbf{y}$ as argument instead of $\mathbf{X}$ here to emphasize that we are describing a relation between the two binary functions ϕ and ϕ^D. This should not be confused with the relation between the two random variables $\phi(\mathbf{X})$ and $\phi^D(\mathbf{X}^D)$.

The dual structure function is then given by:

$$\begin{aligned}\phi^D(\boldsymbol{y}) &= 1 - \phi(\boldsymbol{1} - \boldsymbol{y}) \\ &= 1 - \prod_{i=1}^{n}(1 - y_i) \\ &= \coprod_{i=1}^{n} y_i.\end{aligned}$$

Thus, (C^D, ϕ^D) is a parallel system of order n.

Example 2.3.4 *Let ϕ be the structure of a parallel system of order n. That is, ϕ is given by:*

$$\phi(\boldsymbol{y}) = \coprod_{i=1}^{n} y_i.$$

The dual structure function is then given by:

$$\begin{aligned}\phi^D(\boldsymbol{y}) &= 1 - \phi(\boldsymbol{1} - \boldsymbol{y}) \\ &= 1 - \coprod_{i=1}^{n}(1 - y_i) \\ &= 1 - (1 - \prod_{i=1}^{n}(1 - (1 - y_i)) \\ &= \prod_{i=1}^{n} y_i.\end{aligned}$$

Thus, (C^D, ϕ^D) is a series system of order n.

We close this section by a result showing that if we take the dual of the dual system, we get the original system back.

Theorem 2.3.5 *Let ϕ be the structure function of a binary monotone system, and let ϕ^D be the corresponding dual structure function. Then we have:*

$$(\phi^D)^D = \phi.$$

That is, the dual of the dual system is equal to the original system.

Proof: For all $\mathbf{y} \in \{0,1\}^n$ we have:

$$\begin{aligned}(\phi^D)^D(\mathbf{y}) &= 1 - \phi^D(\mathbf{1} - \mathbf{y}) \\ &= 1 - [1 - \phi(\mathbf{1} - (\mathbf{1} - \mathbf{y}))] \\ &= \phi(\mathbf{y}).\end{aligned}$$

□

2.4 Reliability of binary monotone systems

If (C, ϕ) is a binary monotone system, the *reliability* of a component $i \in C$, denoted by p_i, is defined as the probability that component i is functioning. That is, $p_i := P(X_i = 1)$, for all $i \in C$. Obviously, a component with high reliability is likely to be functioning, while a component with low reliability is more likely to fail. Note that since X_i is binary, we have:

$$E[X_i] = 0 \cdot P(X_i = 0) + 1 \cdot P(X_i = 1) = P(X_i = 1) = p_i, \text{ for all } i \in C. \quad (2.5)$$

Thus, the reliability of component i is equal to the expected value of its component state variable, X_i.

Similarly, the *reliability* of the system, denoted by h, is defined as the probability that the system is functioning. That is, $h := P(\phi = 1)$. Again, since ϕ is binary, we have:

$$E[\phi(\mathbf{X})] = 0 \cdot P(\phi(\mathbf{X}) = 0) + 1 \cdot P(\phi(\mathbf{X}) = 1) = P(\phi(\mathbf{X}) = 1) = h. \quad (2.6)$$

Thus, the reliability of the system is equal to the expected value of the structure function, $\phi(\mathbf{X})$. From this it immediately follows that the reliability of a system, at least in principle, can be calculated as:

$$h = E[\phi(\mathbf{X})] = \sum_{\mathbf{x} \in \{0,1\}^n} \phi(\mathbf{x}) P(\mathbf{X} = \mathbf{x}) \quad (2.7)$$

In the case where the component state variables are dependent, the system reliability h will depend on the full joint distribution of the component state vector. Hence, if we only know the component reliabilities, the best we can do is to establish upper and lower bounds for h. See Section 6.2 for more on this. In the remaining part of this section we instead consider the case where the component state variables can be assumed to be independent. In order to do so we introduce the vector of component reliabilities, $\mathbf{p} := (p_1, p_2, \ldots, p_n)$, and note that:

$$P(X_i = x_i) = \begin{cases} p_i & \text{if } x_i = 1, \\ 1 - p_i & \text{if } x_i = 0. \end{cases}$$

Since x_i is either 0 or 1, this probability can be written in the following more compact form:

$$P(X_i = x_i) = p_i^{x_i}(1 - p_i)^{1 - x_i}.$$

If the component state variables $X_1, X_2, \ldots, X_n$ are independent, we can use this in order to write the joint distribution of $\mathbf{X}$ in terms of $\mathbf{p}$ as follows:

$$P(\mathbf{X} = \mathbf{x}) = \prod_{i=1}^{n} P(X_i = x_i) = \prod_{i=1}^{n} p_i^{x_i}(1 - p_i)^{1 - x_i}. \quad (2.8)$$

Thus, when the component state variables are independent, we may insert (2.8) into (2.7) and get the following expression for the system reliability:

$$h = E[\phi(\mathbf{X})] = \sum_{\mathbf{x} \in \{0,1\}^n} \phi(\mathbf{x}) \prod_{i=1}^{n} p_i^{x_i}(1 - p_i)^{1 - x_i} \quad (2.9)$$

Note that in this case the reliability of the system is a function of the component reliabilities, so we may write

$$h = h(\mathbf{p}). \tag{2.10}$$

The function $h(\mathbf{p})$ is called *the reliability function* of the system.

Example 2.4.1 *Consider a series system of order n. Assuming that the component state variables are independent, the reliability of this system is given by:*

$$h(\boldsymbol{p}) = E[\phi(\boldsymbol{X})] = E[\prod_{i=1}^{n} X_i] = \prod_{i=1}^{n} E[X_i] = \prod_{i=1}^{n} p_i,$$

where the third equality follows by using that $X_1, X_2, \ldots, X_n$ are assumed to be independent.

Example 2.4.2 *Consider a parallel system of order n. Assuming that the component state variables are independent, the reliability of this system is given by:*

$$\begin{aligned} h(\boldsymbol{p}) = E[\phi(\boldsymbol{X})] = E[\coprod_{i=1}^{n} X_i] = E[1 - \prod_{i=1}^{n}(1 - X_i)] \\ = 1 - \prod_{i=1}^{n}(1 - E[X_i]) = \coprod_{i=1}^{n} E[X_i] = \coprod_{i=1}^{n} p_i, \end{aligned}$$

where the fourth equality follows by using that $X_1, X_2, \ldots, X_n$ are assumed to be independent.

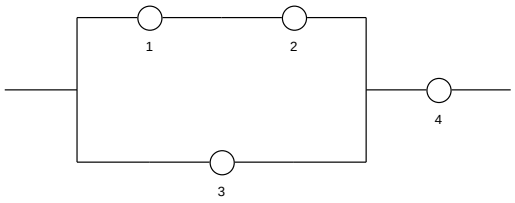

Figure 2.6: A mixed parallel and series system

Example 2.4.3 *The structure function for the system in Figure 2.6 is*

$$\phi(\boldsymbol{X}) = \Big((X_1 \cdot X_2) \amalg X_3\Big) \cdot X_4, \tag{2.11}$$

because for the system to function ((components 1 and 2 must function) or (component 3 must function)) and component 4 must function.

Alternatively, we see that components 1 and 2 are in series implying that the structure function depends on components 1 and 2 through $X_1 \cdot X_2$. Furthermore,

the part of the system containing components 1, 2 *is in parallel with component* 3 *implying that the structure function depends on components* 1, 2 *and* 3 *through* $(X_1 \cdot X_2) \amalg X_3$. *Finally, the part of the system containing components* 1, 2 *and* 3 *is in series with component* 4, *implying that the structure function is as in equation* (2.11).

Assuming that the component state variables are independent it is easy to verify, using a similar argument as we used for the structure function, that the reliability of the system is given by:

$$h(\boldsymbol{p}) = \Big((p_1 \cdot p_2) \amalg p_3\Big) \cdot p_4.$$

Theorem 2.2.5 can be extended to reliability functions as well. That is, we have the following result:

Theorem 2.4.4 *Let* $h(\boldsymbol{p})$ *be the reliability function of a monotone system* (C, ϕ) *of order* n. *Then for all* $\boldsymbol{p}, \boldsymbol{p}' \in [0,1]^n$ *we have:*

(i) $h(\boldsymbol{p} \amalg \boldsymbol{p}') \geq h(\boldsymbol{p}) \amalg h(\boldsymbol{p}')$,

(ii) $h(\boldsymbol{p} \cdot \boldsymbol{p}') \leq h(\boldsymbol{p}) \cdot h(\boldsymbol{p}')$

If in addition (C, ϕ) *is coherent, equality holds in* (i) *for all* $\boldsymbol{p}, \boldsymbol{p}' \in [0,1]^n$ *if and only if* (C, ϕ) *is a parallel structure. Similarly, equality holds in* (ii) *for all* $\boldsymbol{p}, \boldsymbol{p}' \in [0,1]^n$ *if and only if* (C, ϕ) *is a series structure.*

Proof: Note that the theorem implicitly assumes independence, since we can write $h(\mathbf{p})$. Then, for independent $\mathbf{X}, \mathbf{Y}$ (corresponding to $\mathbf{p}, \mathbf{p}'$), we have:

$$\begin{aligned} h(\mathbf{p} \amalg \mathbf{p}') - h(\mathbf{p}) \amalg h(\mathbf{p}') \\ &= E[\phi(\mathbf{X} \amalg \mathbf{Y})] - E[\phi(\mathbf{X})] \amalg E[\phi(\mathbf{Y})] \\ &= E[\phi(\mathbf{X} \amalg \mathbf{Y}) - \phi(\mathbf{X}) \amalg \phi(\mathbf{Y})], \end{aligned}$$

where the last expectation must be non-negative since by Theorem 2.2.5 we know that:

$$\phi(\mathbf{x} \amalg \mathbf{y}) - \phi(\mathbf{x}) \amalg \phi(\mathbf{y}) \geq 0,$$

for all $\mathbf{x}, \mathbf{y} \in \{0,1\}^n$. This completes the proof of (i). The proof of (ii) is similar and is left as an exercise.

We now show that equality in (i) holds for all $\mathbf{p}, \mathbf{p}' \in [0,1]^n$ if and only if (C, ϕ) is a parallel system. In order to do so, we may choose $\mathbf{p}$ and $\mathbf{p}'$ arbitrary such that $0 < p_i < 1, 0 < p_i' < 1$ for $i = 1, \ldots, n$. This implies that:

$$P(\mathbf{X} = \mathbf{x}, \mathbf{Y} = \mathbf{y}) > 0, \text{ for all } \mathbf{x} \in \{0,1\}^n \text{ and } \mathbf{y} \in \{0,1\}^n.$$

Hence, $E[\phi(\mathbf{X} \amalg \mathbf{Y}) - \phi(\mathbf{X}) \amalg \phi(\mathbf{Y})] = 0$ if and only if $\phi(\mathbf{x} \amalg \mathbf{y}) - \phi(\mathbf{x}) \amalg \phi(\mathbf{y}) = 0$ for all $\mathbf{x} \in \{0,1\}^n$ and $\mathbf{y} \in \{0,1\}^n$. By Theorem 2.2.5 this holds if and only if (C, ϕ) is a parallel system. The other equivalence is proved similarly, and left as an exercise. □

Example 2.4.5 *Consider the system* (C, ϕ) *of order* 3 *where:*

$$\phi(\boldsymbol{x}) = x_1 \cdot (x_2 \amalg x_3),$$

and where the component state variables are independent and $P(X_i = 1) = p$ *for all* $i \in C$*. Then it is easy to verify that* $h(p) = p \cdot (p \amalg p) = 2p^2 - p^3$*. From Theorem 2.4.4, it follows that for all* $0 \leq p \leq 1$*, we have:*

$$\begin{aligned} h(p \amalg p) &= 2(p \amalg p)^2 - (p \amalg p)^3 \\ &\geq h(p) \amalg h(p) = (2p^2 - p^3) \amalg (2p^2 - p^3) \end{aligned}$$

In Figure 2.7 we have plotted the reliabilities of the two systems as functions of p*. We see that indeed* $h(p \amalg p)$ *is greater than or equal to* $h(p) \amalg h(p)$ *for all* $0 \leq p \leq 1$*. In fact the inequality is strict when* $0 < p < 1$*.*

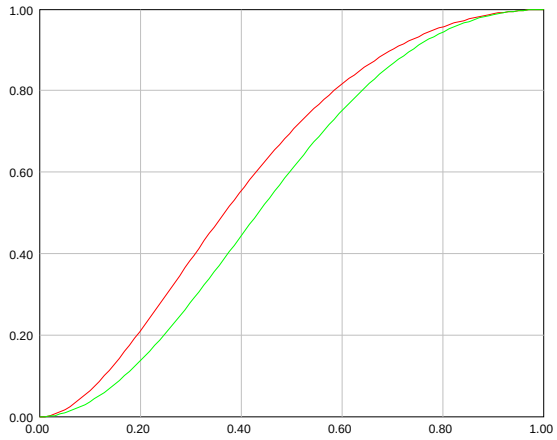

Figure 2.7: $h(p \amalg p)$ (red curve) versus $h(p) \amalg h(p)$ (green curve).

Example 2.4.6 *Consider the system* (C, ϕ) *of order* 3 *where:*

$$\phi(\boldsymbol{x}) = x_1 \amalg (x_2 \cdot x_3),$$

and where the component state variables are independent and $P(X_i = 1) = p$ *for all* $i \in C$*. In this case* $h(p) = p \amalg p^2 = p + p^2 - p^3$*. From Theorem 2.4.4, it follows that for all* $0 \leq p \leq 1$*, we have:*

$$\begin{aligned} h(p \cdot p) &= p^2 + p^4 - p^6 \\ &\leq h(p) \cdot h(p) = (p + p^2 - p^3)^2 \end{aligned}$$

In Figure 2.8 we have plotted the reliabilities of the two systems as functions of p*. We see that indeed* $h(p \cdot p)$ *is less than or equal to* $h(p) \cdot h(p)$ *for all* $0 \leq p \leq 1$*, and that the inequality is strict when* $0 < p < 1$*.*

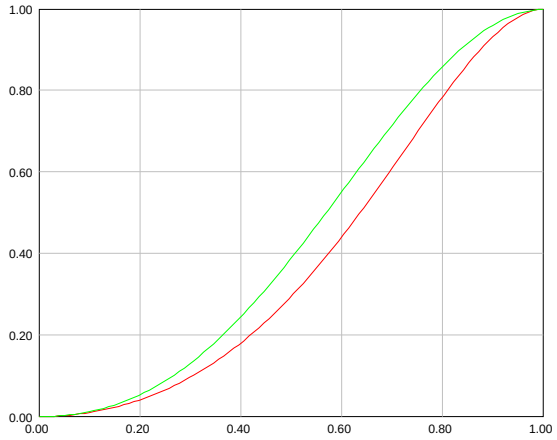

Figure 2.8: $h(p \cdot p)$ (red curve) versus $h(p) \cdot h(p)$ (green curve).

2.5 k-out-of-n systems

A k-out-of-n system is a binary monotone system (C, ϕ) where $C = \{1, \ldots, n\}$ which functions if and only if at least k out of the n components are functioning. Note that an n-out-of-n system is a series structure, while a 1-out-of-n system is a parallel structure.

We observe that in order to evaluate the structure function of a k-out-of-n system, we need to count the number of functioning components and check if this number is greater than or equal to k. The number of functioning components is simply the sum of the component state variables. Hence, the structure function can be written as:

$$\phi(\mathbf{X}) = \begin{cases} 1 & \text{if } \sum_{i=1}^{n} X_i \geq k \\ 0 & \text{otherwise.} \end{cases} \tag{2.12}$$

In order to evaluate the reliability of a k-out-of-n system it is convenient to introduce the following random variable:

$$S = \sum_{i=1}^{n} X_i.$$

Thus, S is the number of functioning components. This implies that:

$$h = P(\phi(\mathbf{X}) = 1) = P(S \geq k), \quad i = 0, 1, \ldots, n.$$

If the component state variables are independent, it is easy to derive the distribution of S. In particular, if $p_1 = \cdots = p_n = p$, S is a binomially distributed random variable. Thus, we have:

$$P(S = i) = \binom{n}{i} p^i (1-p)^{n-i}.$$

Hence, the reliability of the system is given by:

$$h(\mathbf{p}) = h(p) = P(S \geq k) = \sum_{i=k}^{n} \binom{n}{i} p^i (1-p)^{n-i}$$

In the general case with unequal component reliabilities, an explicit analytical expression for the distribution of S is not so easy to derive. Still this distribution can be calculated efficiently using generating functions. See Exercise 2.3.3.

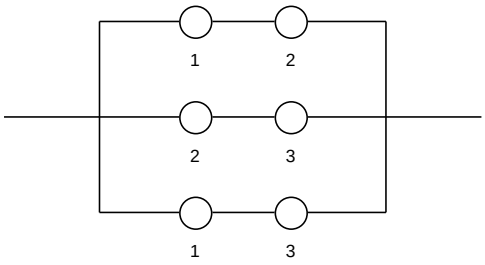

Figure 2.9: A reliability block diagram of a 2-out-of-3 system.

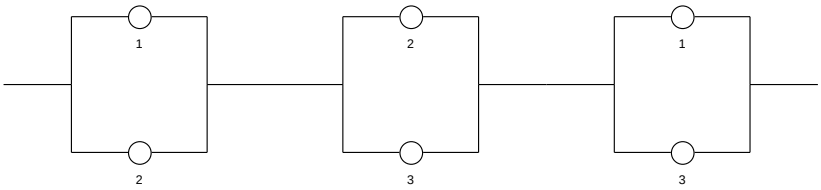

Figure 2.10: An alternative reliability block diagram of a 2-out-of-3 system.

A k-out-of-n system can represented by a reliability block diagram in several ways. The first type of diagram is shown for a 2-out-of-3 system in Figure 2.9. Here, we consider a parallel connection of several series structures. Each of the series structures corresponds to a minimal way in which a signal can pass through the system, i.e. any way two out of three components are functioning. The second type of diagram is shown for the same 2-out-of-3 system in Figure 2.10. In this case, we draw a series connection of several parallel structures. Each of the parallel structures corresponds to a minimal way to make the system fail. Again, this happens if any two out of the three components are not working.

2.6 Exercises

Exercise 2.1: What is the reliability function of a series system of order n where the component states are assumed to be independent?

Exercise 2.2: Consider the parallel structure of order 2 in Figure 2.2. Assume that the component states are independent.

a) If you know that $p_1 = P(X_1 = 1) = 0.5$ and $p_2 = P(X_2 = 1) = 0.7$, what is the reliability of the parallel system?

b) What is the system reliability if $p_1 = 0.9$ and $p_2 = 0.1$?

c) Can you give an interpretation of these results?

Exercise 2.3: Consider a binary monotone system (C, ϕ), where the component set is $C = \{1, \ldots, 4\}$ and where ϕ is given by:

$$\phi(\mathbf{X}) = X_1 \cdot X_2 \cdot (X_3 \amalg X_4).$$

a) Draw a reliability block diagram of this system.

b) Assume that the components in the system are independent. What is the corresponding reliability function?

Exercise 2.4: Consider the system shown in Figure 2.11.

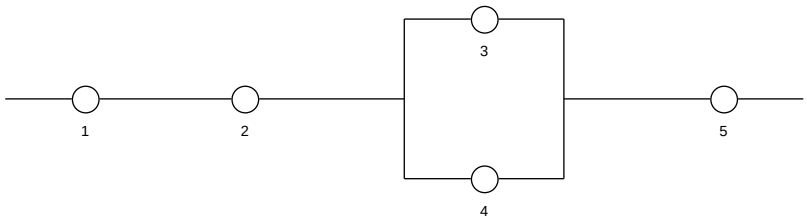

Figure 2.11: A mixed parallel and series system.

a) What is the structure function of this system?

b) Assume that the components in the system are independent. What is the corresponding reliability function?

Exercise 2.5: There are 8 different coherent systems of order less than or equal to 3 (not counting permutations in the numbering of components). What are they?

(*Hint:* Here are a couple to get you started $\phi(\mathbf{X}) = X_1 \cdot X_2 \cdot X_3$, $\phi(\mathbf{X}) = X_1 \amalg X_2$.)

Exercise 2.6: Consider a monotone system (C, ϕ) of order n, and let $A \subset C$ the set of irrelevant components. Furthermore, let $(C \setminus A, \phi')$, be a binary

monotone system of order $m = n - |A|$, where ϕ' is a binary non-decreasing function defined for all m-dimensional binary vectors $\mathbf{x} \in \{0,1\}^m$ such that:

$$\phi'(\mathbf{x}) = \phi(\mathbf{1}^A, \mathbf{x}^{\bar{A}})$$

Show that $(C \setminus A, \phi')$ is coherent.

Exercise 2.7: Prove the right-hand inequality in Theorem 2.2.4.

Exercise 2.8: Prove item (ii) in Theorem 2.2.5 as well as the equivalence which is not proved in the text.

Exercise 2.9: Prove that the dual structure of a k-out-of-n structure is an $(n-k+1)$-out-of-n structure.

Exercise 2.10: Let S be a stochastic variable with values in $\{0,1,\ldots,n\}$. We then define the *generating function* of S as:

$$G_S(y) = E[y^S] = \sum_{s=0}^{n} y^s P(S=s). \tag{2.13}$$

a) Explain why $G_S(y)$ is a polynomial, and give an interpretation of the coefficients of this polynomial.
b) Let T be another non-negative integer valued stochastic variable with values in $\{0,1,\ldots,m\}$ which is independent of S. Show that:

$$G_{S+T}(y) = G_S(y) \cdot G_T(y). \tag{2.14}$$

c) Let $X_1,\ldots,X_n$ be independent binary variables with $P(X_i=1) = p_i$ and $P(X_i=0) = 1-p_i = q_i$, $i = 1,\ldots,n$. Show that:

$$G_{X_i}(y) = q_i + p_i y, \quad i = 1,\ldots,n. \tag{2.15}$$

d) Introduce:

$$S_j = \sum_{i=1}^{j} X_i, \quad j = 1,2,\ldots,n,$$

and assume that we have computed $G_{S_j}(y)$. Thus, all the coefficients of $G_{S_j}(y)$ are known at this stage. We then compute:

$$G_{S_{j+1}}(y) = G_{S_j}(y) \cdot G_{X_{j+1}}(y).$$

How many algebraic operations (addition and multiplication) will be needed to complete this task?

e) Explain how generating functions can be used in order to calculate the reliability of a k-out-of-n system. What can you say about the order of this algorithm[4]?

[4]By the *order* of an algorithm we mean a suitable function of the size of the problem such that the number of basic operations needed to complete the calculations grows (approximately) proportionally to this function. In the context of reliability calculations the size of the problem will typically be the number of components in the system.

Chapter 3

Basic reliability calculation methods

In this chapter we will present various methods for simplifying reliability calculations by structuring the problems in certain ways. The first method involves conditioning on the state of a given component. This enables us to decompose the system into simpler building blocks. Furthermore, will see that the structure function of monotone systems can be represented via its so-called paths and cuts; roughly, this is the components which ensure that the system functions or fails. Another useful way to represent monotone systems is via modules. The idea here is to divide a complicated system into parts which may then be analysed separately. Finally, we will comment briefly on how to include time into this framework by looking at dynamic system analysis.

3.1 Pivotal decompositions

Sometimes reliability calculations can be simplified by dividing the problem into two simpler problems. By considering these simpler problems and combining the results afterwards, even very complex systems can be analysed successfully. The following result shows how a structure and reliability function of order n can be expressed via two corresponding functions of order $n-1$. This is called *pivotal decomposition*, and allows us to reduce the order of the system at hand. This is particularly useful when computing the exact system reliability, see Chapter 4:

Theorem 3.1.1 *Let (C, ϕ) be a binary monotone system of order n. Then we have:*

$$\phi(\boldsymbol{x}) = x_i\phi(1_i, \boldsymbol{x}) + (1 - x_i)\phi(0_i, \boldsymbol{x}), \quad i = 1, 2, \ldots, n. \tag{3.1}$$

Similarly, for the reliability function of a monotone system where the component

state variables are independent, we have

$$h(\boldsymbol{p}) = p_i h(1_i, \boldsymbol{p}) + (1 - p_i)h(0_i, \boldsymbol{p}), \quad i = 1, 2, \ldots, n. \tag{3.2}$$

Proof: Let $i \in C$. Since x_i is binary, we consider two cases: $x_i = 1$ and $x_i = 0$. If $x_i = 1$, then the right-hand side of (3.1) becomes $\phi(1_i, \mathbf{x})$. Hence (3.1) holds in this case. On the other hand, if $x_i = 0$, the right-hand side of (3.1) becomes $\phi(0_i, \mathbf{x})$, so (3.1) holds in this case as well. Equation (3.2) is proved by replacing the vector $\mathbf{x}$ by the stochastic vector $\mathbf{X}$ in (3.1), and taking the expected value on both sides of this equation. □

Note that (3.1) and (3.2) can alternatively be written respectively as:

$$\phi(\mathbf{x}) = \phi(0_i, \mathbf{x}) + [\phi(1_i, \mathbf{x}) - \phi(0_i, \mathbf{x})]x_i, \quad i = 1, \ldots n, \tag{3.3}$$

$$h(\mathbf{p}) = h(0_i, \mathbf{p}) + [h(1_i, \mathbf{p}) - h(0_i, \mathbf{p})]p_i \quad i = 1, \ldots n. \tag{3.4}$$

From these expressions it follows that ϕ is a linear function of x_i, while h is a linear function of p_i, $i = 1, \ldots, n$. A function which is linear in each argument is said to be *multilinear*.

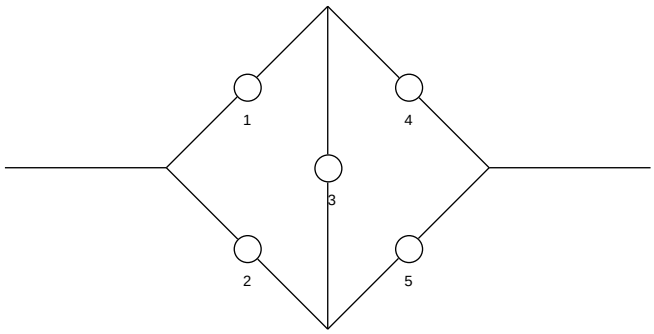

Figure 3.1: A bridge structure.

Example 3.1.2 *Let* (C, ϕ) *be the* bridge structure *shown in Figure 3.1, where* $C = \{1, \ldots, 5\}$. *In order to derive the structure function of this system, we note that:*

$\phi(1_3, \boldsymbol{X})$ = *The state of the system given that component* 3 *is functioning,*

$\phi(0_3, \boldsymbol{X})$ = *The state of the system given that component* 3 *is failed.*

Given that component 3 *is functioning, the system becomes a series connection of two parallel systems, the first one is the parallel system of component* 1 *and*

2, while the second one is the parallel system of component 4 *and* 5*. Hence, we get that:*

$$\phi(1_3, \boldsymbol{X}) = (X_1 \amalg X_2) \cdot (X_4 \amalg X_5).$$

If on the other hand, component 3 *is failed, the system becomes a parallel connection of two series systems, the first one is the series system of component* 1 *and* 4*, while the second one is the series system of component* 2 *and* 5*. From this it follows that:*

$$\phi(0_3, \boldsymbol{X}) = (X_1 \cdot X_4) \amalg (X_2 \cdot X_5).$$

By Theorem 3.1.1 it follows that ϕ *can be written as:*

$$\phi(\boldsymbol{X}) = X_3 \cdot \phi(1_3, \boldsymbol{X}) + (1 - X_3) \cdot \phi(0_3, \boldsymbol{X}).$$

Combining all this we get that ϕ *is given by:*

$$\phi(\boldsymbol{X}) = X_3 \cdot (X_1 \amalg X_2)(X_4 \amalg X_5) + (1 - X_3) \cdot (X_1 \cdot X_4 \amalg X_2 \cdot X_5). \tag{3.5}$$

Assuming that the component state variables are independent, we get that:

$$\begin{aligned} h(\boldsymbol{p}) &= E[\phi(\boldsymbol{X})] \\ &= E[X_3(X_1 \amalg X_2)(X_4 \amalg X_5) + (1 - X_3)(X_1X_4 \amalg X_2X_5)] \\ &= p_3(p_1 \amalg p_2)(p_4 \amalg p_5) + (1 - p_3)(p_1p_4 \amalg p_2p_5) \end{aligned}$$

A consequence of Theorem 3.1.1 is that for coherent systems, the reliability function is strictly increasing if the reliabilities are strictly between 0 and 1:

Theorem 3.1.3 *Let* $h(\boldsymbol{p})$ *be the reliability function of a binary monotone system* (C, ϕ) *of order* n*, and assume that* $0 < p_j < 1$ *for* $j \in C$*. If component* i *is relevant, then* $h(\boldsymbol{p})$ *is strictly increasing in* p_i*.*

Proof: From equation (3.2), it follows that

$$\begin{aligned} \frac{\partial h(\mathbf{p})}{\partial p_i} &= h(1_i, \mathbf{p}) - h(0_i, \mathbf{p}) \qquad (3.6) \\ &= E[\phi(1_i, \mathbf{X})] - E[\phi(0_i, \mathbf{X})] = E[\phi(1_i, \mathbf{X}) - \phi(0_i, \mathbf{X})] \\ &= \sum_{(\cdot_i, \mathbf{x}) \in \{0,1\}^{n-1}} \big[\phi(1_i, \mathbf{x}) - \phi(0_i, \mathbf{x})\big] P\big((\cdot_i, \mathbf{X}) = (\cdot_i, \mathbf{x})\big) \end{aligned}$$

Since (C, ϕ) is a monotone system, ϕ is non-decreasing in each argument, and hence all the terms in this sum are non-negative. We then assume that

component i is relevant. That is, there exists $(\cdot_i, \mathbf{y})$ such that $\phi(1_i, \mathbf{y}) - \phi(0_i, \mathbf{y}) = 1$. Moreover, since we have assumed that $0 < p_j < 1$ for $j \in C$ we have:

$$P\big((\cdot_i, \mathbf{X}) = (\cdot_i, \mathbf{y})\big) = \prod_{j=1, j\neq i}^{n} p_j^{y_j}(1-p_j)^{1-y_j} > 0$$

Thus, at least one of the terms in the last sum in (3.6) is positive which implies that

$$\frac{\partial h(\mathbf{p})}{\partial p_i} > 0.$$

Hence, we conclude that $h(\mathbf{p})$ is strictly increasing in each p_i. □

Note that if (C, ϕ) is *coherent* and $0 < p_j < 1$ for $j \in C$, it follows from Theorem 3.1.3 that h is strictly increasing in *each argument*, since in this case, all components are relevant.

3.2 Representation of binary monotone systems by paths and cuts

Now, we will look at a way to represent monotone systems via certain subsets of the component set, called path and cut sets. This turns out to be very useful in order to compute the exact system reliability, which will be done in Chapter 4.

In the following, we consider deterministic vectors $\mathbf{x}, \mathbf{y}$. We say that a vector $\mathbf{y}$ is smaller than another vector $\mathbf{x}$, that is $\mathbf{y} < \mathbf{x}$ if $y_i \leq x_i$ for $i = 1, \ldots, n$ and $y_i < x_i$ for at least one i.

Notation 3.2.1 *For any binary monotone system $(C, \phi\}$ of order n, and vector $\boldsymbol{x} \in \{0,1\}^n$ the component set C can be divided into two subsets*

$$C_0(\boldsymbol{x}) = \{i : x_i = 0\}, \qquad C_1(\boldsymbol{x}) = \{i : x_i = 1\}.$$

Thus, if $\mathbf{X} = \mathbf{x}$, the subset $C_1(\mathbf{x})$ denotes the set of components which are functioning, while $C_0(\mathbf{x})$ denotes the set of components which are failed.

Definition 3.2.2 *Let (C, ϕ) be a binary monotone system.*

- *A vector $\boldsymbol{x}$ is a* path vector *if and only if $\phi(\boldsymbol{x}) = 1$. The corresponding* path set *is $C_1(\boldsymbol{x})$.*
- *A* minimal path vector *is a path vector, $\boldsymbol{x}$, such that $\boldsymbol{y} < \boldsymbol{x}$ implies that $\phi(\boldsymbol{y}) = 0$. The corresponding* minimal path set *is $C_1(\boldsymbol{x})$.*
- *A vector $\boldsymbol{x}$ is a* cut vector *if and only if $\phi(\boldsymbol{x}) = 0$. The corresponding* cut set *is $C_0(\boldsymbol{x})$.*

- *A* minimal cut vector *is a cut vector,* $\boldsymbol{x}$*, such that* $\boldsymbol{x} < \boldsymbol{y}$ *implies that* $\phi(\boldsymbol{y}) = 1$*. The corresponding* minimal cut set *is* $C_0(\boldsymbol{x})$*.*

Thus, a path vector is a vector $\mathbf{x}_1$ such that the system is functioning if $\mathbf{X} = \mathbf{x}_1$. A cut vector is a vector $\mathbf{x}_0$ such that the system is failed if $\mathbf{X} = \mathbf{x}_0$. A minimal path vector is a path vector which cannot be decreased in any way and still be a path vector. Similarly, a minimal cut vector is a cut vector which cannot be increased in any way and still be a cut vector. A minimal path set is a minimal set of components which ensure that the system is working if these components are working. A minimal cut set is a minimal set of components such that if these are sabotaged, the system will be sabotaged.

Example 3.2.3 *Consider the bridge structure in Figure 3.2. To find the minimal path sets, look at Figure 3.2 and try to locate paths where it is possible to send a signal through the system. For instance, one such path is through components* 1 *and* 4*. Hence,* $\{1, 4\}$ *is a path set. We then see that this set is in fact minimal, because if only components* 1 *and* 4 *are functioning, and one of them fails, the system will fail as well. Similarly, we can argue for e.g. the path through components* 2 *and* 5*. Continuing like this, we find that the minimal path sets are:*

$$P_1 = \{1, 4\}, \quad P_2 = \{2, 5\}, \quad P_3 = \{1, 3, 5\} \text{ and } P_4 = \{2, 3, 4\}.$$

(Verify that these are in fact minimal!)

To find the minimal cut sets, look at Figure 3.2 and see what combination of components not functioning will sabotage the whole system. For example, if components 1 *and* 2 *are not functioning, the system is sabotaged because there is no way for a signal to make it past the initial point. The set* $\{1, 2\}$ *is in fact a minimal cut set, because if all other components are functioning except* 1 *and* 2*, and one of them is fixed, the system will start to function. A similar argument shows that* $\{4, 5\}$ *is a minimal cut set. Continuing in the same way, we find that the minimal cut sets are:*

$$K_1 = \{1, 2\}, \quad K_2 = \{4, 5\}, \quad K_3 = \{1, 3, 5\} \text{ and } K_4 = \{2, 3, 4\}.$$

(Make sure you can explain why all of these are minimal cut sets!)

Now, we are ready to explain how structures can be represented via their minimal path and cut sets. In order to do so, we consider a binary monotone system (C, ϕ), and let $P_j \subseteq C$ be the jth minimal path set of this system. Then we can define the following binary monotone system (P_j, ρ_j), where $\rho_j = \rho_j(\mathbf{x}^{P_j})$ is given by:

$$\rho_j(\mathbf{x}^{P_j}) = \prod_{i \in P_j} x_i. \tag{3.7}$$

(P_j, ρ_j) is referred to as the jth *minimal path series structure* of the system, $j = 1, \ldots, p$. From this, we find that:

$$\phi(\mathbf{x}) = \coprod_{j=1}^{p} \rho_j(\mathbf{x}^{P_j}) = \coprod_{j=1}^{p} \prod_{i \in P_j} x_i. \tag{3.8}$$

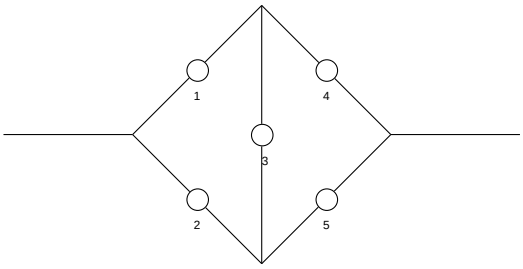

Figure 3.2: A bridge structure.

That is, the system (C, ϕ) can be represented as a parallel structure of the minimal path series structures. See Figure 3.2 for an illustration of how the bridge structure in Example 3.2.3 can be represented as a parallel structure of its minimal path series structures. The reason for this is that the system functions if and only if at least one of the minimal path series structures are functioning.

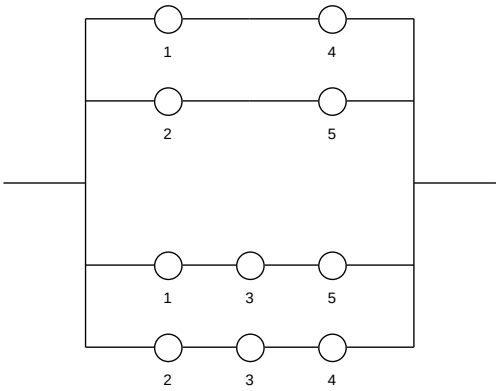

Figure 3.3: The reliability block diagram of a bridge structure in Example 3.2.3 represented as a parallel structure of its minimal path series structures.

Similarly, let $K_j \subseteq C$ be the jth minimal cut set of this system. Then we can define the following binary monotone system (K_j, κ_j), where $\kappa_j = \kappa_j(\mathbf{x}^{K_j})$ is given by:

$$\kappa_j(\mathbf{x}^{K_j}) = \coprod_{i \in K_j} x_i. \tag{3.9}$$

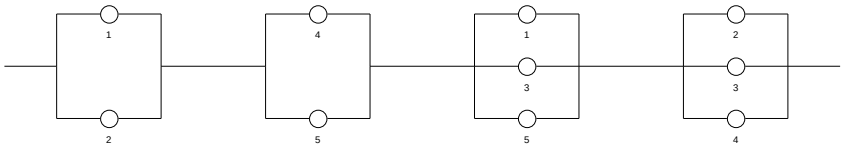

Figure 3.4: The reliability block diagram of a bridge structure in Example 3.2.3 represented as a series structure of its minimal cut parallel structures.

(K_j, κ_j) is referred to as the jth *minimal cut parallel structure* of the system, $j = 1, \ldots, k$. From this, we find that:

$$\phi(\mathbf{x}) = \prod_{j=1}^{k} \kappa_j(\mathbf{x}^{K_j}) = \prod_{j=1}^{k} \coprod_{i \in K_j} x_i. \tag{3.10}$$

That is, the system (C, ϕ) can be represented as a series structure of the minimal cut parallel structures. See Figure 3.4. The reason for this is that for the system to be failed, it suffices for one of the minimal cut parallel structures is failed.

These representations of the structure function are very useful, and will be applied in Chapter 4 to compute the exact system reliability. However, a major challenge is that computing the minimal path and cut sets for a large system is computationally very time consuming. Therefore, it is important to derive fast algorithms for finding the minimal path and cut sets. This is an important challenge in applied reliability theory, but beyond the scope of these notes.

The proof of the following theorem is left as an exercise:

Theorem 3.2.4 *Let (C, ϕ) be a binary monotone system, and let (C^D, ϕ^D) be its dual. Then the following statements hold:*

(i) $\boldsymbol{x}$ is a path vector (alternatively, cut vector) for (C, ϕ) if and only if $\boldsymbol{x}^D$ is a cut vector (path vector) for (C^D, ϕ^D).

(ii) A minimal path set (alternatively, cut set) for (C, ϕ) is a minimal cut set (path set) for (C^D, ϕ^D).

3.3 Modules of monotone systems

Sometimes, it is possible to divide a monotone system into subsystems which are analysed separately before finally analysing how these subsystems are connected.

Let $A \subseteq C$. Then, the complement set of A, i.e., $C \setminus A$, is denoted by $\bar{A}$. We have the following formal definition of a module:

Definition 3.3.1 *Let (C, ϕ) be a monotone system, and $A \subseteq C$. The monotone system (A, χ) is a* module *of (C, ϕ) if and only if the structure function ϕ can be written as:*

$$\phi(\boldsymbol{x}) = \psi(\chi(\boldsymbol{x}^A), \boldsymbol{x}^{\bar{A}}), \quad \text{for all } \boldsymbol{x} \in \{0,1\}^n,$$

where ψ is a monotone structure function. The set A is called a modular set *of (C, ϕ).*

The structure function χ expresses the state of the module as a function of the states of the components in the modular set, while ψ expresses how the state of the system depends on the state of the module as well as on the states of componentsoutside of the module. You can think of a module as a "large component" which can be circled in and separated out from the rest of the system. More precisely, a module consists of a subset of the component set which is such that the structure function of the system depends on the states of the components within this set only through the structure function of the module. Thus, in order to determine the state of the system, we only need to know the state of the module, not the states of the individual components within the module.

Definition 3.3.2 *A* modular decomposition *of a monotone system (C, ϕ) is a set of modules $\{(A_j, \chi_j)\}_{j=1}^r$ connected by a binary monotone organisation structure function ψ. The following conditions must be satisfied:*

(i) $C = \bigcup_{j=1}^r A_j$, *where* $A_j \cap A_k = \emptyset$ *for* $j \neq k$.

(ii) $\phi(\boldsymbol{X}) = \psi[\chi_1(\boldsymbol{X}^{A_1}), \ldots, \chi_r(\boldsymbol{X}^{A_r})]$.

We observe that a modular decomposition is a disjoint partition of the component set into modules such that the structure function of the whole system is a function of the structure functions of these modules.

Example 3.3.3

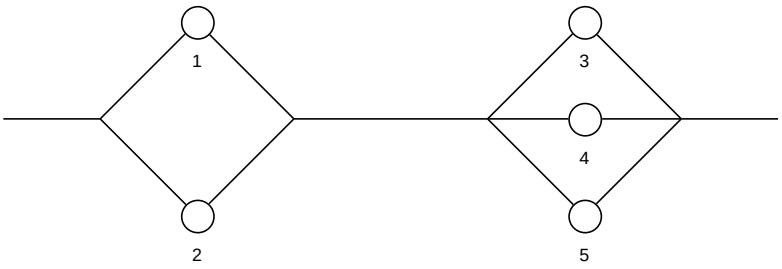

Figure 3.5: A monotone system which can be modularly decomposed.

In order to find a modular decomposition for the system in Figure 3.5, we can let $A_1 = \{1, 2\}$, $A_2 = \{3, 4, 5\}$,

$$\chi_1(\boldsymbol{X}^{A_1}) = X_1 \amalg X_2, \ \chi_2(\boldsymbol{X}^{A_2}) = X_3 \amalg X_4 \amalg X_5$$

and

$$\psi(\chi_1(\boldsymbol{X}^{A_1}), \chi_2(\boldsymbol{X}^{A_2})) = \chi_1(\boldsymbol{X}^{A_1}) \cdot \chi_2(\boldsymbol{X}^{A_2}).$$

If the component state variables are independent, we may calculate the reliability of the system in a two-step process. In the first step we compute the reliability of the modules. Then in the second step we treat the modules as components and compute the system relaibility by using the reliabilities of the modules as input. Thus, whenever one can identify modules in a system, this can be utilized in order to simplify and speed up the reliability calculations.

Modular decompositions can also be used in cases where it is not possible to compute the system reliability exactly as it is too time consuming. In such cases, we may have to resort to finding upper and lower bounds for the system reliability. It turns out that modules may be used to improve bounds. We will return to this in Section 6.2.

3.4 Dynamic system analysis

So far, we have only considered components and systems at one specific time. Now, we will also consider the development over time. Let $t \geq 0$, $i = 1, \ldots, n$:

$$X_i(t) = \text{ the state of component } i \text{ at time } t, \tag{3.11}$$

Here, $X_i(t)$ is a random variable (corresponding to X_i of the previous sections). Hence, the development of the state of component i over time, $\{X_i(t)\}_{t\geq 0}$, is a stochastic process. We assume that the stochastic processes $\{X_i(t), t \geq 0\}_{i=1}^n$ are independent. This corresponds to assuming that the components are independent. Also, let

$$\phi(\mathbf{X}(t)) = \text{ the state of the system at time } t. \tag{3.12}$$

As for the components, $\phi(\mathbf{X}(t))$ is a random variable, and the development of the system state over time, $\{\phi(\mathbf{X}(t))\}_{t\geq 0}$ is a stochastic process. We also introduce

$$\begin{aligned} p_i(t) = P(X_i(t) = 1) &= \text{ the reliability of component } i \text{ at time } t, \\ h(\mathbf{p}(t)) = P(\phi(\mathbf{X}(t) = 1)) &= \text{ the reliability of the system at time } t. \end{aligned} \tag{3.13}$$

In the following, we do not consider the possibility of repairing components, and let (for $i = 1, \ldots, n$)

$$\begin{aligned} T_i &= \text{ the lifetime of component } i, \\ S &= \text{ the lifetime of the system.} \end{aligned} \tag{3.14}$$

Now, let T_i have cumulative distribution function F_i, $i = 1, \ldots, n$ and S cumulative distribution function G. We assume that the F_i's are known, and would like to determine G. Then we have the following relations:

$$p_i(t) = P(X_i(t) = 1) = P(T_i > t) = 1 - F_i(t) =: \bar{F}_i(t), \tag{3.15}$$

and

$$G(t) = 1 - P(S > t) = P(\phi(\mathbf{X}(t)) = 0) = 1 - h(\mathbf{p}(t)). \tag{3.16}$$

Then we have the following theorem:

Theorem 3.4.1 *For a monotone system (C, ϕ) with minimal path sets $P_1, \ldots, P_p$ and minimal cut sets $K_1, \ldots, K_k$ we have:*

$$S = \begin{cases} \max_{1 \le j \le p} \min_{i \in P_j} T_i \\ \min_{1 \le j \le k} \max_{i \in K_j} T_i \end{cases}$$

Proof: To prove the first equality: The lifetime of the system equals the lifetime of the minimal path series structure which lives the longest (from the definition of minimal path series structures). The lifetime of the longest living minimal path series structure equals the lifetime of the shortest living component in this path set.

The second equality can be proved similarly. This is left as an exercise.

□

3.5 Exercises

Exercise 3.1: Prove equation (3.2) in another way than what is done in the proof of Theorem 3.1.1.

(*Hint:* Condition on the state of component i.)

Exercise 3.2: Prove Theorem 3.2.4.

Exercise 3.3: Find all of the path and cut sets of the brigde structure in Figure 3.2.

Exercise 3.4: Find the representations via the minimal path sets and the minimal cut sets of the following systems:

(i) The 2-out-of-3 system.

(ii) The 3-out-of-4 system.

(iii) The series system of 3 components.

(iv) The parallel system of 4 components.

Exercise 3.5: Consider a coherent structure (C, ϕ) with minimal path sets $P_1, \ldots, P_p$ and minimal cut sets $K_1, \ldots, K_k$. Prove that

$$\bigcup_{j=1}^{p} P_j = C = \bigcup_{j=1}^{k} K_j.$$

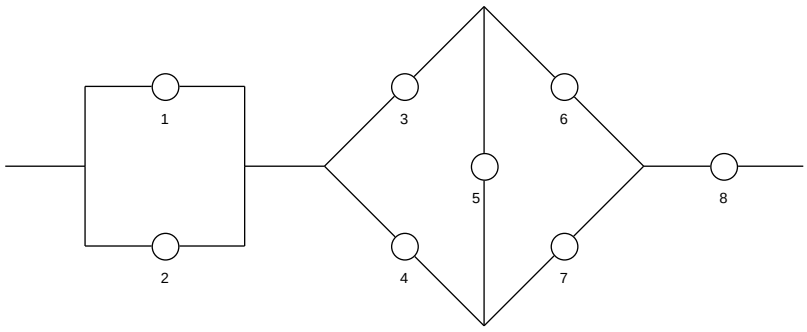

Figure 3.6: A series connection of a parallel structure, a bridge structure and a single component.

Exercise 3.6: Find the minimal path sets and the minimal cut sets of the system in Figure 3.6. Find two different expressions for the structure function of this system.

Exercise 3.7: Show that if (C, ϕ) is a bridge structure (see Figure 3.1), then (C^D, ϕ^D) is a bridge structure as well. [*Hint*: Compare the minimal path and cut sets of (C, ϕ).]

Exercise 3.8: Let (A, χ) be a module of (C, ϕ). Assume that $\mathbf{x}_1$ and $\mathbf{x}_0$ are such that $\chi(\mathbf{x}_1^A) = 1$ and $\chi(\mathbf{x}_0^A) = 0$. Prove that

$$\phi(\mathbf{x}_1^A, \mathbf{x}) \equiv \phi(\mathbf{1}^A, \mathbf{x}) \text{ and } \phi(\mathbf{x}_0^A, \mathbf{x}) \equiv \phi(\mathbf{0}^A, \mathbf{x}).$$

Exercise 3.9: Find all the modules of the following structure function:

$$\phi(\mathbf{X}) = (X_1(X_2 \amalg X_3)) \amalg (X_4 \amalg X_5).$$

Exercise 3.10: What are the modules of a k-out-of-n structure where $1 < k < n$? Give a reason for your answer.

Does this mean that a k-out-of-n system is well suited for modular decomposition?

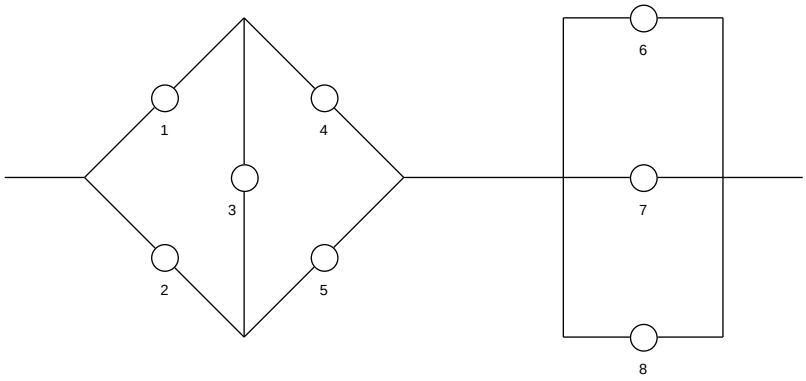

Figure 3.7: A system suited for modular decompostition.

Exercise 3.11: Consider the system shown in Figure 3.7.

a) Find the structure function of the system by dividing the system into modules.

b) Find the structure function of the system *without* dividing the system into modules.

c) Compare the computational efforts in a) and b).

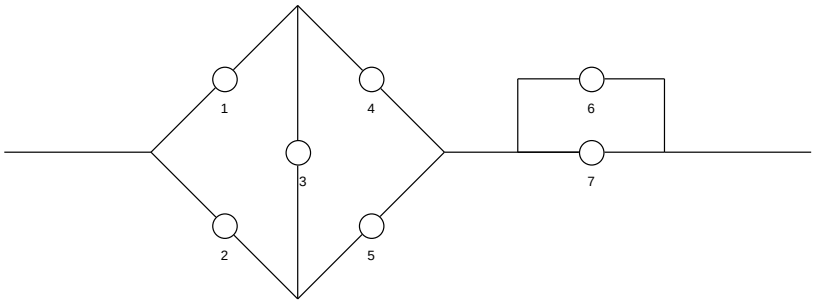

Figure 3.8: A mixed bridge and parallel system.

Exercise 3.12: Consider the system shown in Figure 3.8.

a) What is the structure function of this system?

b) Assume that all of the component states are independent. What is the reliability function of the system?

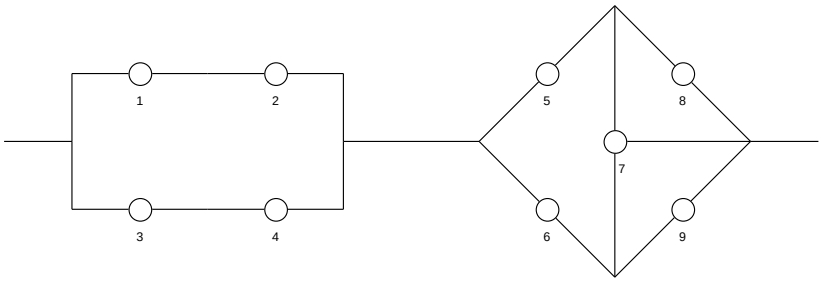

Figure 3.9: A twist on a mixed bridge and parallel system.

Exercise 3.13: Consider the system shown in Figure 3.9.
a) What is the structure function of this system.

b) Assume that all of the component states are independent. What is the reliability function of the system?

Exercise 3.14: Prove the second equality of Theorem 3.4.1.

Exercise 3.15: Consider a binary monotone system (C, ϕ) with minimal path sets $P_1, \ldots, P_p$ and minimal cut sets $K_1, \ldots, K_k$ and let f be a non-negative function defined over the positive integers $\{1, \ldots, n\}$. Prove that

$$\min_{1 \leq j \leq k} \max_{i \in K_j} f(i) = \max_{1 \leq j \leq p} \min_{i \in P_j} f(i).$$

Chapter 4

Exact computation of the reliability of binary monotone systems

In this chapter, we will look at various methods for computing the exact reliability of a binary monotone system. In the general case this is a *hard* problem in the sense that the running time of all known algorithms grows exponentially in the size of the problem, i.e., in the order of the system.

If the running time of an algorithm grows approximately proportionally to a positive function $f(n)$, where n is measuring the size of the problem (i.e, the number of components in a binary monotone system), we say that the algorithm is of *order* $O(f(n))$. More formally, let $t(n)$ denote the worst case running time of the algorithm as a function of the size n. Then the order of the algorithm is $O(f(n))$ if and only if there exists a positive constant M and a positive integer n_0 such that:

$$t(n) \leq Mf(n), \text{ for all } n \geq n_0.$$

If f is a polynomial in n, we say that the algorithm is a *polynomial time* algorithm, while if f is an exponential function of n, we say that the algorithm is an *exponential time* algorithm.

In computational complexity theory, NP (for nondeterministic polynomial time) is a complexity class used to describe certain types of problems. NP contains many important problems, the hardest of which are called NP-complete problems. The most important open question in complexity theory is whether polynomial time algorithms actually exist for solving NP-complete problems. It is widely believed that this is not the case. The class of *NP-hard* problems is a class of problems that are, informally, *at least as hard as the hardest problems in NP*. A comprehensive treatment of complexity is beyond the scope of this book. See Garey and Johnson [17] for more on this.

Unfortunately, the problem of computing the reliability of a binary monotone system is known to be NP-hard in the general case. See Rosenthal [40] and Ball [2] for details. Thus, all known algorithms for analytical reliability calculations typically has a computational complexity which grows exponentially with the order of the system. Still for moderately sized systems it is possible to calculate the reliability using standard methods. Moreover, for certain classes of systems faster algorithms are available.

In Exercise 2.6 we saw that the reliability of a k-out-of-n system can be calculated in polynomal time using generating functions, even when the components have unequal reliability. In fact this particular algorithm is of order $O(n^2)$, which is very fast compared to an exponential time algorithm. From this result one might think that it would be easy to obtain similar performance for a larger class of systems. The following example shows that this is not the case.

Example 4.0.1 *A* threshold system *is a binary monotone system (C, ϕ), where the structure function has the following form:*

$$\phi(\boldsymbol{x}) = I(\sum_{i=1}^{n} a_i x_i \geq b),$$

where $a_1, \ldots, a_n$ and b are non-negative real numbers, and $I(\cdot)$ denotes the indicator function, i.e., a function defined for any event A which is 1 *if A is true and zero otherwise.*

We observe that if (C, ϕ) is a threshold system where $a_1 = \cdots = a_n = 1$ and $b = k$, (C, ϕ) is a k-out-of-n system. Thus, the class of threshold systems is a generalization of the class of k-out-of-n systems. However, despite the similarity between threshold systems and k-out-of-n systems, it can be shown that calculating the reliability of a threshold system is NP-hard. See Winther [48].

In the next sections we present both general algorithms as well as specialized algorithms tailored for certain classes of systems.

4.1 State space enumeration

The idea of this method is simple. We know that the reliability of a binary monotone system (C, ϕ) is simply the expected value of $\phi(\mathbf{X})$. Thus, by standard probability theory this can be calculated as:

$$h = E[\phi(\mathbf{X})] = \sum_{\mathbf{x} \in \{0,1\}^n} \phi(\mathbf{x}) P(\mathbf{X} = \mathbf{x}) \tag{4.1}$$

Note that in order to calculate the reliability using (4.1) we must enumerate all possible states of the component vector (i.e., all $\mathbf{x} \in \{0,1\}^n$). Moreover, if the components are dependent, we must know the entire joint distribution of $\mathbf{X}$ in

order to compute the system reliability using (4.1). As this distribution is usually not known, this method is unsuitable for the case of dependent components. However, if the components are independent, we have:

$$P(\mathbf{X} = \mathbf{x}) = \prod_{i=1}^{n} p_i^{x_i}(1 - p_i)^{1-x_i}$$

Inserting this into equation (4.1) we get:

$$h(\mathbf{p}) = \sum_{\mathbf{x}\in\{0,1\}^n} \phi(\mathbf{x}) \prod_{i=1}^{n} p_i^{x_i}(1 - p_i)^{1-x_i}. \tag{4.2}$$

To compute the system reliability via (4.2), it suffices to know the individual component reliabilities. However, since we need to enumerate all the possible states of the component state vector $\mathbf{x}$, the number of terms in the sum (4.2) is 2^n. Thus, this algorithm is at least of order $O(2^n)$. In fact the order may be even greater since we have not included the work of computing $\phi(\cdot)$ for each of the possible states of the component state vector $\mathbf{x}$. For some structure functions this may also represent a considerable task.

While the order of the algorithm cannot be changed, there are still ways to improve the method. Sometimes it is possible to run through the states in a clever order. For example, if we come across $\mathbf{x}_0$ such that $\phi(\mathbf{x}_0) = 0$, we do not have to include any of the states $\mathbf{x}$ such that $\mathbf{x} < \mathbf{x}_0$ since ϕ is non-decreasing (as it is assumed to be monotone), so $\phi(\mathbf{x})$ must be 0 as well. Note also that since $\phi(\mathbf{0}) = 0$, the maximum amount of terms we need to compute in (4.2) is $2^n - 1$.

Example 4.1.1 *Consider a binary monotone system (C, ϕ) where $C = \{1, 2, 3\}$, and where the structure function ϕ is given by:*

$$\phi(\boldsymbol{x}) = (x_1 \cdot x_2) \amalg x_3.$$

We assume that the component state variables are independent, and that the component reliabilities are $P(X_i = 1) = p_i$, $i = 1, 2, 3$. By using (4.2) we find that:

$$\begin{aligned} h(\boldsymbol{p}) &= \phi(0,0,0) \cdot (1-p_1)(1-p_2)(1-p_3) \\ &\quad + \phi(1,0,0) \cdot p_1(1-p_2)(1-p_3) + \cdots + \phi(1,1,1) \cdot p_1p_2p_3 \\ &= p_1p_2(1-p_3) \\ &\quad + (1-p_1)(1-p_2)p_3 + p_1(1-p_2)p_3 + (1-p_1)p_2p_3 + p_1p_2p_3. \end{aligned}$$

Note that the last expression contains just 5 terms, not $2^3 = 8$ terms. The reason for this is that there are exactly 5 vectors such that $\phi(\boldsymbol{x}) = 1$. The other 3 terms vanish in the final formula. Moreover, we observe that the last 4 terms

have p_3 as a common factor. Thus, the expression for $h(\boldsymbol{p})$ can be simplified to:

$$\begin{aligned} h(\boldsymbol{p}) &= p_1p_2(1-p_3) \\ &\quad + [(1-p_1)(1-p_2) + p_1(1-p_2) + (1-p_1)p_2 + p_1p_2] \cdot p_3 \\ &= p_1p_2(1-p_3) + p_3 \end{aligned}$$

This example shows that the expansion obtained by state space enumeration may not be the most efficient expression for the reliability function.

4.2 The multiplication method

The first step of the multiplication method for computing the exact system reliability, it to find the minimal path sets or alternatively, the minimal cut sets of the system. Denote these by respectively $P_1, \ldots, P_p$ and $K_1, \ldots, K_k$. Now, from the formula (3.8) and (3.10), we know that:

$$\begin{aligned} \phi(\mathbf{X}) &= \coprod_{j=1}^{p} \prod_{i \in P_j} X_i = 1 - [\prod_{j=1}^{p} (1 - \prod_{i \in P_j} X_i)] \\ &= \prod_{j=1}^{k} \coprod_{i \in K_j} X_i = \prod_{j=1}^{k} [1 - \prod_{i \in K_j} (1 - X_i)]. \end{aligned}$$

By expanding either the formula based on the minimal path sets, or the formula based on the minimal cut sets, and using that $X_i^r = X_i$, $i = 1, \ldots, n$, $r = 1, 2, \ldots$, we eventually get an expression of the form:

$$\phi(\mathbf{X}) = \sum_{A \subseteq C} \delta(A) \prod_{i \in A} X_i \tag{4.3}$$

where for all $A \subseteq C$, $\delta(A)$ denotes the coefficient of the term associated with $\prod_{i \in A} X_i$. The δ-function is called the *signed domination function* of the structure. The fact that we end up with an expression of the form (4.3) follows since ϕ is multilinear (see the comments after Theorem 3.1.1), and all multilinear functions can be written in this form.

By taking the expectation on both sides of (4.3), and assuming that the component state variables are independent, we obtain the following expression:

$$h(\mathbf{p}) = E[\phi(\mathbf{X})] = \sum_{A \subseteq C} \delta(A) \prod_{i \in A} E[X_i] = \sum_{A \subseteq C} \delta(A) \prod_{i \in A} p_i \tag{4.4}$$

As for the state space enumeration method, we again end up with a sum containing 2^n terms. Thus, calculating the reliability of the system using (4.4) also has the order $O(2^n)$. In a given case, however, there will typically be far less terms in this expansion since $\delta(A) = 0$ for many of the sets A.

Note, that when using this method, we also need to determine the signed domination function before we can use (4.4). It turns out that identifying all

the minimal path or cut sets is another operation which has the order $O(2^n)$. Moreover, depending on the number of minimal path or cut sets, the task of expanding the structure function into the form (4.3) is yet another complicated task.

The multiplication method can be improved by avoiding identifying the minimal path or cut sets as well as the multiplication step. In order to explain this, we introduce the following alternative way of expressing the structure function of a binary monotone system (C, ϕ):

$$\phi(B) = \phi(\mathbf{1}^B, \mathbf{0}^{\bar{B}}), \text{ for all } B \subseteq C.$$

Thus, for all $B \subseteq C$, $\phi(B)$ denotes the state of the system given that all the components in the set B are functioning, while all the components in the set $\bar{B} = C \setminus B$ are failed. In Huseby [22] the following formula was proved:

$$\delta(A) = \sum_{B \subseteq A} (-1)^{|A|-|B|} \phi(B), \text{ for all } A \subseteq C. \tag{4.5}$$

This formula allows us to compute the signed domination function without using the minimal path and cut sets. In its most general form the sum in (4.5) still contains 2^n terms. Thus, evaluating this formula is yet another complex operation of order 2^n. However, for certain classes of systems (4.5) can be used as basis for deriving simpler formulas allowing much faster calculations. We will return to this in Section 4.4.

Example 4.2.1 *Consider a binary monotone system (C, ϕ) where $C = \{1, 2, 3\}$, and where the minimal path sets of the system are $P_1 = \{1, 2\}$ and $P_2 = \{1, 3\}$. Using the multiplication method we obtain:*

$$\begin{aligned}
\phi(\boldsymbol{X}) &= (X_1 X_2) \amalg (X_1 X_3) = 1 - (1 - X_1 X_2)(1 - X_1 X_3) \\
&= 1 - (1 - X_1 X_2 - X_1 X_3 + X_1^2 X_2 X_3) \\
&= X_1 X_2 + X_1 X_3 - X_1 X_2 X_3,
\end{aligned}$$

where we have used that $X_1^2 = X_1$. Thus, we see that ϕ is of the form (4.3), where $\delta(\{1, 2\}) = \delta(\{1, 3\}) = 1$, $\delta(\{1, 2, 3\}) = -1$, while $\delta(A) = 0$ for all other subsets of C.

Note that these coefficients can be obtained using (4.5) as well since:

$$\begin{aligned}
\delta(\{1,2\}) &= (-1)^{|\{1,2\}|-|\{1,2\}|} \phi(\{1,2\}) = 1, \\
\delta(\{1,3\}) &= (-1)^{|\{1,3\}|-|\{1,3\}|} \phi(\{1,3\}) = 1, \\
\delta(\{1,2,3\}) &= (-1)^{|\{1,2,3\}|-|\{1,2\}|} \phi(\{1,2\}) + (-1)^{|\{1,2,3\}|-|\{1,3\}|} \phi(\{1,2\}) \\
&\quad + (-1)^{|\{1,2,3\}|-|\{1,2,3\}|} \phi(\{1,2,3\}) = -1 - 1 + 1 = -1.
\end{aligned}$$

Having derived the formula for the structure function ϕ, we immediately obtain the reliability function:

$$h(\boldsymbol{p}) = p_1 p_2 + p_1 p_3 - p_1 p_2 p_3.$$

4.3 The inclusion-exclusion method

Just like for the multiplication method, the inclusion exclusion method is based on first finding the minimal path sets and the minimal cut sets of the system. Let these sets be denoted respectively by $P_1, \ldots, P_p$ and $K_1, \ldots, K_k$. We then introduce the events

$$E_j = \{\text{All of the components in } P_j \text{ are functioning}\}, \quad j = 1, \ldots, p.$$

Since the system is functioning if and only if at least one of the minimal path sets is functioning, we have:

$$\phi = 1 \quad \text{if and only if} \quad \bigcup_{j=1}^{p} E_j \text{ holds true.} \tag{4.6}$$

Calculating the probability of a union of events can be done by using the well-known *inclusion-exclusion* formula. That is we have:

$$\begin{aligned} h &= P(\bigcup_{j=1}^{p} E_j) \\ &= P(E_1) + P(E_2) + \cdots + P(E_p) \\ &\quad - P(E_1 \cap E_2) - P(E_1 \cap E_3) - \cdots - P(E_{p-1} \cap E_p) \\ &\quad + P(E_1 \cap E_2 \cap E_3) + \cdots + P(E_{p-2} \cap E_{p-1} \cap E_p) \\ &\quad \cdots \\ &\quad + (-1)^{p-1} P(E_1 \cap \cdots \cap E_p). \end{aligned} \tag{4.7}$$

The initial number of terms in (4.7) is $2^p - 1$. However, typically many of the terms can be merged as they correspond to the same component set. Thus, after simplifying the expression we usually end up with much fewer terms.

Note that an event of form $E_{i_1} \cap \cdots \cap E_{i_r}$ occurs if and only if all the components in the set $P_{i_1} \cup \cdots \cup P_{i_r}$ are functioning. Thus, when the component state variables are independent, we get that:

$$P(E_{i_1} \cap \cdots \cap E_{i_r}) = \prod_{i \in P_{i_1} \cup \cdots \cup P_{i_r}} p_i.$$

Example 4.3.1 *Consider a binary monotone system (C, ϕ) where $C = \{1, 2, 3, 4\}$ with minimal path sets $P_1 = \{1, 2\}$, $P_2 = \{1, 3\}$, $P_3 = \{2, 3, 4\}$. We then let E_j denote the event that all components in P_j are functioning, $j = 1, 2, 3$. Assuming that the component state variables are independent, we then get the following probabilities:*

$$\begin{aligned} P(E_1) = p_1 p_2, \quad P(E_2) = p_1 p_3, \quad P(E_3) = p_2 p_3 p_4 \\ P(E_1 \cap E_2) = p_1 p_2 p_3, \quad P(E_1 \cap E_3) = P(E_2 \cap E_3) = p_1 p_2 p_3 p_4 \\ P(E_1 \cap E_2 \cap E_3) = p_1 p_2 p_3 p_4. \end{aligned}$$

Hence, the reliability of the system is:

$$\begin{aligned} h(\boldsymbol{p}) &= p_1p_2 + p_1p_3 + p_2p_3p_4 - p_1p_2p_3 - 2p_1p_2p_3p_4 + p_1p_2p_3p_4 \\ &= p_1p_2 + p_1p_3 + p_2p_3p_4 - p_1p_2p_3 - p_1p_2p_3p_4. \end{aligned}$$

We observe that the final expression for the reliability function is exactly the same as the one we get using the multiplication methods. That is, in both cases we end up with a formula of the form:

$$h(\mathbf{p}) = \sum_{A \subseteq C} \delta(A) \prod_{i \in A} p_i,$$

where δ denotes the signed domination function. However, the steps we take in order to get this expression is different.

When using the inclusion-exclusion formula we see that all terms in the expansion correspond to sets $A \subseteq C$ which are unions of minimal path sets. If $P_{i_1}, \ldots, P_{i_r}$ is a collection of minimal path sets such that:

$$\bigcup_{j=1}^{r} P_{i_j} = A,$$

the collection is said to be a *formation* of the set A. The formation is *odd* if r is an odd number, and *even* if r is an even number. We note that odd formations produce terms with a coefficient $+1$ in the inclusion-exclusion formula, while even formations produce terms with a coefficient -1 in the inclusion-exclusion formula. When we simplify the expansion, all terms corresponding to formations of the same set are merged. From this observation it follows that we have the following result:

Theorem 4.3.2 *Let (C, ϕ) be a binary monotone system with minimal path sets $P_1, \ldots, P_p$, and let δ denote the signed domination function of the system. Then for all $A \subseteq C$ we have:*

$$\begin{aligned} \delta(A) = &\textit{ The number of odd formations of } A \\ &- \textit{ The number of even formations of } A. \end{aligned} \tag{4.8}$$

In particular $\delta(A) = 0$ if A is not a union of minimal path sets.

As a corollary we obtain the following result:

Corollary 4.3.3 *If (C, ϕ) is a binary monotone system which is not coherent, then $\delta(C) = 0$.*

The proof is left as an exercise.

A nice feature of the inclusion-exclusion formula is that it can be used to produce a simple upper bound for the system reliability. Thus, if we skip all higher

order terms and only keep the probabilities of the individual events $E_1, \ldots, E_p$, we get:

$$h \leq \sum_{j=1}^{p} P(E_j). \tag{4.9}$$

By including higher order terms, it is possible to obtain both lower bounds and improved upper bounds as well. However, as we introduce the higher order terms, the total number of terms in these expansions grows fast. Thus, only the simple bound (4.9) is considered here. In order to obtain a lower bound we instead use a *dual* approach. That is, we consider the minimal cut sets of the system $K_1, \ldots, K_k$, and introduce the events:

$$F_j = \{\text{All the components in } K_j \text{ are failed}\}, \quad j = 1, \ldots, k.$$

Since the system is failed if and only if all the components in at least one cut set are failed, we have:

$$1 - h = P(\bigcup_{j=1}^{k} F_j). \tag{4.10}$$

The quantity $1 - h$ is referred to as the *unreliability* of the system. By the same argument as we used above, an upper bound on this quantity is given by:

$$1 - h \leq \sum_{j=1}^{k} P(F_j). \tag{4.11}$$

By combining (4.9) and (4.11), we see that

$$1 - \sum_{j=1}^{k} P(F_j) \leq h \leq \sum_{j=1}^{p} P(E_j). \tag{4.12}$$

If the components are independent, (4.12) implies that:

$$1 - \sum_{j=1}^{k} \prod_{i \in K_j} (1 - p_i) \leq h(\mathbf{p}) \leq \sum_{j=1}^{p} \prod_{i \in P_j} p_i. \tag{4.13}$$

In many real life applications the component reliabilities are close to 1. In such cases the lower bound given in (4.13) turns out to be very good, while the upper bound is extremely poor. In fact it may easily happen that the upper bound becomes greater than 1. If on the other hand the component reliabilites are close to 0, the lower bound may be less than 0, while the upper bound turns out to be very good. In order to avoid bounds outside of the interval $[0, 1]$, one could replace (4.13) by:

$$\max(1 - \sum_{j=1}^{k} \prod_{i \in K_j} (1 - p_i),\, 0) \leq h(\mathbf{p}) \leq \min(\sum_{j=1}^{p} \prod_{i \in P_j} p_i,\, 1). \tag{4.14}$$

4.4 Computing the reliability of directed network systems

For some particular kinds of systems, the multiplication method and the inclusion-exclusion method can be significantly improved. The most important class of such systems are Source-to-K terminal systems, or *SKT-systems*:

Definition 4.4.1 *A Source-to-K-terminal-system (SKT-system) is a system defined relative to a directed network where the system functions if and only if a node S (called the source) can send information to a given set of K nodes $T_1, \ldots, T_K$ (called the terminals). The components of the system are the directed edges of the network, while the nodes are assumed to be functioning perfectly with probability one.*

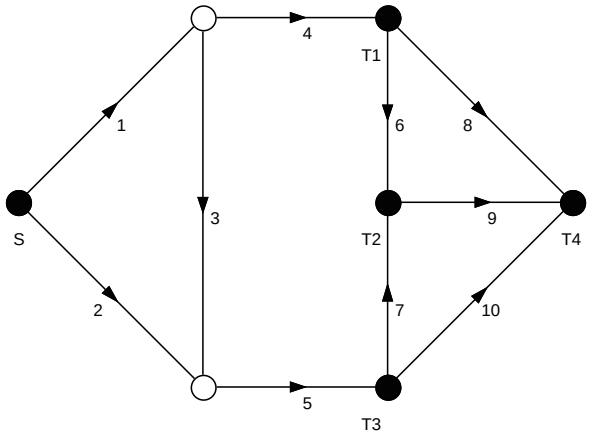

Figure 4.1: An S4T system

Example 4.4.2 *Figure 4.1 shows an example of an S4T system with components $1, 2, \ldots, 10$. The node S is the source, while the nodes T_1, T_2, T_3, T_4 are the terminals.*

In the following theorem we recall that any structure function can be written on the form (4.3).

Theorem 4.4.3 *Let $\phi(\boldsymbol{X}) = \sum_{A \subseteq C} \delta(A) \prod_{i \in A} X_i$ be the structure function of an SKT system.*

(i) If A can be expressed as a union of minimal path sets, and the subgraph spanned by A does not contain any directed cycle, we have:

$$\delta(A) = (-1)^{|A|-v(A)+1}$$

where $v(A)$ denotes the number of nodes in the subgraph spanned by A.

(ii) In the opposite case we have:

$$\delta(A) = 0$$

Theorem 4.4.3 was first proved by Satyanarayana [41] using Theorem 4.3.2. However, a simpler proof based on the formula (4.5) is given in Huseby [22].

Prabhakar and Satyanarayana [43] also developed an efficient algorithm for identifying all sets $A \subseteq C$ such that $\delta(A) \neq 0$ directly, without using the minimal path sets. By using this algorithm in combination with Theorem 4.4.3, we get a very fast method for computing the reliability of an SKT system.

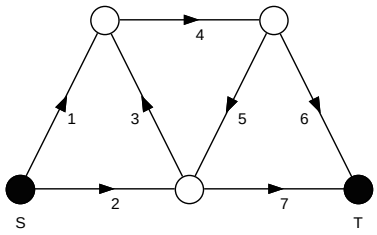

Figure 4.2: An S1T system

Example 4.4.4 *Consider the S1T-system shown in Figure 4.2. The component set $C = \{1, \ldots, 7\}$ consists of the directed edges in the networks. The system is functioning if the source S can send signals to the terminal T through the network. We assume that the component state variables are independent and that $P(X_i = 1) = p_i$ for $i \in C$.*

The minimal path sets of this system are $P_1 = \{1, 4, 6\}$, $P_1 = \{1, 4, 5, 7\}$, $P_1 = \{2, 3, 4, 6\}$ and $P_1 = \{2, 7\}$. We calculate the reliability of this sysem by using the inclusion-exclusion formula (4.7). Since there are 4 minimal path sets, this formula will consist of $2^4 - 1 = 15$ terms before we simplify:

$$\begin{aligned} h(\boldsymbol{p}) = {} & p_1p_4p_6 + p_1p_4p_5p_7 + p_2p_3p_4p_6 + p_2p_7 \\ & - p_1p_4p_5p_6p_7 - p_1p_2p_3p_4p_6 - p_1p_2p_4p_6p_7 \\ & - p_1p_2p_3p_4p_5p_6p_7 - p_1p_2p_4p_5p_7 - p_1p_2p_4p_5p_7 \\ & + p_1p_2p_3p_4p_5p_6p_7 + p_1p_2p_4p_5p_6p_7 + p_1p_2p_3p_4p_6p_7 + p_1p_2p_3p_4p_5p_6p_7 \\ & - p_1p_2p_3p_4p_5p_6p_7. \end{aligned}$$

By merging similar terms we obtain:

$$\begin{aligned} h(\boldsymbol{p}) &= p_1p_4p_6 + p_1p_4p_5p_7 + p_2p_3p_4p_6 + p_2p_7 \\ &\quad - p_1p_4p_5p_6p_7 - p_1p_2p_3p_4p_6 - p_1p_2p_4p_6p_7 \\ &\quad - p_1p_2p_4p_5p_7 - p_1p_2p_4p_5p_7 \\ &\quad + p_1p_2p_4p_5p_6p_7 + p_1p_2p_3p_4p_6p_7. \end{aligned}$$

We observe that $\delta(A)$ is either $+1$, -1 or zero for all $A \subseteq C$. In particular, since the network contains a directed cycle $\{3,4,5\}$, the highest order term will have $\delta(C) = 0$. Moreover, we can easily verify that the formula for $\delta(A)$ given in Theorem 4.4.3 is correct:

$$\begin{aligned} \delta(\{1,4,6\}) &= (-1)^{3-4+1} = +1, \\ \delta(\{1,4,5,7\}) &= (-1)^{4-5+1} = +1, \\ &\cdots\cdots \\ \delta(\{1,4,5,6,7\}) &= (-1)^{5-5+1} = -1, \\ \delta(\{1,2,3,4,6\}) &= (-1)^{5-5+1} = -1, \\ \delta(\{1,2,4,6,7\}) &= (-1)^{5-5+1} = -1, \\ &\cdots\cdots \\ \delta(\{1,2,4,5,6,7\}) &= (-1)^{6-5+1} = +1, \\ \delta(\{1,2,3,4,6,7\}) &= (-1)^{6-5+1} = +1. \end{aligned}$$

We close this section by noting that SKT-systems are not the only types of systems where the signed domination is either $+1$, -1 or zero. In fact this class can be extended to a much larger class called *oriented matroid systems*. See Huseby [25] for details. Moreover, there are also other classes of systems with similar properties:

Example 4.4.5 *Let (C, ϕ) be a linear consecutive 2-out-of-5 system. That is, $C = \{1, \ldots, 5\}$, and the minimal path sets are $P_1 = \{1,2\}$, $P_2 = \{2,3\}$, $P_3 = \{3,4\}$, $P_4 = \{4,5\}$. By using e.g., the inclusion-exclusion formula it is easy to see that the reliability of this system, assuming independent component state variables, is:*

$$\begin{aligned} h(\boldsymbol{p}) &= p_1p_2 + p_2p_3 + p_3p_4 + p_4p_5 \\ &\quad - p_1p_2p_3 - p_2p_3p_4 - p_3p_4p_5 - p_1p_2p_4p_5 \\ &\quad + p_1p_2p_3p_4p_5. \end{aligned}$$

It can be shown that this system is not an SKT-system. The proof is left as an exercise. Still, linear consecutive k-out-of-n systems share the property with SKT-systems (and the more general class of oriented matroid systems) that the signed domination function is either $+1$, -1 or zero for all subsets of the component set. See Calkin et. al [9].

4.5 Computing the reliability of undirected network systems

In principle this algorithm can be applied to any binary monotone system where the component state varaiables are independent. However, for optimal performance we assume that the system can be represented as an *undirected network system*, where the components of the system are the *edges* of the systems, and where the system is functioning if a set of terminals can communicate through the network. In an undirected network signals can be sent both ways along the edges, see Figure 4.3.

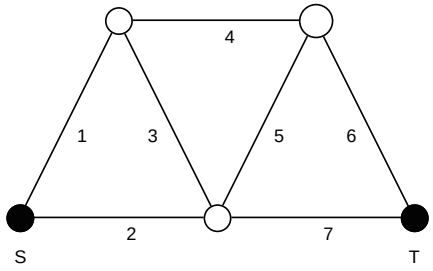

Figure 4.3: An undirected network with components $1, 2, \ldots, 7$ and terminal nodes S and T.

The method is based on pivoting (see Theorem 3.1.1) and the following result:

Theorem 4.5.1 *Let $i, j \in C$, $i \neq j$.*

(i) If i and j are connected in series, then $h(\boldsymbol{p})$ will only depend on p_i and p_j through $p_i \cdot p_j$. Hence, i and j can be replaced by a single component with reliability $p_i \cdot p_j$ without altering the system reliability. Such a reduction is called a series reduction.

(ii) If i and j are connected in parallel, then $h(\boldsymbol{p})$ will only depend on p_i and p_j through $p_i \amalg p_j$. Hence, i and j can be replaced by a single component with reliability $p_i \amalg p_j$ without altering the system reliability. Such a reduction is called a parallel reduction.

The common term for either series or parallel reductions is *s-p-reductions.*

Definition 4.5.2 *We say that a system is* s-p-reducable *if there are components in either series or parallel in the system. If not, the system is said to be* complex. *A system which can be s-p-reduced to a single component is called an* s-p system.

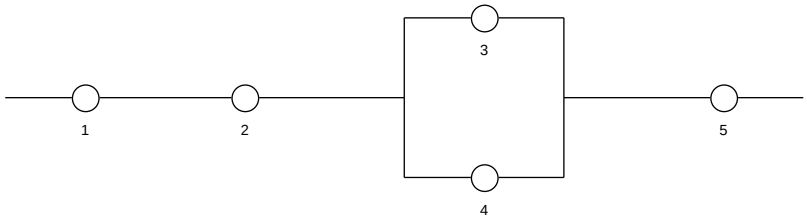

Figure 4.4: A reliability block diagram of a mixed parallel and series system.

The bridge structure shown in Figure 3.1 is an example of a complex system because there are no components in series or parallel.

The system in Figure 4.4 is an example of an s-p system because components 1 and 2 (in series) can be replaced with a component $1'$ with reliability $p_{1'} = p_1 \cdot p_2$. Similarly, components 3 and 4 can be replaced by a component $3'$ with reliability $p_{3'} = p_3 \amalg p_4$. The resulting system is a series structure of components $1', 3'$ and 5. These three components can be replaced by a single component with reliability $p_{1'} \cdot p_{3'} \cdot p_5$, which is the reliability of the whole system.

The factoring algorithm can be described as follows:

Algorithm 4.5.3 *Let (C, ϕ) be a monotone system, that is, assume that at least one of its components is relevant. To compute the reliability $h(\boldsymbol{p})$, proceed as follows:*

Step 1: *Perform all possible s-p-reductions. Let the reduced system be denoted by (C^r, ϕ^r). Then, (C^r, ϕ^r) must also have at least one relevant component (make sure you understand why).*

Step 2: *Now, one of the two following cases can happen:*

Case 1. *(C^r, ϕ^r) containts precisely one relevant component with updated reliability p_e. Then, $h(\boldsymbol{p}) = p_e$.*

Case 2. *(C^r, ϕ^r) contains several relevant components. In this case, choose a component $e \in C^r$ and do a pivotal decomposition. That is, compute $h(\boldsymbol{p})$ by using Theorem 3.1.1:*

$$h(\boldsymbol{p}^{C^r}) = p_e h(1_e, \boldsymbol{p}^{C^r}) + (1 - p_e) h(0_e, \boldsymbol{p}^{C^r}).$$

Then, compute $h(1_e, \boldsymbol{p}^{C^r})$ and $h(0_e, \boldsymbol{p}^{C^r})$ by repeated use of the algorithm.

Note that there is a choice involved in Case 2. of Algorithm 4.5.3. In general, the efficiency of the factoring algorithm depends on the choice of pivoting component. For the bridge structure in Example 3.1.2, we actually used the factoring algorithm. There, we chose to pivot with respect to component 3. For this particular system the choice does not matter. In general, however, it can be shown that when the factoring algorithm is applied to undirected network

systems, one should always pivot such that both resulting substructures are coherent[1]. This was first proved by Satyanarayana and Chang [42], and later generalized to a more abstract class by Huseby [22].

Despite being a very easy and convenient method for reliability calculations, the factoring algorithm is still of order $O(2^n)$. Hence, in that sense it is not significantly better than e.g., the state space enumeration method. Still, in many cases, it is very efficient.

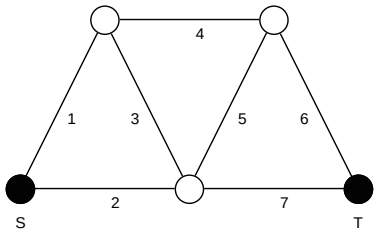

Figure 4.5: An undirected network system

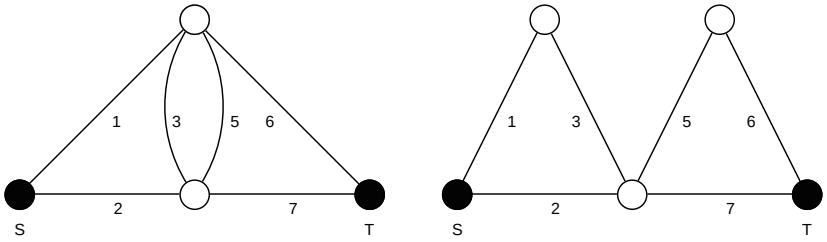

Figure 4.6: Subsystems $(C \setminus 4, \phi_{+4})$ and $(C \setminus 4, \phi_{-4})$ obtained by pivotal decomposition with respect to component 4

Example 4.5.4 *Consider the undirected network system (C, ϕ) shown in Figure 4.5. The component set $C = \{1, \ldots, 7\}$ consists of the undirected edges in the networks. The system is functioning if the terminals S and T can communicate through the network. We assume that the component state variables are independent and that $P(X_i = 1) = p_i$ for $i \in C$.*

[1]Note that the assumtion that the system is an undirected network system is essential here. While the result can be generalized to a larger class of systems, see Huseby [22], there exists many other types of systems where this rule is *not* optimal.

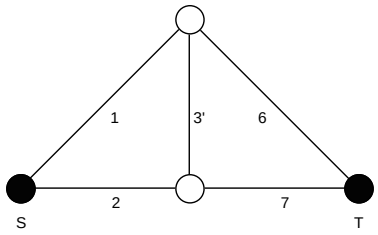

Figure 4.7: The subsystem $((C \setminus 4)^r, (\phi_{+4})^r)$ obtained by a parallel reduction of $(C \setminus 4, \phi_{+4})$ with respect to components 3 and 5

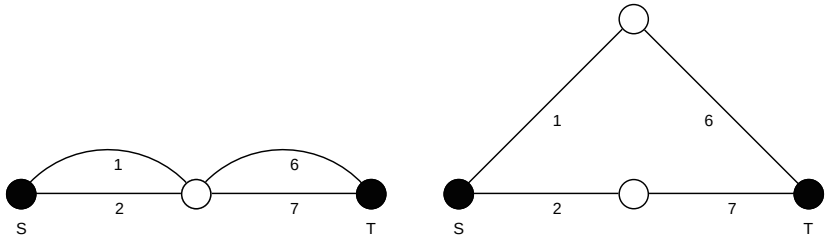

Figure 4.8: Subsystems $((C \setminus 4)^r \setminus 3', (\phi_{+4})^r_{+3'})$ and $((C \setminus 4)^r \setminus 3', (\phi_{+4})^r_{-3'})$ obtained from the subsystem $((C \setminus 4)^r, (\phi_{+4})^r)$ by a pivotal decomposition with respect to component $3'$

This system is not *s-p-reducable, so we need to do a pivotal decomposition. For this purpose we choose component* 4*. The resulting subsystems are shown in Figure 4.6, where the left-hand system, denoted* $(C \setminus 4, \phi_{+4})$ *and with reliability* h_{+4}*, is obtained by conditioning on that component* 4 *is functioning, while the right-hand system, denoted* $(C \setminus 4, \phi_{-4})$ *and with reliability* h_{+4}*, is obtained by conditioning on that component* 4 *is failed. The system reliability is then given by:*

$$h = p_4 \cdot h_{+4} + (1 - p_4) \cdot h_{-4}.$$

We observe that $(C \setminus 4, \phi_{-4})$ *is s-p-reducable with reliability* h_{-4} *given by:*

$$h_{-4} = [(p_1 \cdot p_3) \amalg p_2] \cdot [(p_5 \cdot p_6) \amalg p_7].$$

The subsystem $(C \setminus 4, \phi_{+4})$ *is s-p-reducable as well, but the only possible s-p-reduction is the parallel reduction of components* 3 *and* 5*. The resulting component is denoted* $3'$ *and has component reliability* $p_{3'} = p_3 \amalg p_5$*. The resulting system is shown in Figure 4.6, and is denoted by* $((C \setminus 4)^r, (\phi_{+4})^r)$*.*

In order to calculate the reliability of this system, we need to do another pivotal decomposition. This time we choose component $3'$*. The resulting subsystems, denoted respectively* $((C \setminus 4)^r \setminus 3', (\phi_{+4})^r_{+3'})$ *and* $((C \setminus 4)^r \setminus 3', (\phi_{+4})^r_{-3'})$ *are shown in Figure 4.8. Both these subsystems are s-p-reducable, and we get that:*

$$h_{+4} = (h_{+4})^r = p_{3'} \cdot (h_{+4})^r_{+3'} + (1 - p_{3'}) \cdot (h_{+4})^r_{-3'},$$

where:

$$(h_{+4})^r_{+3'} = (p_1 \amalg p_2) \cdot (p_6 \amalg p_7)$$
$$(h_{+4})^r_{-3'} = (p_1 \cdot p_6) \amalg (p_2 \cdot p_7)$$

This completes the calculations.

Note that it is obviously possible to combine all the various expressions in Example 4.5.4 into one unified large formula for the reliability h. However, when implementing the factoring algorithm, this is done recursively. This implies that only the resulting numbers, not the symbolic expressions, are passed backwards through the algorithm. This is actually an essential point as this saves both time and space.

4.6 Exercises

Exercise 4.1: Consider the bridge structure from Example 3.1.2 with independent component state variables and reliability function $h(\mathbf{p})$ as given in this example. Assume that the reliability of all five components is 0.9. What do the bounds (4.12) reduce to in this case? Comment on the result.

Exercise 4.2: Draw an example of the following systems:

(i) An S3T system.

(ii) An S5T system.

Exercise 4.3: Consider the S1T system in Figure 4.9.

a) Find the minimal path sets of the system.

b) How many terms will there be in the inclusion-exclusion formula for the reliability of this system before simplification?

c) How many of these terms will vanish in the final simplified expression?

Exercise 4.4: Show that the linear consecutive 2-out-of-5 system given in Example 4.4.5 cannot be an SKT-system. [*Hint:* Let δ denote the signed domination function of the system. Compare $\delta(\{1,2,3,4\})$ and $\delta(\{1,2,3,4,5\})$ and interpret the result in the light of Theorem 4.4.3.

Exercise 4.5: Compute the reliability function of the undirected network system in Figure 4.10 which functions if and only if the nodes S and T can communicate through the network. Use the factoring algorithm.]

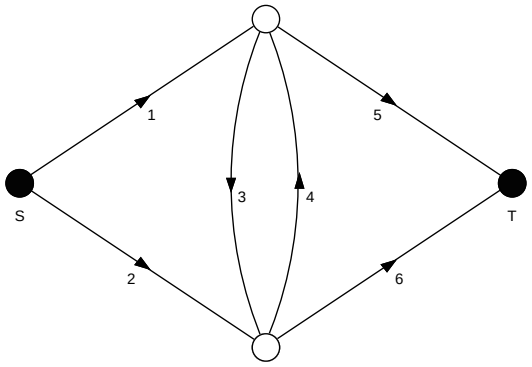

Figure 4.9: An S1T system.

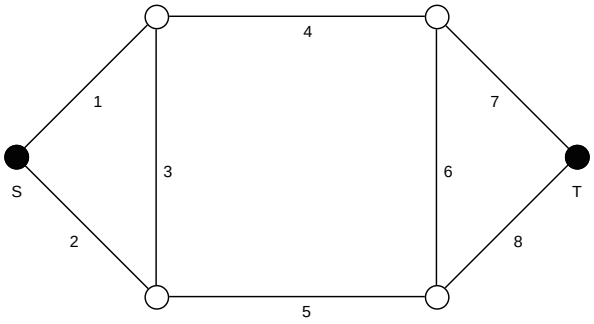

Figure 4.10: An undirected network system with components $1, 2, \ldots, 8$ and terminals S and T.

Exercise 4.6: Compute the reliability function of the undirected network system in Figure 4.11 by using the factoring algorithm.

Since the system is complex, it has to be factored with respect to some component. Which component? Compare the computational work for different component choices.

Exercise 4.7: Consider the series connection of bridge structures in Figure 4.12.

a) Compute the reliability function of the system by using the factoring algorithm.

b) Instead, compute this by taking the product of the three reliability func-

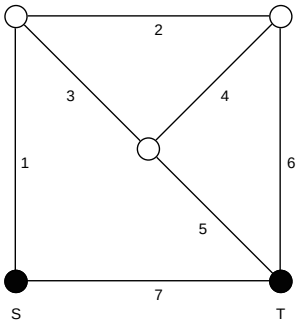

Figure 4.11: An undirected network system with components $1, 2, \ldots, 7$ and terminals S and T.

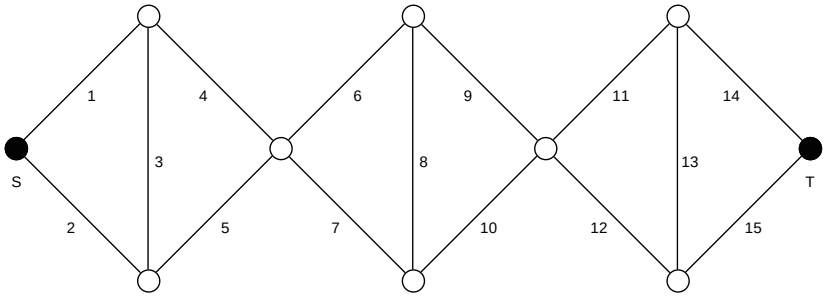

Figure 4.12: A series system of bridge structures.

tions of the three bridge structures. Compare the computational effort with the one in a).

c) How many terms are there in the inclusion-exclusion formula for the reliability of this system? Comment this result.

Exercise 4.8: Prove Corollary 4.3.3. [*Hint*: Use the result given in Exercise 3.5.]

Chapter 5

Structural and reliability importance for components in binary monotone systems

By considering the structure- or reliability function of a binary monotone system, it is evident that not all the components are equally important for the system to function. For instance, a component connected in series or parallel with the rest of the system will typically be more important than the other components in the system. In this section, we will define various measures for importance of the components in a system. There are many reasons for considering such measures, but among the most important ones are the following:

1. A measure of importance can be used to identify components that should be improved in order to increase the system reliability.

2. A measure of importance can be used to identify components that most likely have failed, given that the system has failed.

5.1 Structural importance of a component

The measures of importance which we are about to introduce are all based on the notion of *criticality*. This is defined as follows:

Definition 5.1.1 *Let (C, ϕ) be a binary monotone system, and let $i \in C$. We say that component i is* critical for the system *if:*

$$\phi(1_i, \boldsymbol{x}) = 1 \text{ and } \phi(0_i, \boldsymbol{x}) = 0.$$

If this is the case, we also say that $(\cdot_i, \boldsymbol{x})$ is a critical vector *for component i.*

We observe that criticality is strongly related to the notion of relevance. Thus, a component i in a binary monotone system (C, ϕ) is relevant if and only if there exists at least one critical vector for i.

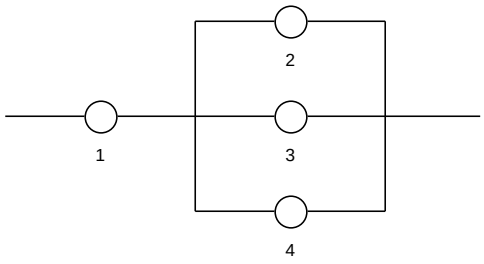

Figure 5.1:

Example 5.1.2 *Consider the binary monotone system (C, ϕ) shown in Figure 5.1, where $C = \{1, 2, 3, 4\}$, and where the structure function is given by:*

$$\phi(\boldsymbol{x}) = x_1(x_2 \amalg x_3 \amalg x_4)$$

We start out by considering component 1, and we see that this component is critical if and only if at least one of the other components are functioning. Thus, there are seven critical vectors for i:

$$\begin{array}{lll} (\cdot, 1, 0, 0), & (\cdot, 0, 1, 0), & (\cdot, 0, 0, 1) \\ (\cdot, 1, 1, 0), & (\cdot, 1, 0, 1), & (\cdot, 0, 1, 1) \\ (\cdot, 1, 1, 1). & & \end{array}$$

We then compare this to any of the other components, e.g., component 2. For this component to be critical component 1 must function, while both component 3 and 4 must be failed. Hence, the only critical vector for component 2 is $(1, \cdot, 0, 0)$.

By Definition 5.1.1 it follows that a component i of a binary monotone system (C, ϕ) is critical if and only if:

$$\phi(1_i, \mathbf{x}) - \phi(0_i, \mathbf{x}) = 1.$$

Hence, the number of critical vectors for i can be calculated by summing this expression over all possible vectors $(\cdot_i, \mathbf{x})$:

$$\sum_{(\cdot_i, \mathbf{x})} [\phi(1_i, \mathbf{x}) - \phi(0_i, \mathbf{x})].$$

Based on this Birnbaum [7] suggested the following measure of structural importance of a component in a binary monotone system:

Definition 5.1.3 *Let (C, ϕ) be a binary monotone system of order n, and let $i \in C$. The* Birnbaum measure for the structural importance *of component i, denoted $J_B^{(i)}$, is defined as:*

$$J_B^{(i)} := \frac{1}{2^{n-1}} \sum_{(\cdot_i, \boldsymbol{x})} [\phi(1_i, \boldsymbol{x}) - \phi(0_i, \boldsymbol{x})].$$

Note that the denominator, 2^{n-1} is the total number of states for the $n-1$ other components. Thus, $J_B^{(i)}$ can be interpreted as the fraction of all states for the $n-1$ other components where component i is critical. By computing $J_B^{(i)}$ for all of the components in a system, we can rank the components based on their respective structural importance.

Example 5.1.4 *Let ϕ be a 2-out-of-3 system. To compute the structural importance of component 1, we note that the critical vectors for this component are $(\cdot, 1, 0)$ and $(\cdot, 0, 1)$. Hence, we have:*

$$J_B^{(1)} = \frac{2}{2^{3-1}} = \frac{1}{2}.$$

By similar arguments, we find that:

$$J_B^{(2)} = J_B^{(3)} = \frac{1}{2}.$$

So in a 2-out-of-3 system, all of the components have the same structural importance. This is intuitively obvious since the structure function is symmetrical with respect to the components.

Example 5.1.5 *Consider the system ϕ with structure function:*

$$\phi(\boldsymbol{x}) = (x_1 \amalg x_2)x_3.$$

To find the structural importance of component 1, note that the only critical vector for this component is $(\cdot, 0, 1)$. Hence, we have:

$$J_B^{(1)} = \frac{1}{2^{3-1}} = \frac{1}{4}.$$

Since components 1 and 2 play completely similar roles in the system, it follows that we must have:

$$J_B^{(2)} = \frac{1}{4}$$

as well. Thus, components 1 and 2 have the same structural importance. However, this is not the case for component 3: The critical path vectors for component 3 are $(1, 0, \cdot)$, $(0, 1, \cdot)$ and $(1, 1, \cdot)$, so:

$$J_B^{(3)} = \frac{3}{2^{3-1}} = \frac{3}{4}.$$

From this, we see that:

$$J_B^{(1)} = J_B^{(2)} < J_B^{(3)}.$$

Hence, component 3, which is in series with the rest of the system, is more important than components 1 and 2.

5.2 Reliability importance of a component

While structural importance only depends on the structure function of the system, *reliability importance* takes into account the joint distribution of the component state variables as well. Thus, instead of considering the fraction of components state vectors where a given component is critical, we compute the probability that a given component is critical. This motivates the following definition introduced by Birnbaum [7]:

Definition 5.2.1 *Let (C, ϕ) be a binary monotone system, and let $i \in C$. Moreover, let $\boldsymbol{X}$ be the vector of component state variables. The Birnbaum measure for the reliability importance of component i, denoted $I_B^{(i)}$ is defined as:*

$$\begin{aligned} I_B^{(i)} &:= P(\textit{Component } i \textit{ is critical for the system}) \\ &= P(\phi(1_i, \boldsymbol{X}) - \phi(0_i, \boldsymbol{X}) = 1). \end{aligned}$$

Since the difference $\phi(1_i, \mathbf{X}) - \phi(0_i, \mathbf{X})$ is a binary variable, it follows that we may write:

$$I_B^{(i)} = E[\phi(1_i, \mathbf{X}) - \phi(0_i, \mathbf{X})] = E[\phi(1_i, \mathbf{X})] - E[\phi(0_i, \mathbf{X})]. \tag{5.1}$$

In particular, if the component state variables of the system are independent, and $P(X_i = 1) = p_i$ for $i \in C$, we get that:

$$I_B^{(i)} = h(1_i, \mathbf{p}) - h(0_i, \mathbf{p}). \tag{5.2}$$

By using the last identity we can show the following result which is sometimes taken as the definition of reliability importance in the case of independent components:

Theorem 5.2.2 *Let (C, ϕ) be a binary monotone system where the component state variables are independent, and $P(X_i = 1) = p_i$ for $i \in C$. Then:*

$$I_B^{(i)} = \frac{\partial h(\boldsymbol{p})}{\partial p_i}, \quad \textit{for all } i \in C.$$

Proof: By Theorem 3.1.1 we have that:

$$h(\mathbf{p}) = p_i h(1_i, \mathbf{p}) + (1 - p_i) h(0_i, \mathbf{p})$$

By differentiating this identity with respect to p_i we get:

$$\frac{\partial h(\mathbf{p})}{\partial p_i} = h(1_i, \mathbf{p}) - h(0_i, \mathbf{p}).$$

Hence, the result follows by (5.2). □

We observe that Theorem 5.2.2 shows that the reliability importance of a component indicates how much the system reliability grows with respect to the reliability of this component.

Now, we will prove some further properties of the Birnbaum measure.

Theorem 5.2.3 *For a binary monotone system, (C, ϕ), we always have*

$$0 \leq I_B^{(i)} \leq 1. \tag{5.3}$$

Assume that the component state variables are independent, and $P(X_j = 1) = p_j$, where $0 < p_j < 1$ for all $j \in C$. If component i is relevant, we have:

$$0 < I_B^{(i)}. \tag{5.4}$$

Furthermore, if there exists at least one other relevant component, we also have:

$$I_B^{(i)} < 1. \tag{5.5}$$

Proof: We note that (5.3) follows directly from Definition 5.2.1 since the reliability importance is a probability.

We then assume that the component state variables are independent, and $P(X_j = 1) = p_j$, where $0 < p_j < 1$ for all $j \in C$. If component i is relevant, we know from Theorem 3.1.3 that h is strictly increasing in p_i. That is, we must have:

$$\frac{\partial h(\mathbf{p})}{\partial p_i} > 0.$$

Combining this with Theorem 5.2.2, we get (5.4).

Finally, we assume that there exists at least one other relevant component, say $k \in C$. In order to prove that this implies that $I_B^{(i)} < 1$, we assume conversely that $I_B^{(i)} = 1$. We will then show that this leads to a contradiction. From this assumption, it follows by Definition 5.2.1 that :

$$P(\phi(1_i, \mathbf{X}) - \phi(0_i, \mathbf{X}) = 1) = 1$$

Since $0 < p_j < 1$, for all $j \in C$, it follows that $P((\cdot_i, \mathbf{X}) = (\cdot_i, \mathbf{x})) > 0$ for all $(\cdot_i, \mathbf{x})$. Hence, we must have that:

$$\phi(1_i, \mathbf{x}) = 1 \text{ and } \phi(0_i, \mathbf{x}) = 0 \text{ for all } (\cdot_i, \mathbf{x}).$$

At the same time, since component k is relevant, there exists a vector $(\cdot_k, \mathbf{y})$ such that:

$$\phi(1_k, \mathbf{y}) = 1 \text{ and } \phi(0_k, \mathbf{y}) = 0.$$

If $y_i = 1$, it follows that $\phi(1_i, 0_k, \mathbf{y}) = 0$, contradicting that $\phi(1_i, \mathbf{x}) = 1$ for all $(\cdot_i, \mathbf{x})$. Similarly, if $y_i = 0$, it follows that $\phi(0_i, 1_k, \mathbf{y}) = 1$, contradicting that $\phi(0_i, \mathbf{x}) = 0$ for all $(\cdot_i, \mathbf{x})$. Hence, we conclude that for both possible values of y_i we end up with contradictions. Thus, the only possibility is that $I_B^{(i)} < 1$. □

We recall from Section 5.1 that $J_B^{(i)}$ can be interpreted as the fraction of all states for the $n-1$ other components where component i is critical. By Definition 5.2.1 $I_B^{(i)}$ is the probability that component i is critical. Thus, $J_B^{(i)}$ is just a special case of $I_B^{(i)}$ corresponding to the situation where all component state vectors have the same probability. This occurs when $P(X_i = 1) = \frac{1}{2}$ for all $i \in C$. Thus, we have the following result

Theorem 5.2.4 *Consider a binary monotone system (C, ϕ) where the component state variables are independent, and where $P(X_i = 1) = \frac{1}{2}$ for all $i \in C$. Then we have:*

$$I_B^{(i)} = J_B^{(i)}$$

We close this section by considering some examples. In all these examples we consider binary monotone systems (C, ϕ) where $C = \{1, \ldots, n\}$, where the component state variables are independent, and where $P(X_i = 1) = p_i$ for all $i \in C$. Without loss of generality we assume that the components are ordered so that:

$$p_1 \leq p_2 \leq \ldots \leq p_n. \tag{5.6}$$

Example 5.2.5 *Let (C, ϕ) be a series system. Then for all $i \in C$ we have:*

$$I_B^{(i)} = \frac{\partial \coprod_{j=1}^{n} p_j}{\partial p_i} = \prod_{j \neq i} p_j.$$

Hence, by the ordering (5.6)*, we get that:*

$$I_B^{(1)} \geq I_B^{(2)} \geq \cdots \geq I_B^{(n)}.$$

Thus, in a series system the worst *component, i.e., the one with the smallest reliability, has the greatest reliability importance.*

Example 5.2.6 *Let (C, ϕ) be a parallel system. Then for all $i \in C$ we have:*

$$I_B^{(i)} = \frac{\partial \coprod_{j=1}^{n} p_j}{\partial p_i} = \prod_{j \neq i} (1 - p_j).$$

Hence, from the ordering (5.6)

$$I_B^{(1)} \leq I_B^{(2)} \leq \cdots \leq I_B^{(n)}.$$

Thus, in a parallel system the best *component, i.e., the one with the greatest reliability, has the greatest reliability importance.*

Example 5.2.7 *Let (C, ϕ) be a 2-out-of-3 system. It is then easy to show that:*

$$h(\boldsymbol{p}) = p_1 p_2 + p_1 p_3 + p_2 p_3 - 2 p_1 p_2 p_3.$$

See Exercise 5.4. This implies that

$$\begin{aligned} I_B^{(1)} &= p_2 + p_3 - 2 p_2 p_3, \\ I_B^{(2)} &= p_1 + p_3 - 2 p_1 p_3, \\ I_B^{(3)} &= p_1 + p_2 - 2 p_1 p_2. \end{aligned}$$

We then consider the function $f(p,q) = p + q - 2pq$ *and note that* $I_B^{(1)} = f(p_2, p_3)$, $I_B^{(2)} = f(p_1, p_3)$, *and* $I_B^{(3)} = f(p_1, p_2)$. *Moreover, the partial derivatives of* f *are respectively:*

$$\frac{\partial f}{\partial p} = 1 - 2q, \qquad \frac{\partial f}{\partial q} = 1 - 2p.$$

If $p, q \leq \frac{1}{2}$, f *is non-decreasing in* p *and* q. *Thus, if* $p_1 \leq p_2 \leq p_3 \leq \frac{1}{2}$, *we have:*

$$f(p_1, p_2) \leq f(p_1, p_3) \leq f(p_2, p_3).$$

Hence, in this case we have:

$$I_B^{(3)} \leq I_B^{(2)} \leq I_B^{(1)}. \tag{5.7}$$

If $p, q \geq \frac{1}{2}$, f *is non-increasing in* p *and* q. *Thus, if* $\frac{1}{2} \leq p_1 \leq p_2 \leq p_3$, *we have:*

$$f(p_2, p_3) \leq f(p_1, p_3) \leq f(p_1, p_2).$$

Hence, in this case we have:

$$I_B^{(1)} \leq I_B^{(2)} \leq I_B^{(3)}. \tag{5.8}$$

However, other orderings are possible as well. Assume e.g., that $p_1 = \frac{1}{2} - z$, $p_2 = \frac{1}{2}$ *and* $p_1 = \frac{1}{2} + z$, *where* $z \in (0, \frac{1}{2})$. *In this case we get:*

$$\begin{aligned}
I_B^{(1)} &= (\frac{1}{2}) + (\frac{1}{2} + z) - 2 \cdot (\frac{1}{2})(\frac{1}{2} + z) = \frac{1}{2}, \\
I_B^{(2)} &= (\frac{1}{2} - z) + (\frac{1}{2} + z) - 2 \cdot (\frac{1}{2} - z)(\frac{1}{2} + z) = \frac{1}{2} + 2z^2, \\
I_B^{(3)} &= (\frac{1}{2} - z) + (\frac{1}{2}) - 2 \cdot (\frac{1}{2} - z)(\frac{1}{2}) = \frac{1}{2},
\end{aligned}$$

Hence in this case we have:

$$I_B^{(1)} = I_B^{(3)} \leq I_B^{(2)}. \tag{5.9}$$

Note that this result holds also if $z \in (-\frac{1}{2}, 0)$ *in which case* $p_1 > p_2 > p_3$.

Example 5.2.7 shows that the reliability ordering of the components depends strongly on how reliable the components are. Thus, predicting the final ranking can often be difficult if we rely on our intuition only. Still some intuitive explanation can be obtained in special cases. Note e.g., that if $p_1 \leq p_2 \leq p_3 \leq \frac{1}{2}$, the components fail with a high probability. In this case the system behaves almost like a series system, which is reflected in (5.7) where the worst component is ranked as the most important component. Thus, in this case we have the same ordering as in Example 5.2.5. Similarly, if $\frac{1}{2} \leq p_1 \leq p_2 \leq p_3$, the components function with a high probability. In this case the system behaves almost like a parallel system, which is reflected in (5.8) where the best component is ranked on top. In this case we have the same ordering as in Example 5.2.6.

5.3 The Barlow-Proschan measure of reliability importance

One disadvantage of the Birnbaum measure for the reliability importance is that it is time dependent. In order to avoid this problem, we consider the following measure introduced by Barlow and Proschan [4]:

Definition 5.3.1 *Consider a binary monotone system (C, ϕ), where the components are never repaired. Moreover, let T_i denote the lifetime of component i, $i \in C$, and let S denote the lifetime of the system. Then for all $i \in C$ the* Barlow-Proschan measure of the reliability importance *of component i is defined as:*

$$\begin{aligned} I_{B-P}^{(i)} &:= P(\textit{Component } i \textit{ fails at the same time as the system}) \\ &= P(T_i = S). \end{aligned}$$

Note that if component i fails at the same time as the system, this may be interpreted as i is the component eventually causing the system to fail.

Before we proceed we need to introduce the concept of *absolute continuity*. We say that a real-valued stochastic variable, T has an absolutely continuous distribution if $P(T \in A) = 0$ for all measurable sets $A \subseteq \mathbb{R}$ such that $m_1(A) = 0$, where m_1 denotes the Lebesgue measure[1] in $\mathbb{R}$. It can be shown that if $T_1, \ldots, T_n$ are independent and absolutely continuously distributed, then the vector $\mathbf{T} = (T_1, \ldots, T_n)$ is absolutely continuously distributed in $\mathbb{R}^n$. That is, $P(\mathbf{T} \in A) = 0$ for all (measurable) sets $A \subseteq \mathbb{R}^n$ such that $m_n(A) = 0$, where m_n denotes the Lebesgue measure in $\mathbb{R}^n$. In particular, if $A = \{\mathbf{t} : t_i = t_j\}$, where $i \neq j$, then $m_n(A) = 0$. Hence, $P(T_i = T_j) = 0$ when $i \neq j$.

The following theorem uses the concept of absolute continuity in order to obatin some fundamental properties of the Barlow-Proschan measure.

Theorem 5.3.2 *Let (C, ϕ) be a non-trivial binary monotone system where the components are never repaired where $C = \{1, \ldots, n\}$, and let T_i denote the lifetime of component i, $i = 1, \ldots, n$, while S denotes the lifetime of the system. Moreover, assume that $T_1, \ldots, T_n$ are independent, absolutely continuously distributed. Then S is absolutely continuously distributed as well, and we have:*

$$\sum_{i=1}^{n} I_{B-P}^{(i)} = 1.$$

[1] The Lebesgue measure in $\mathbb{R}$, denoted m_1, is a function defined on the family of all measurable subsets A of $\mathbb{R}$. If $A = [a, b]$, then $m_1(A) = b - a$. Thus, the Lebesgue measure m_1 generalizes the length of intervals to arbitrary measurable subsets of $\mathbb{R}$. Similarly, the Lebesgue measure in $\mathbb{R}^n$, denoted m_n, is a function defined on the family of all measurable subsets A of $\mathbb{R}^n$. If $A = [a_1, b_1] \times \cdots \times [a_n, b_n]$, then $m_n(A) = \prod_{i=1}^{n}(b_n - a_n)$. Thus, the Lebesgue measure m_n generalizes the volume of n-dimensional rectangular sets to arbitrary measurable subsets of $\mathbb{R}^n$.

Proof: Since we have assumed that the system is non-trivial, the lifetime of the system, S can be expressed as:

$$S = \max_{1 \leq j \leq p} \min_{i \in P_j} T_i, \tag{5.10}$$

where $P_1, \ldots, P_p$ are the minimal path sets of the system. This implies that:

$$P(\bigcup_{i=1}^{n} \{T_i = S\}) = 1. \tag{5.11}$$

Let $A \subseteq \mathbb{R}$ be an arbitrary measurable set such that $m_1(A) = 0$. Since we have assumed that $T_1, \ldots, T_n$ are absolutely coninuously distributed, we get that:

$$0 \leq P(S \in A) \leq P(\bigcup_{i=1}^{n} \{T_i \in A\}) \leq \sum_{i=1}^{n} P(T_i \in A) = 0,$$

which implies that S is absolutely continuously distributed as well.

Since $T_1, \ldots, T_n$ are absolutely continuously distributed, it follows that the probability of having two or more components failing at the same time is zero. This implies e.g., that $P(\{T_i = S\} \cap \{T_j = S\}) = 0$ for $i \neq j$. Thus, when calculating the probability of the union of the events $\{T_i = S\}$, $i = 1, \ldots, n$, all intersections can be ignored as they have zero probability of occurring. Hence, by (5.11) we get:

$$1 = P(\bigcup_{i=1}^{n} \{T_i = S\}) = \sum_{i=1}^{n} P(T_i = S) = \sum_{i=1}^{n} I_{B-P}^{(i)},$$

where the second equality follows by ignoring all intersections of the events $\{T_i = S\}$, $i = 1, \ldots, n$. The last equality follows by the definition of $I_{B-P}^{(i)}$. Hence, the proof is complete. □

The next result provides a convenient formula for computing the Barlow-Proschan measure.

Theorem 5.3.3 *Let (C, ϕ) be a non-trivial binary monotone system where the components are never repaired where $C = \{1, \ldots, n\}$, and let T_i denote the lifetime of component i, $i = 1, \ldots, n$. Moreover, assume that $T_1, \ldots, T_n$ are independent, absolutely continuously distributed with densities $f_1, \ldots, f_n$ respectively. Then, we have:*

$$I_{B-P}^{(i)} = \int_0^\infty I_B^{(i)}(t) f_i(t) dt,$$

where $I_B^{(i)}(t)$ denotes the Birnbaum measure of the reliability importance of component 1 at time t, i.e., the probability that component i is critical at time t.

Proof: From the definitions of the Barlow-Proschan measure and the Birnbaum measure, it follows that:

$$\begin{aligned} I_{B-P}^{(i)} &= P(\text{Component } i \text{ fails at the same time as the system}) \\ &= \int_0^\infty P(\text{Component } i \text{ is critical at time } t) \cdot f_i(t)dt \\ &= \int_0^\infty I_B^{(i)}(t) f_i(t)dt. \end{aligned}$$

□

We observe that according to Theorem 5.3.3 the Barlow-Proschan measure, $I_{B-P}^{(i)}$, can be expressed as a weighted average of the Birnbaum measure $I_B^{(i)}(t)$ with respect to the component lifetime densities, $f_i(t)$. Hence, according to the Barlow-Proschan measure, the points of times where $f_i(t)$ is large are viewed as important. The Barlow-Proschan measure implies that components with long lifetimes compared to the system lifetime will have a large reliability importance.

5.4 The Natvig measure of reliability importance

As mentioned, the Barlow-Proschan measure implies that components with long lifetimes compared to the system lifetime will have a large reliability importance. An alternative measure can be defined by instead saying components which greatly reduce the remaining system lifetime by failing, are the most important components. In order to reflect this Natvig [34] introduced the following measure:

Definition 5.4.1 *Consider a monotone structure ϕ consisting of n components which are not repaired and let*

$Z_i :=$ *reduction of remaining lifetime for the system due to comp. i failing.*

Then, the Natvig measure for the reliability importance of the i'th component*, $I_N^{(i)}$, is defined by*

$$I_N^{(i)} = \frac{E[Z_i]}{\sum_{i=1}^n E[Z_j]}$$

where we assume that $E[Z_i]$ is finite.

From this definition, we see that $0 \leq I_N^{(i)} \leq 1$ and $\sum_{i=1}^n = 1$. It is shown in Natvig [35], that $E[Z_i]$ can be interpreted as the reduction in remaining system lifetime due to the i'th component failing. Natvig [35] also shows the following connection between the Natvig measure and the Birnbaum measure:

Theorem 5.4.2 *Consider a monotone structure ϕ of n components which are not repaired, where the state processes of the components $\{X_i(t), t \geq 0\}_{i=1}^n$ are independent and their lifetimes are absolutely continuously distributed. Then,*

$$E[Z_i] = \int_0^\infty \bar{F}_i(t)(-\ln(\bar{F}_i(t)))I_B^{(i)}(t)dt.$$

We omit the proof, and refer interested readers to Natvig [35] Theorem 3.4.10. Note that Theorem 5.4.2 says that $E[Z_i]$, which is the building block of the Natvig measure, is a weighted average of the Birnbaum measure, $I_B^{(i)}(t)$ with $\bar{F}_i(t)(-\ln(\bar{F}_i(t)))$ as weight (recall the definition of $\bar{F}_i(t)$ in (3.15)). It is shown in Natvig [35] that this weight is the improvement of the survival probability of the i'th component at time t by allowing *one* repair. Hence, the Natvig measure implies that times where this improvement is large are of great importance.

There are also other examples of measures of component reliability importance. One example is the so-called Vesely-Fussell measure, which has been omitted here due to several critical weaknesses (for instance, if a component fails after the system, it still contributes to the measure). For more on this, see Natvig [35].

To conclude our review of reliability importance measures, we remark that no measure can be said to be universially best. They all have strengths and weaknesses, which are important to be aware of when applying such measures in practice. Note for instance that none of the measures presented in these sections take into consideration the cost of improving the different components.

5.5 Exercises

Exercise 5.1: Let ϕ be a k-out-of-n system. Prove that all the components of this system have the same Birnbaum measure for structural importance.

Exercise 5.2: Compute $J_B^{(i)}$ for the components of the bridge structure in Example 3.1.2 and compare their structural importance.

Exercise 5.3: Let the i'th component be in series with the rest of a monotone structure ϕ, while the j'th component is not. Prove that

$$J_B^{(i)} > J_B^{(j)}.$$

Exercise 5.4: Prove that the reliability function of a 2-out-of-3 structure is

$$h(\mathbf{p}) = p_1p_2 + p_1p_3 + p_2p_3 - 2p_1p_2p_3.$$

Exercise 5.5: Compute $I_B^{(i)}$ for the components of the bridge structure in Example 3.1.2 and compare their reliability importance.

Exercise 5.6: Assume that the component lifetimes have so-called *proportional hazards*, that is:

$$\bar{F}_i(t) = \exp(-\lambda_i R(t)), \quad \lambda_i > 0, t \geq 0, \quad i = 1, \ldots, n,$$

where R is a strictly increasing, differentiable function such that $R(0) = 0$, and $\lim_{t\to\infty} R(t) = \infty$. Prove that for a series structure, we have:

$$I_{B-P}^{(i)} = I_N^{(i)} = \frac{\lambda_i}{\sum_{i=1}^n \lambda_j}.$$

Exercise 5.7: Assume that the i'th component is irrelevant for the system ϕ. Then, what is $I_B^{(i)}(t)$, $I_{B-P}^{(i)}$ and $I_N^{(i)}$?

Chapter 6

Association and bounds for the system reliability

In this chapter, we begin by introducing a natural kind of positive dependency between random variables (for instance component states in a system). This positive dependency is called association. We show that associated random variables have some nice properties, which can be used in order to derive upper and lower bounds of the reliability of a system. As previously mentioned, it is not always possible to derive the exact reliability of a system due to e.g. computational time or complexity of the system in question. Nevertheless, we would like to find out something about how reliable a system is. This is the purpose of reliability bounds. In Section 6.2, we will derive several different such bounds, where some assume independence of components, while others just assume that the component states are associated.

6.1 Associated random variables

So far, we have usually assumed that the components in a system are independent, and hence that the corresponding random variables are independent. However, in many real-life situations, this is not the case. Some examples of such dependencies are components which are exposed to a common source of outer stress (for instance parts of a windmill) and structures where several components share a common load, so if one component fails, the load on the other components increases (for instance parts of a bridge or a house). In this section, we define a concept of non-negative dependence between random variables called association.

A fairly natural definition of non-negative association is that two random variables S and T are associated if and only if $\mathrm{Cov}(S,T) \geq 0$. Though this definition is intuitive, the actual definition of association is stricter and applicable

to any number of random variables.

Definition 6.1.1 *Let $T_1, \ldots, T_n$ be random variables, and let $\boldsymbol{T} = (T_1, \ldots, T_n)$. We say that $T_1, \ldots, T_n$ are* associated *if*

$$\mathrm{Cov}(\Gamma(\boldsymbol{T}), \Delta(\boldsymbol{T})) \geq 0,$$

for all binary non-decreasing functions Γ and Δ.

Note that according to this definition we only require $\mathrm{Cov}(\Gamma(\mathbf{T}), \Delta(\mathbf{T})) \geq 0$ for all *binary* non-decreasing functions. This property may seem somewhat strange. It turns out, however, that association implies the following apparently stronger result:

Theorem 6.1.2 *Let $T_1, \ldots, T_n$ be associated random variables, and f and g functions which are non-decreasing in each argument such that $\mathrm{Cov}(f(\boldsymbol{T}), g(\boldsymbol{T}))$ exists, i.e.,*

$$E[|f(\boldsymbol{T})|] < \infty, E[|g(\boldsymbol{T})|] < \infty, E[|f(\boldsymbol{T})g(\boldsymbol{T})|] < \infty.$$

Then we have:

$$\mathrm{Cov}(f(\boldsymbol{T}), g(\boldsymbol{T})) \geq 0.$$

We skip the proof here and refer to Barlow and Proschan [5] for details.

Associated random variables have some very useful properties. In order to establish these properties we need the following lemma:

Lemma 6.1.3 *Let X, Y and Z be random variables, and assume that $\mathrm{Cov}(X, Y) < \infty$. Then we have:*

$$\mathrm{Cov}(X, Y) = E[\mathrm{Cov}(X, Y|Z)] + \mathrm{Cov}[E(X|Z), E(Y|Z)].$$

Proof: By the definition of covariance we know that:

$$\mathrm{Cov}(X, Y) = E(XY) - E(X)E(Y).$$

We then apply the law of total expectation and condition with respect to the variable Z. This yields:

$$\mathrm{Cov}[X, Y] = E[E(XY|Z)] - E[E(X|Z)]E[E(Y|Z)]$$

Now we rewrite the first term using the definition of covariance once again:

$$\begin{aligned} E[E(XY|Z)] &= E[E(XY|Z) - E(X|Z)E(Y|Z)] + E[E(X|Z)E(Y|Z)] \\ &= E[\mathrm{Cov}(X, Y|Z)] + E[E(X|Z)E(Y|Z)] \end{aligned}$$

Inserting this in the expression for $\mathrm{Cov}(X, Y)$ we get:

$$\begin{aligned} \mathrm{Cov}[X, Y] &= E[\mathrm{Cov}(X, Y|Z)] + E[E(X|Z)E(Y|Z)] - E[E(X|Z)]E[E(Y|Z)] \\ &= E[\mathrm{Cov}(X, Y|Z)] + \mathrm{Cov}[E(X|Z), E(Y|Z)], \end{aligned}$$

which concludes the proof of the lemma. □

Theorem 6.1.4 *Associated random variables have the following properties:*
(i) Any subset of a set of associated random variables also consists of associated random variables.
(ii) A single random variable is always associated.
(iii) Non-decreasing functions of associated random variables are associated.
(iv) If two sets of associated random variables are independent, then their union is a set of associated random variables.

Proof: We note that *(i)* follows from *(iii)*. However, we can also prove this property directly. Thus, let $T_1, \ldots, T_n$ be a set of associated random variables, and let $A \subset \{1, \ldots, n\}$. We would like to prove that $\{T_i\}_{i \in A}$ is a set of associated random variables. To do so, let Γ_A, Δ_A be arbitrary, binary functions which are non-decreasing in all of their arguments T_i, $i \in A$. We then define:

$$\Gamma(\mathbf{T}) = \Gamma_A(\mathbf{T}^A), \quad \Delta(\mathbf{T}) = \Delta_A(\mathbf{T}^A).$$

From this it follows that:

$$\mathrm{Cov}(\Gamma_A(\mathbf{T}^A), \Delta_A(\mathbf{T}^A)) = \mathrm{Cov}(\Gamma(\mathbf{T}), \Delta(\mathbf{T})) \geq 0,$$

where the inequality follows from Definition 6.1.1 because we have assumed that $T_1, \ldots, T_n$ is a set of associated random variables. Hence, *(i)* is proved.

To prove *(ii)* we let T be a random variable, and let Γ, Δ be arbitrary, binary functions which are non-decreasing in T. Then, since Γ, Δ are binary and non-decreasing in T, there are two possible cases:

Case 1. $\Gamma(T) \leq \Delta(\mathbf{T})$ for all T,

Case 2. $\Gamma(T) \geq \Delta(\mathbf{T})$ for all T.

We consider Case 1 only, as Case 2 can be handled similarly. Then, we have:

$$\begin{aligned}
\mathrm{Cov}(\Gamma(T), \Delta(T)) &= E[\Gamma(T)\Delta(T)] - E[\Gamma(T)]E[\Delta(T)] \\
&= E[\Gamma(T)] - E[\Gamma(T)]E[\Delta(T)] \\
&= E[\Gamma(T)](1 - E[\Delta(T)]) \geq 0,
\end{aligned}$$

where the second equality follows from the fact that we are in case 1. This means that either $\Gamma(T) = \Delta(T)$, in which case (since Γ, Δ are binary) $E[\Gamma(T)\Delta(T)] = E[\Gamma(T)]$. Or, alternatively, $\Gamma(T) = 0$, $\Delta(T) = 1$, in which case $E[\Gamma(T)\Delta(T)] = 0 = E[\Gamma(T)]$. The inequality holds because $\Gamma(T), \Delta(T) \in \{0, 1\}$ for all T. Hence, *(ii)* is proved as well.

We now turn to the proof of *(iii)* and let $T_1, \ldots, T_n$ be associated, and let $\mathbf{T} = (T_1, \ldots, T_n)$. Moreover, we let $S_i = f_i(\mathbf{T})$, $i = 1, \ldots, m$, where $f_1, \ldots, f_m$ are non-decreasing functions, and let $\mathbf{S} = (S_1, \ldots, S_m)$. Finally, let $\Gamma = \Gamma(\mathbf{S})$ and $\Delta = \Delta(\mathbf{S})$ be binary non-decreasing functions. Then $\Gamma(\mathbf{S}) = \Gamma(f_1(\mathbf{T}), \ldots, f_m(\mathbf{T}))$ and $\Delta(\mathbf{S}) = \Delta(f_1(\mathbf{T}), \ldots, f_m(\mathbf{T}))$ are non-decreasing functions of $\mathbf{T}$ as well. Hence, by Definition 6.1.1 it follows that:

$$\mathrm{Cov}(\Gamma(\mathbf{S}), \Delta(\mathbf{S})) = \mathrm{Cov}(\Gamma(f_1(\mathbf{T}), \ldots, f_m(\mathbf{T})), \Delta(f_1(\mathbf{T}), \ldots, f_m(\mathbf{T}))) \geq 0.$$

Hence, we conclude that $S_1, \ldots, S_m$ are associated as well.

Finally, we prove (iv). Thus, we let $\mathbf{X}$ and $\mathbf{Y}$ be two vectors of associated random variables, and assume that $\mathbf{X}$ and $\mathbf{Y}$ are independent of each other. Moreover, we assume that $\Gamma = \Gamma(\mathbf{X}, \mathbf{Y})$ and $\Delta = \Delta(\mathbf{X}, \mathbf{Y})$ are binary and non-decreasing functions in both $\mathbf{X}$ and $\mathbf{Y}$. Then by Lemma 6.1.3 we have:

$$\mathrm{Cov}(\Gamma, \Delta) = E[\mathrm{Cov}(\Gamma, \Delta|\mathbf{X})] + \mathrm{Cov}[E(\Gamma|\mathbf{X}), E(\Delta|\mathbf{X})]. \tag{6.1}$$

We then note that for any $\mathbf{x}$, $\Gamma(\mathbf{x}, \mathbf{Y})$ and $\Delta(\mathbf{x}, \mathbf{Y})$ are binary non-decreasing functions of $\mathbf{Y}$. Hence, we must have:

$$\mathrm{Cov}(\Gamma(\mathbf{X}, \mathbf{Y}), \Delta(\mathbf{X}, \mathbf{Y})|\mathbf{X} = \mathbf{x}) = \mathrm{Cov}(\Gamma(\mathbf{x}, \mathbf{Y}), \Delta(\mathbf{x}, \mathbf{Y})) \geq 0, \text{ for all } \mathbf{x},$$

where the equality follows since $\mathbf{Y}$ is independent of $\mathbf{X}$, while the inequality follows since $\mathbf{Y}$ is associated. This implies that:

$$E[\mathrm{Cov}(\Gamma, \Delta|\mathbf{X})] \geq 0. \tag{6.2}$$

Moreover, $E[\Gamma(\mathbf{x}, \mathbf{Y})]$ and $E[\Delta(\mathbf{x}, \mathbf{Y})]$ are non-decreasing (but not necessarily binary) functions of $\mathbf{x}$. Hence, since $\mathbf{X}$ is associated, it follows by Theorem 6.1.2 that:

$$\mathrm{Cov}[E(\Gamma|\mathbf{X}), E(\Delta|\mathbf{X})] \geq 0. \tag{6.3}$$

Note that since Γ and Δ are binary, we must have that $E(\Gamma|\mathbf{X}) \in [0, 1]$ and $E(\Gamma|\mathbf{X}) \in [0, 1]$ with probability one. Thus, obviously $\mathrm{Cov}[E(\Gamma|\mathbf{X}), E(\Delta|\mathbf{X})]$ exists.

Combining (6.1), (6.2) and (6.3) implies that:

$$\mathrm{Cov}(\Gamma, \Delta) \geq 0.$$

Thus, we conclude that $(\mathbf{X}, \mathbf{Y})$ is associated. □

Theorem 6.1.5 *Let $T_1, \ldots, T_n$ be independent random variables. Then, they are also associated.*

Proof: The proof is by induction on n. The result obviously holds for $n = 1$ by Theorem 6.1.4 (ii). Assume that the theorem holds for $n = m - 1$. That is, $\{T_1, \ldots, T_{m-1}\}$ is a set of associated random variables. Moreover, by Theorem 6.1.4 (ii), $\{T_m\}$ is associated as well. By the assumption, these two sets are independent. Hence, it follows from Theorem 6.1.4 (iv) that their union $\{T_1, \ldots, T_{m-1}, T_m\}$ is a set of associated random variables. Thus, the result is proved by induction. □

At the completely opposite end of the scale, we have the following result:

Theorem 6.1.6 *Let $T_1, \ldots, T_n$ be completely positively dependent random variables, i.e.,*

$$P(T_1 = T_2 = \ldots = T_n) = 1.$$

Then they are associated.

Proof: Let Γ, Δ be binary functions which are non-decreasing in each argument and let $\mathbf{T} = (T_1, \ldots, T_n)$ and $\mathbf{T}_1 = (T_1, \ldots, T_1)$. By the assumption it follows that $\mathbf{T}$ and $\mathbf{T}_1$ must have the same distribution. Hence, we get that:

$$\mathrm{Cov}(\Gamma(\mathbf{T}), \Delta(\mathbf{T})) = \mathrm{Cov}(\Gamma(\mathbf{T}_1), \Delta(\mathbf{T}_1)) \geq 0$$

where the final inequality follows by of Theorem 6.1.4 (*ii*) and Definition 6.1.1. □

In the next section, we will establish upper and lower bounds for the system reliability. In order to do this, the following theorem is useful:

Theorem 6.1.7 *Let $X_1, \ldots, X_n$ be the associated or independent component state variables of a monotone system (C, ϕ). Moreover, let the minimal path series structures of the system be $(P_1, \rho_1), \ldots, (P_p, \rho_p)$, where:*

$$\rho_j(\boldsymbol{X}^{P_j}) = \coprod_{i \in P_j} X_i, \quad j = 1, \ldots, p. \tag{6.4}$$

Then, $\rho_1, \ldots, \rho_p$ are associated. Similarly, let the minimal cut parallel structures of the system be $(K_1, \kappa_1), \ldots, (K_k, \kappa_k)$, where:

$$\kappa_j(\boldsymbol{X}^{K_j}) = \coprod_{i \in K_j} X_i, \quad j = 1, \ldots, k. \tag{6.5}$$

Then, $\kappa_1, \ldots, \kappa_k$ are associated.

Proof: The result follows since $\rho_1, \ldots, \rho_p$ in (6.4) and $\kappa_1, \ldots, \kappa_k$ in (6.5) are non-decreasing functions of associated random variables, and hence associated by Theorem 6.1.4, (*iii*). □

The following result extends Theorem 6.1.4, (*iii*) to non-increasing functions as well.

Theorem 6.1.8 *Let $\boldsymbol{T} = (T_1, \ldots, T_n)$ be associated, and let:*

$$U_i = g_i(\boldsymbol{T}), \quad i = 1, \ldots, m,$$

where g_i, $i = 1, \ldots, m$ are non-increasing functions. Then, $\boldsymbol{U} = (U_1, \ldots, U_m)$ is associated.

Proof: Let Γ, Δ be binary non-decreasing functions, and introduce $\mathbf{g}(\mathbf{T}) = (g_1(\mathbf{T}), \ldots, g_m(\mathbf{T}))$. Thus, $\mathbf{U} = \mathbf{g}(\mathbf{T})$. We then introduce

$$\bar{\Gamma}(\mathbf{T}) = 1 - \Gamma(\mathbf{g}(\mathbf{T})) = 1 - \Gamma(\mathbf{U})$$

$$\bar{\Delta}(\mathbf{T}) = 1 - \Delta(\mathbf{g}(\mathbf{T})) = 1 - \Delta(\mathbf{U})$$

It follows directly that $\bar{\Gamma}$ and $\bar{\Delta}$ are binary and non-decreasing in T_i, $i = 1, \ldots, n$. Hence, by the assumption that $\mathbf{T}$ is associated, it follows that:

$$\begin{aligned}\mathrm{Cov}(\Gamma(\mathbf{U}), \Delta(\mathbf{U}))) &= \mathrm{Cov}(1 - \bar{\Gamma}(\mathbf{T}), 1 - \bar{\Delta}(\mathbf{T})) \\ &= \mathrm{Cov}(1,1) + \mathrm{Cov}(1, -\bar{\Delta}(\mathbf{T})) + \mathrm{Cov}(-\bar{\Gamma}(\mathbf{T}), 1) \\ &\quad + \mathrm{Cov}(-\bar{\Gamma}(\mathbf{T}), -\bar{\Delta}(\mathbf{T})) \\ &= \mathrm{Cov}(\bar{\Gamma}(\mathbf{T}), \bar{\Delta}(\mathbf{T})) \geq 0.\end{aligned}$$

Hence, we conclude that $\mathbf{U}$ is associated. □

In order to check whether two random variables are associated, the following result can sometimes be used. It states that for two binary random variables, association and non-negative covariance are equivalent.

Theorem 6.1.9 *Let X and Y be two binary random variables. Then, X and Y are associated if and only if*

$$\mathrm{Cov}(X, Y) \geq 0.$$

Proof: Assume first that X and Y are associated. We may then choose $\Gamma(X, Y) = X$ and $\Delta(X, Y) = Y$. Since obviously Γ and Δ are binary and non-decreasing functions, it follows from Definition 6.1.1 that $\mathrm{Cov}(X, Y) = \mathrm{Cov}(\Gamma, \Delta) \geq 0$.

We then assume conversely that $\mathrm{Cov}(X, Y) \geq 0$. If can prove that this implies that $\mathrm{Cov}(\Gamma, \Delta) \geq 0$ for all binary non-decreasing functions, the result is proved. Since there are only two variables in this case, X and Y, there are only a very limited number of choices for Γ and Δ. In fact, the only choices are:

$$\Gamma_1 \equiv 0, \quad \Gamma_2 = X \cdot Y, \quad \Gamma_3 = X, \quad \Gamma_4 = Y, \quad \Gamma_5 = X \amalg Y, \quad \Gamma_6 \equiv 1.$$

Moreover, we have the following partial ordering:

$$\Gamma_1 \leq \Gamma_2 \leq \left\{ \begin{array}{c} \Gamma_3 \\ \Gamma_4 \end{array} \right\} \leq \Gamma_5 \leq \Gamma_6.$$

Assume first that Γ and Δ from the set $\{\Gamma_1, \ldots, \Gamma_6\}$ such that $\Gamma(X, Y) \leq \Delta(X, Y)$. We then have:

$$\begin{aligned}\mathrm{Cov}(\Gamma, \Delta) &= E(\Gamma \cdot \Delta) - E(\Gamma) \cdot E(\Delta) \\ &= E(\Gamma) - E(\Gamma) \cdot E(\Delta) = E(\Gamma)[1 - E(\Delta)] \geq 0.\end{aligned}$$

The only possibility left is $\Gamma = \Gamma_3 = X$ and $\Delta = \Gamma_4 = Y$. However, in this case we get that:

$$\mathrm{Cov}(\Gamma, \Delta) = \mathrm{Cov}(X, Y) \geq 0,$$

where the last inequality follows by the assumption. Hence, we conclude that $\mathrm{Cov}(\Gamma, \Delta) \geq 0$ for all binary non-decreasing functions, and thus the result is proved. □

6.2 Upper and lower bounds for the reliability of monotone systems

As mentioned, it is often the case that we are unable to compute the exact reliability of a system. This may be the case for large, complex systems where the computations simply take too much time, but also for systems where the components are not independent (recall that we assumed independence of the components throughout Section 4). In this section, we derive several different upper and lower bounds for the system reliability. These bounds can be useful whenever computing the exact reliability is impossible or too computationally costly.

Note that in this section, we will not always assume that the component state variables are independent. However, to get useful bounds, we still need to assume something about the dependencies between the components. In most of the results of this section, this assumption will be that the component state variables are associated (see Section 6.1).

The following theorem is a key result for deriving reliability bounds.

Theorem 6.2.1 *Let $T_1, \ldots, T_n$ be associated random variables such that $0 \leq T_i \leq 1$, $i = 1, \ldots, n$. Then,*

$$E[\prod_{i=1}^{n} T_i] \geq \prod_{i=1}^{n} E[T_i] \tag{6.6}$$

$$E[\coprod_{i=1}^{n} T_i] \leq \coprod_{i=1}^{n} E[T_i] \tag{6.7}$$

Proof: We note that since $0 \leq T_i \leq 1$, both T_i and $S_i = 1 - T_i$ are non-negative random variables, $i = 1, \ldots, n$. Hence, the product functions $\prod_{i=1}^{n} T_i$ and $\prod_{i=1}^{n} S_i$ are both non-decreasing in each argument.

By using Theorem 6.1.2, we find:

$$E[\prod_{i=1}^{n} T_i] - E[T_1]E[\prod_{i=2}^{n} T_i] = \mathrm{Cov}(T_1, \prod_{i=2}^{n} T_i) \geq 0,$$

since the product function is non-decreasing in each argument because $T_i \geq 0$, $i = 2, \ldots, n$. This implies that:

$$E[\prod_{i=1}^{n} T_i] \geq E[T_1]E[\prod_{i=2}^{n} T_i].$$

By repeated use of this inequality, we get (6.6).

From Theorem 6.1.8, $S_1, \ldots, S_n$ are associated random variables. Moreover, $0 \leq S_i \leq 1$, $i = 1, \ldots, n$, so we can apply (6.6) to these variables. From this it

follows that:

$$\begin{aligned} E[\coprod_{i=1}^{n} T_i] &= 1 - E[\prod_{i=1}^{n}(1 - T_i)] = 1 - E[\prod_{i=1}^{n} S_i] \\ &\leq 1 - \prod_{i=1}^{n} E(S_i) = 1 - \prod_{i=1}^{n}(1 - E[T_i]) \\ &= \coprod_{i=1}^{n} E[T_i], \end{aligned}$$

so (6.7) is proved as well.

□

Theorem 6.2.1 has an important interpretation: If we apply the theorem to the binary component state variables $X_1, \ldots, X_n$, (6.6) says that for a series structure of associated components, an incorrect assumption of independence will lead to an *underestimation* of the system reliability. Correspondingly, (6.7) says that for a parallel structure, an incorrect assumption of independence between the components will lead to an *overestimation* of the system reliability. Clearly, overestimating the system reliability can have serious consequences in applications, and should be avoided. Since most systems are not purely series or purely parallel, we conclude that for an arbitrary structure, we *cannot say for certain what the consequences of an incorrect assumption of independence will be.* This means that in any application where we do not know for sure that the components are independent, simply assuming independence in order to use one of the methods in Section 4 to compute the exact system reliability can be dangerous. Fortunately, it is still possible to obtain bounds on the system reliability.

In the next results we interpret T_i, $i = 1, \ldots, n$ as the lifetimes of components in a binary monotone system.

Theorem 6.2.2 *Let $T_1, \ldots, T_n$ be non-negative associated random variables. Then, for all t_i, $i = 1, \ldots, n$:*

$$P[\cap_{i=1}^{n}(T_i > t_i)] \geq \prod_{i=1}^{n} P[T_i > t_i], \tag{6.8}$$

$$P[\cap_{i=1}^{n}(T_i \leq t_i)] \geq \prod_{i=1}^{n} P[T_i \leq t_i]. \tag{6.9}$$

Proof: Interpreting T_i as the lifetime of the ith component in a binary monotone system, we may introduce the corresponding binary component state process $\{X_i(t)\}$, $i = 1, \ldots, n$, where:

$$X_i(t) = \begin{cases} 1, & t < T_i \\ 0, & t \geq T_i. \end{cases}$$

In Figure 6.1 we have plotted $X_i(t)$ as a function of $t \geq 0$ for a given value of T_i. The plot shows that $X_i(t)$ is a non-increasing function of t. For a fixed point of time t we may alternatively view $X_i(t)$ as a function of the lifetime T_i. The resulting plot is shown in Figure 6.2. From this plot we see that $X_i(t)$ is a non-decreasing function of T_i. Hence, by Theorem 6.1.4, $X_1(t), \ldots, X_n(t)$ are associated for all $t \geq 0$.

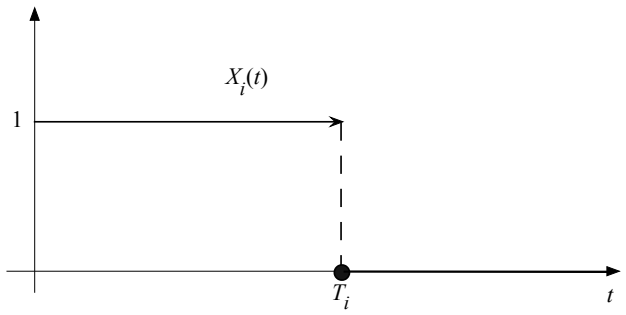

Figure 6.1: The binary component state process $X_i(t)$ plotted as a function of the time $t \geq 0$.

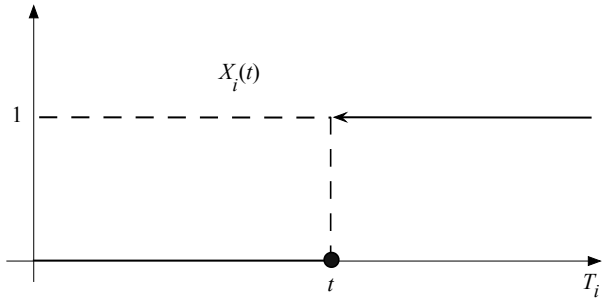

Figure 6.2: The binary component state process $X_i(t)$ plotted as a function of the lifetime T_i for a given $t \geq 0$.

Hence, by applying (6.6):

$$\begin{aligned} P[\cap_{i=1}^{n}(T_i > t_i)] &= P[\cap_{i=1}^{n}(X_i(t_i) = 1)] \\ &= P[\prod_{i=1}^{n} X_i(t_i) = 1] = E[\prod_{i=1}^{n} X_i(t_i)] \\ &\geq \prod_{i=1}^{n} E[X_i(t_i)] = \prod_{i=1}^{n} P[X_i(t_i) = 1] \\ &= \prod_{i=1}^{n} P[T_i > t_i], \end{aligned}$$

so (6.8) is proved. The proof of (6.9) is similar by using (6.7) instead of (6.6) and left as an exercise to the reader.

□

Note that the left hand side of (6.9) is the joint cumulative distribution of $T_1, \ldots, T_n$, while the right hand side is the corresponding cumulative distribution in the case where $T_1, \ldots, T_n$ are independent. As a corollary to Theorem 6.2.2, we have:

Corollary 6.2.3 *Let $T_1, \ldots, T_n$ be non-negative associated random variables. Then:*

$$P(\min_{1\leq i\leq n} T_i > t) \geq \prod_{i=1}^{n} P(T_i > t), \tag{6.10}$$

$$P(\max_{1\leq i\leq n} T_i > t) \leq \coprod_{i=1}^{n} P(T_i > t). \tag{6.11}$$

Proof: This follows from Theorem 6.2.2 by letting $t_i = t$ for $i = 1, \ldots, n$. Alternatively, this can also be proved by applying Theorem 6.2.1 to the binary component states at time t, since $\min_{1\leq i\leq n} T_i$ is the lifetime of a series system and $\max_{1\leq i\leq n} T_i$ is the lifetime of a parallel system. □

Now, we are ready to derive our first (crude) bounds of the reliability of a monotone system. In words, these bounds say that for any monotone system, the reliability is lower bounded by the reliability of a series system and upper bounded by the reliability of a parallel system. This corresponds to our previous observation (see the comments after Theorem 2.2.4) that the series system is the *worst* non-trivial monotone system, while the parallel system is the *best*.

Theorem 6.2.4 *Let $X_1, \ldots, X_n$ be the component state variables of a non-trivial binary monotone structure (C, ϕ) with component reliabilities $p_1, \ldots, p_n$, and assume that $X_1, \ldots, X_n$ are associated. Then we have:*

$$\prod_{i=1}^{n} p_i \leq h \leq \coprod_{i=1}^{n} p_i.$$

Proof: We get:

$$\begin{aligned}\prod_{i=1}^{n} p_i &= \prod_{i=1}^{n} E[X_i] \\ &\leq E[\prod_{i=1}^{n} X_i] \leq E[\phi(\mathbf{X})] \\ &\leq E[\coprod_{i=1}^{n} X_i] \leq \coprod_{i=1}^{n} E[X_i] \\ &= \coprod_{i=1}^{n} p_i.\end{aligned}$$

Here, the first inequality follows from (6.6), the second and third inequalities from Theorem 2.2.4 and the fourth inequality follows from (6.7). □

As mentioned, these bounds are very crude, and usually not of much practical use. Luckily, we are always able to establish better bounds than those of Theorem 6.2.4 by including information about the minimal path and cut sets of the system. These improved bounds are a corollary to the following result:

Theorem 6.2.5 *Consider a monotone structure (C, ϕ) with minimal path sets $P_1, \ldots, P_p$, and minimal cut sets $K_1, \ldots, K_k$. Then we have:*

$$\max_{1 \leq j \leq p} P[\min_{i \in P_j} X_i = 1] \leq h \leq \min_{1 \leq j \leq k} P[\max_{i \in K_j} X_i = 1].$$

Proof: From (3.8) and (3.10), we get

$$\min_{i \in P_r} X_i \leq \max_{1 \leq r \leq p} \min_{i \in P_r} X_i = \phi(\mathbf{X}) = \min_{1 \leq s \leq k} \max_{i \in K_s} X_i \leq \max_{i \in K_s} X_i,$$

for all $r = 1, \ldots, p$ and all $s = 1, \ldots, k$. This implies that:

$$P(\min_{i \in P_r} X_i = 1) \leq h \leq P(\max_{i \in K_s} X_i = 1)$$

for all $r = 1, \ldots, p$ and all $s = 1, \ldots, k$. This proves the theorem. □

In Theorem 6.2.5, we made no assumptions about the dependence between the components. In order to obtain an explicit expression for the lower bound, we need to compute $P(\min_{i \in P_r} X_i = 1)$, $j = 1, \ldots, p$. Similarly, to obtain an explicit expression for the upper bound, we need to compute $P(\max_{i \in K_r} X_i = 1)$, $j = 1, \ldots, k$. This is difficult without making further assumptions about the joint distribution of the component state variables. In the following corollary we make the assumption that the component state variables are associated. This enables us to obtain explicit bounds.

Corollary 6.2.6 *Consider the same monotone system (C, ϕ) as in Theorem 6.2.5. Moreover, assume that the component state variables are associated with component reliabilities $p_1, \ldots, p_n$. Then we have:*

$$\max_{1\leq j\leq p} \prod_{i\in P_j} p_i \leq h \leq \min_{1\leq j\leq k} \coprod_{i\in K_j} p_i.$$

Proof: The result follows immediately from Theorem 6.2.5 by using (6.6) and (6.7) because:

$$\prod_{i\in P_j} p_i \leq E[\prod_{i\in P_j} X_i] = P[\min_{i\in P_j} X_i = 1]$$

and

$$\coprod_{i\in K_j} p_i \geq E[\coprod_{i\in K_j} X_i] = P[\max_{i\in K_j} X_i = 1].$$

□

The assumptions in Theorem 6.2.4 and Corollary 6.2.6 are the same. However, the bounds in Corollary 6.2.6 are obviously better than those in Theorem 6.2.4 because:

$$\prod_{i=1}^{n} p_i = \prod_{i\in P_j} p_i \prod_{i\notin P_j} p_i \leq \prod_{i\in P_j} p_i \leq \max_{1\leq j\leq p} \prod_{i\in P_j} p_i$$

since $p_i \in [0, 1]$ for $i = 1, \ldots, n$. A similar argument proves that the upper bound in Corollary 6.2.3 is better than the one in Theorem 6.2.4.

If we in addition to the assumptions in Corollary 6.2.6 assume that the components are independent, we get bounds which are sometimes better than those in Corollary 6.2.6 and sometimes worse. These bounds are a corollary to the following theorem:

Theorem 6.2.7 *Let $X_1, \ldots, X_n$ be the associated component states of a binary monotone system (C, ϕ) with minimal path series structures $(P_1, \rho_1), \ldots, (P_p, \rho_p))$ and minimal cut parallel structures $(K_1, \kappa_1), \ldots, (K_k, \kappa_k))$. Then,*

$$\prod_{j=1}^{k} P(\kappa_j(\boldsymbol{X}^{K_j}) = 1) \leq h \leq \coprod_{j=1}^{p} P(\rho_j(\boldsymbol{X}^{P_j}) = 1). \tag{6.12}$$

Proof: From the comments after Theorem 6.1.4, it follows from Theorem 6.1.4 (*iii*) that the minimal path series structures, and the minimal cut parallel struc-

tures, are associated. Hence, we get that:

$$\begin{aligned}\prod_{j=1}^{k} P(\kappa_j(\mathbf{X}^{K_j}) = 1) &\leq E[\prod_{j=1}^{k} \kappa_j(\mathbf{X}^{K_j}] \\ &= h \\ &= E[\coprod_{j=1}^{p} \rho_j(\mathbf{X}^{P_j})] \\ &\leq \coprod_{j=1}^{p} P(\rho_j(\mathbf{X}^{P_j}) = 1),\end{aligned}$$

where the first inequality follows from (6.6), the first and second equalities follow from (3.10) and (3.8) respectively and the final inequality follows from (6.7). □

The problem with Theorem 6.2.7 is the same as the problem with Theorem 6.2.5. Without having explicit expressions for the reliabilities of the path series structures and cut parallel structures the bounds are difficult to use. However, by assuming independent component state variables, we can derive the following corollary.

Corollary 6.2.8 *Let (C, ϕ) be a binary monotone system of independent component states and where the component reliabilities are $p_1, \ldots, p_n$. Let $P_1, \ldots, P_n$ and $K_1, \ldots, K_n$ be respectively the minimal path and cut sets of the system. Then we have:*

$$\prod_{j=1}^{k} \coprod_{i \in K_j} p_i \leq h(\boldsymbol{p}) \leq \coprod_{j=1}^{p} \prod_{i \in P_j} p_i. \tag{6.13}$$

Proof: For independent components, the lower bound in (6.13) is equal to the one in (6.12) because

$$P(\kappa_j(\mathbf{X}^{K_j}) = 1) = E[\coprod_{i \in K_j} X_i] = \coprod_{i \in K_j} p_i.$$

The upper bound in (6.13) is proved in the same way. □

We conclude this section with an example.

Example 6.2.9 *Consider a 3-out-of-4 system with $p_i = p$, $i = 1, 2, 3, 4$. We assume that all the component state variables are independent. We shall compare the bounds from Corollary 6.2.6 to those from Corollary 6.2.8.*

The minimal path sets for a 3-out-of-4 system are:

$$\{1, 2, 3\}, \{1, 2, 4\}, \{1, 3, 4\}, \{2, 3, 4\},$$

and the minimal cut sets are:

$$\{1, 2\}, \{1, 3\}, \{1, 4\}, \{2, 3\}, \{2, 4\}, \{3, 4\}.$$

In the calculations below we will repeatedly use the formula for the reliability of a parallel system of two components with independent component state variables. Denoting the reliability of this system by h_2, and assumin that the components have reliability q, we have:

$$h_2(q) = q \amalg q = 1 - (1-q)(1-q) = 2q - q^2.$$

We denote the lower and upper bounds in Corollary 6.2.6 by $l_1(p)$ and $u_1(p)$ respectively. Since all the components have the same reliability, these bounds are given by:

$$l_1(p) = \max_{1 \le j \le 4} \coprod_{i \in P_j} p = \max_{1 \le j \le 4} p^3 = p^3,$$
$$u_1(p) = \min_{1 \le j \le 6} \coprod_{i \in K_j} p = \min_{1 \le j \le 6} h_2(p) = 2p - p^2.$$

Similarly, we denote the bounds in Corollary 6.2.8 by $l_2(p)$ and $u_2(p)$ respectively. Again since all the components have the same reliability, and since we have assumed that the component state variables are independent, these bounds are given by:

$$l_2(p) = \coprod_{j=1}^{6} \coprod_{i \in K_j} p = \coprod_{j=1}^{6} h_2(p) = (2p - p^2)^6,$$
$$u_2(p) = \coprod_{j=1}^{4} \prod_{i \in P_j} p = h_2(p^3) \amalg h_2(p^3) = 2(2p^3 - p^6) - (2p^3 - p^6)^2.$$

From Example **??**, *we find that the reliability of the* 3*-out-of-*4 *system is:*

$$h(p) = 4p^3(1-p) + p^4.$$

The bounds $l_1(p)$, $u_1(p)$, $l_2(p)$, $u_2(p)$, as well as the reliability function $h(p)$, are plotted in Figure 6.3.

From Figure 6.3, we see that in this case, the bounds from Corollary 6.2.8 are better than those from Corollary 6.2.6. However, note that the lower bound from Corollary 6.2.6 is actually best for small values of p, while the opposite is true for larger p-values. Correspondingly, the upper bound from Corollary 6.2.6 is best for large p-values, while the upper bound from Corollary 6.2.8 is best for smaller values of p.

By defining a new lower bound as the largest of the two lower bounds from Corollary 6.2.6 and Corollary 6.2.8, and a new upper bound as the smallest of the two upper bounds from the same corollaries, we get a new set of bounds which is better than the previous ones. It is also possible to improve these bounds further by using a modular decomposition of the system at hand. We will not go into details on this, but refer to Natvig [35] for more on this topic.

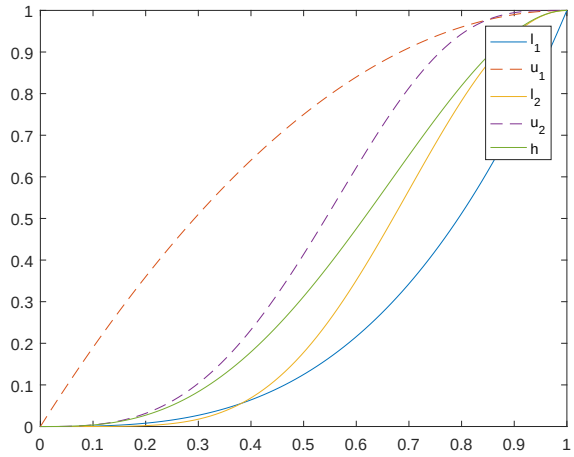

Figure 6.3: Illustration of the bounds in Example 6.2.9.

6.3 Exercises

Exercise 6.2.1: Prove (6.9) of Theorem 6.2.2 by using equation (6.7).

Exercise 6.2.2: Prove that the upper bound in Corollary 6.2.6 is better than the one in Theorem 6.2.4.

Exercise 6.2.3: Prove the upper bound of Corollary 6.2.8.

Exercise 6.2.4: Prove the upper bound in Corollary 6.2.6 by applying the lower bound on the dual structure function ϕ^D.

Exercise 6.2.5: Prove the upper bound in Corollary 6.2.8 by applying the lower bound on the dual structure function ϕ^D.

Exercise 6.2.6: Consider the network in Figure 6.4 consisting of independent components with component reliabilities p.

a) What is the reliability $h(p)$ for this system?

b) For $p \in [0, 1]$, write a program to compute the reliability $h(p)$. Plot $h(p)$.

c) In the same plot, illustrate the bounds from Corollary 6.2.6 and 6.2.8. Comment on the result.

d) Is it possible to improve these bounds further?

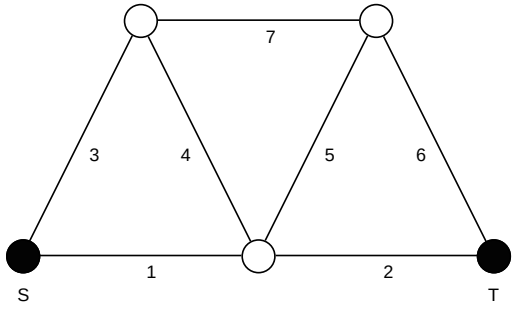

Figure 6.4: Illustration of the network for Exercise 6.2.6.

Chapter 7

Computing reliability using conditional Monte Carlo methods

As a result of the availability of high-speed computers, the use of Monte Carlo methods have accelerated. In many cases, however, it is necessary to use various clever tricks in order to make the methods converge faster. In particular this is true in situations where one needs to estimate something that involves events with very low probabilities. One such area is system reliability estimation. Since failure events often have very low probability, a large number of simulations is needed in order to obtain stable results. Sometimes, however, conditioning can be used to improve the convergence.

7.1 Monte Carlo simulation and conditioning

In the general case the principle of conditioning can be stated as follows: Let $\mathbf{X} = (X_n, \ldots, X_n)$ be a random vector with a known distribution, and assume that we want to calculate the expected value of $\phi = \phi(\mathbf{X})$. We denote this expectation by h. Assume furthermore that the distribution of ϕ cannot be derived analytically in polynomial time with respect to n. Using Monte Carlo simulations, however, we can proceed by generating a sample of size N, of independent vectors $\mathbf{X}_1, \ldots, \mathbf{X}_N$, all having the same distribution as $\mathbf{X}$, and then estimate h with the simple *Monte Carlo* estimate:

$$\hat{h}_{MC} = \frac{1}{N} \sum_{r=1}^{N} \phi(\mathbf{X}_r). \tag{7.1}$$

The variance of this estimate is given by:

$$\mathrm{Var}\left(\hat{h}_{MC}\right) = \frac{1}{N^2} \sum_{r=1}^{N} \mathrm{Var}(\phi) = \frac{1}{N} \mathrm{Var}(\phi) \tag{7.2}$$

Now, assume that $S = S(\mathbf{X})$ is some other function whose distribution can be calculated analytically in polynomial time with respect to n. More specifically, we assume that S is a *discrete* variable with values in the set $\{s_1, \ldots, s_k\}$. We also introduce:

$$\theta_j = E[\phi \mid S = s_j], \quad j = 1, \ldots, k. \tag{7.3}$$

By the formula for conditional expectation we then have:

$$h = \sum_{j=1}^{k} E[\phi \mid S = s_j] P(S = s_j) = \sum_{j=1}^{k} \theta_j P(S = s_j) \tag{7.4}$$

Instead of generating N samples from the distribution of $\mathbf{X}$ as in (7.1), we divide the set into k groups, one for each θ_j, where the j-th group has size N_j, $j = 1, \ldots, k$, and $N_1 + \cdots + N_k = N$. The data in the j-th group are sampled from the conditional distribution of $\mathbf{X}$ given that $S = s_j$, and are used to estimate the conditional expectation θ_j, $j = 1, \ldots, k$. Denoting the data in the j-th group by $\{\mathbf{X}_{r,j} : r = 1, \ldots, N_j\}$, θ_j is estimated by:

$$\hat{\theta}_j = \frac{1}{N_j} \sum_{r=1}^{N_j} \phi(\mathbf{X}_{r,j}), \quad j = 1, \ldots, k. \tag{7.5}$$

The variances of these estimates are:

$$\mathrm{Var}(\hat{\theta}_j) = \frac{1}{N_j^2} \sum_{r=1}^{N_j} \mathrm{Var}(\phi | S = s_j) = \frac{1}{N_j} \mathrm{Var}(\phi | S = s_j), \quad j = 1, \ldots, k. \tag{7.6}$$

By combining $\hat{\theta}_1, \ldots, \hat{\theta}_k$, we get the *conditional Monte Carlo estimate*:

$$\hat{h}_{CMC} = \sum_{j=1}^{k} \hat{\theta}_j P(S = s_j). \tag{7.7}$$

Since $\hat{\theta}_1, \ldots, \hat{\theta}_k$ are independent, the variance of the conditional Monte Carlo estimate is:

$$\begin{aligned} \mathrm{Var}(\hat{h}_{CMC}) &= \mathrm{Var}[\sum_{j=1}^{k} \hat{\theta}_j P(S = s_j)] = \sum_{j=1}^{k} \mathrm{Var}(\hat{\theta}_j)[P(S = s_j)]^2 \\ &= \sum_{j=1}^{k} \frac{1}{N_j} \mathrm{Var}(\phi | S = s_j)[P(S = s_j)]^2. \end{aligned}$$

We observe that the variance of the conditional estimate depends on the choices of the N_j's. Ideally we would like N_j to be large if the product of the conditional variance and the squared probability is large, and small otherwise, as this would yield the most efficient partition of the total sample with respect to minimizing the total variance. In practice, however, this is difficult, since the

conditional variance is typically not known. In order to compare the result with the variance of the original Monte Carlo estimate, we instead let:

$$N_j \approx N \cdot P(S = s_j), \quad j = 1, \ldots, k,$$

so that:

$$\sum_{j=1}^{k} N_j \approx \sum_{j=1}^{k} N \cdot P(S = s_j) = N \cdot \sum_{j=1}^{k} P(S = s_j) = N.$$

With this choice we get:

$$\begin{aligned} Var(\hat{h}_{CMC}) &\approx \sum_{j=1}^{k} \frac{1}{N \cdot P(S = s_j)} \operatorname{Var}(\phi|S = s_j)[P(S = s_j)]^2 \qquad (7.8) \\ &= \frac{1}{N} \sum_{j=1}^{k} \operatorname{Var}(\phi|S = s_j) P(S = s_j) = \frac{1}{N} E[\operatorname{Var}(\phi|S)] \\ &= \frac{1}{N} \big(\operatorname{Var}(\phi) - \operatorname{Var}[E(\phi|S)] \big) \leq \frac{1}{N} \operatorname{Var}(\phi) = \operatorname{Var}(\hat{h}_{MC}). \end{aligned}$$

From the above equation we see that the conditional estimate has smaller variance than the original Monte Carlo estimate provided that $\operatorname{Var}[E(\phi \mid S)]$ is positive. This quantity can be interpreted as a measure of how much information S contains relative to ϕ. Thus, when looking for good choices for S we should look for variables containing as much information about ϕ as possible. However, there are some other important points that need to be considered. First of all S must have a distribution that can be derived analytically in polynomial time. Next, the number of possible values of S, i.e., k, must be polynomially limited by n. Finally, it must be possible to sample efficiently from the distribution of $\mathbf{X}$ given S.

The problem of sampling from conditional distributions has been studied recently by several authors. Lindqvist and Taraldsen [32] proposes a general method for sampling from conditional distributions in cases where S is a sufficient statistic. Of special relevance to the present paper, is Broström and Nilsson [8] who consider the problem of sampling independent Bernoulli variables given their sum. See also Nilsson [38]. In the remaining part of this paper we shall focus on applications in reliability theory, where $\mathbf{X}$ typically is a vector of independent Bernoulli variables. In relation to this we shall derive our own conditional sampling methods.

We now apply the above ideas in the context of system reliability. That is, we assume that $\mathbf{X}$ is a vector of independent Bernoulli variables, interpreted as the component state vector relative to a system of n components. A component is *failed* if its state variable is 0, and *functioning* if the state variable is 1. The function ϕ is also assumed to take values 0, 1, and interpreted as the *system state*. If ϕ is 0, the system is *failed*, and if ϕ is 1, the system is *functioning*. The function ϕ is referred to as the structure function of the system, and is assumed

to be a nondecreasing function of the component state vector, $\mathbf{X}$. In general the calculation of ϕ for a given component state vector depends very much on the representation of this function. If e.g., the system is represented as some sort of communication network its state can usually be evaluated in $O(n)$ time or better. If on the other hand the system is specified in terms of a list of minimal path (or cut) sets, the evaluation can be very slow since the number of such sets in the worst cases grows exponentially with the number of components. For this study, however, we assume that ϕ can be calculated in $O(n)$ time for a given component state vector.

There are in general many different choices for the variable S which can be used in a reliability setting. In the next sections we will consider one possible choice. For other choices see Huseby et al. [24].

7.2 Conditioning on the sum of the component state variables

For the remaining part of the paper we focus on the case where S is the sum of all the component state variables. This approach was taken in in Naustdal [37]. In this situation the random variable S takes values in the set $\{0, 1, \ldots, n\}$. Thus, the uncertain quantities we need to estimate, are:

$$\theta_s = E[\phi | S = s], \qquad s = 0, 1, \ldots, n. \tag{7.9}$$

As in the previous section, we need to have an efficient way of sampling from the conditional distribution of $\mathbf{X}$ given $S = s$, $s = 0, 1, \ldots, n$. In relation to this we need the use of the partial sums, $S_1, \ldots, S_n$, defined as:

$$S_m = \sum_{i=m}^{n} X_i, \quad m = 1, \ldots, n.$$

Using these quantities, we can develop the algorithm for sampling from the distribution of $\mathbf{X}$ given $S = s$. The idea is simply to start out by sampling X_1 from the conditional distribution of $X_1 \mid S = s$. We then continue by sampling X_2 from $X_2 \mid S = s$, $X_1 = x_1$, where x_1 denotes the sampled outcome of X_1, and so on. This turns out to be easy noting that:

$$\begin{aligned}
&P(X_m = x_m \mid X_1 = x_1, \ldots, X_{m-1} = x_{m-1}, S = s) \\
&\quad = \frac{P(X_m = x_m, S = s \mid X_1 = x_1, \ldots, X_{m-1} = x_{m-1})}{P(S = s \mid X_1 = x_1, \ldots, X_{m-1} = x_{m-1})} \\
&\quad = \frac{P(X_m = x_m, S_{m+1} = s - \sum_{j=1}^{m} x_j)}{P(S_m = s - \sum_{j=1}^{m-1} x_j)} \\
&\quad = \frac{p_m^{x_m}(1 - p_m)^{1-x_m} P(S_{m+1} = s - \sum_{j=1}^{m} x_j)}{P(S_m = s - \sum_{j=1}^{m-1} x_j)},
\end{aligned}$$

Assuming that the distributions of $S_1, \ldots, S_n$ are all calculated before running the simulations, we see that all the necessary conditional probabilities can be calculated when needed during the simulations without imposing additional computational complexity. Note that we only calculate those conditional probabilities we really need along the way during the simulations, not the entire set of all possible conditional probabilities corresponding to all possible combinations of values of the X_j's. Thus, in each simulation run we calculate n probabilities, one for each X_j. Moreover, each probability can be calculated using a fixed number of operations (independent of n). Hence, it follows that sampling from the conditional distribution of $\mathbf{X}$ given $S = s$, can be done in $O(n)$ time.

The CMC-algorithm can now be summarized as follows: Divide the N simulations into $(n+1)$ groups, one for each θ_s, where the s-th group has size N_s, $s = 0, 1, \ldots, n$. The data in the s-th group is sampled from the conditional distribution of $\mathbf{X}$ given that $S = s$, and is used to estimate the conditional expectation θ_s, $s = 0, 1, \ldots, n$. Denoting the data in the s-th group by $\{\mathbf{X}_r, s : r = 1, \ldots, N_s\}$, θ_s is estimated essentially by using (7.5), and combined into the CMC-estimate by using (7.7).

A remaining problem is of course how to choose the sample sizes for each of the $(n+1)$ groups. One possibility is to proceed as we did in Section 1 and let $N_s \approx N \cdot P(S = s)$, $s = 0, 1, \ldots, n$. In this case, however, it is possible to improve the results slightly. As in the previous section we denote the size of the smallest path set by d, and the size of the smallest cut set by c. By examining the system, it is often easy to determine d and c. Given these two numbers it is easy to see that $\theta_s = 0$ for $s < d$, and $\theta_s = 1$ for $s > n - c$. Moreover, $\mathrm{Var}(\phi \mid S = s) = 0$ for $s < d$, or $s > n - c$. Hence, there is no point in spending simulations on estimating θ_s for $s < d$, or $s > n - c$, so we let $N_s = 0$ for $s < d$, or $s > n - c$. As a result, we have more simulations to spend on the remaining quantities.

An extreme situation occurs when the system is a k-out-of-n-system, i.e., when $\phi(\mathbf{X}) = I(S \geq k)$. For such systems $d = k$, and $c = (n - k + 1)$. In this situation *all* the θ_s's are known. Specifically, $\theta_s = 0$ for $s < k$ and $\theta_s = 1$ for $s \geq k$. Thus, the CMC-estimate isequal to the true value of the reliability, and can be calculated without doing any simulations at all. The reason for this is of course that in this case ϕ depends on $\mathbf{X}$ only through S.

For all other nontrivial systems, however, it is easy to see that we always have that $d \leq n - c$. In order to ensure that we get improved results, we assume that $P(d \leq S \leq n - c) > 0$, and let:

$$N_s \approx N \cdot P(S = s)/P(d \leq S \leq n - c), \quad s = d, \ldots, n - c\,. \tag{7.10}$$

Using a similar argument as we did in (7.8) we now get that:

$$\mathrm{Var}(\hat{h}_{CMC}) \leq P(d \leq S \leq n - c)\,\mathrm{Var}(\hat{h}_{MC}) \tag{7.11}$$

Hence, we see that if d and $(n - c)$ are close, i.e., there are few unknown θ_s's, the variance is reduced considerably.

7.3 System reliability when all the component state variables have identical reliabilities

If all the components in the system have the same reliability, i.e., $p_1 = \cdots = p_n = p$, it is possible to improve things even further. In this case we note that S has a binomial distribution. Moreover, the conditional distribution of $\mathbf{X}$ given S is given by:

$$P(\mathbf{X} = \mathbf{x} \mid S = s) = \frac{p^{\sum_{i=1}^n x_i}(1-p)^{n-\sum_{i=1}^n x_i}}{\binom{n}{s} p^s (1-p)^{n-s}} = \frac{1}{\binom{n}{s}}\,, \tag{7.12}$$

for all $\mathbf{x}$ such that $\sum_{i=1}^n x_i = s$. From this it follows that:

$$\theta_s = E[\phi \mid S = s] = \sum_{\left\{\mathbf{x} \mid \sum_{i=1}^n x_i = s\right\}} \phi(\mathbf{x}) P(\mathbf{X} = \mathbf{x} \mid S = s) = \frac{b_s}{\binom{n}{s}}\,, \tag{7.13}$$

where b_s = the number of path sets with s components, $s = 0, \ldots, n$. Finally, the system reliability, h, expressed as a function of p, is given by:

$$\begin{aligned} h(p) &= \sum_{s=0}^n \theta_s P(S = s) \\ &= \sum_{s=0}^n \frac{b_s}{\binom{n}{s}} \binom{n}{s} p^s (1-p)^{n-s} = \sum_{s=0}^n b_s p^s (1-p)^{n-s}. \end{aligned}$$

Note that the unknown quantities, $\theta_0, \ldots, \theta_n$, do not depend on p. Thus, by estimating these quantities, we get an estimate of the entire $h(p)$-function for all $p \in [0, 1]$. Moreover, we see that θ_s can be interpreted as the fraction of path sets of size s among all sets of size s, for $s = 0, 1, \ldots, n$. Thus, θ_s can be estimated by sampling random sets of size s and calculating the frequency of path sets among the sampled sets. It turns out that this can be done very efficiently as follows:

The idea is to sample sequences of sets of increasing size by sampling from the component set, $C = \{1, \ldots, n\}$, without replacements. The only set of size 0 is of course $\emptyset$. If $\emptyset$ is a path set, we have a trivial system which is always functioning. Obviously, $\theta_0 = 1$ in this case. Moreover, the reliability of such a system is 1. If $\emptyset$ is not a path set, it follows that $\theta_0 = 0$. In both cases we do not need to sample anything to determine the value of θ_0. Thus, we can focus on estimating $\theta_1, \ldots, \theta_n$ by sampling sets of size s, for $s = 1, \ldots, n$.

Algorithm 7.3.1 *Let $\boldsymbol{T} = (T_1, \ldots, T_n)$ be a vector for storing results from the simulations. Before we start the simulations, this vector is initialized as $(0, \ldots, 0)$. Then for $i = 1, \ldots, N$ do:*

1. *Sample a component from the set C, say component i_1, and define $A_1 = \{i_1\}$. If A_1 is a path set, T_1 is incremented with 1.*

2. *Sample a component from the set* $C \setminus A_1$*, say component* i_2*, and define* $A_2 = \{i_1, i_2\}$*. If* A_2 *is a path set,* T_2 *is incremented with 1.*

   ...

n. *Sample the last remaining component, say component* i_n*, and define* $A_n = C$*. If* A_n *is a path set,* T_n *is incremented with 1.*

When all the simulations are carried out, the vector T *contains the number of observed path sets of sizes* $1, \ldots, n$*. From this we get the resulting estimates of* $\theta_1, \ldots, \theta_n$ *simply as:*

$$\theta_s = T_s/N, \quad s = 1, \ldots, n.$$

It is easy to see that sampling components randomly without replacements is equivalent to sampling the components according to a random permutation, $(i_1, \ldots, i_n)$, of the component set $\{1, \ldots, n\}$. Such a permutation can easily be generated using a well-known algorithm described in Knuth [31]. This algorithm works as follows:

Assume that we start out with a vector representing an initial arbitrary permutation of the components, say $(i_1, \ldots, i_n)$. One may e.g., simply use $(1, \ldots, n)$. We then generate a random permutation vector by modifying this initial permutation vector by running through n steps. In the k-th step we consider the element currently being the k-th element of the vector. We then sample a random index, j, in the interval $k, \ldots, n$, and let the k-th and the j-th elements switch places in the permutation vector. When all the n steps are completed, it is easy to see that the resulting vector is indeed a random permutation, regardless of the initial state of the vector.

By using this algorithm, we are able to generate a complete sequence of n sets in just $O(n)$ time. However, since all sets in a sequence are derived from the same permutation, the θ_s-estimates become correlated. Still, the effect of this is more than compensated for since each θ_s-estimate now can be based on all N simulations, compared to just a subset of size N_s as was the case in the previous section.

When running this algorithm, much of the time is spent on the steps where the system state is evaluated. Thus, in order to minimize the running time of the algorithm, one would like to reduce the number of such evaluations. Since $A_1 \subset A_2 \subset \cdots \subset A_n$, it follows by the monotonicity of the structure function, ϕ, that if A_k is a path set, then so is A_{k+1}. Thus, we can stop evaluating the state of the system as soon as we have generated a path set, since we then know that all the remaining sets will be path sets as well. Still it is obviously important to carry out the system state evaluations as efficient as possible. The methodology for doing this may typically depend strongly on the representation of the given system.

In order to show how this can be done, we consider two different classes of systems. The first class we consider is the class of *threshold systems* (see Example

4.0.1), i.e., binary monotone systems (C, ϕ) where the structure function can be written in the following form:

$$\phi(\mathbf{X}) = I\Big(\sum_{i=1}^{n} a_i X_i \geq b\Big), \tag{7.14}$$

where $a_1, \ldots, a_n$ and b are positive numbers. The a_i's are called the *weights* associated with the components, while the number b is called the *threshold value*. Notice that if $a_1 = \cdots = a_n = 1$ and $b = k$, we get a standard k-out-of-n-system. For such systems it is very easy to carry out all the system state evaluations. To do so we keep track of the sum of the weights associated with the sampled components. When a new component is sampled, the sum is updated by adding the weight associated with this component. When the sum of weights is greater than or equal to the threshold value, we know that we have generated a path set. Using this method, each system evaluation is carried out in constant time. Thus, in this case the total computational complexity of the simulation algorithm is just $O(nN)$, where n is the number of components in the system, and N is the number of simulations.

The second class is the class of undirected network systems (see Section 4.5). The components of such a system are the edges of an undirected network. An edge is functioning if its *end-nodes* can "communicate" through it. The system is functioning if a subset of the nodes (with K elements), called the *terminals*, can communicate through the network.

If all the edges are functioning, we assume that the network is connected, i.e., all nodes can communicate with each other. If only a subset of the edges are functioning, however, the node set is partitioned into a set of equivalence classes such that a pair of nodes can communicate if and only if they belong to the same equivalence class. Moreover, the system is functioning if and only if all the terminals belong to the same equivalence class.

In order to minimize the running time spent on system state evaluation, we maintain lists of nodes belonging to each equivalence class as we sample the edges. For each list we also keep track of the number of terminals contained in the list. When a new edge is sampled, we investigate its end-nodes. If they belong to the same equivalence class, the system state is not changed. On the other hand, if the end-nodes belong to different equivalence classes, we merge the two classes and calculate the number of terminals included in the merged class. When we arrive at an equivalence class containing K terminals, (i.e., all the terminals), we know that we have generated a path set.

The most time-consuming part of this algorithm is the merging of equivalence classes. Assuming that the classes are stored as linked lists, the actual merging of the lists can be done in constant time. However, in order to minimize the time spent on checking whether or not two nodes belong to the same class, the nodes should be marked with a reference to its class list. When two classes, say A and B, are merged, all the nodes in the smallest class, say B must be marked with a reference to the class list of A instead. Updating such references can be done in $O(v)$ time, where v is the number of nodes in the network. It

should be noted, however, that this is a very pessimistic estimate, as in most cases the number of updating operations is much smaller. Anyway, the total computational complexity of the simulation algorithm is $O(nvN)$ in this case.

We close this section by illustrating the effect of the improved method on a specific example. The system we consider is the 2-terminal undirected network system illustrated in Figure 7.1. We assume that all the 7 components have the same reliability p, and estimate the system reliability $h(p)$ for all $p \in [0, 1]$. Moreover, we let $N = 100$. With so few iterations it is easier to see how much better the conditional Monte Carlo method is compared to crude Monte Carlo method.

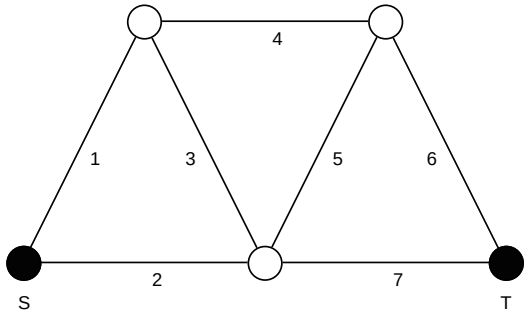

Figure 7.1: A 2-terminal undirected network system

For the conditional Monte Carlo method we know that we get an estimate of the entire $h(p)$-function by estimating the fractions $\theta_1, \ldots, \theta_n$. For the crude Monte Carlo estimate one possible approach is to proceed by obtaining estimates of $h(p)$ for each individual value of p. This would imply that we would need $N = 100$ iterations for each chosen value of p. Thus, with a suitable resolution of p-values, say $p \in \{0.01, 0.02, \ldots, 1.00\}$ we need $100N = 10000$ samples. Fortunately, it is possible to avoid this issue. In order to do so, the following sampling method can be used. In the jth iteration we generate uniformly distributed random variables $U_{1j}, \ldots, U_{7j}$. For a given value of $p \in [0, 1]$ we then define:

$$X_{ij}(p) = I(U_{ij} \leq p), \quad i = 1, \ldots, 7, \quad j = 1, \ldots N,$$

where $I(U_{ij} \leq p) = 1$ if $U_{ij} \leq p$ and zero otherwise. This implies that:

$$P(X_{ij}(p) = 1) = P(U_{ij} \leq p) = p, \quad i = 1, \ldots, 7, \quad j = 1, \ldots N.$$

We denote the structure function of the system by ϕ. Moreover, we denote the jth sampled component state vector, expressed as a function of the common component reliability p, by $\mathbf{X}_j(p) = (X_{1,j}(p), \ldots, X_{7,j}(p))$, $j = 1, \ldots, N$. We

can then estimate $h(p)$ as follows:

$$\hat{h}(p) = \frac{1}{N} \sum_{j=1}^{N} \phi(\mathbf{X}_j(p)).$$

By varying p in the interval $[1, 0]$, we obtain an estimate of $h(p)$ for all p. Hence, we can use the *same* sample of uniformly distributed variables to obtain an estimate of the system reliability for each value of p. Thus, we only need $N = 100$ samples for this method as well, which saves us a lot of time. Moreover, it is easy to see that $\mathbf{X}_j(p)$ is a non-decreasing function of p for all j. Hence, since ϕ is a non-decreasing function of $\mathbf{X}_j$, the estimated curve, $\hat{h}(p)$, is non-decreasing as well, a property which the true function $h(p)$ also has.

Using a program called *CMCsim*[1] we can run the conditional Monte Carlo simulations. This program has an intuitive graphical user interface, and can be used to estimate reliability and component importance of any undirected network system.

In Figure 7.2 we have plotted the two estimates for $h(p)$ along with the true reliability function. The red curve is the crude Monte Carlo estimate, the green curve is the conditional Monte Carlo curve, while the blue curve represents the true reliability function. As we can see, the green and the blue curves are almost identical. Thus, even with as few as $N = 100$ iterations, we get a very precise estimate of the entire $h(p)$-function. The crude Monte Carlo estimate is far more irregular and deviates considerably from the true curve.

[1] CMCsim is a java program developed at the Department of Mathematics, University of Oslo. The program is freely available at http://www.riscue.org/cmcsim/.

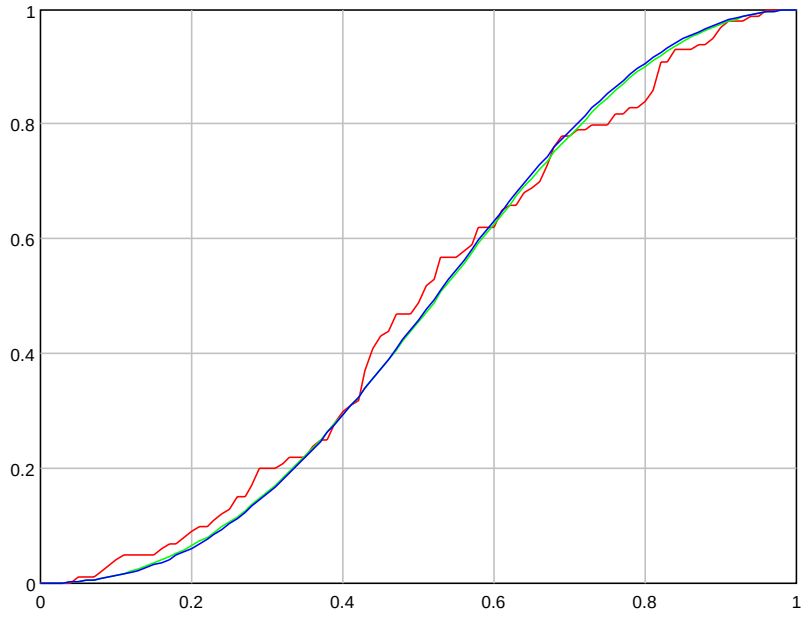

Figure 7.2: System reliability as a function of the common component reliability p. Crude Monte Carlo estimate (red curve), conditional Monte Carlo estimate (green curve) and true reliability (blue curve).

Chapter 8

Discrete event simulation

Discrete event models are typically used in simulation studies to model and analyze *pure jump processes*. For an extensive introduction to discrete event models we refer to Glasserman [18]. See also Klebaner [30]. A discrete event model can be viewed as a system consisting of a collection of stochastic processes, where the states of the individual processes change as results of various kinds of *events* occurring at random points of time. Between these events the states of the processes are assumed to be constant. We refer to the processes included in the collection, as the *elementary processes* of the system. In our context we always assume that each event only affects *one* of the elementary processes.

Asymptotic system availability and component criticality, can easily be estimated by running a single discrete event simulation on the system over a sufficiently long time horizon, or by working directly on the stationary component availabilities. Sometimes, however, one needs to estimate how these properties evolve over time. In such cases it is necessary to run many simulations to obtain stable availability and criticality estimates. One possible approach to this problem is to sample the system state at fixed points of time, and then use the mean values of the states at these points as curve estimates. With this approach all information about the process between the sampling points is thrown away. In the present paper we propose an alternative sampling procedure where we utilize process data between the sampling points as well. This simulation method is particularly useful when estimating various kinds of component importance measures for repairable systems.

8.1 Pure jump processes

In this section we formalize the consept of a pure jump process. Thus, let $\{S(t)\}$ be a stochastic process where $S(t)$ denotes the state of the process at time $t \geq 0$. Moreover, we let $T_1 < T_2 < \cdots$ denote the points of time of the events affecting the process, and let $T_0 = 0$. Following Klebaner [30] a pure jump process is a

process where $S(t)$, can be written as:

$$S(t) = S(0) + \sum_{j=1}^{\infty} I(T_j \leq t) J_j, \qquad t \geq 0, \tag{8.1}$$

where $I(\cdot)$ denotes the indicator function, and J_j denotes the change in the state (positive or negative) of the process at time T_j. The representation (8.1) implies that the state function $S(t)$ is piecewise constant and right-continuous in t, with jumps at $T_1 < T_2 < \cdots$. In particular, for $k = 0, 1, \ldots$, we have:

$$S(t) = S(0) + \sum_{j=1}^{k} J_j = S(T_k), \qquad \text{for all } t \in [T_k, T_{k+1}). \tag{8.2}$$

Thus, in order to keep track of how the process evolves and update the value of the state function, only the points of time where the events happen need to be considered. This property is convenient during simulations.

The infinite sum in (8.1) indicates that the number of events occurring in the interval $[0, t]$ is unbounded. The possibility of having an infinite number of events in $[0, t]$, however, may cause various technical difficulties. In particular, this may cause simulations to break down since an infinite number of events need to be generated and handled. See Glasserman [19] for a further discussion of this issue. To avoid these difficulties, we always assume that the number of events occurring in any finite interval is finite with probability one. A pure jump process satisfying this assumption is said to be *regular*. Note, however, that the regularity assumption does not imply the existence of a fixed number $N < \infty$ such that the number of events in the interval $[0, t]$ is bounded by N with probability one. The number of events in $[0, t]$ will typically be a random variable with a distribution that ranges over all non-negative integers. When the simulation procedure is chosen, this must be taken into account.

We now present a few basic results on regularity of pure jump processes which we need later. We consider a pure jump process $\{S(t)\}$ with jumps at $T_1 < T_2 < \cdots$. We also introduce the times between the events defined as:

$$\Delta_j = T_j - T_{j-1}, \qquad j = 1, 2, \ldots. \tag{8.3}$$

Using these quantities the event times can be expressed as:

$$T_k = \sum_{j=1}^{k} \Delta_j, \qquad k = 1, 2, \ldots. \tag{8.4}$$

Obviously, the process $\{S(t)\}$ is regular if and only if $T_\infty = \infty$ almost surely. Thus, it follows that a necessary and sufficient criterion for regularity is that the series $\sum_{j=1}^{\infty} \Delta_j$ is *divergent* with probability one. This condition can often be verified using the following simple result:

Proposition 8.1.1 *Let $\{S(t)\}$ be a pure jump process with jumps at $T_1 < T_2 < \cdots$. Moreover, we let $T_0 = 0$ and introduce the non-negative random variables $\Delta_j = T_j - T_{j-1}$, $j = 1, 2, \ldots$. Assume then that the sequence $\{\Delta_j\}$ contains an infinite subsequence $\{\Delta_{k_j}\}$ of independent, identically distributed random variables such that $E[\Delta_{k_j}] = d > 0$. Then S is regular.*

Proof: By the strong law of large numbers it follows that:

$$P(\lim_{n\to\infty} n^{-1} \sum_{j=1}^{n} \Delta_{k_j} = d) = 1.$$

This implies that the series $\sum_{j=1}^{\infty} \Delta_{k_j}$ is divergent with probability one. Hence, since obviously $\sum_{j=1}^{\infty} \Delta_{k_j} \leq \sum_{j=1}^{\infty} \Delta_j$, the result follows. □

The regularity property implies that the set of points where the process jumps almost surely does not have any accumulation points. The following result utilizes this to show the existence of left limits of the state function of a regular pure jump process.

Proposition 8.1.2 *Let $\{S(t)\}$ be a regular pure jump process with jumps at $T_1 < T_2 < \cdots$. Then $\lim_{t\to s^-} S(t)$ exists for every $s > 0$ with probability one.*

Proof: Let $0 \leq t < s < \infty$. We then consider the set $\mathcal{T} = \{T_j : t \leq T_j < s\} \cup \{t\}$. Since S is assumed to be regular, the number of elements in $\mathcal{T}$ is finite with probability one. Moreover, $\mathcal{T}$ is non-empty since $t \in \mathcal{T}$. Thus, this set contains a maximal element, which we denote by t'. Moreover, since every element in $\mathcal{T}$ is less than s, then so is t'. From this it follows that the interval (t', s) is nonempty. At the same time (t', s) does not contain any jumps, so $S(t)$ is constant throughout this interval. Hence, $\lim_{t\to s^-} S(t)$ exists. Since s was arbitrary chosen, this holds for any $s > 0$ □

Regularity is also of importance when considering the integral of a pure jump process:

Proposition 8.1.3 *Let $\{S(t)\}$ be a regular pure jump process with jumps at $T_1 < T_2 < \cdots$, and let $0 \leq u < v < \infty$. Assume that $\{T_j : u < T_j < v\} = \{T^{(1)}, \ldots, T^{(k)}\}$, where $T^{(1)} < \cdots < T^{(k)}$. Moreover, we let $T^{(0)} = u$ and $T^{(k+1)} = v$. Then we have:*

$$\int_u^v S(t)dt = \sum_{j=0}^{k} S(T^{(j)})(T^{(j+1)} - T^{(j)}).$$

Proof: We first note that since S is assumed to be regular, the number of elements in the set $\{T_j : u < T_j < v\}$ is finite with probability one. Thus, this set can almost surely be written in the form $\{T^{(1)}, \ldots, T^{(k)}\}$, for some suitable

$k < \infty$. Since S is right-continuous and piecewise constant, it follows that $S(t) = S(T^{(j)})$ for all $t \in [T^{(j)}, T^{(j+1)})$, $j = 0, 1, \ldots, k$. Thus, we have:

$$\int_{T^{(j)}}^{T^{(j+1)}} S(t)dt = S(T^{(j)})(T^{(j+1)} - T^{(j)}), \quad j = 0, 1, \ldots, k.$$

The result then follows by adding up the contributions to the integral from each of the $k+1$ intervals $[T^{(0)}, T^{(1)}), \ldots, [T^{(k)}, T^{(k+1)})$ □

We then consider a system consisting of a collection of n regular pure jump processes, $S_1, \ldots, S_n$. The state of the system is then typically expressed as a function of the states of the elementary processes. It is easy to see that the system state also evolves as a regular pure jump process. That is, we have:

Proposition 8.1.4 *Let $\{S_1(t)\}, \ldots, \{S_n(t)\}$ be n regular pure jump processes, and let $H(t) = H(\boldsymbol{S}(t))$, where $\boldsymbol{S}(t) = (S_1(t) \ldots, S_n(t))$, $t \geq 0$. Then $\{H(t)\}$ is a regular pure jump process as well. That is, $H(t) = H(\boldsymbol{S}(t))$ is piecewise constant and right-continuous in t, and the number of jumps in any finite interval is finite with probability one.*

Proof: Let $\mathcal{T}_i$ be the set of time points corresponding to the jumps of the process S_i, $i = 1, \ldots, n$, and let $\mathcal{T}$ be the set of time points corresponding to the jumps of the process $\{H(t)\}$. Since the state value of H cannot change unless there is a change in the state value of at least one of the elementary processes, it follows that $\mathcal{T} \subseteq (\mathcal{T}_1 \cup \cdots \cup \mathcal{T}_n)$. Thus, $H(t)$ is piecewise constant and right-continuous in t. Moreover, for any finite interval $[t, s]$ we also have:

$$\mathcal{T} \cap [t, s] \subseteq [(\mathcal{T}_1 \cap [t, s]) \cup \cdots \cup (\mathcal{T}_n \cap [t, s])].$$

Since by regularity $(\mathcal{T}_i \cap [t, s])$ is finite for $i = 1, \ldots, n$, it follows that $\mathcal{T} \cap [t, s]$ is finite as well. Hence, we conclude that $\{H(t)\}$ is regular. □

8.2 Binary monotone system of repairable components

As we shall see, stationary statistical properties of a pure jump process, can often be calculated by working directly on the stationary probability distributions of the elementary processes. Sometimes, however, one needs to estimate how the expected value of the process evolves over time. In such cases it is often necessary to use some sort of Monte Carlo simulation method.

In order to illustrate how this can be done we consider a binary monotone system (C, ϕ) of n *repairable* components. The elementary processes are the component state processes, denoted $\{X_1(t)\}, \ldots, \{X_n(t)\}$, where $X_i(t)$ represents the state of component i at time $t \geq 0$, $i = 1, \ldots, n$. That is, $X_i(t) = 1$ if

component i is functioning at time t, and zero otherwise, $i = 1, \ldots, n$. We also introduce the component state vector at time $t \geq 0$, $\mathbf{X}(t) = (X_1(t), \ldots, X_n(t))$. Thus, $\phi = \phi(\mathbf{X}(t)) = 1$ if the system is functioning at time t and zero otherwise.

We now introduce the lifetimes and repair times for the components. That is, for $i = 1, \ldots, n$ and $j = 1, 2, \ldots$ we let:

$$U_{ij} = \text{The } j\text{th lifetime of the } i\text{th component.} \tag{8.5}$$

$$D_{ij} = \text{The } j\text{th repair time of the } i\text{th component.} \tag{8.6}$$

We assume that U_{ij} has an absolutely continuous distribution F_i with a positive mean value $\mu_i < \infty$, while D_{ij} has an absolutely continuous distribution G_i with a positive mean value $\nu_i < \infty$, $i = 1, \ldots, n$, $j = 1, 2, \ldots$. All lifetimes and repair times are assumed to be independent. Thus, in particular the component states $X_1(t), \ldots, X_n(t)$ are independent for each $t \geq 0$.

We then introduce the *availability* of the ith component at time t, denoted $A_i(t)$, as the probability that the component is functioning at time t. That is, for $i = 1, \ldots, n$ we have:

$$A_i(t) = \Pr(X_i(t) = 1) = E[X_i(t)].$$

It can then be shown (see Barlow and Proschan [5]) that the corresponding *stationary availabilities* can be expressed as:

$$A_i = \lim_{t \to \infty} A_i(t) = \frac{\mu_i}{\mu_i + \nu_i}, \qquad i = 1, \ldots, n. \tag{8.7}$$

Introduce $\mathbf{A}(t) = (A_1(t), \ldots, A_n(t))$ and $\mathbf{A} = (A_1, \ldots, A_n)$. The system availability at time t is given by:

$$A_\phi(t) = \Pr(\phi(\mathbf{X}(t)) = 1) = E[\phi(\mathbf{X}(t))] = h(\mathbf{A}(t)),$$

where h is the system's reliability function. The corresponding stationary availability is given by:

$$A_\phi = \lim_{t \to \infty} A_\phi(t) = h(\mathbf{A}). \tag{8.8}$$

We now extend Definition 5.1.1 to repairable components in the obvious way by stating that component i is *critical* at time t if:

$$\psi_i(\mathbf{X}(t)) = \phi(1_i, \mathbf{X}(t)) - \phi(0_i, \mathbf{X}(t)) = 1. \tag{8.9}$$

We will refer to $\psi_i(\mathbf{X}(t))$ as the *criticality state* of component i at time t. Moreover, we extend Definition 5.2.1 to this context by defining the Birnbaum measure of importance of a repairable component i at time t, as the probability that component i is critical at time t. This is denoted $I_B^{(i)}(t)$ and given by:

$$\begin{aligned} I_B^{(i)}(t) &= \Pr(\psi_i(\mathbf{X}(t)) = 1) = E[\psi_i(\mathbf{X}(t))] \\ &= h(1_i, \mathbf{A}(t)) - h(0_i, \mathbf{A}(t)). \end{aligned} \tag{8.10}$$

The corresponding stationary measure is given by:

$$I_B^{(i)} = \lim_{t \to \infty} I_B^{(i)}(t) = h(1_i, \mathbf{A}) - h(0_i, \mathbf{A}). \tag{8.11}$$

8.3 Simulating repairable systems

Discrete event simulation is a very well established technique for analyzing systems of pure jump processes with applications in many different areas such as queueing and inventory systems and reliability. For an introduction to this field see Banks et al. [3].

As in the previous section we consider a binary monotone system (C, ϕ) with component state processes, denoted $\{X_1(t)\}, \ldots, \{X_n(t)\}$. The events affecting the ith component are denoted by $E_{i1}, E_{i2}, \ldots$, listed in chronological order, $i = 1, \ldots, n$. Since we assumed that all lifetimes and repair times have absolutely continuous distributions, all these events happen at distinct points of time almost surely. We let $T_{i1} < T_{i2}, \ldots$ be the corresponding points of time for these events. We also let $T_{i0} = 0$, $i = 1, \ldots, n$. As in (8.1) the component state processes can then be expressed as:

$$X_i(t) = X_i(0) + \sum_{j=1}^{\infty} I(T_{ij} \leq t) J_{ij}, \qquad t \geq 0, ^i = 1, \ldots, n, \tag{8.12}$$

where the jumps J_{ij} are either -1 if E_{ij} is a failure event, or $+1$ if E_{ij} is a repair event. We assume that all components start out by being functioning. Thus, we have $X_i(0) = 1$, and $J_{ij} = (-1)^j$, for $i = 1, \ldots, n$ and $j = 1, 2, \ldots$. Finally, for $i = 1, \ldots, n$ we introduce the times between the events defined as:

$$\Delta_{ij} = T_{ij} - T_{ij-1}, \qquad i = 1, \ldots, n, ^j = 1, 2, \ldots. \tag{8.13}$$

Then for $i = 1, \ldots, n$ we have:

$$\Delta_{i1} = U_{i1}, \quad \Delta_{i2} = D_{i1}, \quad \Delta_{i3} = U_{i2}, \quad \ldots \tag{8.14}$$

Since $U_{i1}, U_{i2}, \ldots$ are independent and identically distributed with positive mean value μ_i, it follows by Proposition 8.1.1 that X_i is a regular pure jump process, $i = 1, \ldots, n$. Hence, by Proposition 8.1.4 the system state $\phi = \phi(\mathbf{X})$ as well as the criticality states $\psi_1(\mathbf{X}), \ldots, \psi_n(\mathbf{X})$ are regular pure jump processes.

At the system level the event set is the *union* of all the component event sets. Let $E^{(1)}, E^{(2)}, \ldots$ denote these events sorted with respect to their respective points of time, and let $T^{(1)} < T^{(2)} < \cdots$ be the corresponding points of time. Note that since we assumed that all lifetimes and repair times have absolutely continuous distributions, each system event corresponds almost surely to a unique component event.

In order to simulate such a system, we use an *object oriented approach* where the components as well as the system are represented as *objects*. The component objects are equipped with methods for generating failure and repair events according to their respective life- and repair time distributions. The system object determines the state of the system as a function of the component states. To keep track of the events and process them in the correct order, they are organized in a dynamic queue sorted with respect to the points of time of the

events. The component processes place their upcoming events into the queue where they stay until they are processed.

More specifically, at time zero each component starts out by being functioning, and places its first failure event into the queue. As soon as all these failure events have been placed into the queue, the first event in the queue is processed. That is, the *system time* is set to the time of the first event, and the event is taken out of the queue and passed on to the component responsible for handling this event. The component then updates its state, generates a new event, in this case a repair event, which is placed into queue, and notifies the system about its new state so that the system state can be updated as well. Then the next event in the queue is processed in the same fashion, and so forth until the system time reaches a certain predefined point of time. Note that since the component events are generated as part of the event processing, the number of events in the queue stays constant.

Although the system state and component states stay constant between events, it may still be of interest to log the state values at predefined points of time. In order to facilitate this, we introduce yet another type of event, called a *sampling event*. Such sampling events will typically be spread out evenly on the timeline. Thus, if $e_1, e_2, \ldots$ denote the sampling events, and $t_1 < t_2 < \cdots$ are the corresponding points of time, we would typically have $t_j = j \cdot \Delta$ for some suitable number $\Delta > 0$.

The sampling events will be placed into the queue in the same way as for the ordinary events. As a sampling event is processed, the next sampling event will be placed into the queue. Thus, at any time only one sampling event needs to be in the queue.

In principle one must update the system state every time there is a change in the component states. For large complex systems, these updates may slow down the simulations considerably. Thus, whenever possible one should avoid computing the system state. Fortunately, since the structure function of a binary monotone system is non-decreasing in each argument, it is possible to reduce the updating to a minimum. To explain this in detail, we consider the event E_{ij} affecting component i. Let T_{ij} be the corresponding point of time, and let $\mathbf{X}(T_{ij}^-)$ denote the value of the component state vector immediately before E_{ij} occurs, i.e., $\mathbf{X}(T_{ij}^-) = \lim_{t \to T_{ij}^-} \mathbf{X}(t)$. Note that by Proposition 8.1.2 these limits exist since the component state processes are regular.

If E_{ij} is a failure event of component i, i.e., $X_i(T_{ij}^-) = 1$ and $X_i(T_{ij}) = 0$, then the event cannot change the system state if the system is already failed, i.e., $\phi(\mathbf{X}(T_{ij}^-)) = 0$. Similarly, if E_{ij} is a repair event of component i, i.e., $X_i(T_{ij}^-) = 0$ and $X_i(T_{ij}) = 1$, this event cannot change the system state if the system is already functioning, i.e., $\phi(\mathbf{X}(T_{ij}^-)) = 1$. Thus, we see that we only need to recalculate the system state whenever:

$$\phi(\mathbf{X}(T_{ij}^-)) \neq X_i(T_{ij}). \tag{8.15}$$

Hence, the number of times we need to recalculate the system state is drastically reduced.

In cases where we keep track of the criticality state of each of the components, we can simplify the calculations even further by noting that the system state is changed as a result of the event E_{ij} if and only if component i is critical at the time of the event. Moreover, if i is critical, and E_{ij} is a failure event, it follows that the system fails as a result of this event, i.e., $\phi(\mathbf{X}(T_{ij})) = 0$. If on the other hand i is critical, and E_{ij} is a repair event, it follows that the system becomes functioning as a result of this event, i.e., $\phi(\mathbf{X}(T_{ij})) = 1$. Thus, we see that in this setup all the calculations we need to carry out, are related to the updating of the criticality states.

A similar technique can be used when updating the criticality states of the components. Thus, we consider the event E_{ij} affecting the state of component i. We first note that the criticality state function of component i, $\psi_i(\mathbf{X}(t)) = \phi(1_i, \mathbf{X}(t)) - \phi(0_i, \mathbf{X}(t))$ does not depend on the state of component i. Thus, the event E_{ij} does not have any impact on the criticality state of i. However, E_{ij} may still change the criticality state of other components in the system even when the system state remains unchanged. Thus, let $k \neq i$ be another component, and consider its criticality state function $\psi_k(\mathbf{X}(T_{ij}))$.

If $X_k(T_{ij}) = 1$ and $\phi(\mathbf{X}(T_{ij})) = 0$, it follows that:

$$\phi(1_k, \mathbf{X}(T_{ij})) = \phi(0_k, \mathbf{X}(T_{ij})) = 0. \tag{8.16}$$

Thus, in this case we must have $\psi_k(\mathbf{X}(T_{ij})) = 0$. On the other hand, if $X_k(T_{ij}) = 0$ and $\phi(\mathbf{X}(T_{ij})) = 1$, it follows that:

$$\phi(1_k, \mathbf{X}(T_{ij})) = \phi(0_k, \mathbf{X}(T_{ij})) = 1. \tag{8.17}$$

Thus, we must have $\psi_k(\mathbf{X}(T_{ij})) = 0$ in this case as well. Hence, we see that a necessary condition for component k to be critical at time T_{ij} is that:

$$\phi(\mathbf{X}(T_{ij})) = X_k(T_{ij}). \tag{8.18}$$

Utilizing these observations reduces the need to recalculate the criticality states.

8.4 Estimating availability and importance

Stationary availability and importance measures are typically easy to derive. If the system under consideration is not too complex, these quantities can be calculated analytically using (8.7), (8.8) and (8.11). For larger complex systems one may estimate the availability and importance using Monte Carlo simulations. A fast simulation algorithm for this is provided in [**?**]. Alternatively, estimates can be obtained by running a *single* discrete event simulation on the system over a sufficiently long time horizon.

Here, however, we focus on the problem of estimating the system availability $A_\phi(t)$ and the component importance measures $I_B^{(1)}(t), \ldots, I_B^{(n)}(t)$ as functions of t. Ideally we would like to estimate these quantities for any $t \geq 0$. For practical purposes, however, we have to limit the estimation to a finite set of

points. More specifically, we will estimate $A_\phi(t)$ for $t \in \{t_1, \ldots, t_N\}$, i.e., the set of the N first sampling points. For the points of time between the sampling points, we just use linear interpolation to obtain the curve estimate.

A simple approach to this problem is to run M simulations on the system, where each simulation covers the time interval $[0, t_N]$. In each simulation we sample the values of ϕ and $\psi_1, \ldots, \psi_n$ at each sampling point $t_1, \ldots, t_N$. We denote the sth simulated value of the component state vector process at time $t \geq 0$ by $\mathbf{X}_s(t)$, $s = 1, \ldots, M$, and obtain the following estimates for $j = 1, \ldots, N$:

$$\hat{A}_\phi(t_j) = \frac{1}{M} \sum_{s=1}^{M} \phi(\mathbf{X}_s(t_j)), \tag{8.19}$$

$$\hat{I}_B^{(i)}(t_j) = \frac{1}{M} \sum_{s=1}^{M} \psi_i(\mathbf{X}_s(t_j)). \tag{8.20}$$

We will refer to these estimates as *pointwise estimates*. It is easy to see that for $j = 1, \ldots, N$, $\hat{A}_\phi(t_j)$ and $\hat{I}_B^{(i)}(t_j)$ are unbiased and strongly consistent estimates of $A_\phi(t_j)$ and $I_B^{(i)}(t_j)$ respectively. In order to estimate $A_\phi(t)$ and $I_B^{(i)}(t)$ between the sampling points, one may use interpolation. Using a sufficiently high sampling rate, i.e., a small value of Δ, a satisfactory estimate of the full curve can be obtained. Still, all information about the process between the sampling points is thrown away.

We now present an alternative approach where we utilize process data between the sampling points as well. As above we assume that the system is simulated M times over the interval $[0, t_N]$, and let $\mathbf{X}_s(t)$ denote the sth simulated value of the component state vector process at time $t \geq 0$, $s = 1, \ldots, M$. Then let $E_s^{(1)}, E_s^{(2)}, \ldots$ denote the events in the interval $[0, t_N]$ in the sth simulation, *including* sampling events at times $t_1, \ldots, t_N$, and let $T_s^{(1)} < T_s^{(2)} < \cdots$ be the corresponding points of time, $s = 1, \ldots, M$. In this case we also include an extra sampling event in each simulation at time $t_0 = 0$, denoted $E_s^{(0)}$, and let $T_s^{(0)} = 0$, $s = 1, \ldots, M$.

The idea now is to use average simulated availability and criticalities from each interval $[t_{j-1}, t_j)$, $j = 1, \ldots, N$ as respective estimates for the availability and criticalities at the midpoints of these intervals. By using Proposition 8.1.3, we obtain the following estimates for $j = 1, \ldots, N$:

$$\tilde{A}_\phi(\bar{t}_j) = \frac{1}{M} \sum_{s=1}^{M} \frac{1}{\Delta} \sum_{k \in \mathcal{E}_{sj}} \phi(\mathbf{X}_s(T_s^{(k)}))(T_s^{(k+1)} - T_s^{(k)}), \tag{8.21}$$

$$\tilde{I}_B^{(i)}(\bar{t}_j) = \frac{1}{M} \sum_{s=1}^{M} \frac{1}{\Delta} \sum_{k \in \mathcal{E}_{sj}} \psi_i(\mathbf{X}_s(T_s^{(k)}))(T_s^{(k+1)} - T_s^{(k)}), \tag{8.22}$$

where $\mathcal{E}_{sj}$ denotes the index set of the events in $[t_{j-1}, t_j)$ in the sth simulation, and where we have introduced the interval midpoints $\bar{t}_j = (t_{j-1} + t_j)/2$, $j = 1, \ldots, N$. We will refer to these estimates as *interval estimates*.

By Proposition 8.1.3, it follows that for $j = 1, \ldots, N$, $\tilde{A}_\phi(\bar{t}_j)$ and $\tilde{I}_B^{(i)}(\bar{t}_j)$ are unbiased and strongly consistent estimates of the corresponding average availability and criticality in the intervals $[t_{j-1}, t_j)$ respectively. By choosing Δ so that the availabilities and criticalities are relatively stable within each interval, the interval estimates are approximately unbiased estimates for $A_\phi(\bar{t}_j)$ and $I_B^{(i)}(\bar{t}_j)$ as well. In fact the resulting interval estimates tend to stabilize much faster than the pointwise estimates. In order to estimate $A_\phi(t)$ and $I_B^{(i)}(t)$ between the interval midpoints, one may again use interpolation. Note that since all process information is used in the estimates, satisfactory curve estimates can be obtained for a much higher value of Δ than the one needed for the pointwise estimates.

In order to illustrate this we consider a simple bridge system shown in Figure 8.1. The components of this system are the five edges in the graph, labeled $1, \ldots, 5$. The system is functioning if the source node s can communicate with the terminal node t through the graph. All the components in the system have exponential lifetime and repair time distributions with mean values 1 time unit. The objective of the simulation is to estimate $A_\phi(t)$ and $I_B^{(1)}(t), \ldots, I_B^{(5)}(t)$ for $t \in [0, t_N]$, where $t_N = 1000$.

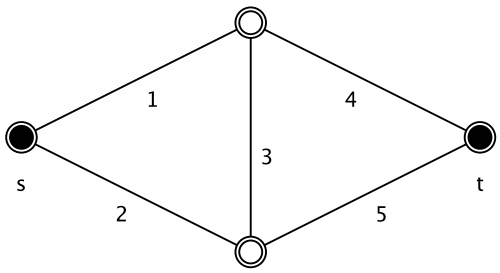

Figure 8.1: A bridge system.

All the simulations were carried out using a program called *Eventcue*[1]. This program has an intuitive graphical user interface, and can be used to estimate availability and criticality of any undirected network system.

Since all the lifetimes and repair times are exponentially distributed with the *same mean*, it is easy to derive explicit analytical expressions for the component availabilities. To see this, we consider the ith component at a given point of time t and introduce $N_i(t)$ as the number of failure and repair events affecting component i in $[0, t]$. With times between events being independent

[1]Eventcue is a java program developed at the Department of Mathematics, University of Oslo. The program is freely available at http://www.riscue.org/eventcue/.

and exponentially distributed with mean 1 it follows that $N_i(t)$ has a Poisson distribution with mean t. Moreover, component i is functioning at time t if and only if $N_i(t)$ is even. Thus, the ith component availability at time t is given by:

$$A_i(t) = \sum_{k=0}^{\infty} \Pr(N_i(t) = 2k) = \sum_{k=0}^{\infty} \frac{t^{2k}}{(2k)!} e^{-t}. \tag{8.23}$$

Using (8.23) one can verify numerically that all the component availabilities converge very fast towards their common stationary value, 0.5. As a result of this the system availability, $A_\phi(t)$, converges very fast towards its stationary value, 0.5, as well. In fact, for $t > 20$, numerical calculations show that $|A_\phi(t) - 0.5| < 10^{-15}$. Similarly, the Birnbaum measures of importance converges so that for $t > 20$, $|I_B^{(i)}(t) - 0.375| < 10^{-15}$, $i = 1, 2, 4, 5$, while $|I_B^{(3)}(t) - 0.125| < 10^{-15}$. Thus, for $t > 20$ the true values of all the curves are approximately constant. This makes it easy to evaluate and compare the quality of the different Monte Carlo estimates in this particular case.

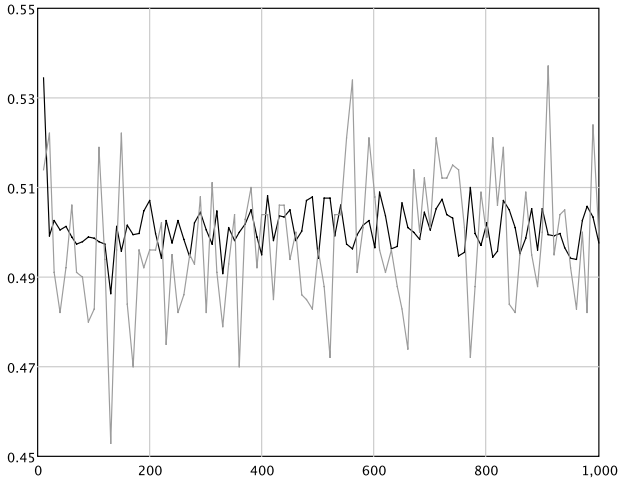

Figure 8.2: Interval estimate (black curve) and pointwise estimate (gray curve) of the availability curve.

Figure 8.2 and Figure 8.3 show respectively the availability curve and the criticality curve of component 1. The black curves are obtained using the interval estimates, while the gray curves show the corresponding pointwise estimate curves. In all cases we have used $M = 1000$ simulations and $N = 100$ sample points.

The plots clearly show the difference between the two methods. The black interval estimate curves are much more stable, and thus much closer to the true curve values, compared to the gray pointwise estimates.

One may think that increasing the number of sampling points would make the pointwise curve estimate better as more information is sampled. However, it

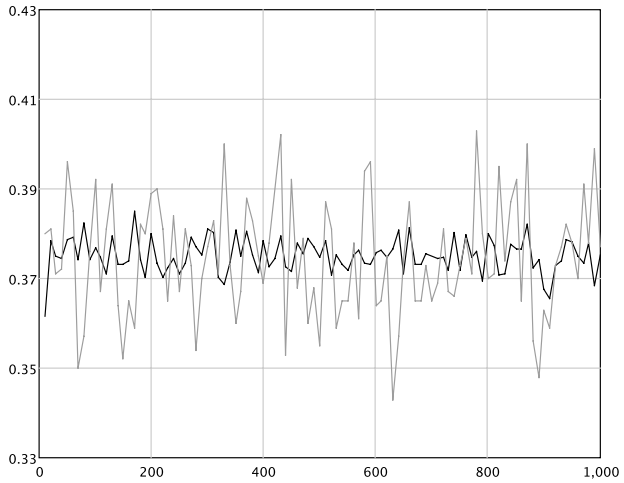

Figure 8.3: Interval estimate (black curve) and pointwise estimate (gray curve) of the importance curve.

turns out that the main effect of this is that the curve jumps more and more up and down. In fact with shorter intervals between sampling points the interval estimate becomes more unstable as well, and in the limit where the interval lengths go to zero, the two methods become equivalent. The only effective way of stabilizing the results for the pointwise curve estimate is to increase the number of simulations, i.e., M.

Table 8.1: Standard deviations for the pointwise curve estimates

M	2000	4000	6000	8000
St.dev.	0.0121	0.0076	0.0062	0.0054

In Table 8.1 we have listed estimated standard deviations for pointwise curve estimates for different values of M. We see that the standard deviation shows a steady decline as M increases. The corresponding numbers for $M = 1000$ are 0.0055 for the interval curve estimate and 0.0148 for the pointwise estimate. Thus, in this particular case we see that to obtain a pointwise curve estimate with a comparable stability to the interval curve estimate, one needs about eight times as many simulations.

For the interval curve estimate it is possible to obtain an even smoother curve simply by increasing Δ. Still, in general Δ should not be made too large, as this could produce a curve where important effects are obscured. Thus, in order to obtain optimal results, one should try out different values for Δ, and balance smoothness against the need of capturing significant oscillation properties of the curve.

Now, if smoothness is important, it is of course possible to apply some stan-

dard smoothing technique, such as moving averages or exponential smoothing, to the pointwise curve estimate. While such post-smoothing would clearly make the curve smoother, this technique does not add any new information to the estimate. The main advantage with the interval curve estimates is that such estimates actually use information about all events. Especially in cases where events occur at a very high rate, this turns out to be a great advantage.

Chapter 9

Applications

In this chapter, we study two different applications of the theory presented in the previous chapters. In Section 9.1, we show how to perform a reliability analysis of two different kinds of fishing boat engines, while in Section 9.2, we compute the reliability of a network for transmission of electronic pulses.

9.1 Case study: Reliability analysis and comparison of two fishing boat engines

This section is based on a collaboration from 1977 between Bent Natvig and what was then called Fiskeriteknologisk Forkningsinstitutt (the Norwegian department of fishery technology and research).

We will perform a reliability analysis and comparison of two different engines for fishing boats. In the following, we say that the boat engine is functioning if and only if the propeller and the power supply are functioning.

In 1977, Fiskeriteknologisk Forskningsinstitutt only approved one of the two engines we will study. The critical parts of this engine are a propulsion engine, an auxiliary engine and a hydraulic operated clutch. In Natvig [35], you can find a detailed illustration of the engine. We omit the details, and instead focus on the parts which are essential to our reliability analysis.

We define the following fault-events which may happen in this engine:

$$\begin{aligned} F1 &: \text{The propulsion engine fails} \\ K1 &: \text{The hydraulic clutch fails} \\ H1 &: \text{The auxiliary engine fails} \end{aligned} \tag{9.1}$$

Corresponding to the engine, we can make a fault tree to illustrate what may cause engine failure. This fault tree is shown in Figure 9.1.

The "OR" part of the fault tree means that at least one of the faults under the "OR" must occur for for the above fault to occur. For instance, in order for

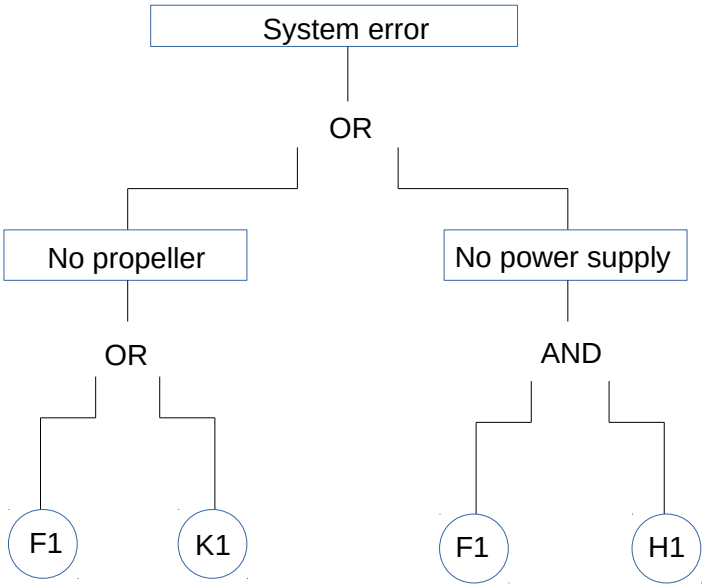

Figure 9.1: Fault tree for the first boat engine.

there to be a system failure, either the propeller is not functioning or there is no power supply. Correspondingly, the "AND" means that both of the faults under the "AND" must occur for the fault to occur. That is, in order for there to be no power supply, both $F1$ and $H1$ must happen. A more detailed description of fault trees can be found in Barlow and Proschan [5].

The auxiliary engine is only used when the boat is at harbour, while the propulsion engine and the clutch are used while the boat is running at sea. The time at sea per year is 3000 hours, while the harbour time is 2000 hours. Hence, the run time for the propulsion engine and the clutch is 3000 hours, while the run time for the auxiliary engine is 2000 hours (per year). We assume that the lifetime distributions of the components are exponential. Due to the memoryless property of the exponential distribution and the fact that we are considering mechanical components, this may be unrealistic. The Weibull distribution (which takes aging of components into account) would be preferable. However, to simplify, we assume exponential lifetimes.

Det Norske Veritas supplied point estimates for the failure rates, measured in failures per hour run. For a diesel engine (such as the propulsion engine and the auxiliary engine), this failure rate is estimated to be $2.96 \cdot 10^{-4}$. For a magnetic clutch, the estimate is $3.88 \cdot 10^{-5}$, while for a mechanic clutch it is $6.0 \cdot 10^{-6}$ (we will use this later in the section). Due to lack of data for hydraulic clutches, the

error estimate for the magnetic clutch was used as an approximation.

Based on this, we can compute the following estimates for the probability of the basic fault events in (9.1) happening in one year:

$$\begin{aligned} P(F1) &= 1 - \exp(-2.96 \cdot 10^{-4} \cdot 3000) = 0.58852, \\ P(K1) &= 1 - \exp(-3.88 \cdot 10^{-5} \cdot 3000) = 0.10988, \\ P(H1) &= 1 - \exp(-2.96 \cdot 10^{-4} \cdot 2000) = 0.44678. \end{aligned} \tag{9.2}$$

This gives the following estimates:

$$\begin{aligned} &P(\text{propeller failure in a year}) = 1 - (1 - P(F1)) \cdot (1 - P(K1)) = 0.63373, \\ &P(\text{no power supply in a year}) = P(H1) \cdot P(F1) = 0.26294. \end{aligned}$$

From this, Fiskeriteknologisk Forskningsinstitutt concluded that:

$$P(\text{system failure in a year}) = 1 - (1 - 0.63373)(1 - 0.26294) = 0.73004. \tag{9.3}$$

However, this method is incorrect because it does not take into account the dependence which the basic fault $F1$ creates in the fault tree because it occurs twice. In order to derive the correct reliability (or equivalently, the probability of system failure within a year), we introduce the following binary variables:

$$X_1 = \begin{cases} 1 & \text{if } F1 \text{ does } \textit{not} \text{ occur within a year} \\ 0 & \text{otherwise} \end{cases} \tag{9.4}$$

$$X_2 = \begin{cases} 1 & \text{if } K1 \text{ does } \textit{not} \text{ occur within a year} \\ 0 & \text{otherwise} \end{cases}$$

$$X_3 = \begin{cases} 1 & \text{if } H1 \text{ does } \textit{not} \text{ occur within a year} \\ 0 & \text{otherwise} \end{cases}$$

The fault tree 9.1 is equivalent to reliability block diagram shown in Figure 9.2. In the reliability block diagram component 1 with state variable X_1 corresponds to the propulsion engine, component 2 with state variable X_2 corresponds to the clutch, while component 3 with state variable X_3 corresponds to the auxiliary engine.

From Figure 9.2, we see that component 3, i.e., the auxiliary engine, is an irrelevant component. It does not matter whether the auxiliary engine is working for the system to be working. This can also be seen by deriving the structure function of the system in Figure 9.2:

$$\begin{aligned} \phi_1(\mathbf{X}) &= X_1X_2(1 - (1 - X_1)(1 - X_3)) \\ &= X_1X_2 + X_1X_2X_3 - X_1X_2X_3 \\ &= X_1X_2. \end{aligned}$$

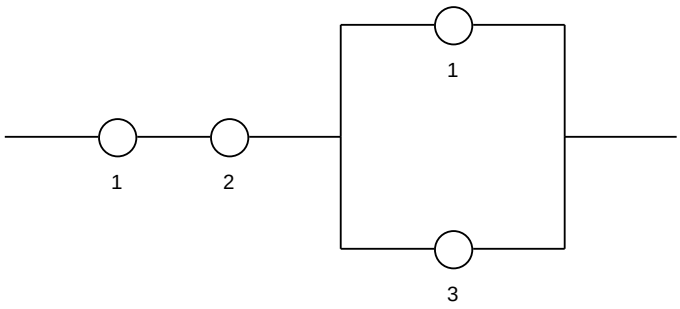

Figure 9.2: System equivalent to the fault tree.

Hence, we get the following correct estimate of the probability of system error within a year.

$$\begin{aligned} &P(\text{System failure within a year}) \\ &\quad = P(\phi_1(\mathbf{X}) = 0) = 1 - P(\phi_1(\mathbf{X}) = 1) \\ &\quad = 1 - (1 - P(F1))(1 - P(K1)) = 0.63373 \end{aligned} \tag{9.5}$$

By comparing this to the estimate in equation (9.3), we see that Fiskeriteknologisk Forskningsinstutt ended up with a failure probability which is larger than the actual one.

As mentioned in the beginning of the section, we would like to perform a reliability analysis for two different engines, and compare them. Let us now turn to the second engine, which is a two-engine system consisting of right and left propulsion engines, right and left hydraulic clutches and a mechanical clutch. Again, we will omit a detailed illustration of the system (this can be found in Natvig [35], Figure 5.1.4). Instead, we focus on the corresponding fault tree shown in Figure 9.3 consisting of the following basic fault-events:

$$\begin{aligned} F2: &\ \text{Error in the left propulsion engine} \\ K2: &\ \text{Error on left hydraulic clutch} \\ F3: &\ \text{Error in the right propulsion engine} \\ K3: &\ \text{Error on right hydraulic clutch} \\ K4: &\ \text{Error on mechanical clutch} \end{aligned} \tag{9.6}$$

For this system, both left and right engines are used at sea. At the harbour, only the right engine is used (this runs the generator which produces power when docking). Hence, the run time per year for the right engine is 5000 hours, and for the mechanical clutch it is 2000 hours (this is only used when docking). For the left propulsion engine and the two hydraulic clutches, the run time per year is 3000 hours, as they are only used at sea.

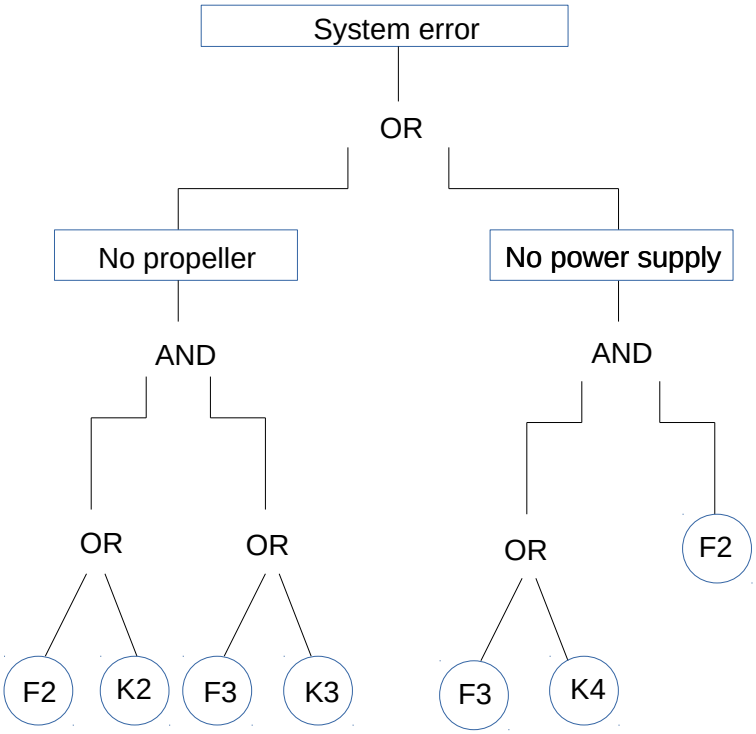

Figure 9.3: Fault tree for the second boat engine.

Based on the point estimates from Det Norske Veritas (see the comments before equation (9.2)), we compute the following estimates for the probabilities of the basic fault-events:

$$
\begin{aligned}
&P(F2) = P(F1) = 0.58852, \\
&P(K2) = P(K3) = P(K1) = 0.10988, \\
&P(F3) = 1 - \exp(-2.96 \cdot 10^{-4} \cdot 5000) = 0.77236, \\
&P(K4) = 1 - \exp(-6 \cdot 10^{-6} \cdot 2000) = 0.01192.
\end{aligned}
\tag{9.7}
$$

This leads to the following estimates:

$$\begin{aligned} &P(\text{propeller failure in a year}) \\ &\quad = [1-(1-P(F2))(1-P(K2))][1-(1-P(F3))(1-P(K3))] \\ &\quad = 0.63372 \cdot 0.79737 \\ &\quad = 0.50532, \end{aligned}$$

$$\begin{aligned} &P(\text{no power supply in a year}) \\ &\quad = [1-(1-P(F3))(1-P(K4))]P(F2) \\ &\quad = 0.77508 \cdot 0.58852 \\ &\quad = 0.45615. \end{aligned}$$

If we make the same kind of error as was done for the one-engine system in equation (9.3), we find that:

$$\begin{aligned} &P(\text{system error in a year}) \\ &\quad = 1-(1-0.50532)(1-0.45615) = 0.73097. \end{aligned} \tag{9.8}$$

In order to correct this mistake, we introduce some binary random variables:

$$X_1 = \begin{cases} 1 & \text{if } F2 \text{ does not occur within a year} \\ 0 & \text{if } F2 \text{ occurs within a year} \end{cases} \tag{9.9}$$

$$X_2 = \begin{cases} 1 & \text{if } K2 \text{ does not occur within a year} \\ 0 & \text{if } K2 \text{ occurs within a year} \end{cases}$$

$$X_3 = \begin{cases} 1 & \text{if } F3 \text{ does not occur within a year} \\ 0 & \text{if } F3 \text{ occurs within a year.} \end{cases}$$

$$X_4 = \begin{cases} 1 & \text{if } K3 \text{ does not occur within a year} \\ 0 & \text{if } K3 \text{ occurs within a year.} \end{cases}$$

$$X_5 = \begin{cases} 1 & \text{if } K4 \text{ does not occur within a year} \\ 0 & \text{if } K4 \text{ occurs within a year.} \end{cases}$$

Again, by letting component 1 with state variable X_1, correspond to the left propulsion engine, component 2 with state variable X_2 correspond to the left hydraulic clutch and so on, we see that the fault tree for the two-engine system is equivalent to the reliability block diagram shown in Figure 9.4.

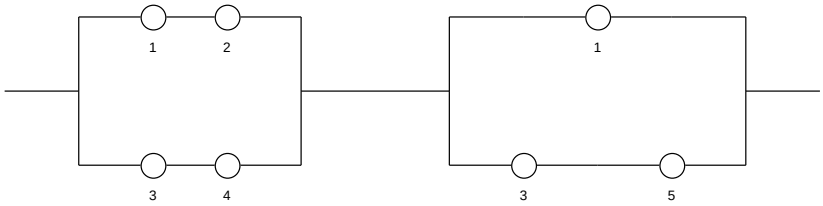

Figure 9.4: System equivalent to the fault tree for the two-engine system.

From this, we get that the structure function for the two-engine system is:

$$\begin{aligned}
\phi_2(\mathbf{X}) &= [1-(1-X_1X_2)(1-X_3X_4)][1-(1-X_1)(1-X_3X_5)] \\
&= (X_1X_2 + X_3X_4 - X_1X_2X_3X_4)(X_1 + X_3X_5 - X_1X_3X_5) \\
&= X_1X_2 + X_1X_3X_4 - X_1X_2X_3X_4 + X_1X_2X_3X_5 + X_3X_4X_5 \\
&- X_1X_2X_3X_4X_5 - X_1X_2X_3X_5 - X_1X_3X_4X_5 + X_1X_2X_3X_4X_5 \\
&= X_1X_2 + X_1X_3X_4 - X_1X_2X_3X_4 + X_3X_4X_5 - X_1X_3X_4X_5 \\
&= X_1X_2 + X_3X_4[X_1(1-X_2) + X_5(1-X_1)].
\end{aligned}$$

Hence, all of the components are relevant since they are all part of the structure function. From this, we derive the correct estimate for the probability of the two-engine system failing:

$$\begin{aligned}
&P(\text{System error within a year}) \qquad (9.10)\\
&\quad = 1 - P(\phi_2(\mathbf{X}) = 1) \\
&\quad = 1 - E[\phi_2(\mathbf{X})] \\
&\quad = 1 - \big((1-P(F2)) \cdot (1-P(K2)) + (1-P(F3)) \cdot (1-P(K3)) \\
&\quad \cdot [(1-P(F2))P(K2) + (1-P(K4))P(F2)]\big) \\
&\quad = 0.50666.
\end{aligned}$$

To conclude, we see from equations (9.3) and (9.8) that when we are not careful with taking the system structure into account in the correct way, the two systems are (almost) equally reliable. However, if we do the reliability analysis in the correct way, considering the actual structure of the system, we find from equations (9.5) and (9.10) that the two-engine system is actually more reliable than the one-engine system (which was the system that was approved by Fiskeriteknologisk Forskningsinstitutt). Intuitively, this is actually quite obvious, since in the two-engine system, both engines can be used for both running the propeller and supplying power. In the one-engine system, the auxiliary engine is just an irrelevant component.

9.2 Case study: Reliability analysis of a network for transmission of electronic pulses

In this case study, we will perform a reliability analysis of a system for transmission of electronic pulses. This system consists of several identical parts. As we shall see, this property can be utilized in order to simplify the calculations.

In this case study twe consider a *multistate* system. This means that the system is not just functioning or failed. Instead the set of possible system states is a set of non-negative integers. This approach provides a more detailed description of the system. As before, we let:

$$X_i(t) = \text{ The state of component } i \text{ at time } t,$$

and assume that the stochastic processes $\{X_i(t), t \geq 0\}_{i=1}^n$ are independent. The structure function is given by:

$$\phi(t) = \phi(\mathbf{X}(t)) = \text{ The state of the system at time } t.$$

We assume that $\phi(t) \in \{\phi_1, \ldots, \phi_k\}$. In principle, one can find the distribution of the state of a multistate system by using the approach introduced in Section 4.1, i.e., by enumerating all possible component states:

$$P(\phi(t) = \phi_j) = \sum_{\mathbf{x}} I(\phi(\mathbf{x}) = \phi_j) \cdot P(\mathbf{X}(t) = \mathbf{x}),\ j = 1, \ldots, k. \tag{9.11}$$

where the indicator function $I(\phi(\mathbf{x}) = \phi_j)$ is 1 if $\phi(\mathbf{x}) = \phi_j$ and 0 otherwise. However, as in Section 4.1, we typically need to reduce the number of terms in this sum because the calculations quickly become too time-consuming. We now assume that there exists variables $Y_1 = Y_1(\mathbf{X}), \ldots, Y_m = Y_m(\mathbf{X})$ such that $\phi(t)$ can be written:

$$\phi(t) = \phi(\mathbf{X}(t)) = \phi(\mathbf{Y}(\mathbf{X}(t))),$$

where $\mathbf{Y}(\mathbf{X}(t)) = (Y_1(\mathbf{X}(t)), \ldots, Y_m(\mathbf{X}(t)))$. In this case, the probability distribution of ϕ can be found from the formula:

$$P(\phi(t) = \phi_j) = \sum_{\mathbf{y}} I(\phi(\mathbf{y}) = \phi_j) P(\mathbf{Y}(\mathbf{X}(t)) = \mathbf{y}),\ j = 1, \ldots, k. \tag{9.12}$$

In order to use (9.12) to compute the system reliability, we have to compute $P(\mathbf{Y}(\mathbf{X}(t)) = \mathbf{y})$ for all $\mathbf{y}$. Hence, if the number of possible values of $\mathbf{Y}$ is large, little is gained by using (9.12) instead of (9.11). In some cases, however, we can achieve a large reduction in the computational load by a clever choice of $Y_1, \ldots, Y_m$. In particular, this turns out to be the case for the network in Figure 9.5.

The purpose of the system shown in Figure 9.5 is to ensure communication between a control room and 5 production units. In Figure 9.5 the components in the system are represented by blue circles. In addition to these components, the system consists of a control room, 5 production units and two connection

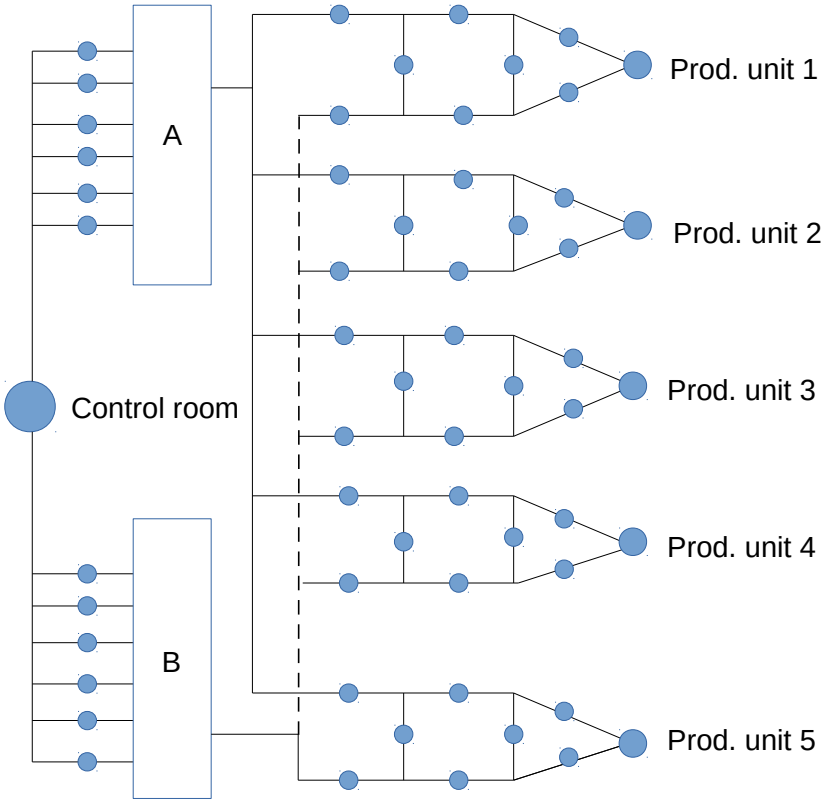

Figure 9.5: A network for transmission of electronic pulses.

units called A and B. To simplify the example somewhat, we will not consider potential errors in the control room, the production units and the connection units. Hence, the system consists of $n = 52$ components.

Each of the connection units, A and B, receive signals from the control room via 6 input wires. The number of production units which can be controlled by a connection unit is bounded by the number of functioning input wires to the respective connection unit. That is, if only 3 of the input wires to connection unit A are functioning, A can control at most 3 production units. If all 6 of the input wires to A are functioning, then all of the 5 production units can be controlled via connection unit A.

We assume that all of the components are stochastically independent, and

we also assume that all of the 12 input wires to the connection units A and B have the same lifetime distributions. In particular, we assume that for any wire between the control room and a connection unit we have:

$$P(\text{The wire is functioning at time } t) = w(t).$$

We observe that the part of the system which ensures communication between the connection units A and B and the production units consists of 5 identical subsystems. Each of these subsystems consists of 8 components. We assume that all of these subsystems have the same stochastic properties, in the sense that the corresponding components in the different subsystems have the same lifetime distributions.

Now, we define the state of the system, represented by the function ϕ, as:

$$\begin{aligned}\phi(t) = &\text{ The number of production units which can be controlled from} \\ &\text{ the control room at time } t.\end{aligned}$$

Thus, the set of possible values for ϕ is $\{0, 1, \ldots, 5\}$. The components are assumed to be either functioning or failed, so the component states are binary. Hence, the system in Figure 9.5 is a multistate system of binary components. The number of terms in the sum (9.11) for computing the distribution of ϕ is $2^{52} = 4.504 \cdot 10^{15}$, so finding this by enumerating all of the states is very time-consuming. However, by introducing appropriate variables, we can obtain a significant reduction of terms in (9.12) compared to (9.11). More specifically, we introduce the following variables:

$$\begin{aligned}Y_1(t) = &\text{ The number of intact input wires to connection unit } A \text{ at time } t \\ Y_2(t) = &\text{ The number of intact input wires to connection unit } B \text{ at time } t \\ Y_3(t) = &\text{ The number of production units connected to } A \text{ and } B \text{ at time } t \\ Y_4(t) = &\text{ The number of production units only connected to } A \text{ at time } t \\ Y_5(t) = &\text{ The number of production units only connected to } B \text{ at time } t.\end{aligned}$$

The state of the system can now be expressed as:

$$\phi(t) = W_1(t) + W_2(t) + W_3(t), \tag{9.13}$$

where:

$$\begin{aligned}&W_1(t) = \min\{Y_1(t), Y_4(t)\}, \quad W_2(t) = \min\{Y_2(t), Y_5(t)\}, \\ &W_3(t) = \min\{(Y_1(t) - W_1(t)) + (Y_2(t) - W_2(t)), Y_3(t)\}.\end{aligned} \tag{9.14}$$

In order to derive equation (9.13), we distribute the $Y_1(t)$ functioning input wires to A and the $Y_2(t)$ functioning input wires to B between the 5 production units in such a way that as many production units as possible are running. To do this, we first distribute input wires to the production units which are only connected to one connection unit, i.e., the $Y_4(t)$ production units only connected

to A and the $Y_5(t)$ production units only connected to B. In this way, we establish connection with $W_1(t)$ production units via A and $W_2(t)$ connection units via B. When this is done, there are $Y_1(t) - W_1(t)$ input wires available via connection unit A, and $Y_2(t) - W_2(t)$ input wires available via B. These are then used to establish connection to as many as possible of the $Y_3(t)$ remaining production units. The number of production units which can be reached in this way is then $W_3(t)$. Note that both $Y_1(t) - W_1(t)$ an $Y_2(t) - W_2(t)$ may be 0. We conclude that $W_1(t) + W_2(t) + W_3(t)$ is the number of production units which can be controlled from the control room. Thus, equation (9.13) is indeed correct.

Next, we need to find the probability distributions of $Y_1(t), \ldots, Y_5(t)$. Note that $Y_1(t)$ is a function of the state variables of the wires into connection unit A, $Y_2(t)$ is a function of the state variables of the wires into connection unit B, while the vector $(Y_3(t), Y_4(t), Y_5(t))$ is a function of state variables of the other 40 components. Since we have assumed that the components are independent, this implies that Y_1, Y_2 and the vector $(Y_3(t), Y_4(t), Y_5(t))$ are independent. Note, however that $Y_3(t)$, $Y_4(t)$ and $Y_5(t)$ are dependent. In fact we have $0 \leq Y_3(t) + Y_4(t) + Y_5(t) \leq 5$ for all $t \geq 0$.

Both $Y_1(t)$ and $Y_2(t)$ are sums of 6 independent, identically distributed binary random variables. Hence, from standard probability theory, $Y_1(t)$ and $Y_2(t)$ are binomially distributed:

$$P(Y_i(t) = y) = \binom{6}{y}[w(t)]^y[1 - w(t)]^{6-y}, \quad y = 0, 1, \ldots, 6, \quad i = 1, 2, \quad (9.15)$$

We then consider the probability distribution of $(Y_3(t), Y_4(t), Y_5(t))$. This depends on the 5 subsystems which ensure communication between the connection units and the production units. Each of these subsystems can be in one out of four possible states, denoted S_{AB}, S_A, S_B and $S_\emptyset$, where we define:

$$\begin{aligned} S_{AB} &= \text{The prod. unit can communicate with both connection units} \quad (9.16)\\ S_A &= \text{The prod. unit can only communicate with connection unit } A\\ S_B &= \text{The prod. unit can only communicate with connection unit } B\\ S_\emptyset &= \text{The prod. unit cannot communicate with any connection unit.} \end{aligned}$$

Furthermore, we have assumed that the subsystems are independent and that each of the subsystems have the same stochastic properties. This implies that the probability of being in either of the four states S_{AB}, S_A, S_B or $S_\emptyset$ at a given time t is the same for all the subsystems.

We observe that at time t the variables $Y_3(t), Y_4(t)$ and $Y_5(t)$ are the number of production units in states S_{AB}, S_A and S_B respectively. It follows from standard probability theory that for $t \geq 0$ the vector $(Y_3(t), Y_4(t), Y_5(t))$ is

multinomially distributed,. That is, we have:

$$P(Y_3(t) = y_3 \cap Y_4(t) = y_4 \cap Y_5(t) = y_5) = \frac{5!}{y_3!y_4!y_5!(5-y_3-y_4-y_5)!} \cdot [p_3(t)]^{y_3}[p_4(t)]^{y_4}[p_5(t)]^{y_5}[1-p_3(t)-p_4(t)-p_5(t)]^{5-y_3-y_4-y_5} \quad (9.17)$$

where $y_3 = 0, 1, \ldots, 5$, $y_4 = 0, 1, \ldots, 5 - y_3$, $y_5 = 0, 1, \ldots, 5 - y_3 - y_4$ and $p_3(t), p_4(t), p_5(t)$ are the probabilities of a subsystem being in states S_{AB}, S_A and S_B respectively.

In order to compute the distribution of ϕ, all that remains is to find $p_3(t), p_4(t)$ and $p_5(t)$. To do so, we study the structure of the subsystems, as shown in Figure 9.6, more closely.

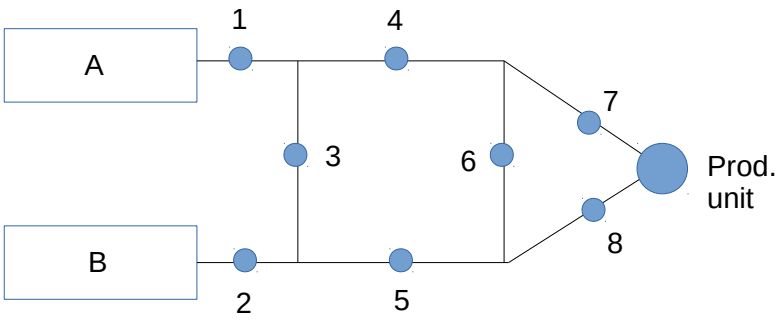

Figure 9.6: Subsystem of a network for transmission of electronic pulses.

In Figure 9.6, we have numbered the components of the subsystem under consideration from 1 through 8. We denote the corresponding component state variables by $X_1, \ldots, X_8$, and their respective reliabilities by $q_1, \ldots, q_8$, where we have simplified the notation by omitting the time t. Moreover, we introduce the events:

$$E_A = \{\text{The production unit communicates with connection unit } A\}$$
$$E_B = \{\text{The production unit communicates with connection unit } B\},$$

and observe that

$$\begin{aligned} p_3(t) &= P(E_A \cap E_B), \\ p_4(t) &= P(E_A) - P(E_A \cap E_B), \\ p_5(t) &= P(E_B) - P(E_A \cap E_B). \end{aligned} \quad (9.18)$$

Computing $P(E_A \cap E_B)$

To find this probability, we condition with respect to the two bridges in the structure, i.e., components 3 and 6, and get:

$$\begin{aligned} P(E_A \cap E_B) = & P(E_A \cap E_B | X_3 = 1, X_6 = 1) q_3 q_6 \\ & + P(E_A \cap E_B | X_3 = 1, X_6 = 0) q_3 (1 - q_6) \\ & + P(E_A \cap E_B | X_3 = 0, X_6 = 1) (1 - q_3) q_6 \\ & + P(E_A \cap E_B | X_3 = 0, X_6 = 0) (1 - q_3)(1 - q_6). \end{aligned} \tag{9.19}$$

The four conditional probabilities in (9.19) can be computed by series- and parallel reductions (this is left as an exercise). They are given by:

$$\begin{aligned} &P(E_A \cap E_B | X_3 = 1, X_6 = 1) = q_1 q_2 [q_4 + q_5 - q_4 q_5][q_7 + q_8 - q_7 q_8], \\ &P(E_A \cap E_B | X_3 = 1, X_6 = 0) = q_1 q_2 [q_4 q_7 + q_5 q_8 - q_4 q_5 q_7 q_8], \\ &P(E_A \cap E_B | X_3 = 0, X_6 = 1) = q_1 q_2 q_4 q_5 [q_7 + q_8 - q_7 q_8], \\ &P(E_A \cap E_B | X_3 = 0, X_6 = 0) = q_1 q_2 q_4 q_5 q_7 q_8. \end{aligned} \tag{9.20}$$

Computing $P(E_A)$ and $P(E_B)$

Note that in order to compute $P(E_A)$, component 2 is irrelevant (see Figure 9.6). If we remove this component, we see that components 3 and 5 are in series. By series redulction, we end up with a bridge structure (see Example 3.1.2) connected in series with component 1. Hence, $P(E_A)$ is given by the following (again, see Example 3.1.2):

$$\begin{aligned} P(E_A) = & q_1 q_6 [q_4 + q_3 q_5 - q_3 q_4 q_5][q_7 + q_8 - q_7 q_8] \\ & + q_1 (1 - q_6)[q_4 q_7 + q_3 q_5 q_8 - q_3 q_4 q_5 q_7 q_8]. \end{aligned} \tag{9.21}$$

Similarly, $P(E_B)$ is given by:

$$\begin{aligned} P(E_B) = & q_2 q_6 [q_5 + q_3 q_4 - q_3 q_4 q_5][q_7 + q_8 - q_7 q_8] \\ & + q_2 (1 - q_6)[q_5 q_8 + q_3 q_4 q_7 - q_3 q_4 q_5 q_7 q_8]. \end{aligned} \tag{9.22}$$

By inserting the expressions (9.19), (9.21) and (9.22) into (9.18), we have all the probabilities we need in order to calculate the distribution of the system state ϕ using equation (9.12).

What have we gained by this approach? Instead of enumerating all possible states of the 52 binary state variables of the components in Figure 9.5, we now have to enumerate the possible states of just the five multinary variables $Y_1, \ldots Y_5$. We observe that Y_1 and Y_2 each attain 7 values in the set $\{0, 1, \ldots 6\}$. The vector (Y_3, Y_4, Y_5) can attain any value in the set $\{(Y_3, Y_4, Y_5) : 0 \leq Y_3 + Y_4 + Y_5 \leq 5, 0 \leq Y_3, Y_4, Y_5\}$. In order to count the number of values in this set, we consider 6 different cases corresponding to the possible values of Y_3, and count the number of possible values for each case, noting that if $Y_3 = y$, then

$Y_4 + Y_5 \leq 5 - y$, $y = 0, 1, \ldots, 5$. Hence, we get:

$$\begin{array}{lllll} \text{Case 0.} & Y_3 = 0, & Y_4 + Y_5 \leq 5 - 0 = 5: & \text{21 possible values,} \\ \text{Case 1.} & Y_3 = 1, & Y_4 + Y_5 \leq 5 - 1 = 4: & \text{15 possible values,} \\ \text{Case 2.} & Y_3 = 2, & Y_4 + Y_5 \leq 5 - 2 = 3: & \text{10 possible values,} \\ \text{Case 3.} & Y_3 = 3, & Y_4 + Y_5 \leq 5 - 3 = 2: & \text{6 possible values,} \\ \text{Case 4.} & Y_3 = 4, & Y_4 + Y_5 \leq 5 - 4 = 1: & \text{3 possible values,} \\ \text{Case 5.} & Y_3 = 5, & Y_4 + Y_5 \leq 5 - 5 = 0: & \text{1 possible value.} \end{array} \tag{9.23}$$

Adding these counts up, we see that there are 56 values in total. Hence, the sum in (9.12) for computing $P(\phi(t) = \phi_j)$ contains $7 \cdot 7 \cdot 56 = 2744$ terms (since there are 7 possible values for Y_1 and Y_2, as well as 56 possible values for (Y_3, Y_4, Y_5)). This is very easy to compute, as opposed to the $4.504 \cdot 10^{15}$ terms of the original method by simply enumerating the states. We see that by introducing the new variables $Y_1, \ldots, Y_5$, the computational load has been significantly reduced. The same approach is useful in other systems where the structure consists of many identical components and where there is a strong symmetry. A difficulty with the method is that in practice, it may not be possible to find efficient ways to introduce new variables. However, as there is so much to be gained from the method when it works, it may be worth spending some time trying to figure out which variables to introduce.

9.3 Exercises

Exercise 9.2.1:

Use series- and parallel reductions to show that the conditional probabilities in (9.19) are given by the expressions in (9.20).

Exercise 9.2.2:

Explain why the 6 different values in (9.23) hold true (for example, the only way $Y_4 + Y_5 \leq 0$ for non-negative integers is for $Y_4 = 0, Y_5 = 0$).

Chapter 10

Notations

A^C, the complement of a set. See Section 3.3.

:=, marks definitions.

≡, identically equal.

$\phi(\cdot)$, the structure function of a binary system. See equation (2.2).

$h(\cdot)$, the reliability function of a binary system. See equation (2.6).

$\prod_{i=1}^{n} a_i := a_1 \cdot a_2 \ldots \cdot a_n$, the product.

$\coprod_{i=1}^{n} a_i := 1 \prod_{i=1}^{n}(1 - a_i)$, the ip notation.

$\sum_{\mathbf{y} \in \{0,1\}^n} \ldots$, summing over all possible vectors in $\mathbb{R}^n$ where all components are either 0 or 1.

ϕ^D, the dual structure function. See Definition 2.3.1.

0, a vector where all components are 0.

1, a vector where all components are 1.

Bibliography

[1] Baarholm, G.S., Haver, S. and Økland, O.D. Combining contours of significant wave height and peak period with platform response distributions for predicting design response. *Marine Structures* (23): 147–163, 2010.

[2] Ball, M. O., The complexity of network reliability computations. Networks, 10, 253–278, 1977.

[3] Banks, J., Carson, J., Nelson, B.L., and Nicol, D. *Discrete-Event System Simulation (4th Edition).* Prentice-Hall International Series in Industrial and Systems, 2004.

[4] Barlow, R. E. and Proschan, F., Importance of system components and fault tree events, *Stochastic Processes and their Applications*, 3, 153–173, 1975.

[5] Barlow, R. E. and Proschan, F., *Statistical Theory of Reliability and Life Testing*, To Begin With, Silver Spring, MD, 1981.

[6] Barlow, R. E. and Proschan, F., *Mathematical Theory of Reliability*, SIAM, Philadelphia, 1996.

[7] Birnbaum, Z. W., On the Importance of Different Components in a Multicomponent System, Multivariate Analysis - II, Edited by P. R. Krishnaiah, Academic Press, pp. 581–592, 1969.

[8] Broström, G. and L. Nilsson, L., Acceptance/rejection sampling from the conditional distribution of independent discrete random variables, given their sum, Statistics, 34, 247, 2000.

[9] Calkin, N. J., Edds, J. D. and Shier, D. R. Cancellation in cyclic consecutive systems, Journ. of Comp. and Appl. Math. 142, 13–26, 2002.

[10] Cancela, H. and El Khadari, M., A recursive variance-reduction algorithm for estimating communication-network reliability, IEEE Transactions on Reliability, R-44, 1995.

[11] Cancela, H. and El Khadari, M., Series-parallel reductions in Monte Carlo network reliability evaluation, IEEE Transactions on Reliability, R-47, 1998.

[12] Cancela, H. and El Khadari, M., The recursive variance-reduction simulation algorithm for network reliability evaluation, IEEE Transactions on Reliability, R-52, 2003.

[13] Ditlevsen, O. Stochastic model for joint wave and wind loads on offshore structures. *Structural Safety* (24): 139–163, 2002.

[14] Fontaine, E., Orsero, P., Ledoux, A., Nerzic, R., Prevesto, M. and Quiniou, V. Reliability analysis and response based design of a moored FPSO in West Africa. *Structural Safety* (41): 82–96, 2013.

[15] Fishman, G. S., A Monte Carlo sampling plan for estimating network reliability, Operations Research, 34, 1986.

[16] Fishman, G. S., A comparison of four Monte Carlo methods for estimating the probability of $s - t$ connectedness, IEEE Transactions on Reliability, R-35, 1986.

[17] Garey, M. R. and Johnson, D. S., Computers and Intractibility: A Guide to the theory of NP-Completeness, W. H. Freeman, San Francisco, 1979.

[18] Glasserman, P. and Yao, D. D. *Monotone Structure in Discrete-event Systems.* John Wiley and Sons, Inc., 1994.

[19] Glasserman, P. *Monte Carlo Methods in Financial Engineering.* Springer Verlag, 2004.

[20] Haver, S. On the joint distribution of heights and periods of sea waves. *Ocean Engineering* (14): 359–376, 1987.

[21] Haver, S. and Winterstein, S. Environmental contour lines: A method for estimating long term extremes by a short term analysis. *Transactions of the Society of Naval Architects and Marine Engineers* (116): 116–127, 2009.

[22] Huseby, A. B., A unified theory of domination and signed domination with application to exact reliability computations, Statistical Research Report No. 3, Department of Mathematics, University of Oslo, 1984.

[23] Huseby, A. B., Domination theory and the Crapo β-invariant, Networks, 19, 135–149, 1989.

[24] Huseby, A. B. , Naustdal, M. and Vårli, I. D., System reliability evaluation using conditional Monte Carlo methods, Department of Mathematics, University of Oslo, 2004.

[25] Huseby, A. B., Oriented matroid systems, Discrete Applied Mathematics, 159 1, 31–45, 2011.

[26] Huseby, A. B., Vanem, E., Natvig, B. A new approach to environmental contours for ocean engineering applications based on direct Monte Carlo simulations. *Ocean Engineering*, (60): 124–135, 2013.

[27] Huseby, A. B., Vanem, E., Natvig, B. Alternative environmental contours for structural reliability analysis. *Structural Safety*, (54): 32–45, 2015.

[28] Huseby, A. B., Vanem, E., Natvig, B. A new Monte Carlo method for environmental contour estimation, In *Safety and Reliability : Methodology and Applications*, Proceedings of the European safety and reliability Conference, Taylor & Francis. ISBN 978-1-138-02681-0. Chapter 270. 2091–2098, 2015.

[29] Jonathan, P., Ewans, K. and Flynn, J. On the estimation of ocean engineering design contours. In: *Proc. 30th International Conference on Offshore Mechanics and Arctic Engineering (OMAE 2011)* American Society of Mechanical Engineers (ASME), 2011.

[30] Klebaner, F. C. *Introduction to stochastic calculus with applications.* Imperial College Press, 2005.

[31] Knuth, D. E., The art of computer programming, Vol. 2, Seminumerical algorithms, Addison-Wesley, Reading, Mass. second edition, 1981.

[32] Lindqvist, B. H. and Taraldsen, G., Monte Carlo Conditioning on a Sufficient Statistics, Preprint Statistics 9, Norwegian University of Science and Technology, 2001.

[33] Moan, T. Development of accidental collapse limit state criteria for offshore structures. *Structural Safety* (31): 124–135, 2009.

[34] Natvig, B., A suggestion for a new measure of importance of system components, Stochastic Processes and their Applications, 9, 319–330, 1979.

[35] Natvig, B., Pålitelighetsanalyse med teknologiske anvendelser, lecture notes, Department of Mathematics, University of Oslo, 3rd ed. 1998.

[36] Natvig, B., *Multistate Systems Reliability Theory with Applications*, John Wiley and Sons Ltd., Chichester, 2011.

[37] Naustdal, M., Forbedrede betingede Monte Carlo metoder i pålitelighetsberegninger, (in Norwegian) Cand. Scient. thesis, Department of Mathematics, University of Oslo. 2001.

[38] Nilsson, L., On the simulation of conditional distributions in the Bernoulli case, Research report 14, Department of Mathematical Statistics, Umeå University, Sweden 1997.

[39] Rosenblatt, M. Remarks on a Multivariate Transformation. *The Annals of Mathematical Statistics* (23): No 3, 470–472, 1952.

[40] Rosenthal, A., Computing the reliability of complex networks, SIAM, Journal of Applied Math., 32, 384–393, 1977.

[41] Satyanarayana, A., A unified formula for analysis of some network reliability problems, IEEE Trans. Reliability, 31, 23–32, 1982.

[42] Satyanarayana, A. and Chang, M. K., Network Reliability and the Factoring Theorem, Networks, 13 107–120, 1983.

[43] Prabhakar, A. and Satyanarayana, A., New topological formula and rapid algorithm for reliability analysis of complex networks, IEEE Trans. Reliability, 27, 82–100, 1978.

[44] Vanem, E. and Bitner-Gregersen, E. Stochastic modelling of long-term trends in wave climate and its potential impact on ship structural loads. *Applied Ocean Research*, (37): 235–248, 2012.

[45] Vanem, E. and Bitner-Gregersen, E. Alternative Environmental Contours for Marine Structural Design – A Comparison Study. *Journal of Offshore Mechanics and Arctic Engineering*, (137): 051601-1–051601-8, 2015.

[46] Vårli, I. D. Forbedrede Monte Carlo metoder i pålitelighetsberegninger, (in Norwegian), Cand. Scient. thesis, Department of Mathematics, University of Oslo, 1996.

[47] Winterstein, S., Ude, T., Cornell, C., Bjerager, P. and Haver, S. Environmental parameters for extreme response: Inverse FORM with omission factors. In: *Proc. 6th International Conference on Structural Safety and Reliability*, 1993.

[48] Winther, R., Threshold systems. PhD-thesis, University of Oslo, 1996.