Reliability Investigations on Large-scale Technical Networks: Applications to Electric Power Systems

A dissertation submitted to
ETH ZURICH

for the degree of
Doctor of Sciences

presented by
MARKUS SCHLÄPFER
Dipl. Ing. ETH
born March 26, 1977
citizen of Rehetobel and St. Gallen, Switzerland

accepted on the recommendation of
Prof. Dr. Wolfgang Kröger, examiner
Prof. Dr. Göran Andersson, co-examiner

Zurich, 2010
Acknowledgments

This thesis is the result of the research work I have carried out at the Laboratory for Safety Analysis (LSA), Swiss Federal Institute of Technology (ETH) Zurich, between 2006 and 2010.

First of all, I wish to express my sincere gratitude to my supervisor, Prof. Dr. Wolfgang Kröger, for giving me the opportunity to conduct this multidisciplinary Ph.D. project as a member of his research team. I am particularly grateful for the provided freedom to explore a variety of related scientific fields as well as for his great guidance, inspiration and continuous encouragement.

I owe special thanks to Prof. Dr. Göran Andersson for being my co-supervisor and for supporting me with his precious expertise in the field of electric power engineering.

With pleasure I acknowledge the generous funding support by Swiss-electric Research. Special thanks go to Hans Achermann, Dr. Martin Kauert and Dr. Michael Paulus for all the motivating discussions and valuable feedbacks on the progress reports.

The Ph.D. project greatly benefited from the supervision of several semester and master theses, realized in close collaboration with industry partners. I thus would like to express my thanks to the former students David Henzi, Tom Kessler and Pascal Kienast for their great work, as well as to the Ph.D. students Michèle Arnold and Florian Kienzle for the co-supervision of the theses. I am thankful to Philippe Huber, Dr. Walter Sattinger and Matthias Schmid from Swissgrid AG as well as to Dr. Jürg Bader, Dr. Lukas Küng and Dr. Hans-Heinrich Schiesser from EWZ for the successful collaboration and for providing the empirical system data. Dr. Walter Sattinger further promoted my work with
his outstanding knowledge about the operational concerns of electric power grids.

I am very grateful to Dr. Lubos Buzna, Sven Dietz, Dr. Manuel Kaegi, Dr. Pierluigi Mancarella, Dr. Jonathan L. Shapiro and Konstantinos Trantopoulos. Writing research papers together and sharing our views and ideas was an indispensable enrichment for my scientific experience.

I am indebted to Prof. Dr. Henrik Jeldtoft Jensen for hosting my stay within his research team at the Imperial College in London and for his constructive suggestions in the field of statistical physics. With his Ph.D. students Giovanni Petri and Paul Expert I have enjoyed many stimulating discussions on the topic.

Many thanks also to Roberto Nibali, Patrick Probst, Manuel Oetiker and Tobias Oetiker for their efficient help with soft- and hardware technical questions and for ensuring a troublefree operation of the computers, to Monika Mortimer for proofreading my research papers and this thesis, to Sigrid Wagner and Evangelos Bilis for being such delightful officemates and to Dr. Thilo Krause for providing me with the very helpful $\LaTeX$ template used for this document.

I dedicate this thesis to my family, in particular to my parents Rosmarie and Peter Schläpfer who are always there for me. It is your love, trust and unwavering support which made this work possible.

Zurich, February 2010

Markus Schläpfer
Abstract

Transportation networks, telecommunication systems, energy and water supply facilities form an essential backbone of today’s societies. Their malfunction or even breakdown carry detrimental effects ranging from inconvenient supply interruptions to failures of vital importance with high economic impacts, eventually endangering the stability of entire regions.

This thesis addresses the reliability of such large-scale technical networks with a focus on electric power systems. It introduces probabilistic modeling and simulation approaches and conducts investigations on two fundamentally different yet complementary levels of system abstraction. According to these abstraction levels the thesis is divided into two main parts.

The first part aims at quantifying the reliability of specific systems based on replicating their real-world behavior in a detailed manner. A probabilistic modeling and simulation framework is developed, combining object-oriented programming with Monte Carlo techniques. Its major advantage is the possible inclusion of highly non-linear and time-dependent responses as well as non-technical factors such as human operational behavior. This allows calculating expected frequencies of adverse events (e.g., power outages) versus their size (e.g., energy not supplied), representing the reliability in a differentiated way. Investigations on both a “virtual” electric power system and on a model of the Swiss high-voltage grid confirm the functionality. For example, the simulation results demonstrate how already a small increase of the network load or a delayed operator response to contingencies might significantly increase the risk of wide-area power outages. For accelerating such extensive stochastic simulations the thesis demonstrates the capability of the RESTART technique, reaching time savings of several orders of
magnitude. Its algorithmic adaptation is used for the investigation of fluctuating power flows as induced by renewable energy sources, and their impact on transmission line loadings. A realistic case study discussing the integration of wind power in a transmission network shows how the thermal inertia on the lines effectively decreases the overload probability.

The second part of the thesis uncovers basic mechanisms leading to network breakdowns. These mechanisms are investigated on synthetic networks at a high level of system abstraction, making use of recent advances in the field of statistical physics. The application of a cascading failure model to degree-correlated networks reveals how a different encapsulation of the nodes with a high connectivity alters the propagation velocity. The results suggest that many real-world spreading processes, ranging from cascading failures in electric power systems to epidemic diseases, can be categorized into two general types for which either positive or negative degree-degree correlations decelerate the spreading. This result contributes to the basic understanding of spreading phenomena, being relevant for solving a broad range of problems in a larger spectrum of scientific fields. In the same line, cascading failures in networks subject to fluctuating flows are analyzed. The insights gained have potential implications for future electric power systems with a large share of renewable and distributed energy sources. For instance, the results suggest that even in a highly distributed system a considerable transmission capacity is still needed, or that the thermal inertia on the transmission lines significantly increases the network robustness.

The elaborated advantages and limitations of the two opposed levels of system abstraction eventually lay the groundwork for defining their specific application areas as well as their conceptual interplay within a comprehensive reliability study of large-scale technical networks.
Kurzfassung


Die in dieser Dissertation erarbeiteten Vor- und Nachteile der beiden Abstraktionsebenen bilden schliesslich die Grundlage zur Bestimmung der spezifischen Anwendungsmöglichkeiten, sowie des konzeptionellen Zusammenspiels im Rahmen einer gesamtheitlichen Zuverlässigskeitsbeurteilung grossräumiger technischer Netzwerke.
Contents

Acknowledgements .................................................. iii
Abstract ......................................................... v
Kurzfassung ....................................................... vii

1 Introduction

1.1 Background and Motivation ................................. 1
1.2 Contributions ................................................ 5
1.3 Definition of Key Terms ..................................... 7
1.4 Outline of the Thesis ........................................ 8
1.5 List of Publications .......................................... 9

I Applied Reliability Investigations on Detailed System Models 13

2 Reliability Analysis of Electric Power Systems Using Object-oriented Modeling 15

2.1 Abstract ..................................................... 15
2.2 Introduction ................................................ 16
2.3 Modeling Framework ...................................... 17
  2.3.1 Conceptual basics ..................................... 17
  2.3.2 Component models .................................... 18
  2.3.3 System splitting ...................................... 26
  2.3.4 Blackout frequency distributions .................... 26
2.4 Case Studies .................................................. 27
   2.4.1 Application to the IEEE Reliability Test System 1996 ........................................... 27
   2.4.2 Application to the Swiss high-voltage grid ...................................................... 31
2.5 Conclusions ...................................................... 35

3 Modeling and Simulation of Line Temperatures under Fluctuating Power Flows 37
   3.1 Abstract ...................................................... 37
   3.2 Introduction .................................................. 38
   3.3 Electrothermal Model ........................................ 39
   3.4 Line Temperature Assessment ................................. 41
      3.4.1 Conceptual basics ........................................ 41
      3.4.2 Accelerated Monte Carlo simulation for dynamic line temperature estimation ................. 42
      3.4.3 Modeling stochastic behavior using Markov chains ........................................... 46
   3.5 Case Studies ..................................................... 47
      3.5.1 Example A: Stochastic power injections on a single line .................................... 47
      3.5.2 Example B: Line temperatures within a transmission network including wind power .......... 52
   3.6 Conclusions ..................................................... 56

II Theoretical Reliability Investigations on Synthetic Networks 59

4 Failure Propagation Dynamics in Synthetic Networks 61
   4.1 Abstract ...................................................... 61
   4.2 Introduction .................................................. 62
   4.3 Network Degradation Model ................................. 63
      4.3.1 Conceptual basics ........................................ 63
      4.3.2 Nodal stress .............................................. 64
List of Figures

1.1 Cross-border energy exchanges within the UCTE operational area and globally installed wind power capacity 2
1.2 The two parts of the thesis and their classification according to the level of system abstraction and to the generality of the findings 5

2.1 Illustration of the two-layers-concept applied to electric power systems 18
2.2 Finite state machine for the load objects 19
2.3 Finite state machine for the generator objects 21
2.4 Finite state machine for the transmission line objects 22
2.5 Finite state machine for the grid operator 24
2.6 Complementary cumulative blackout frequencies for the IEEE RTS-96 29
2.7 Impact of increasing the system loading on the absolute frequencies of blackouts 30
2.8 Influence of the operator response time on the EENS due to generation inadequacy, operator action and system splitting 31
2.9 Topological representation of the Swiss high-voltage grid 32
2.10 Estimated blackout frequencies for the Swiss system with respect to the unserved energy and histogram indicating the distribution of the outages due to generation inadequacy, system splitting and load shedding 33
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.11</td>
<td>Blackout prevention due to operator response to line overloads</td>
<td>34</td>
</tr>
<tr>
<td>2.12</td>
<td>Relative frequency of transmission line overloads</td>
<td>35</td>
</tr>
<tr>
<td>3.1</td>
<td>Illustration of the RESTART algorithm</td>
<td>43</td>
</tr>
<tr>
<td>3.2</td>
<td>Flowchart of the accelerated algorithm for estimating the probability of line temperature $T^m$</td>
<td>45</td>
</tr>
<tr>
<td>3.3</td>
<td>Single-line layout</td>
<td>47</td>
</tr>
<tr>
<td>3.4</td>
<td>Combined power injection patterns</td>
<td>49</td>
</tr>
<tr>
<td>3.5</td>
<td>Power injection sequence and temperature behavior</td>
<td>50</td>
</tr>
<tr>
<td>3.6</td>
<td>Estimates of $\gamma$ as a function of the fluctuation frequency for different numbers of aggregated generating units</td>
<td>50</td>
</tr>
<tr>
<td>3.7</td>
<td>Decrease of $RE(\hat{\gamma})$ with the simulation time</td>
<td>51</td>
</tr>
<tr>
<td>3.8</td>
<td>Single line diagram of the 5-bus / 7-branch network indicating the values for the peak loads</td>
<td>53</td>
</tr>
<tr>
<td>4.1</td>
<td>Time sequence of the degradation process in Erdős-Rényi random graphs and scale-free networks for different average node degrees</td>
<td>70</td>
</tr>
<tr>
<td>4.2</td>
<td>Time sequence of the average nodal failure rate within scale-free networks with different average node degrees</td>
<td>71</td>
</tr>
<tr>
<td>4.3</td>
<td>Time sequence of the cascading events in scale-free networks broken down into the different failure modes</td>
<td>71</td>
</tr>
<tr>
<td>4.4</td>
<td>Influence of the degradation model parameters on the time sequence of the network collapse for Erdős-Rényi random graphs and scale-free networks</td>
<td>73</td>
</tr>
<tr>
<td>4.5</td>
<td>Comparison of the master equation formalism with Monte Carlo simulations</td>
<td>78</td>
</tr>
<tr>
<td>5.1</td>
<td>Illustration of the stress function</td>
<td>84</td>
</tr>
<tr>
<td>5.2</td>
<td>Relative nodal life expectancy in degree-correlated scale-free networks versus the response parameter. The inset shows the absolute life expectancy in the uncorrelated reference networks.</td>
<td>87</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
<td>------</td>
</tr>
<tr>
<td>5.3</td>
<td>(a) Illustration of the $k$-core decomposition of a network before and after reshuffling towards a higher positive correlation. (b)(c) Life expectancy of the $k$-shells for assortative and disassortative scale-free networks. The inset illustrates the relative size of the $k$-shells.</td>
<td>88</td>
</tr>
<tr>
<td>5.4</td>
<td>Average efficient path length of correlated networks normalized with the uncorrelated reference case.</td>
<td>89</td>
</tr>
<tr>
<td>6.1</td>
<td>Schematic plot of a $12 \times 12$ lattice with two different grouping factors and resulting total flow versus time.</td>
<td>96</td>
</tr>
<tr>
<td>6.2</td>
<td>Average flows versus the grouping factor, being fitted by a power law.</td>
<td>99</td>
</tr>
<tr>
<td>6.3</td>
<td>Collapse of the average flow data.</td>
<td>99</td>
</tr>
<tr>
<td>6.4</td>
<td>Relationship between the link capacity and the expected time until the splitting for two different grouping factors and state transition frequencies.</td>
<td>100</td>
</tr>
<tr>
<td>6.5</td>
<td>Expected time $\langle t_{\text{split}} \rangle$ until the lattice splitting versus the state transition frequency: (a) Effect of different grouping factors, (b) collapse of the splitting time data in lattices of different sizes, (c) collapse of $\langle t_{\text{split}} \rangle$ due to adjusting the link capacities according to the grouping factor, (d) effect of the inertia constant on the splitting times.</td>
<td>102</td>
</tr>
<tr>
<td>6.6</td>
<td>Average size of the larger connected component after the splitting into two parts versus the state transition frequency for various grouping factors and link capacities.</td>
<td>103</td>
</tr>
</tbody>
</table>
List of Tables

1.1 Examples of recent wide-area blackouts in Europe and North America ........................................ 3

2.1 Stages of the overall restoration process and corresponding load restoration rates .................. 20
2.2 Dispatch priorities for the generating units ............. 28
2.3 Number of components of the Swiss system .......... 32

3.1 Parameters for the electrothermal model ................. 48
3.2 Lengths of the transmission lines ......................... 53
3.3 Hourly load levels ......................................... 53
3.4 Power output states of the wind farm and corresponding estimates of the stationary probability .............................. 54
3.5 Transition probability matrix of the implemented wind farm ................................................. 54
3.6 Convected heat loss coefficients ......................... 55
3.7 Estimated probabilities of reaching the transmission line temperature $T_i^\ell$ ............................. 55

4.1 Parameter values of the network degradation model for the base case ........................................ 69

5.1 Examples of models for dynamic processes and their categorization based on the local response of a node to the states of its neighboring nodes ........................................ 83
Chapter 1

Introduction

1.1 Background and Motivation

Large-scale technical networks form the backbone of today’s societies. Electric power systems generate the required electricity and distribute it over wide geographic areas, the Internet enables the modern forms of global communication and railway systems provide an efficient transportation of passengers and goods, only to mention a few examples. Their reliable operation depends on a sophisticated interplay of technical components, human activities, management procedures and even of exogenous influences such as institutional and legal boundary conditions. Moreover, all these factors are subject to continuous changes, whereas innovative technological trends and ongoing liberalization processes are currently leading to an unprecedented physical and operational complexity.

These developments are of particular relevance to electric power systems. In continental Europe, for example, regional and vertically integrated monopolies have been and are being replaced with an intricate decentralized market structure - but still with big country-to-country differences. The liberalization process further stimulates supranational electricity transactions on the grid and thus carries significantly higher power flows on the cross-border transmission lines. This is demonstrated in figure 1.1 (a), showing the increasing energy exchanges within the
Figure 1.1: (a) Cross-border energy exchanges within the UCTE operational area, adopted from [1]. (b) Globally installed wind power capacity (cumulative), adopted from [2].

synchronized UCTE\(^1\) operational area. However, the existing transmission assets originally are not designed for these new emerging using patterns and are now operated closer to their security margins. At the same time, the production of electrical energy from distributed generation and “new renewable” sources such as wind and solar power is increasing rapidly. This development is illustrated in figure 1.1 (b) by depicting the growth of the installed global wind power capacity between 1996 and 2008. Most of these energy sources are stochastic in nature, implying a higher ratio of fluctuating generation and more variable flows on the network. Furthermore, intelligent demand side management systems are being introduced, allowing electricity consumers to react on real-time price signals [3]. These evolutions, after all, are only possible through a pervasive use of information and communication technologies. The grid planners and transmission system operators (TSOs), responsible for the security of supply, consequently are faced with a less predictable system containing a hardly manageable variety of uncertainties. Recent major power outages underpin the resulting risk of cascading failures over wide geographic areas, which might be triggered already by a minor single event. Table 1.1 summarizes the characteristics of selected blackouts in Europe and North America be-

\(^1\)The Union for the Coordination of Transmission of Electricity (UCTE) is a former international non-governmental association of the continental European transmission system operators. In 2009 the European Network of Transmission System Operators for Electricity (ENTSO-E) took over all its operational tasks.
### 1.1. Background and Motivation

Table 1.1: Examples of recent wide-area blackouts in Europe and North America

<table>
<thead>
<tr>
<th>Date</th>
<th>Location</th>
<th>Size [GW]</th>
<th>People affected</th>
<th>Main causes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nov. 2006</td>
<td>Western Europe (&quot;controlled&quot;)</td>
<td>14</td>
<td>15 Mio. households</td>
<td>High load flow D-NL, violation of the N-1 rule, poor inter-TSO coordination [5]</td>
</tr>
</tbody>
</table>

Between August 2003 and the end of 2006. Although being still very low probability events, their occurrence implies significant impacts on the society in terms of costs and inconveniences and potentially endangers the stability of entire regions.

A comprehensive examination of potential failure scenarios and underlying system weaknesses is a prerequisite for identifying ways in which the risk of such detrimental events can be reduced in the future. This is only achievable by reliability assessment methods which are able to cope with the increasing system complexity and associated uncertainties, including hardly predictable interactions of technical components, relevant actors and the operating environments. Deterministic studies become questionable as an impractically high number of system state combinations might have to be assessed in order to sufficiently cover the relevant failure scenarios. This is exemplified by the deterministic N-1 criterion, which is the common security principle for both planning and operation of electric power systems and focuses on predefined, individual component failures. As shown by various wide-area power outages its application might miss the root causes of such events, leading

---

2EMS: Energy Management System
3The N-1 principle states that a contingency, which is considered as being relevant, must not endanger the security of the interconnected system, i.e. not trigger an uncontrollable cascading failure propagation across the borders of a TSO’s responsibility area [6].
to inappropriate measures for reliability improvements [7]. Seen from a
different angle, deterministic approaches may also lead to overly conser-
vantive and costly operating constraints. Probabilistic approaches may
overcome part of these problems by explicitly addressing the inherent
system uncertainties [8, 9]. However, “traditional” probabilistic relia-
bility approaches such as the fault and event tree methods require to
structure the considered failure mechanisms into logic frames [10]. They
thus imply a number of deficits for their direct use on large-scale tech-
nical networks, such as [8]:

- the system complexity and dynamic component interdependencies
cannot be sufficiently handled,
- intricate human behavior (e.g., errors or late reactions of control
room operators) cannot be captured and
- highly non-linear system responses (e.g., time evolution of cascad-
ing component failures) cannot be easily modeled, leading to an
inadequate scope of the generated scenarios.

Simulation approaches based on Monte Carlo techniques offer basically
a higher flexibility in their application and thus are deemed to be more
suitable for tackling this kind of problems [10]. A considerable research
and development effort is however still needed to adapt them to the spe-
cific system characteristics and to overcome technical implementation
problems [9]. Moreover, recent advances in the basic theory of complex
systems are advocated to constitute a promising alternative approach
for reliability investigations on large-scale technical networks [10]. Meth-
ods borrowed from the field of statistical physics, for instance, may help
to improve the fundamental understanding of the interplay between the
underlying network topology and the dynamic processes taking place on
top of it [11]. These methods have just recently started to gain attention
in other scientific fields and their benefits for the present problem area
thus deserve further substantiation.4

Based on these considerations it is the overall objective of this thesis
to develop modeling and simulation approaches which are capable to
overcome some of above-mentioned shortcomings and to deploy them
both for reliability studies on specific systems and to uncover the basic
mechanisms behind breakdowns of large-scale technical networks.

4A scientifically sound examination of the relevant literature is given in the con-
text of the individual reliability studies presented in the subsequent chapters.
1.2 Contributions

The thesis introduces two fundamentally different yet complementary types of modeling and simulation approaches for investigating the reliability of large-scale technical networks, with particular applications to electric power systems. The first type aims at a quantification of specific systems, by replicating their dynamic behavior in a highly detailed manner with close adherence to the real world. While the results shall allow assessing the reliability of a given system in an accurate and differentiated manner, and may be even taken as a decision basis for improvement measures, their immediate applicability to other systems is however limited. Therefore, the second type of approach is characterized by a high abstraction level and does not target a specific system. It rather aims at identifying general underlying mechanisms by using parsimonious models on synthetic networks, being in line with recent achievements in the field of statistical physics. According to these distinct types, the thesis is divided into two parts. Their classification according to the level of system abstraction and to the generality of the expected findings is illustrated in figure 1.2.

![Diagram of two parts of the thesis](image_url)

Figure 1.2: The two parts of the thesis and their classification according to the level of system abstraction and to the generality of the findings, respectively.
Chapter 1. Introduction

The main contributions with respect to the first part of the thesis are summarized as follows:

I-a A novel probabilistic modeling and simulation approach is introduced for the reliability analysis of electric power systems. The approach combines object-oriented modeling methods with Monte Carlo simulation techniques, allowing to integrate a broad spectrum of different deterministic and stochastic phenomena which may occur. Important benefits include the possibility to consider non-technical factors such as transmission grid operators, to simulate the propagation of component failures explicitly in time and to quantify the system reliability in a multifaceted way.

I-b A simulation speed-up algorithm is introduced in order to overcome the sometimes prohibiting problem of high computational costs. The basic method has previously been proposed in the telecommunications field. In this work it is tailored to the specific applications within stochastic simulations of electric power systems.

I-c A number of case studies are discussed, demonstrating the functionality of the proposed modeling and simulation methods for a quantitative reliability analysis of electric power systems. Besides using “virtual” test systems, the Swiss high-voltage grid is analyzed with respect to expected blackout frequencies and critical components. Furthermore, the impact of fluctuating renewable energy sources on the probability of transmission line overloads is revealed.

The second part of this thesis contains the following main contributions:

II-a A parsimonious model is developed in order to gain qualitative insights into the underlying mechanisms leading to cascading failures in networked systems. The resulting spreading process is described by a set of coupled differential equations.

II-b Using this model, it is uncovered how the topology of the network influences the velocity of failure cascades. This finding is not only relevant for electric power systems, but also for controlling a broader spectrum of spreading processes such as epidemic diseases or dissemination of information on social networks.
II-c By further extending the model, the impact of fluctuating flow patterns on the occurrence of cascading events is demonstrated. These findings, in turn, are relevant for ensuring the security of supply in future energy networks with a large share of distributed and renewable energy sources.

1.3 Definition of Key Terms

In the scope of this thesis the following working definitions of key terms shall be applied. The definitions in the field of reliability engineering adhere to commonly used terminologies and are thought to be valid for technical systems in general. With regard to their specific application to electric power systems further extensions are provided, if necessary.

**Definition 1 (Large-scale technical network).** A *large-scale technical network is a system consisting of a large number of interconnected and geographically distributed (physical) components whose interaction enables the performance of its required function.*

In this definition human actions are included in terms of performing operational sub-functions. A large-scale technical network thus represents the physical part of the corresponding overall *infrastructure system* which is defined by further including social elements such as the regulatory, legal or institutional framework [12]. Examples of large-scale technical networks are electric power grids, gas and water supply systems as well as transportation and telecommunication networks.

**Definition 2 (Reliability).** *Reliability is the ability of a system to perform its required function, under specified conditions and for a specified period of time (adapted from [13]).*

The required function of an electric power system is commonly defined as the electric power being delivered to the customers within “accepted standards and in the amount desired” [14–16]. Reliability of electric power systems incorporates the aspects of adequacy (Def. 3) and security (Def. 4), and can be quantified by the frequency, duration, and magnitude of adverse effects on the supply [14, 15].
Definition 3 (Adequacy). Adequacy is the ability of a system to perform its required function at all times, taking into account scheduled and reasonably expected unscheduled outages of system elements (adapted from [14, 15]).

Definition 4 (Security). Security is the ability of a system to perform its required function, when exposed to a sudden disturbance (adapted from [14, 15]). Security incorporates the aspects of robustness (Def. 5) and resilience (Def. 6).

Sudden disturbances in electric power systems include short circuits or the unanticipated loss of system elements [14, 15].

Definition 5 (Robustness). Robustness is the ability of a system to retain its state (structure and operation) unchanged, when exposed to a sudden disturbance (adapted from [17]). Robustness is an aspect of security (Def. 4).

Definition 6 (Resilience). Resilience is the ability of a system to adapt to a disturbance-induced change in its state (structure and operation), in order to retain its required function or to recover its original state (adapted from [18]). Resilience is an aspect of security (Def. 4).

Definition 7 (Risk). Risk is the combination of the probability of an event and its consequence [19].

1.4 Outline of the Thesis

The chapters within the two parts of the thesis are organized as follows:

Part I: Applied reliability investigations on detailed system models

Chapter 2 presents the object-oriented modeling concept for the reliability analysis of electric power systems. Its application oriented functionality is scrutinized by performing simulations on a virtual test system and on a model of the Swiss high-voltage grid.
Chapter 3 introduces a simulation speed-up algorithm based on a variance reduction technique. Its potential field of application is demonstrated by assessing the impact of fluctuating renewable energy sources on the occurrence probability of transmission line thermal overloads.

Part II: Theoretical reliability investigations on synthetic networks

Chapter 4 introduces a parsimonious model in order to assess the interplay between the network topology and cascading failure spreading mechanisms. A formal description of the dynamic model behavior on different network topologies is derived.

Chapter 5 applies this model to degree-correlated networks. The numerical results provide evidence that a different encapsulation of nodes with a high connectivity might effectively decelerate the velocity of cascading failure spreading.

Chapter 6 presents an extension of the model for analyzing the impact of fluctuating flows on the occurrence probability of cascading link outages, eventually leading to a splitting of the network. The chapter concludes with delineating potential applications of the results to electric power systems with a large share of distributed renewable energy sources.

Chapter 7 Concludes the thesis with a discussion of the two basic modeling and simulation approaches and their potential interplay for a comprehensive reliability analysis of large-scale technical networks.

1.5 List of Publications

The thesis is based on the following publications:


The author’s contributions to these publications are as follows:

He initiated the research leading to publications [2; 3; 4; 5; 7; 8; 9; 10] and had a central role in developing the basic ideas for publications [1; 6]. He designed and performed the research presented in all
publications, whereas the substantial contribution to [6] is the part on object-oriented modeling. As the primary author of all publications except [6] he is responsible for their written material, and as co-author of publication [6] significantly contributed to its writing.
Part I

Applied Reliability
Investigations on Detailed System Models
Chapter 2

Reliability Analysis of Electric Power Systems Using Object-oriented Modeling


2.1 Abstract

The ongoing evolution of the electric power systems brings about the need to cope with increasingly complex interactions of technical components and relevant actors. In order to integrate a more comprehensive spectrum of different aspects into a reliability study and to include time-dependent effects, this chapter introduces a probabilistic modeling and simulation framework combining object-oriented modeling methods with Monte Carlo simulation techniques. Objects represent both technical components such as generators and transmission lines and non-technical components such as grid operators. The approach allows for the calculation of conventional reliability indices and for the estimation of blackout frequencies. Furthermore, the influence of the time
needed to remove line overloads on the overall system reliability can be assessed. The applicability is demonstrated by performing simulations on the IEEE Reliability Test System 1996 and on a model of the Swiss high-voltage grid.

2.2 Introduction

The ongoing evolution of the electric power systems due to market liberalization and the integration of distributed generation is leading to increasingly complex and hard-to-predict interactions of technical components, relevant actors and the operating environments [see chapter 1]. Furthermore, recent large-area blackouts in North America and Europe demonstrated the potential consequences of inadequate operator response times to contingencies [e.g., [15]]. In recent years several advanced methods have been developed to assess the reliability of electric power systems in general and to model and analyze cascading blackouts [e.g., [20, 21]]. However, these approaches do not explicitly simulate the evolution of the events in time and represent the operator intervention to contingencies by using highly simplified models not taking into account the time needed for the corrective action. While Anghel et al. [22] introduce a time-dependent probabilistic approach incorporating a model for the utility response to line overloads, the influence of the response time on the occurrence of cascading line outages remains neglected.

The contribution of this chapter is to present a basic modeling framework which allows for the explicit integration of highly non-linear, time-dependent effects and non-technical factors into a probabilistic reliability assessment. Therefore, object-oriented modeling [23] is combined with classical methods such as Monte Carlo simulation [24]. Objects represent both technical components such as generators and transmission lines and non-technical components such as grid operators. They interact with each other directly (e.g., via the generator dispatch) or via the physical power flows on the network. By means of long-term simulations the statistical data is gathered for the calculation of system reliability indices and for the estimation of blackout frequencies.

The chapter is organized as follows. Section 2.3 introduces the conceptual basics of the modeling framework and the derivation of the different component models. Section 2.4 presents the results of applying
the model to the IEEE Reliability Test System 1996 and to the Swiss high-voltage system. Section 2.5 concludes.

2.3 Modeling Framework

2.3.1 Conceptual basics

The conceptual modeling framework consists in the abstraction of the relevant technical and non-technical components of the electric power system as individual interacting objects. Each object is modeled by attributes and rules of behavior. An example for an attribute is a technical component constraint such as the rating of a transmission line. The rules of behavior are represented by using finite state machines (FSM) and include both deterministic and stochastic time-dependent, discrete events. A deterministic event is, for instance, the outage of a component when reaching a failure threshold, while stochastic processes are probabilistic component failure models using Monte Carlo techniques.

Those behavioral rules and interactions of the components which are directly governed by the physical laws of electricity transmission are modeled separately, resulting in a two-layers concept as depicted in figure 2.1. The lower layer represents the modeling of the technical network components by means of conventional, deterministic techniques such as power flow calculations and the upper layer represents the abstraction of the electric power system with all its technical and non-technical components as individual objects.

The integration of non-technical components is demonstrated by modeling the behavior of the grid operators in case of line overloads. Furthermore, the model accounts for the possible division of the power system into several control areas. To each control area a distinct grid operator and a distinct control object are assigned. The control object is not an abstraction of a technical component as such but rather represents an implementation construct which controls the balance between generation and load within the corresponding control area. The model captures the system behavior over an operational period of one year.
2.3.2 Component models

The components of the power system as being modeled here as objects are $n_L$ loads, $n_G$ generators, $n_T$ transmission lines, $n_B$ busbars and $n_K$ grid operators. Owing to the high flexibility of the modeling framework, other components could be readily included as well. An example is the supporting information and communication infrastructure for system monitoring and control [25].

Loads

The power demand trajectory $D_i(t)$ of load $i$ is described by

$$D_i(t) = \gamma(t)D_i^{max}(1 + \rho(t)) - \Delta D_i(t).$$

(2.1)

The demand factor $\gamma(t)$ is the actual time-dependent percentage of the peak demand $D_i^{max}$ and follows a chronologically changing load profile over the predefined time period of one year. The percent deviation $\rho(t)$ represents stochastic demand fluctuations and is sampled hourly,
assuming a normal distribution with $\rho(t) \sim N(0, \sigma^2)$ and standard deviation $\sigma = 0.0192$ according to [26]. The value of $\rho(t)$ is assumed to be equal for all loads within the same control area. The term $\Delta D_i(t)$ represents the actual amount of partially shed load.

Figure 2.2 shows the FSM as implemented for the load objects. With the exception of the restoration process all the transitions of the four-state model are externally governed by the control object $k$ of the corresponding control area. Partial load shedding occurs only when the control object sheds load due to an operator action for removing a line overload. As soon as the transmission system can be operated within its security margins again, the load object receives the signal to cancel the partial load shedding. The load gets totally disconnected if there is not enough generation capacity available within the entire system to cover its demand or as a consequence of system splitting [see section 2.3.3]. If several loads have to be disconnected within one control area all loads are given the same priority to be shed and are therefore selected randomly.

Figure 2.2: Finite state machine for the load objects
Table 2.1: Stages of the overall restoration process and corresponding load restoration rates, adopted from [27]

<table>
<thead>
<tr>
<th>Δ(t_{tot}^m) [min]</th>
<th>(\nu) [MW/min]</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-30</td>
<td>10.0</td>
</tr>
<tr>
<td>30-60</td>
<td>33.3</td>
</tr>
<tr>
<td>60-90</td>
<td>66.6</td>
</tr>
<tr>
<td>&gt; 90</td>
<td>83.3</td>
</tr>
</tbody>
</table>

The restoration process is started once sufficient generation capacity is available again to cover the disconnected demand, and is modeled by a queue technique. The load which has been disconnected first is also restored first, the subsequent one waits until the previous is reconnected. Based on [27] an incremental overall restoration rate \(\nu(\Delta t_{tot}^m)\) is assumed for four different restoration stages according to table 2.1, where \(\Delta t_{tot}^m\) is the elapsed time measured from the start of the overall restoration process \(m\). Hence, the time needed to reconnect a specific load, \(\Delta t_i^m\), is dependent on the actual overall restoration stage.

**Generators**

The commitment of the generating units is continuously governed by the control object in order to cover the actual demand \(D_k(t) = \sum_{i \in \Omega_k} D_i(t)\) within the respective control area \(k\). Being constrained by the maximum power outputs \(P_j^{max}\), the commitment and economic dispatch follows a heuristic priority list method according to [28] and is implemented in the control object. By using a recursive algorithm and starting with the highest priority, \(D_k(t)\) is equally distributed among the units with the same priority. As their maximum capacity is reached the algorithm proceeds to the next lower priority and so forth. In case \(D_k(t)\) is larger than the available generation capacity within a control area, the control object analogously commits available generating units from the other control areas of the system. As a simplification, ramp rates and maintenance are not considered at this stage of the work. The FSM for the generator object is made up of a two-state model for repairable forced failures being treated as random events as shown in figure 2.3. The repairable forced failures are modeled by an independent stochastic up-down-up cycle assuming stationarity and constant failure and repair rates \(\lambda_j = 1/MTTF_j\) and \(\mu_j = 1/MTTR_j\) respectively. Hence,
2.3. Modeling Framework

Figure 2.3: Finite state machine for the generator objects

this alternating renewable process is characterized by the cumulative distribution functions of the failure-free times $\tau_j$ and repair times $\tau'_j$ respectively, and by the probability $p_j$ that the generating unit is in upstate at $t = 0$ [29]:

\[
F_j(t_u) = Pr\{\tau_j \leq t_u\} = 1 - e^{-\lambda_j t_u}, \quad (2.2)
\]

\[
G_j(t_d) = Pr\{\tau'_j \leq t_d\} = 1 - e^{-\mu_j t_d}, \quad (2.3)
\]

\[
p_j = \frac{\mu_j}{\lambda_j + \mu_j}, \quad (2.4)
\]

where $t_u$ and $t_d$ are the time spans measured from the moment of entering the upstate and forced down state respectively. All state transitions are reported to the control object of the corresponding control area.

Transmission lines

The time variant line flows are calculated by the DC power flow approximation with

\[
P_{\ell}(t) = x_{ab}^{-1}\left(\theta_a(t) - \theta_b(t)\right), \quad (2.5)
\]

where $P_{\ell}(t)$ is the active power flow on line $\ell$ connecting busbar $a$ with busbar $b$, having reactance $x_{ab}$ and phase angles $\theta_a(t)$ and $\theta_b(t)$. The approximate solution of the power flow problem does not allow for the analysis of voltage disturbances. Nevertheless, the DC model is assumed
to be accurate enough for investigating cascading events in meshed networks due to line overloads, as well as for showing the feasibility of the proposed modeling concept.

A five-state model for the basic behavior of the transmission line is used considering outages triggered by its protection device and by independent random failures, see figure 2.4. Thereby, the protection device is modeled by a separate FSM. In this model the protection device has two functions. Firstly, it continuously measures the power flow $P_\ell(t)$ and sends an alarm message (“overload”) to the operator of the control area if $P_\ell(t)$ becomes equal to or larger than the line rating $P_\ell^{\text{max}}$. Secondly, if $P_\ell(t)$ reaches $P_\ell^{\text{out}}$ it disconnects the line. However, as a consequence of the stochastic time-dependent system behavior or the intervention of the operator, $P_\ell(t)$ may again fall back to less than $P_\ell^{\text{max}}$ before
reaching $P_{\ell}^{\text{out}}$ and the protection device returns to the idle state. By following the assumptions made by Zima and Andersson [30] the probability for the line outage increases linearly with the power flow, being zero below $P_{\ell}^{\text{max}}$. Thus, $P_{\ell}^{\text{out}}$ is assumed being uniformly distributed in the interval $[P_{\ell}^{\text{max}}, \beta P_{\ell}^{\text{max}}]$ with $\beta=1.4$. The line is either reconnected if the phase angle difference $\theta_{\ell}(t) = \theta_a(t) - \theta_b(t)$ becomes smaller than the preset value $\theta_{\ell}^{\text{max}} = \eta x_{ab} P_{\ell}^{\text{max}}$ or after a time delay of $\Delta t^c_{\ell}$ which models the time until a manual attempt to re-close the breakers. The parameter $\eta$ is used to avoid an immediate recurrence of the overload, potentially resulting in a persistently repeating state change cycle, and is set to $\eta=0.9$. Analogous to the probabilistic failure model of the generating units [equations (2.2-2.4)], the time to permanent outage and the time to repair follow an exponential distribution with failure rate $\lambda_{\ell}$ and $\mu_{\ell}$ respectively.

### Busbars

Every busbar object continuously calculates its phase angle $\theta_a(t)$ relative to its neighboring busbars:

$$
\theta_a(t) = \frac{P_a^{\text{tot}}(t) + \sum_{b \in \Omega_a} (x_{ab}^{-1} \theta_b(t))}{\sum_{b \in \Omega_a} x_{ab}^{-1}},
$$

(2.6)

where $P_a^{\text{tot}}(t) = \sum_{j \in a} P_j(t) - \sum_{i \in a} D_i(t)$ is the net power injection at busbar $a$, to which several loads and generating units might be connected. This exact yet ‘distributed’ formalism allows avoiding time consuming matrix calculations in case of network decompositions and restorations due to line outages and reconnections. Potential random outages of busbars are not considered.\(^1\)

\(^1\)Busbar failures at substations and their impact on the overall system reliability have been investigated in the scope of a semester thesis, initiated and supervised by this Ph.D. project [31].
Grid operator

The grid operator becomes active in case of transmission line overload contingencies, trying to remove the overload by redispatching the generators or by shedding load if necessary. The basic model for the operator behavior is illustrated for the overload of a tie-line between two control areas, see figure 2.5.

If a tie-line becomes overloaded the protection device sends an alarm message to the two operators of both control areas [compare figure 2.4]. Having the alarm received the neighboring operators try to contact each other with a time delay $\Delta t_{cd}$. The operator which has been assigned responsible for the line then needs some time to find a solution to the overload problem, which is modeled by a time delay $\Delta t_{rd}$. The corrective action to remove the overload is subsequently formulated as a conventional optimal power flow (OPF) problem [28] and is implemented within the control object by using the linear programming (LP) method minimizing potential load shedding, $\Delta D_i$, and the change in generation, $\Delta P_j$, subject to the transmission line constraints and the

---

Figure 2.5: Finite state machine for the grid operator
power balance:

$$
\min z = \sum_{a=1}^{n_B} \left( \omega_a \left( \sum_{j \in a} |\Delta P_j| + W \sum_{i \in a} \Delta D_i \right) \right), \quad (2.7)
$$

subject to

$$
\sum_{j=1}^{n_G} \Delta P_j = - \sum_{i=1}^{n_L} \Delta D_i, \quad (2.8)
$$

$$
-P_j(t) \leq \Delta P_j \leq P_j^{max} - P_j(t), \quad (2.9)
$$

$$
0 \leq \Delta D_i \leq D_i(t), \quad (2.10)
$$

$$
\left| P_\ell(t) + \sum_{a=1}^{n_B} \left( a_\ell^a(t) \left( \sum_{j \in a} \Delta P_j + \sum_{i \in a} \Delta D_i \right) \right) \right| \leq \xi P_\ell^{max}, \quad (2.11)
$$

where $\omega_a$ is the busbar specific distance weighting factor and is set to $\omega_a=1$ for the two busbars at each end of the overloaded line, $\omega_a=10$ for the busbars being one line further away and $\omega_a=100$ for all other busbars within the overall system. The introduction of this weighting defines the actions for line overload alleviation as a local problem [32].

The weighting factor $W=10000$ lets partial load shedding be more expensive relative to the generator redispach. The linear line sensitivity factor $a_\ell^a(t) = \frac{d P_\ell}{d P_{\ell i}^{tot}}$ with respect to busbar $a$ is dependent on the network connectivity at the model time $t$ and is calculated using the conventional matrix method as described in [28]. Equation (2.11) holds for all lines within the two neighboring control areas. Similar to the model for the reconnection of a failed line, the parameter $\xi$ is used to delay the potential recurrence of an overload and is set to $\xi=0.8$.

The procedure for line overloads within a single control area is basically the same, but without the interaction of the operators and by restricting the load control variables $\Delta D_i$ to the busbars and equation (2.11) to the transmission lines within the control area. In order to prioritize the generator redispach within the associated control area, the distance weighting factor is set to $\omega_a=1$ for busbars inside and $\omega_a=100$ for busbars outside the control area.
2.3.3 System splitting

The splitting of the network due to transmission line outages usually leads to an imbalance between load and generation within the separated subsystems. Further, depending on the total inertia of the rotating machines within the separated parts, on the frequency control performance and the protection device behavior of the generators, and on implemented automatic load shedding procedures, this imbalance comes along with a frequency deviation potentially leading to stability problems [33]. The consequences range from small load losses to a total collapse of the subsystem [e.g., [15]]. In order to include load outages as a consequence of a network splitting while avoiding a complicated model with a high amount of parameters to be estimated a highly simplified approach is used. Thereby, an excess of demand within a separated subsystem leads to the immediate disconnection of a minimum number of randomly selected loads so that the excess is at least reduced to zero. An excess of generation is modeled to lead to an immediate reduction of the generator outputs in order to re-establish the balance and implies no load outages. This strong simplification might be insufficient to represent the real system behavior and the amount of disconnected load thus has to be viewed as a rather indicative value for the system resilience regarding the splitting of the network.

2.3.4 Blackout frequency distributions

By means of long-term simulations (i.e., repeatedly over the operation period of one year) the necessary statistical data is gathered for the calculation of conventional reliability indices such as the Expected Energy Not Supplied (EENS). Moreover, frequency distributions of expected blackouts per year are derived. Therefore, let $X$ be a random variable counting the number of blackouts per year greater than a specified size $C$. The size is thereby classified by the unserved energy or the maximum amount of demand not being supplied in the course of an event. The expectation $E(X)$ is approximated by generating $N$ realizations of $X$ and calculating their mean, which represents the observed complementary cumulative frequency of events related to one year, denoted by $\hat{F}_c(C)$:

$$E(X) \approx \frac{1}{N} \sum_{i=1}^{N} X_i \equiv \hat{F}_c(C), \quad (2.12)$$
where $N$ is the number of simulated years.

Assuming that $X$ follows a Poisson distribution, the confidence interval for $E(X)$ can be constructed by using the central Chi-square distribution [34]:

$$\gamma = 1 - \alpha = Pr \left[ \frac{1}{2N} \chi^2_{f^*;\alpha/2} \leq E(X) \leq \frac{1}{2N} \chi^2_{f;1-\alpha/2} \right], \quad (2.13)$$

where $\gamma$ is the confidence level, $\alpha$ is the probability of error, and $f^* = 2 \sum_{i=1}^{N} X_i$ and $f = 2(\sum_{i=1}^{N} X_i + 1)$ are the degrees of freedom.

The blackout events can further be classified into the three outage causes as implemented in the model:

- **Generation inadequacy:**
  Loads are disconnected as not enough generation capacity is available to cover the actual demand within the overall system or within a previously separated subsystem.

- **System splitting:**
  Loads are disconnected as a consequence of the separation of the system.

- **Operator intervention:**
  Load is partially shed in order to remove transmission line overloads.

### 2.4 Case Studies

#### 2.4.1 Application to the IEEE Reliability Test System 1996

**System layout and model parameters**

The three-area IEEE Reliability Test System 1996 (RTS-96) has 73 busbars, 120 transmission lines and 96 generating units [35]. The year-long load data with an hourly resolution provided in [35] serves for modeling the demand trajectories $D_i(t)$. The three areas have a base case peak
load $D_{k,0}^{max}$ of 2850 MW each and are treated as three single control areas with three corresponding control objects and grid operators. The priorities given to the different generator types are shown in table 2.2.

The failure and repair rates for the generators and the transmission lines are taken from [35]. The parameter value for the time until the manual attempt to re-close the breaker of a disconnected line is assumed to be $\Delta t^c = 1h$. Regarding the operator model, $\Delta t^d_c$ is set to 2min.

### Computational results

The modeling framework has been implemented by making use of the AnyLogic™ tool [36]. Two parameter variation studies have been selected in order to demonstrate the practical benefits of the object-oriented modeling framework:

1. sensitivity of the blackout frequency to an increase of the system loading without any operator intervention;
2. influence of the operator response time on the Expected Energy Not Supplied (EENS).

Concerning the first experiment, the system loading level is incremented as $L = D_k^{max} / D_{k,0}^{max}$. The maximum generator outputs, $P_j^{max}$, are augmented by the same factor. In order to gain statistically significant results (i.e., $N \approx 1000$) about 50 hours of simulation time are needed on a single desktop computer (Dell Optiplex GX260 with a Pentium 4 CPU

<table>
<thead>
<tr>
<th>Unit Type</th>
<th>$P_j^{max}$ [MW]</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hydro</td>
<td>50</td>
<td>1</td>
</tr>
<tr>
<td>Nuclear</td>
<td>400</td>
<td>2</td>
</tr>
<tr>
<td>Coal/Steam</td>
<td>350</td>
<td>3</td>
</tr>
<tr>
<td>Coal/Steam</td>
<td>155</td>
<td>4</td>
</tr>
<tr>
<td>Coal/Steam</td>
<td>76</td>
<td>5</td>
</tr>
<tr>
<td>Oil/Steam</td>
<td>197</td>
<td>6</td>
</tr>
<tr>
<td>Oil/Steam</td>
<td>100</td>
<td>7</td>
</tr>
<tr>
<td>Oil/Steam</td>
<td>12</td>
<td>8</td>
</tr>
<tr>
<td>Oil/CT</td>
<td>20</td>
<td>9</td>
</tr>
</tbody>
</table>

*Table 2.2: Dispatch priorities for the generating units*
of 2.66 GHz and 512 MB of RAM). This time was considerably reduced by running several simulations in parallel.

Figure 2.6 shows the resulting complementary cumulative blackout frequencies with respect to the unserved energy per event, \( \hat{F}_c(C_E) \), for four different values of \( L \). Regarding the two lower system loading levels (\( L=1.0 \) and \( L=1.1 \)) the observed complementary cumulative frequencies follow approximately an exponential curve. However, increasing \( L \) to 1.2 already leads to a remarkable increase of large events, while the shape of the curve in the range of the smaller events (up to about \( 10^3 \) MWh) stays qualitatively the same. The value of \( L=1.37 \) represents the maximum system loading level where no line overloads would occur without any stochastic component outages. This loading level can be characterized by a high frequency of large blackouts predominantly in the range between \( 10^4 \) MWh and \( 10^5 \) MWh. In order to further analyze the differences between the overall frequency curves the distributions of the underlying power outage causes have to be considered. The logarithmic histograms of figure 2.7 report the impact of increasing the system loading from \( L=1.0 \) to \( L=1.37 \) on the absolute frequency of blackouts \( f(C_E) \).

![Figure 2.6: Complementary cumulative blackout frequencies for four different system loading levels \( L=1.0, 1.1, 1.2 \) and 1.37 (circles, stars, triangles and diamonds, respectively) without operator intervention. The error bars indicate the 90% confidence interval.](image-url)
caused by generation inadequacy (left hand side) and system splitting (right hand side).

System splitting is the predominant cause of the observed blackouts for both loading levels. In comparison to generation inadequacy the absolute frequencies for this outage mode show a stronger increase and a stronger shift towards larger events when it comes to an increase of the system loading. Hence, the substantial increase of large blackouts as shown in figure 2.6 is mainly the result of an increased frequency of line overloads and subsequent system splitting.

The results of the second parameter variation study are presented in figure 2.8, showing the influence of the operator response time $\Delta t_d$ on the EENS broken down into the different outage causes for the system loading level $L=1.37$. For the interpretation of the results it should be reminded that thermal aspects of the line overloads are not taken into consideration.\(^2\) Under the present model assumptions an operator intervention with a delay of 5 hours after the occurrence of the overload still reduces the EENS due to system splitting by about 30%. On the other hand, an increase of the response time from 15 minutes to 30 minutes leads to a significant increase of the EENS due to system splitting of about 26%. The EENS due to generation inadequacy is increasing with

\(^2\)The thermal limitation of transmission lines as defined by a maximum allowable temperature is investigated in chapter 3.
2.4. Case Studies

Figure 2.8: Influence of the operator response time on the EENS due
to generation inadequacy (left, black bar), operator action
(middle, dark-grey bar) and system splitting (right, light-
grey bar) for $L=1.37$.

the response time as the system is more often separated which, in turn,
reduces the redundancy of the generators within the split subsystems.
The values for the EENS due to the operator intervention are negligible.

2.4.2 Application to the Swiss high-voltage grid

System layout and model parameters

The Swiss electric power system consists of a single control area with an
annual energy consumption of $57.8 \cdot 10^3$ GWh in the year 2006 and a peak
load of about 10 GW [37]. In the same year, the electricity production
totaled to $61.2 \cdot 10^3$ GWh, consisting of 42.2% nuclear, 52.4% hydro
and 5.4% conventional thermal generation. The installed capacity was
about 12 GW. The number of components as used in the model for the
380/220 kV transmission grid are shown in table 2.3. Figure 2.9 shows
a topological illustration.

Based on a particular system snapshot taken on a January morning the
fluctuating power injections $P_{a}^{tot}(t)$ at the different nodes are derived by
using publicly available statistical data [37]. For each hydro power gen-
erator a different production capacity is assigned for the winter half-year
Table 2.3: Number of components of the Swiss system

<table>
<thead>
<tr>
<th>$n_L$</th>
<th>$n_G$</th>
<th>$n_T$</th>
<th>$n_B$</th>
<th>$n_K$</th>
</tr>
</thead>
<tbody>
<tr>
<td>99</td>
<td>34</td>
<td>229</td>
<td>161</td>
<td>1</td>
</tr>
</tbody>
</table>

Figure 2.9: Topological representation of the Swiss 220 kV/380 kV grid [38, 39]

and the summer half-year respectively. The failure and repair rates for all hydro generators are equally set to $\lambda_j = 4.42 \text{y}^{-1}$ and $\mu_j = 0.05 \text{h}^{-1}$, and for all nuclear units to $\lambda_j = 3 \text{y}^{-1}$ and $\mu_j = 0.027 \text{h}^{-1}$. Regarding the transmission lines the failure model parameters are chosen as $\lambda_{e} = 0.234 \text{y}^{-1}$ and $\mu_{e} = 0.056 \text{h}^{-1}$. As the phase shifting transformers have considerable influence on the power flows, corrective injections were calculated for the nodes adjacent to a phase shifting transformer. The energy exchange with the neighboring countries is modeled by independent positive or negative power injections at the surrounding boundary nodes. The parameter values for the time until the manual attempt to re-close the breaker of a disconnected line and for the operator response time are assumed to be $\Delta t^{e}_{f}=1 \text{h}$ and $\Delta t^{r}_{d}=15 \text{min}$, respectively.

**Computational results**

It should be noted that the intention of the analysis was primarily to investigate the applicability of the proposed modeling method to a real system. The computational results thus make no claim to quantify the reliability of the Swiss high-voltage grid in absolute terms.
The AnyLogc™ tool again served as the workbench for the implementation of the model. The estimated blackout frequencies and the histogram of the different outage causes both with respect to the unserved energy per event are depicted in figure 2.10.

The model potentially overestimates the duration of the events and thus the unserved energy as switching operations on lower voltage levels for the reconnection of de-energized loads and the possibility to import extra power from neighboring countries to overcome generation shortages are not taken into consideration. The complementary cumulative blackout frequency follows an exponential curve. Generation inadequacy is the dominant factor regarding the larger events while load shedding for line overload relief becomes important in the range of the smaller events. The influence of load disconnections due to system splitting is significant but the frequency of this outage cause never exceeds the frequency of load disconnections due to generation inadequacy or load shedding due to the operator action. Hence, under the present model assumptions, it can be concluded that the system reliability is somewhat more sensitive to generation outages than to transmission line failures.

Figure 2.10: Left: estimated blackout frequencies for the Swiss system with respect to the unserved energy. The error bars indicate the 90% confidence interval. Right: histogram indicating the distribution of the outages due to generation inadequacy (continuous line), system splitting (dotted line) and load shedding for line overload removal (dashed line).
The benefit of the operator response to line overloads is shown in figure 2.11 where the frequencies of the events with and without operator action are compared. The event size is thereby measured by the maximum unserved demand. The impact of the operator intervention becomes significant in the range of the larger events where a high fraction of blackouts with a size greater than 200 MW is prevented. A large number of disconnected loads due to system splitting and thus a high value for the unserved demand generally needs a high number of subsequently disconnected lines due to overload. Such a sequence of events potentially gives the operator a higher chance to intervene in comparison to a disconnection of a single load due to the outage of a few lines without further cascading failures.

The relative overload frequencies for each transmission line, $h_{\ell}$, are reported in figure 2.12. About 15% of all overload contingencies are occurring on only two lines. Furthermore, several groups of adjacent lines can be identified as being prone to overloads, helping to highlight the most critical system regions.

![Figure 2.11](image_url)

Figure 2.11: Blackout prevention due to operator response to line overloads. Triangles: no operator intervention, circles: operator intervention with $\Delta t' = 15$ min. The error bars indicate the 90% confidence interval.
2.5 Conclusions

This chapter presented an object-oriented hybrid modeling framework for a comprehensive reliability analysis of electric power systems. The main advantages are the explicit integration of highly non-linear, time-dependent effects and the possibility to include non-technical factors. The chosen level of modeling detail allows analyzing a multitude of different (time-dependent) reliability aspects such as the identification of weak points and the assessment of system upgrades. Although several model refinements need to be further developed, the results of the case studies performed on the IEEE RTS-96 and on a model of the Swiss high-voltage grid confirm the applicability of the approach with respect to mid-period power system planning purposes. Optimizing the technical implementation of the models together with the evolution of both hard- and software will fasten up the simulation speed. Gaining experience in applying the proposed approach will give insight in the parameters to be used, thus lessen the problem of the high number of parameters to be estimated from empirical data.

Figure 2.12: Relative frequency of transmission line overloads ($N=2242$)
Chapter 3

Modeling and Simulation of Line Temperatures under Fluctuating Power Flows

Parts of this chapter have been approved for publication in M. Schläpfer and P. Mancarella, \textit{Probabilistic modeling and simulation of transmission line temperatures under fluctuating power flows}, IEEE Transactions on Power Delivery.

3.1 Abstract

Increasing shares of fluctuating renewable energy sources induce higher and higher power flow variability in the existing transmission networks. This calls for revisiting classical approaches to line temperature estimation, relevant to reliability aspects such as thermal rating and contingency anticipation. With respect to short-term power system planning, this chapter introduces a probabilistic modeling and simulation approach for estimating the occurrence of maximum allowed line temperatures in the presence of fluctuating power flows. Cumbersome integration of the dynamic thermal equations at each Monte Carlo trial is sped up by formulating a specific algorithm that makes use of a variance reduction technique adopted from the telecommunications field. A case
study performed on a single line model illustrates the substantial reduction in computational time and provides fundamental insights into the probability of hitting critical line temperatures under given power flow fluctuations. A transmission system application shows how the proposed method exploits these characteristics for a fast yet accurate operational security assessment.

3.2 Introduction

Increasing volumes of fluctuating renewable energy sources, as exemplified by wind energy conversion, are leading to more variable and less predictable power flows in networks [40]. This also implies a decrease in the average network utilization and, possibly, an increase in the probability of hitting thermal limits due to peak flows. While on the long run network assets will eventually be upgraded, the expansion of the transmission system is rather slow with time horizons up to several years. The existing assets therefore need to be used as efficiently as possible. As a consequence, the question arises whether present classical concepts for estimating the loading capability of overhead lines are adequate to cope with these rapid developments. In this respect, it has recently been shown that in order to fulfill the sag clearance requirements, the direct use of temperature rather than power limits allows for a significantly more precise and less conservative loadability assessment [41, 42]. New approaches and tools are crucial not only for contingency analysis, but also for releasing available power transfer capability, potentially underestimated by classical line rating methodologies, to increase the amount of fluctuating renewable energy sources that can be securely integrated in the system as well as to increase the volume of energy that can be traded between nodes. The uncertainty given by the stochastic nature of renewable sources thereby calls for approaches based on probability concepts [43]. In this outlook, existing probabilistic models for different renewable energy sources (e.g., [44–46]), theoretical advances in simulation speed-up techniques (e.g., [47, 48]) and the evolution of computational power pave the way to Monte Carlo-based methodologies.

On these premises, the present chapter introduces a probabilistic modeling and simulation approach for assessing the impact of power flow fluctuations on the occurrence probability of maximum allowed transmission line temperatures. The time-varying temperatures are explicitly
modeled and calculated from the heat balance equations. Coupling to and interaction with the AC power flow variables is carried out through the ohmic losses and the temperature dependent line resistances. Monte Carlo simulations are used to generate the probabilistic information on the line temperature dynamics in the presence of uncertainties adherent to, for instance, fluctuating wind turbine generation or forced outages of conventional generating units. The proposed methodology can be applied regardless of the specific probabilistic model used for capturing the relevant stochastic phenomena. In order to overcome the downside of the slow simulation speed when performing massive Monte Carlo extractions with continuous integration of the heat balance equations, a specific algorithm has been developed for the problem under analysis. It is based on a technique for the fast simulation of rare events called RESTART (REpetitive Simulation Trials After Reaching Thresholds), mainly adopted hitherto in the telecommunications field.

The chapter is organized as follows. Section 3.3 introduces the electrothermal model for the dynamic calculation of transmission line temperatures. Section 3.4 discusses the modeling framework and the Monte Carlo accelerated algorithm for the line temperature probabilistic assessment, also introducing the Markov chain formalism to represent stochastic behavior. Section 3.5 reports the simulation results from a single line example, discussing the effects of different flow fluctuation characteristics on the line temperature dynamics, as well as demonstrating the efficiency of the acceleration algorithm. Furthermore, a case study is carried out for a transmission network with fluctuating wind power injections and generator failure events, highlighting the benefits of the proposed methodology for short-term operational purposes. Section 3.6 concludes the chapter.

3.3 Electrothermal Model

The electrothermal model described here is aimed at calculating the time-varying transmission line temperatures as driven by the power flow fluctuations and by the meteorological conditions. Each unit length of a conductor of a given transmission line $\ell$ is heated by its temperature dependent ohmic losses $Q_\ell^I(T_\ell(t)) = I_{\ell,p}(t)R_\ell'(T_\ell(t))$, with $I_{\ell,p}$ and $R_\ell'$ being the phase current and series resistance per unit length, and additionally by the solar heat gain $Q_\ell^C(t)$ [49]. Convection $Q_\ell^c(T_\ell(t))$ and
radiation \( Q_{\ell}^r(T_{\ell}(t)) \) are responsible for cooling. This heat balance yields the following differential equation for \( T_{\ell}(t) \):

\[
\rho_{\ell} \frac{d}{dt} T_{\ell}(t) = Q_{\ell}^j(T_{\ell}(t)) + Q_{\ell}^c(T_{\ell}(t)) - Q_{\ell}^r(T_{\ell}(t)) - Q_{\ell}^t(T_{\ell}(t)), \tag{3.1}
\]

where \( \rho_{\ell} \) is the heat capacity per unit length of the conductor. As all three conductors have the same temperature, \( T_{\ell}(t) \) can be regarded as the transmission line temperature. Its time evolution is derived by numerically solving equation (3.1), applying standard integration methods. Following the notation given in [41] the convection and radiation terms can be calculated by

\[
Q_{\ell}^c(T_{\ell}(t)) = A_c^\ell(\omega_{\ell}(t))[T_{\ell}(t) - T_{a}(t)], \tag{3.2}
\]

\[
Q_{\ell}^r(T_{\ell}(t)) = A_r^\ell(\omega_{\ell}(t))([T_{\ell}(t)]^4 - [T_{a}(t)]^4), \tag{3.3}
\]

where \( A_c^\ell \) and \( A_r^\ell \) are the convection and radiation heat loss coefficients, \( \omega_{\ell}(t) \) is the vector of weather parameters and \( T_{a}(t) \) is the ambient temperature. Assuming a \( \pi \)-equivalent line model and neglecting the shunt conductance, \( Q_{\ell}^j(t) \) is obtained from the three phase Joule losses \( q_{\ell}^j(t) \) as

\[
Q_{\ell}^j(t) = q_{\ell}^j(t)/3L_{\ell}, \tag{3.4}
\]

where \( L_{\ell} \) is the length of the line and \( q_{\ell}^j(t) \) is obtained as

\[
q_{\ell}^j(t) = \frac{1}{L_{\ell}} \frac{R_{\ell}^t(t)}{[R_{\ell}^t(t)]^2 + [X_{\ell}^t]^2} [V_y(t)^2 + V_z(t)^2
\]

\[ -2V_y(t)V_z(t) \cos(\theta_y(t) - \theta_z(t))], \tag{3.5}
\]

with \( X_{\ell}^t \) being the reactance per unit length of the line, connecting node \( y \) with node \( z \). The variables \( V_y(t), V_z(t) \) and \( \theta_y(t), \theta_z(t) \) denote the voltage magnitudes and voltage angles at nodes \( y \) and \( z \), respectively.

The temperature of a conductor, in turn, affects its resistance, with a behavior that can be approximated by a linear model [41] as

\[
R_{\ell}^t(T_{\ell}(t)) = R_{\ell}^{t,ref}[1 + \alpha_{\ell}^{ref}(T_{\ell}(t) - T_{\ell}^{ref})] \tag{3.6}
\]

where \( \alpha_{\ell}^{ref} \) denotes the thermal resistivity coefficient and \( R_{\ell}^{t,ref} \) is the resistance at the reference temperature \( T_{\ell}^{ref} \). This dependence of the resistance on the actual temperatures cannot be neglected for accurate power flow calculations [50]. However, after a change of the power flow
the voltage magnitudes and angles show only small variations during the resulting temperature transients, having time spans typically in the range of 30min [41]. This allows for a temporary decoupling of the power flow variables from the actual line temperatures, meaning that the resistances need to be updated only periodically after defined time steps or after significant temperature changes. The specific updating rules are to be defined on the basis of each given study case. The variation of the line reactance $X'_\ell$ with the temperature is small and can be neglected [41].

In practice, the operation of transmission lines is usually constrained by thermal ratings, whereas the maximum allowed operating temperatures become converted into ampacities (i.e., maximum current carrying capacities). The ampacity $I_{\ell,s}$ of a transmission line is conventionally calculated on the basis of steady-state thermal ratings [51]:

$$I_{\ell,s} = \sqrt{\frac{Q^c_\ell(T^r_\ell) + Q^r_\ell(T^r_\ell) - Q^s_\ell}{R^r_\ell(T^r_\ell)}},$$

where $T^r_\ell$ is the maximum allowed operating temperature of the line. Determining $I_{\ell,s}$ can be either based on conservative assumptions for the weather parameters, or on the actual conditions [51, 52]. The latter usually allows a higher ampacity but requires the monitoring of the line temperatures and meteorological data.

### 3.4 Line Temperature Assessment

#### 3.4.1 Conceptual basics

As shown above, the electrothermal model enables a dynamic calculation of transmission line temperatures following the time-dependent variation of the power flows. It is then possible to estimate the probability that a transmission line reaches a certain temperature within a given time span by running suitable Monte Carlo simulations.

In order to overcome the sometimes prohibitive downside of the slow simulation speed coming along with the extensive simultaneous solution of both the power flow and heat balance equations, a specific variance reduction algorithm has been developed. The algorithm borrows from the RESTART technique for the fast simulation of rare events proposed...
in the telecommunications field. The methodology is detailed below. In order to exemplify its applicability, Markov chains are used to simulate stochastic behavior of fluctuating generation as well as possible failure occurrences of conventional generation. The selection of Markov chains is motivated by their widespread application in the power system field for this kind of problems. However, the probabilistic temperature assessment methodology introduced here is completely general and independent of the model used to simulate fluctuating power flows. In particular, RESTART techniques can be applied to both Markovian and non-Markovian processes [48].

3.4.2 Accelerated Monte Carlo simulation for dynamic line temperature estimation

The objective of the assessment is to estimate the probability $\gamma$ that the temperature $T_{L}(t)$ of a transmission line $L$ reaches the maximum allowed operating temperature $T_{LR}$ within a time period $[t_0, t_e)$. A Monte Carlo simulation therefore repeatedly samples the chronological state transitions of each relevant system component [24]. An example is the time-varying power output state of a wind farm. The state of the overall power system at each time is then given by the combination of all respective component states. The chronological system state transition process, in turn, is needed for determining the time-varying power flows which eventually allows calculating the dynamics of the transmission line temperatures according to equation (3.1). A crude Monte Carlo method repeats these simulation steps $N$ times within $[t_0, t_e)$ and estimates the probability $\gamma$ as

$$\hat{\gamma} = \frac{1}{N} \sum_{i=1}^{N} \chi_{c,i}, \quad (3.8)$$

with $\chi_{c,i}$ being the zero-one indicator that $T_{L}(t)$ reaches $T_{LR}$ within trial $i$. The accuracy of the estimation can be quantified by its relative error

$$RE(\hat{\gamma}) := \sqrt{\frac{Var(\hat{\gamma})}{E(\hat{\gamma})}} = \sqrt{\frac{1 - \hat{\gamma}}{N\hat{\gamma}}}. \quad (3.9)$$

The basic idea of RESTART is to perform a higher number of simulation trials in those regions of the state space, where the event of interest is....
more often provoked [48, 53, 54]. Opposite to other variance reduction methods for Monte Carlo simulations such as the importance sampling technique described in [24], RESTART has no influence on the sequence of the discrete events in absolute time.

In the following, the adaptation of the method for the line temperature assessment is introduced. First, the temperature state space $[T^0, T^r]$ is divided into $m$ intermediate intervals $[T^0, T^1), [T^1, T^2), \ldots, [T^{m-1}, T^m]$ with thresholds $T^0 < T^1 < \ldots < T^m = T^r$. Starting at $t_0$ with $T^0 \leq T^r(t_0) < T^1$, the line temperature has to pass all intervals in order to reach $T^m$. Moreover, let $p_i$ denote the conditional probability that $T^r(t)$ reaches threshold $T^i$ before the time reaches $t_e$, given that $T^r(t)$ has already passed threshold $T^{i-1}$. The occurrence probability of reaching $T^m$ then becomes

$$\gamma = p_1p_2 \cdots p_m.$$ (3.10)

A crude Monte Carlo method repeatedly simulates the system within $[t_0, t_e)$ [see figure 3.1, left]. The higher a threshold $T^i$ the less sample paths are reaching it and the less accurate is the estimation of $p_i$. In order to compensate for this lack of trials in the regions closer to $T^m$, RESTART stores the system state as soon as $T^r(t)$ reaches a thresh-

![Figure 3.1: Illustration of the RESTART algorithm. Left: crude simulation, right: simulation by means of RESTART with $N = 2$, $m = 3, n_1 = 1, n_2 = 3$ and $n_3 = 2$. The numbers correspond to an exemplary sequence of consecutive simulation steps (see main text). The squares indicate the initial states for the retrials at the different temperature thresholds.](image-url)
old $T^i$, and splits the sample path into $n_{i+1}$ retrials for the time during which it stays above this threshold (see figure 3.1, right). The first $n_{i+1} - 1$ paths are stopped when they again fall below $T^i$ in order to avoid simulation time in the regions away from $T^m$. Only the last path is permitted to proceed so that it becomes a continuation of the original path. Consequently, a larger number of trials for accurately estimating each $p_i$ is achieved.

Different implementation schemes of the RESTART technique have been proposed [for a discussion see, e.g., [54]]. This study applies the ‘global-step’ approach for estimating $\gamma$. Its main advantage is the need to store at most $m$ system states only [48], as indicated by the squares in figure 3.1, right. Thereby, the unbiased estimator of $\gamma$ is given by

$$
\hat{\gamma} = \sum_{k=1}^{N} \chi_{a,k} \left[ N \prod_{i=1}^{m} n_i \right]^{-1},
$$

where $N$ is the total number of main trials starting at $t_0$ and $\chi_{a,k}$ counts how many times $T^i(t)$ reaches $T^m$ in trial $k$. The simulation is stopped when the relative error of $\hat{\gamma}$ becomes smaller than a predefined accuracy level $\epsilon$,

$$
RE(\hat{\gamma}) \approx \sqrt{\frac{\sum_{k=1}^{N} \hat{\gamma}_k^2 - N \hat{\gamma}^2}{N \hat{\gamma}}} < \epsilon,
$$

where $\hat{\gamma}_k = \chi_{a,k}/\prod_i n_i$ is the estimate in trial $k$. In order to maximize the computational gain, the thresholds are chosen in such a way that $\hat{p}_i$ and $n_i$ reach their quasi-optimal values [48]

$$
\hat{p}_i = e^{-2}; \quad n_i = \sqrt{1/(p_ip_{i+1})}.
$$

The position of the thresholds and the values of $n_i$ are roughly determined by performing a pilot run.

As shown by the flowchart in figure 3.2, the algorithm consists of several loops. All variables needed for the calculation of $T^i(t)$, such as the actual generator or load states and the actual temperatures of all transmission lines, are stored in the system state vector $s(t)$. The time-varying component states are thereby governed by case-specific deterministic and probabilistic models. The outer loop starts the main trial $N$ at $t_0$ with the initial system state $s_0$. As soon as this main trial and all triggered retrials have been finished the probability of reaching $T^m$ is estimated,
Figure 3.2: Flowchart of the accelerated algorithm for estimating the probability of line temperature $T^m$. 

3.4. Line Temperature Assessment
together with the relative error. The inner loops simulate the time evolu-
tion of \( s(t) \) and perform all the consecutive retrials if \( T(t) \) reaches
a threshold \( T_i \). The heat balance equation is continuously integrated,
whereas the line resistances are periodically updated as described in
section 3.3. The numbers in figure 3.1, right, illustrate an exemplary se-
quence of the resulting simulation steps. The first trial (step ➀) does not
reach the first temperature threshold and, as in a crude Monte Carlo
approach, is only stopped when the simulated time equals \( t_e \). When
reaching \( T^1 \) the second trial (step ➁) becomes split at system state
\( s_1 \), whereas the first retrial (step ➂) is stopped when falling below this
threshold again. The second retrial (step ➃) is again split when reaching
\( T^2 \) at \( s_2 \). Restarting from this stored system state, both retrials (steps ➄
and ➅) fall again below \( T^2 \). As the continuation of the original path only
the second retrial (step ➆) is further simulated, reaching again \( T^2 \). The
respective first retrial (step ➇) eventually reaches \( T^m \) and the subse-
quently second retrial (step ➈) is stopped when the time equals \( t_e \). Having
thus finished all retrials starting from \( T^2 \), the algorithm proceeds with
the remaining retrial starting from \( T^1 \) (step ➉).

### 3.4.3 Modeling stochastic behavior using Markov chains

Markov chains are used here for generating synthetic time series of
stochastic processes such as the chronological power output state of
a wind farm. A Markov chain \( X(t) \) is a random process fulfilling the
(Markov) property that, given the present state, the future state is in-
dependent of the past state [55]:

\[
Pr(X(t_{n+1}) = j | X(t_1) = i_1, ..., X(t_n) = i_n) = Pr(X(t_{n+1}) = j | X(t_n) = i_n).
\]  

(3.14)

The chain is called discrete if \( X(t) \) takes values in the discrete space,
and time-continuous if these values change in continuous time [55, 56].
The transition probability \( p_{ij}(s, t) \) is defined as

\[
p_{ij}(s, t) = Pr(X(t) = j | X(s) = i) \quad \text{for } s \leq t. 
\]  

(3.15)

If \( p_{ij}(s, t) = p_{ij}(0, t - s) \equiv p_{ij}(t - s) \) for all \( i, j, s, t \) the chain is said
to be time-homogenous. The transition probabilities can be estimated
from real-life time series, whereas the states of the studied system are
3.5. Case Studies

recorded with a certain frequency $\nu = 1/(t - s)$. The transition probabilities can then be written in the stochastic matrix $P(\nu)$:

$$
P(\nu) = \begin{pmatrix}
p_{11}(\nu) & p_{12}(\nu) & \cdots & p_{1n}(\nu) \\
p_{21}(\nu) & p_{22}(\nu) & \cdots & p_{2n}(\nu) \\
\vdots & \vdots & \ddots & \vdots \\
p_{n1}(\nu) & p_{n2}(\nu) & \cdots & p_{nn}(\nu)
\end{pmatrix},
$$

(3.16)

where $n$ is the total number of discrete states. For each row $i$ applies $\sum_j p_{ij}(\nu) = 1$. In a time-continuous Markov chain, the holding time in a given state is exponentially distributed with mean

$$
\tau_i = \frac{1}{(1 - p_{ii})\nu}.
$$

(3.17)

The stationary state probability $P_i^s$ represents the probability that the stochastic process is in state $i$ at a given point of time.

3.5 Case Studies

3.5.1 Example A: Stochastic power injections on a single line

A single line example is used to gain fundamental insights into the impact of different flow fluctuation characteristics on the line temperature dynamics. The values of the power flow variables and line temperature at $t_0 = 0$ are depicted in figure 3.3. The line consists of three ‘Drake’ 26/7 ACSR conductors with a length of 20 km. The corresponding parameters for calculating the temperature behavior are given in table 3.1.

![Figure 3.3: Single-line layout](image)
For simplicity, the meteorological parameters in $\omega_{\ell}$ are kept constant during the simulated time span. The power injection at busbar I is generated by $NG = 60$ single units with an individual output of $PG_g^{\text{max}} = 1$ MW real power and zero reactive power. For the purpose of this experimental study the power output $PG_g(t)$ of each single generating unit $g$ is governed by a simple two-state Markov process, whereas $PG_g(t)$ is either $PG_g^{\text{max}}$, being in up state, or zero, being in down state. The stochastic alternation between these two states is determined by the transition rates $\lambda_g$ and $\mu_g$, implying exponentially distributed holding times

$$F_g(t_{\text{up}}) = 1 - e^{-\lambda_g t_{\text{up}}}, \quad F_g(t_{\text{down}}) = 1 - e^{-\mu_g t_{\text{down}}}, \quad (3.18)$$

where $t_{\text{up}}$ and $t_{\text{down}}$ are the time spans measured from the moment of entering the up state and down state respectively. The state transition frequency is given by

$$f_g = \frac{\lambda_g \mu_g}{\lambda_g + \mu_g} \quad (3.19)$$

and corresponds to the average number of up-down-up cycles per time unit. The value of this simple model is the high flexibility to reproduce a large number of combined power output patterns, $PG_{\text{tot}}(t) = \sum_{g=1}^{NG} PG_g(t)$, while the mean power output $E(PG_{\text{tot}}(t))$ stays constant, allowing to systematically study the impact of different fluctuation characteristics. This is achieved by varying $f_g$ while keeping the ratio $\lambda_g/\mu_g = 1$ constant and by aggregating different numbers of generating units into different clusters, within which the units follow simultaneously the same production cycles over time. We denote the size of such a cluster (i.e., the number of aggregated units) as $C$. Comparison between figure 3.4, left, and figure 3.4, right, shows how an increased value of $f_g$ leads to a faster fluctuation of the injection at busbar I. According to figure 3.4, upper part, compared to figure 3.4, lower part, a smaller number of $C$ is leading to smoother time-series, while a large value implies a strong fluctuation around $E(PG_{\text{tot}}(t)) = 30$ MW.
The algorithm (figure 3.2) has been implemented in Matlab [57] and the simulations were performed on an Intel Xeon E5450 quad-core processor with 2.99 GHz CPU speed. The heat balance equation (3.1) is numerically integrated using the ‘Matlab ode23 solver’ for ordinary differential equations. By making use of the temporary decoupling of the power flow variables from the actual line temperature as discussed in section 3.3, the line resistance becomes updated only in case of a generator output change. Figure 3.5 shows an exemplary power injection sequence and the resulting temperature behavior both based on this simplification and without the temporary decoupling. The excellent match confirms the validity of this model assumption. Additionally, figure 3.5 depicts the steady-state temperatures corresponding to each subsequent power output state. These values differ significantly from the actual temperatures due to the thermal inertia, supporting the need to consider the transient behavior.

For this example $\gamma$ is defined as the probability that the line temperature reaches $T^r_\ell = 100 \degree C$ within the time interval $[t_0 = 0h, t_e = 12h]$. In order to uncover the impact of the power flow fluctuation on the probability of reaching $T^r_\ell$, the sensitivity with respect to both the ‘fluctuation frequency’ $f_g$ and the ‘fluctuation magnitude’ $C$ has been analyzed. The resulting values for $\hat{\gamma}$ are depicted in figure 3.6. Starting with a low value of $f_g$, its increase leads to a higher probability to reach 100 $\degree C$. However, as $f_g$ is exceeding a critical value, $\hat{\gamma}$ starts to decline again.
Figure 3.5: Lower part: combined power injection sequence. Upper part: resulting dynamic temperature behavior. The line marked with circles corresponds to the temporary decoupling of the power flow from $T(t)$, being overlapped by the numerical results without the decoupling (line marked with squares). The line without markers corresponds to the steady-state temperature.

Figure 3.6: Estimates of $\gamma$ as a function of the fluctuation frequency $f_g$ for $C=1,2,3,5$ (triangles, circles, squares and diamonds respectively). The error bars indicate $RE(\hat{\gamma})$. 
This result can be explained by the thermal inertia effects according to equation (3.1). While the combined power injection reaches more often higher values, the average residence time of such combined states begins to fall below the minimum time needed to heat the line up to $T_r$. A larger fluctuation magnitude is leading to a higher probability to hit $T_r$.

The reduction of the simulation time by the proposed acceleration algorithm in comparison to a crude Monte Carlo simulation is shown in figure 3.7, where the decreasing value of $RE(\hat{\gamma})$ as a function of the time is plotted for different values of $f_g$ and $C$. The computation time savings for reaching a desired accuracy level are significant, whereas the efficiency of the proposed algorithm becomes considerably larger with a smaller value of $C$ [compare figure 3.7 a) with figure 3.7 c)] and a smaller value of $f_g$ [compare figure 3.7 b) with figure 3.7 c)], due to the

<table>
<thead>
<tr>
<th>$i$</th>
<th>$T_i$ [$^\circ$C]</th>
<th>$n_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>72</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>85</td>
<td>6</td>
</tr>
<tr>
<td>3</td>
<td>96</td>
<td>7</td>
</tr>
<tr>
<td>4</td>
<td>100</td>
<td>7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$i$</th>
<th>$T_i$ [$^\circ$C]</th>
<th>$n_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>72</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>86</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>96</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>100</td>
<td>6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$i$</th>
<th>$T_i$ [$^\circ$C]</th>
<th>$n_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>72</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>80</td>
<td>6</td>
</tr>
<tr>
<td>3</td>
<td>86</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>92</td>
<td>6</td>
</tr>
<tr>
<td>5</td>
<td>96</td>
<td>7</td>
</tr>
<tr>
<td>6</td>
<td>100</td>
<td>6</td>
</tr>
</tbody>
</table>

Figure 3.7: Decrease of $RE(\hat{\gamma})$ with the simulation time. Dashed-dotted line: crude simulation, continuous line: accelerated simulation. a) $f_g = 0.01h^{-1}$, $C = 5$. b) $f_g = 0.1h^{-1}$, $C = 2$. c) $f_g = 0.01h^{-1}$, $C = 2$. The tables show the selected thresholds and the number of retrials.
lower probability to reach $T^m$. In the case of $f_g = 0.01 h^{-1}$ and $C = 2$, for example, an accuracy of $RE(\hat{\gamma}) = 0.1$ was reached after 1050 seconds of simulation time. Even after 10 hours of simulation, corresponding to $N \approx 7 \cdot 10^5$ trials, this accuracy level could not be reached with the crude approach.

From this example it can be concluded that the probability of reaching a given line temperature depends in a non-trivial manner on the underlying power flow patterns. A ‘critical’ fluctuation frequency for which this probability reaches a maximum value can be identified. Furthermore, under fluctuating flows the actual temperatures can significantly differ from the corresponding steady-state values. The practical relevance of these fundamental findings becomes substantiated by the following application example.

### 3.5.2 Example B: Line temperatures within a transmission network including wind power

This case study demonstrates the application and practical benefits of the discussed approach for assessing the probabilities of reaching specified line temperatures within a transmission network. The exemplary 5-bus network with the electrical characteristics of the transmission lines is taken from [58]. Its single line diagram together with the values of the peak loads are shown in figure 3.8. The additionally assigned line lengths $L_{\ell}$ are given in table 3.2. The fluctuation of the power flows is induced by the demand trajectory, by conventional generating units being subject to independent random failures and by fluctuating wind power injections. A typical operational time horizon of 4 hours [42] is chosen to be analyzed with regard to the line temperature dynamics. Thereby, a deterministic demand trajectory is constructed by taking a typical hourly load curve as given in [35]. Its values in percentage of the peak loads are given in table 3.3. All loads in the system follow simultaneously this curve. The total demand is covered by two conventional generating stations at busbar 1 and busbar 2 and a wind farm at busbar 4.

Busbar 1 serves as the slack bus with its generator assumed to be perfectly reliable, as generation adequacy issues are out of the scope of this study. The second generating station at busbar 2 consists of 8 single combustion turbines with a power output of 5 MW and 3.75 MVAR each.
Figure 3.8: Single line diagram of the 5-bus / 7-branch network indicating the values for the peak loads. A wind farm is connected to busbar 4.

Table 3.2: Lengths of the transmission lines

<table>
<thead>
<tr>
<th>Line</th>
<th>1-2</th>
<th>1-3</th>
<th>2-3</th>
<th>2-4</th>
<th>2-5</th>
<th>3-4</th>
<th>4-5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L_\ell$ [km]</td>
<td>13.1</td>
<td>52.3</td>
<td>39.2</td>
<td>39.2</td>
<td>26.2</td>
<td>6.5</td>
<td>52.3</td>
</tr>
</tbody>
</table>

Table 3.3: Hourly load levels, adopted from [35]

<table>
<thead>
<tr>
<th>Hour</th>
<th>0→1</th>
<th>1→2</th>
<th>2→3</th>
<th>3→4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Load level</td>
<td>88%</td>
<td>92%</td>
<td>100%</td>
<td>97%</td>
</tr>
</tbody>
</table>

These units are subject to independent random failures and repair processes, being conventionally modeled as a stochastic up-down-up cycle [24], again making use of equation (3.18). The corresponding failure and repair rates are taken from [35] and set to $\lambda_g = 1/450h^{-1}$ and $\mu_g = 1/50h^{-1}$, respectively. The ramp rates of these generators are assumed to be sufficiently small to become neglected. The synthetic time series for the fluctuating power output of the wind farm is modeled here by using a Markov chain for an existing wind farm, which has been de-
derived in [59]. The power output states in per unit (i.e., as fractions of the base apparent power $S_B = 100$ MVA), their stationary state probabilities and the transition probability matrix are given in table 3.4 and table 3.5, respectively. It should be noted that the application of other commonly used approaches for modeling wind farms such as the ARIMA technique [60] is possible and straightforward.

For all lines the ‘Drake’ 26/7 ACSR conductor type is taken. In order to reduce complexity, all lines are assumed to be exposed to the same wind speeds in time, being perfectly correlated with the output states of the wind farm. All other weather conditions are assumed to be constant and as of example A. This leads to the convection heat loss coefficients given in table 3.6. The remaining parameters for the electrothermal model are taken from table 3.1. The AC power flow equations have been solved in Matlab using the Matpower package [61]. Initially, at $t_0 = 0\text{h}$, the wind farm is in state 5, three combustion turbines are in the down state and all line temperatures are at the corresponding steady-state values.

<table>
<thead>
<tr>
<th>State</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_w$</td>
<td>0.0112</td>
<td>0.1092</td>
<td>0.2385</td>
<td>0.4026</td>
<td>0.5409</td>
<td>0.8152</td>
<td>0.7624</td>
<td>0.9199</td>
</tr>
<tr>
<td>$Q_w$</td>
<td>-0.0036</td>
<td>-0.0061</td>
<td>0.0024</td>
<td>0.0206</td>
<td>0.0435</td>
<td>0.1012</td>
<td>0.0922</td>
<td>0.1354</td>
</tr>
<tr>
<td>$\hat{P}_i^s$</td>
<td>0.379</td>
<td>0.1514</td>
<td>0.1012</td>
<td>0.08</td>
<td>0.0597</td>
<td>0.0177</td>
<td>0.0674</td>
<td>0.1437</td>
</tr>
</tbody>
</table>

| $P_w$ | 0.0112 | 0.1092 | 0.2385 | 0.4026 | 0.5409 | 0.8152 | 0.7624 | 0.9199 |
| $Q_w$ | -0.0036 | -0.0061 | 0.0024 | 0.0206 | 0.0435 | 0.1012 | 0.0922 | 0.1354 |
| $\hat{P}_i^s$ | 0.379 | 0.1514 | 0.1012 | 0.08 | 0.0597 | 0.0177 | 0.0674 | 0.1437 |

Table 3.4: Power output states of the wind farm in per unit (p.u.) and corresponding estimates of the stationary probability $\hat{P}_i^s$, adopted from [59]

Table 3.5: Transition probability matrix $\hat{P}$ of the implemented wind farm, adopted from [59]
Table 3.6: Convected heat loss coefficients

<table>
<thead>
<tr>
<th>Wind farm state</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_c^i$ [W/mK]</td>
<td>0.948</td>
<td>2.398</td>
<td>3.749</td>
<td>5.093</td>
<td>6.063</td>
<td>7.733</td>
<td>7.432</td>
<td>8.309</td>
</tr>
</tbody>
</table>

Table 3.7: Estimated probabilities $\hat{\gamma}$ of reaching temperatures $T_i^c$

<table>
<thead>
<tr>
<th>Line</th>
<th>45 °C</th>
<th>50 °C</th>
<th>55 °C</th>
<th>60 °C</th>
<th>65 °C</th>
<th>95 °C</th>
<th>100 °C</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-2</td>
<td>0.80</td>
<td>0.45</td>
<td>0.27</td>
<td>0.24</td>
<td>0.12</td>
<td>0.047</td>
<td>0.0016</td>
</tr>
<tr>
<td>1-3</td>
<td>0.38</td>
<td>0.12</td>
<td>0.10</td>
<td>0.0029</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2-3</td>
<td>0.26</td>
<td>0.11</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2-4</td>
<td>0.26</td>
<td>0.11</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2-5</td>
<td>0.56</td>
<td>0.25</td>
<td>0.11</td>
<td>0.098</td>
<td>0.019</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3-4</td>
<td>0.26</td>
<td>0.097</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4-5</td>
<td>0.24</td>
<td>0.089</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1-2, I_{t,s}</td>
<td>0.80</td>
<td>0.45</td>
<td>0.27</td>
<td>0.27</td>
<td>0.12</td>
<td>0.11</td>
<td>0.061</td>
</tr>
</tbody>
</table>

Table 3.7 shows the resulting estimates of $\gamma$ for each transmission line and different temperatures. All values have an accuracy level of $\epsilon < 0.05$. As both the line characteristics and weather conditions are assumed to be the same for all lines, the temperature distribution directly reflects the actual power flow distribution on the network. Furthermore, the values of $\hat{\gamma}$ for different temperatures indicate the strength of the power flow fluctuation on each line. Only line 1-2 reaches the maximum allowed operating temperature of 100 °C with a small probability during $[t_0 = 0h, t_e = 4h]$. This line becomes heavily loaded if the generator at busbar 1 has to supply a large amount of the load in the system. This is given under the condition of a high demand level while, at the same time, the wind power output is low and several combustion turbines at busbar 2 are unavailable. The simulation time needed to compute $\hat{\gamma}(95^\circ C)$ for line 1-2 with $\epsilon < 0.05$ was about 3min, for $\hat{\gamma}(100^\circ C)$ with $\epsilon < 0.1$ about 10min. The maximum temperatures of all other lines are found to be significantly below $T_i^c$ during the analysis period, whereas the probabilities decrease sharply. These observations are referring to both a generally low loading and a small fluctuation of the power flow on these lines.

In order to compare the probabilities derived by the time-continuous integration of the heat balance equation with conventional steady-state
line rating methods, the probability of reaching those currents $I_{t,s}$ is calculated, which would lead to the corresponding steady-state line temperatures under the given weather conditions according to equation (3.7). The resulting values are indicated in table 3.7 for line 1-2. The steady-state approach significantly overestimates the probability of reaching higher temperatures. At the maximum allowed temperature $T_r = 100 \degree C$ the estimation of $\gamma$ differs by about a factor 40 in comparison to the proposed, more accurate approach. Indeed, at this relatively low probabilities the thermal inertia effect considerably reduces the probability to reach such high line temperatures within the selected time horizon, as the fluctuation frequency of the power flow exceeds the critical value identified in example A. This confirmation of the basic insights through a more realistic scenario further shows the relevance of explicitly considering the temperature dynamics for adequately estimating the overload probability of transmission lines under highly fluctuating flows.

### 3.6 Conclusions

The integration of fluctuating renewable energy sources and increasing short-term energy trading over large distances are leading to both a higher variability of the power flows and a higher operational uncertainty within the existing transmission networks. In order to adapt current line capacity assessment methods to this changing situation, this chapter presented a probabilistic approach for estimating line temperatures being subject to various degrees of variability of the power flow. The temperature transients are explicitly taken into account and coupled to power flow solutions in the presence of stochastic behavior. The probabilistic approach makes use of Monte Carlo simulations and has been made computationally efficient by formulating a specific simulation algorithm that deploys a variance reduction technique borrowed from telecommunication applications. The computational results demonstrate that the proposed approach allows for a fast yet accurate assessment of the operational line temperatures and associated overload occurrences. The presented case studies show evidence that the thermal inertia on the transmission lines can significantly reduce the probability of reaching maximum allowed operating temperatures. Due to its generality, the methodology can be combined with a broad spectrum of different probabilistic component models and therefore offers numerous
applications, including a periodic operational security assessment and an accurate network capacity estimation. The practical implementation may be particularly relevant for coping with power fluctuations from large wind farms or for assessing economic transactions within a market framework. The use of online data acquired from existing SCADA\textsuperscript{1} systems, such as loadings, line thermal performance and actual meteorological conditions, as well as incorporating numerical weather predictions would thereby significantly increase the accuracy of the assessment.

\textsuperscript{1}SCADA: Supervisory Control and Data Acquisition
Part II

Theoretical Reliability Investigations on Synthetic Networks
Chapter 4

Failure Propagation Dynamics in Synthetic Networks

Parts of this chapter are published in

and in

4.1 Abstract

Wide-area breakdowns of large-scale technical networks often result from an initial, relatively slow system degradation that eventually evolves into a fast and uncontrollable failure propagation sequence, as has been observed on cascading line outages within electric power grids. In order to gain qualitative insights into the underlying dynamics of such failure processes, this chapter introduces a parsimonious model which - at a high level of system abstraction - represents the technical components and their interdependencies by a synthetic network. It further considers two stress induced component outage types differing from each other in
their time-scale. In the first case the failure is due to an increase in the nodal stress above a critical level and occurs after a small time delay. In analogy to an accelerated “wearout” process, the second failure type takes the whole history of stress into account. The two failure modes are coupled by simple differential equations. The focus of this study is set on the influence of different topologies on the failure propagation and on the resulting node outage sequence leading to network collapse. As a formal description of the decaying networks a master equation approach is proposed and compared with Monte Carlo simulations.

4.2 Introduction

Large-scale technical networks such as electric power grids or telecommunication networks are highly integrated and interdependent systems. Breakdowns of such complex networks are often the result of a relatively slow system degradation escalating into a fast avalanche of component failures, potentially leading to a complete loss of service. While the first few outages might even be independent of each other, the causal failure chains usually become more pronounced in the course of the events, ending up in a fully cascading regime. A prime example is the wide-area power outage on the 14th August 2003 in the Northeastern United States and Canada [15]. The slow degradation started around noon with the outage of a system monitoring tool, further progressed during the afternoon through the independent outage of a generator and several transmission lines and finally evolved into the full cascade at around 16:00.

Understanding the mechanisms of such network collapses is of great importance, as it can help to predict the overall system behavior and to optimize maintenance strategies. It is therefore not surprising that in the last years much effort has been made to understand the roots and the basic characteristics of cascading failures within complex networks. Several studies have analyzed the impact of random or malicious (selective) removal of nodes and links by focusing on structural network robustness [62, 63], by modeling time-delayed node interactions [64], by including simple nodal load models in order to account for static [65–68] or transient [69] load redistribution after a single component failure. One of the main conclusions was that networks with a scale-free
degree distribution\(^1\) are more robust in regard to random node failures than networks with an exponential degree distribution, but exhibit a significantly higher vulnerability regarding deliberate attacks on highly connected nodes [70].

The contribution of this chapter is to present a generic, time-dependent model which captures both the slow degradation process due to independent failure events and the subsequent cascading regime of a collapse within synthetically created network topologies. The probabilistic approach comprises two coupled component outage types. The first failure type occurs, with only a small time delay after an event, when the actual stress raises above a critical level. In analogy to an accelerated wearout process, the second failure type takes the whole history of stress into account. It is not the intention to model the actual behavior of a specific large-scale technical network, but rather to capture some important properties which have to be further substantiated by the investigation on real-world systems. Yet, the model is kept general enough in order to be applicable to a wider range of complex networks where the outage of a node shortens the life expectancy of its neighboring nodes.

The chapter is organized as follows. The next section introduces the abstraction of these two failure modes by a temperature - wearout analogon. Section 4.4 describes the simulation procedure. Section 4.5 presents the simulation results of applying the model to different network topologies. In section 4.6 the degradation process is further simplified so as to become formally describable by a master equation approach, which is subsequently compared with Monte Carlo simulations. Section 4.7 concludes.

### 4.3 Network Degradation Model

#### 4.3.1 Conceptual basics

This study represents a large-scale technical network mathematically by an undirected network (or graph) \(G(N, K)\) with \(N\) nodes being interconnected by \(K\) links. The network is described by its \(N \times N\) adjacency

\(^1\)The degree of a node equals to the total number of links attached to it. The degree distribution \(P(k)\) of undirected networks is then defined as the probability that a randomly chosen node has degree \(k\) [11].
matrix $A(G)$ with elements $\{a_{ij}\}$ equal to 1 if a link connecting node $i$ and node $j$ exists and to 0 otherwise [71].

The modeling framework for the network degradation (or decay) consists of a parsimonious model for the stress on a node, a model for the fast node outage mechanism and a model for the slow node outage mechanism. The distinct models are kept as simple as possible in order to allow capturing the basic spreading behavior of failure events.

### 4.3.2 Nodal stress

At time $t=0$ the network consists of $N_0$ nodes. The model considers independent and stochastic node outages which increase the “stress” on the remaining nodes as they are assumed to take over the load. Therefore, the stress $s_i(t)$ on node $i$ is assumed to increase with the ratio $x_i$ of failed neighboring nodes to the initial node degree. By measuring the internal state of the node by a boolean variable $I_i(t)$, where $I_i(t) = 0$ represents an operating node and $I_i(t) = 1$ a failed node, $x_i(t)$ is given as

$$x_i(t) = \frac{\sum_{j \in \Omega_{0,i}} I_j(t)}{k_{0,i}} = \frac{k_{f,i}}{k_{0,i}},$$

where $k_{0,i}$ is the initial nodal degree, i.e. the number of links from node $i$ to nodes $j \in \Omega_{0,i}$, with $\Omega_{0,i}$ being the set of its initially connected nodes. The variable $k_{f,i}$ thus denotes the number of failed neighbors. Equation 4.1 implies that the strength of influence between two nodes is supposed to be the same for each pair of connected nodes.

The stress $s_i(t)$ is then defined as an arbitrary function of $x_i(t)$,

$$s_i(t) = s_i(x_i(t)).$$

### 4.3.3 Fast outage mechanism

The model for the fast node outage mechanism is based on the theoretical concept that the stress $s_i(t)$ influences a certain physical quantity of node $i$ with a time delay due to inertia effects. The component fails if this quantity reaches a critical value. By using a temperature analogon this physical quantity is referred to as the “operating temperature” $T_i(t)$. 
4.3. Network Degradation Model

The differential equation for the dependence on the current stress is then given by

\[
\frac{dT_i(t)}{dt} = \frac{1}{a} s_i(t) - \frac{1}{\tau} T_i(t). \tag{4.3}
\]

The stress \(s_i(t)\) thus can be interpreted as the actual heat inflow, \(a\) as the heat capacity of the node and \(a/\tau T_i(t)\) as the temperature dependent heat outflow. In order to simplify equation (4.3), the temperature \(T_i(t)\) is normalized with its maximum \(T_i^{\text{max}} = \tau s_i^{\text{max}}/a = 1\). The maximum stress is assumed to be \(s_i^{\text{max}} = 1\) resulting in \(a = \tau\), where \(\tau\) represents the characteristic time constant. The node fails when reaching its “critical” operating temperature \(T_i^{\text{crit}}\) and therefore changes its state from \(I_i = 0\) to \(I_i = 1\). Although only nodes of the same type are considered, it is assumed that \(T_i^{\text{crit}} \in \{0, \ldots, 1\}\) are beta-distributed with probability density function

\[
f(T_i^{\text{crit}}) = \frac{1}{B(p, q)} (T_i^{\text{crit}})^{p-1} (1 - T_i^{\text{crit}})^{q-1}. \tag{4.4}
\]

The parameters \(p\) and \(q\) can be represented as functions of the mean value \(\langle T_i^{\text{crit}} \rangle\) and the variance \(\sigma^2\) of the beta distribution.

4.3.4 Slow outage mechanism

The model for the slow node outage mechanism is based on the concept, that the history of the stress level, as being reflected by the temperature sequence \(T_i(t)\), has a strong influence on the “wearout” of the component. This relationship between the stress and the long-term failure time of a component is for instance used for accelerated component life testings on electronic devices [72]. To include this kind of failure mode, a 0th order kinetic model is adopted where the “contingent” \(A_i(t)\) of node \(i\) is decomposed by the rate \(r_i(t)\):

\[
r_i(t) = -\frac{dA_i(t)}{dt} = qT_i(t) + r_B. \tag{4.5}
\]

The decomposition rate has a part being proportional to the actual operating temperature with \(q\) as a linear factor and a constant part \(r_B\) being independent from \(T_i(t)\) in order to account for spontaneous failures. At \(t = 0\) every node has an initial contingent \(A_i^0\), so that the node fails if all of its contingent is decomposed, i.e. if \(A_i(t) = 0\). Further
Chapter 4. Failure Propagation Dynamics

assuming an exponential distribution for the temperature independent failure times, the mean value is given as

\[ \langle A^0_i \rangle = MTTF \cdot r_B, \quad (4.6) \]

where \( MTTF \) stands for mean time to failure \([29]\) and can be equated to the life expectancy without additional node interaction. The probability density function of \( A^0_i \) is then calculated as

\[ f(A^0_i) = \frac{1}{\langle A^0_i \rangle} e^{-A^0_i / \langle A^0_i \rangle}. \quad (4.7) \]

4.4 Simulation Procedure

With respect to the implementation of the proposed modeling framework a discrete-event based simulation approach has been developed. This allows describing the time evolution of the network topology with basic and very efficient matrix operations. Furthermore, the algorithm has to perform only \( N_0 \) (initial number of nodes) calculations for the complete simulation until the network totally collapses, making numerical integration over the complete simulation range unnecessary. The last point is important due to a mixture of fast and slow dynamics in the model, with relatively small \( \tau \) values (chosen in the order of \( 10^3 \) model time units), and large \( MTTF \) values (chosen in the order of \( 10^7 \) model time units).

The simulation algorithm comprises the following main steps:

1. Construct the \( N \times N \) network adjacency matrix with \( N = N_0 \). Set \( t_{(0)} = 0 \) and initialize the temperature and contingent of each node to \( T_i(t_{(0)}) = 0 \) and \( A_i(t_{(0)}) = A^0_i \) respectively. Determine for every node the point in time \( t^w_i = A^0_i / r_B \) when its contingent is totally decomposed and the node fails. The time of the first failure event is then given as \( t^\text{out} = \min(t^w) \) with \( t^w = [t^w_1 \ t^w_2 \ \cdots \ t^w_N] \). Build the vector \( t^\text{temp} \) whose \( N \) elements represent the time values when the nodes reach \( T_i^{\text{crit}} \) and thus are initially set to \( t^{\text{temp}}_i >> \max(t^w) \). Set the simulation step to \( n = 1 \).

2. Proceed the simulation to \( t_{(n)} = t^\text{out} \). Remove the failed node with index \( j \) from the network. Update the number of operating
Simulation Procedure

Nodes to $N = N - 1$. For the set of formerly adjacent nodes re-calculate the actual values of stress, temperature and contingent. The stress $s_i(t_{(n)})$ of node $i$ is evaluated using equation (4.2) and remains constant at least until the next node failure. Based on the corresponding values at the time of the previous event, $t_{(n-1)}$, the temperature $T_i(t_{(n)})$ and the contingent $A_i(t_{(n)})$ are given by

$$T_i(t_{(n)}) = s_i(t_{(n-1)}) \left[ 1 - e^{-\frac{1}{\tau} \Delta t_{(n)}} \right] + T_i(t_{(n-1)}) e^{-\frac{1}{\tau} \Delta t_{(n)}}, \quad (4.8)$$

$$A_i(t_{(n)}) = A_i(t_{(n-1)}) - q \left[ s_i(t_{(n-1)}) \Delta t_{(n)} + \frac{\tau s_i(t_{(n-1)}) e^{-\frac{1}{\tau} \Delta t_{(n)}} - \tau s_i(t_{(n-1)})}{\tau T_i(t_{(n-1)})} - \frac{-\tau T_i(t_{(n-1)}) e^{-\frac{1}{\tau} \Delta t_{(n)}} + \tau T_i(t_{(n-1)})}{-r_B \Delta t_{(n)}} \right], \quad (4.9)$$

where $\Delta t_{(n)} = t_{(n)} - t_{(n-1)}$ and $i \in \Omega_{n,j}$.

3. If $s_i(t_{(n)}) \geq T_{i,\text{crit}}$ determine the subsequent time span $\Delta t_{i,(n+1)}$ until the outage due to reaching the critical temperature,

$$\Delta t_{i,(n+1)}^{\text{temp}} = -\tau \ln \left( \frac{T_{i,\text{crit}} - s_i(t_{(n)})}{T_i(t_{(n)}) - s_i(t_{(n)})} \right). \quad (4.10)$$

Recalculate the point in time when these nodes fail as $t_{i,\text{temp}} = t_{(n)} + \Delta t_{i,(n+1)}^{\text{temp}}$ and update $t_{\text{temp}} = [t_{1,\text{temp}}, t_{2,\text{temp}}, \ldots, t_{N,\text{temp}}]$ accordingly. Determine the time of the next node outage due to reaching $T_{i,\text{crit}}$ as $t_{\text{out,temp}} = \min(t_{\text{temp}})$.

For calculating the time span until the contingent is decomposed solve equation (4.9) for $\Delta t_{i,(n+1)}^{\text{w}}$. Neglect the exponential terms if their relative influence on the contingent decomposition becomes smaller than a preset value,

$$\left( \tau q \left( s_i(t_{(n)}) - T_i(t_{(n)}) \right) e^{-\Delta t_{i,(n+1)}^{\text{w}}/\tau} / A_i(t_{(n)}) \right) \leq \alpha, \quad (4.11)$$

so that $\Delta t_{i,(n+1)}^{\text{w}}$ becomes

$$\Delta t_{i,(n+1)}^{\text{w}} = \frac{A_i(t_{(n)}) + \tau q \left( s_i(t_{(n)}) - T_i(t_{(n)}) \right)}{q s_i(t_{(n)}) + r_B}. \quad (4.12)$$
Otherwise, solve equation (4.9) for $\Delta t_{w,i,(n+1)}$ numerically. For each affected node recalculate the point in time when it fails due to wearout as $t_{w,i} = t_{(n)} + \Delta t_{w,i,(n+1)}$ and update $t^w = [t^w_1 \ t^w_2 \ \cdots \ t^w_i]$ accordingly. Determine the time of the next node outage due to wearout as $t^{out,w} = \min(t^w)$.

4. Select the node which fails next and determine the time of the event as $t^{out} = \min\left(t^{out,temp}, t^{out,w}\right)$.

5. While $N > 1$ increment the simulation step to $n = n + 1$ and go back to step 2.

### 4.5 Computational Results

#### 4.5.1 Experimental setup

The described failure propagation model is applied on two basic types of network topologies: random graphs of the Erdős-Rényi type [73] and scale-free networks generated by the preferential attachment method according to Barabási and Albert [74]. The random graphs have a Poissonian degree distribution

$$P(k_0) \approx \langle k_0 \rangle^{k_0} e^{-\langle k_0 \rangle}/\langle k_0 \rangle!, \quad (4.13)$$

where $\langle k_0 \rangle$ is the initial average degree. Scale-free networks have an algebraic degree distribution

$$P(k_0) \propto k_0^{-\gamma}, \quad (4.14)$$

where $\gamma$ is the scale-free exponent. The network generation routines were adapted from the implementation in [75]. Due to their different statistical characteristics the two network types have been chosen for capturing the influence of the topology on the failure spreading process. As the aim is to be as generic as possible they do not necessarily represent real-world topologies of large-scale technical networks. Both synthetic topologies have $N_0 = 2000$ nodes.

As a starting point, the stress $s_i(t)$ is assumed to be equal to the ratio of failed neighboring nodes,

$$s_i(t) = x_i(t). \quad (4.15)$$
Regarding the parameters of the network degradation model a base case is defined with the parameters according to table 4.1. The accuracy factor for the calculation of the wearout time [equation (4.11)] is set to \( \alpha = 0.01 \).

### 4.5.2 Degradation dynamics

Two parameter variation studies are presented and discussed:

1. The sensitivity of the failure propagation process with respect to the connectivity of the networks by using the degradation model parameters of the base case.

2. The influence of the degradation model parameters on the speed of the network degradation.

In both studies the results of the stochastic simulations performed on Erdős-Rényi graphs and scale-free networks are compared. A total of 50 simulation runs have been performed for each parameter variation experiment in order to obtain statistically significant results.

#### Network connectivity

The influence of the network connectivity was analyzed by varying the average nodal degree \( \langle k_0 \rangle \). Figure 4.1 depicts the resulting time sequence of the expected ratio of operating nodes to the total number of initial nodes, \( E(N(t)/N_0) \), for three different values of \( \langle k_0 \rangle \). At the beginning, the degradation speed within networks with a high connectivity is slightly lower than within networks with a rather small number of connections. However, the speed of the full cascade is significantly higher in networks with a large value of \( \langle k_0 \rangle \). Obviously, the failure propagation

Table 4.1: Parameter values of the network degradation model for the base case

<table>
<thead>
<tr>
<th>( \tau )</th>
<th>( \langle T^\text{crit} \rangle )</th>
<th>( MTTF )</th>
<th>( r_B )</th>
<th>( q )</th>
</tr>
</thead>
<tbody>
<tr>
<td>3600</td>
<td>0.5</td>
<td>3.15E (^7)</td>
<td>0.1</td>
<td>1</td>
</tr>
</tbody>
</table>
Chapter 4. Failure Propagation Dynamics

Figure 4.1: Time sequence of the degradation process in Erdős-Rényi random graphs (left) and scale-free networks (right) for different average node degrees. Circles: $\langle k_0 \rangle \approx 10$, triangles: $\langle k_0 \rangle \approx 4$, stars: $\langle k_0 \rangle \approx 2$.

over time is similar in both Erdős-Rényi random graphs and scale-free networks.

The starting point and the speed of the cascading regime becomes more apparent by plotting the average nodal failure rate $\lambda(t) = -\frac{1}{N(t)} \frac{dN(t)}{dt}$ over the time as presented in figure 4.2 for scale-free networks. While for the weakly connected networks with $\langle k_0 \rangle \approx 2$ the rate, and thus the probability that a survived node fails in the next infinitesimal time-instant, is rather small and increases slowly over time. However, it rapidly rises to a peak within networks having an average degree of $\langle k_0 \rangle \approx 4$. The sharpest peak is observed in the networks with the highest connectivity ($\langle k_0 \rangle \approx 10$) whereas the probability that a number of nodes is able to survive the full cascade decreases significantly.

Figure 4.3 shows the time-dependent interplay between the two different node outage mechanisms for scale-free networks with an average node degree of $\langle k_0 \rangle \approx 4$ and $\langle k_0 \rangle \approx 10$ respectively.

A higher degree makes a node inherently more stable in regard to stress generated by the outage of single neighboring nodes in the initial stage, where $N(t)/N_0 \approx 1$. The average consumption of contingent in dense networks is faster, as the failure of a single node will affect a higher
4.5. Computational Results

Figure 4.2: Time sequence of the average nodal failure rate $\lambda(t)$ within scale-free networks with different average node degrees (semi-logarithmic plot). Circles: $\langle k_0 \rangle \approx 10$, triangles: $\langle k_0 \rangle \approx 4$, stars: $\langle k_0 \rangle \approx 2$.

Figure 4.3: Time sequence of the cascading events in scale-free networks broken down into the two different failure modes, whereas $N_f(t) = N_0 - N(t)$. Outages due to reaching $T_i^{\text{crit}}$ (circles for $\langle k_0 \rangle = 4$ and triangles for $\langle k_0 \rangle = 10$), and due to wearout (crosses for $\langle k_0 \rangle \approx 4$ and inverted triangles for $\langle k_0 \rangle \approx 10$).
number of neighboring nodes, increasing their stress. This, in turn, leads to a higher average temperature in the network and a higher number of nodes are becoming prone to failures at the same time. This explains the fact that as soon as the average temperature reaches a certain level the cascade due to the fast node outage mechanism evolves abruptly and at a significantly higher speed. On the other hand, networks with a lower average degree will cascade sooner due to nodes reaching their critical temperature $T_{i}^{\text{crit}}$ as the influence of losing one connection in these types of networks is more dramatic. During the cascading regime the number of node outages due to the decomposition of the contingent increases as well, as the wearout process becomes considerably accelerated according to equation (4.5). At the end of the network degradation process the proportion of the two types of failure events is approximately the same for both average node degrees.

Degradation model parameters

The influence of varying the different parameters of the degradation model on the dynamics of the network collapse is depicted in figure 4.4 for both Erdős-Rényi random graphs and scale-free networks. Only the parameter of interest has been varied within each experiment, while keeping the other model parameters constant according to the base case [see table 4.1].

As expected, reducing the average critical operating temperature $\langle T_{i}^{\text{crit}} \rangle$ implies an earlier and more apparent starting point of the cascading regime [see figure 4.4 a)]. Increasing the characteristic time constant $\tau$ of the fast failure mechanism increases the inertia effect according to equation (4.3) and thus elongates the time spans between the subsequent node outages, resulting in a delayed and slower decay of the degradation curve [see figure 4.4 b)]. A similar delay is observed by decreasing the accelerating effect of the temperature on the wearout process which is achieved by reducing the value of the proportionality factor $q$ in relation to the base rate $r_{B}$ [see figure 4.4 c)]. However, the fast cascading regime persists, as the degradation dynamics due to reaching $T_{i}^{\text{crit}}$ remains unchanged. Regarding the last parameter variation experiment the mean time to failure of the stress independent outage mechanism is increased from $MTTF=864'000$ to $MTTF= 2'592'000$ [see figure 4.4 d)]. The resulting time sequences show that this parameter seems to have no influence regarding the basic degradation mechanics, as the curves seem to be scaled only and their shape is not affected.
Figure 4.4: Influence of the degradation model parameters on the time sequence of the network collapse for Erdős-Rényi random graphs (ER) and scale-free networks (SF) with $\langle k_0 \rangle \approx 4$. a) $T_{\text{crit}}=0.4$ (ER: circles, SF: triangles) and $T_{\text{crit}}=0.6$ (ER: stars, SF: inverted triangles) b) $\tau=900$ (ER: circles, SF: triangles) and $\tau=864'000$ (ER: stars, SF: inverted triangles) c) $q=0.01$ (ER: circles, SF: triangles) and $q=10$ (ER: stars, SF: inverted triangles) d) $MTTF=864'000$ (ER: circles, SF: triangles) and $MTTF=2'592'000$ (ER: stars, SF: inverted triangles).

For the chosen parameter combinations no significant differences in the time sequences of the Erdős-Rényi random graph and the Barabási-Albert scale-free networks can be observed.
Chapter 4. Failure Propagation Dynamics

4.6 Master Equation Formalism

A master equation approach is used to derive a formal description of the dynamic network degradation process. The formalism is restricted to the “wearout” process [equation (4.5)] without inertia effect (i.e., \( s_i = T_i \)). Furthermore, the failure threshold as given by equation (4.4) is not taken into account. Based on these simplifications and in order to derive the intended formalism the overall failure rate \( \lambda_{k_0,k} \) for all nodes having initial degree \( k_0 \) and actual degree \( k \) is introduced,

\[
\lambda_{k_0,k} = cs_{k_0,k}(x_{k_0,k}) + \lambda_B, \tag{4.16}
\]

where \( \lambda_{k_0,k} \) corresponds to the probability of a node with initial degree \( k_0 \) and actual degree \( k \) to fail in a (sufficiently small) time interval \( (t + \delta t] \), given that it is operational at time \( t \). The first term on the RHS, \( s_{k_0,k}(x_{k_0,k}) \), is again the stress, being an arbitrary function of \( x_{k_0,k} = (k_0 - k)/k_0 \) [according to equations (4.1) and (4.2)]. The relative influence of the nodal stress on the overall failure rate is determined by the linear factor \( c \). The second term denotes the base rate. Throughout the remainder of the chapter its value is set to \( \lambda_B = 1 \) (per time unit).

In order to account for different network topologies the time dependent degree-degree correlation \( P(k'|k,t) \) is considered. The degree-degree correlation represents the conditional probability that at time \( t \) a node of degree \( k \) is connected to a node of degree \( k' \). Therefore, a master equation describing the evolution of the nodes is combined with a master equation describing the evolution of the links, whereby the derivation of the formalism follows in a similar manner to [76].

As the failure of a node is a discrete event the approach is based on the following assumption:

**Assumption 1** The total number of nodes \( N(t) \) is large. The number of nodes is sufficiently large so that every event occurs with the frequency given by the probability of that event.

Consider a network with an arbitrary initial degree distribution \( P(k_0) \). At some time \( t \) there exist \( X_{k_0,k}(t) \) nodes with initial degree \( k_0 \) and actual degree \( k \). At the same time the network consists of \( L_{k,k'}(t) \) links connecting nodes of actual degree \( k \) with nodes of actual degree \( k' \). The relation between the expected number of nodes with actual degree \( k \),
\[ X_k(t) = N(t)P(k, t), \] and the expected number of links connected to nodes with actual degree \( k \), \( L_k(t) \), is given by:

\[ L_k(t) = \sum_{k' > 0} L_{k,k'}(t) = k \sum_{k_0 > 0} X_{k_0,k}(t) = kX_k(t). \quad (4.17) \]

The total number of operating nodes \( N(t) \) is then given by:

\[ N(t) = \sum_k X_k(t) = \sum_k \frac{L_k(t)}{k}. \quad (4.18) \]

To satisfy the symmetric characteristic of the link-space matrix, which contains the elements \( L_{k,k'}(t) \), and in accordance with [76] the number of links in the case \( k = k' \) is double counted.

### 4.6.1 Node-space master equation

The expected number of nodes \( X_{k_0,k}(t) \) can change in three ways:

1. A node with initial degree \( k_0 \) and \( k \) neighbors can fail. This decreases \( X_{k_0,k}(t) \) and occurs at rate \( \lambda_{k_0,k} \) according to equation (4.16).

2. One of the neighbors of a node with initial degree \( k_0 \) and actual degree \( k \) can fail, again decreasing \( X_{k_0,k}(t) \). By making use of assumption 1 the expected failure rate of a neighboring node is given by the conditional probability of the actual degree of the neighboring nodes, \( P(k' | k, t) \), times the conditional probability of the initial degree of that node, \( P(k_0 | k', t) \), times the resulting stress-dependent failure rate, \( \lambda_{k_0,k'} \). Hence, the expected (or average) failure rate of a neighboring node becomes:

\[ \langle \lambda_{nn}(t) \rangle_k = \frac{1}{L_k(t)} \sum_{k' > 0} \left[ L_{k,k'}(t) \frac{1}{X_{k'}(t)} \sum_{k_0 > 0} X_{k_0,k'}(t) \lambda_{k_0,k'} \right]. \quad (4.19) \]

The rate of losing a neighbor is then given by multiplying \( \langle \lambda_{nn}(t) \rangle_k \) with the total number of neighbors \( k \).

3. One of the neighbors of a node with actual degree \( k + 1 \) and initial degree \( k_0 \) can fail, increasing \( X_{k_0,k}(t) \). The expected failure rate of such a node is derived analogously to 2 by calculating the average failure rate of neighbors \( \langle \lambda_{nn}(t) \rangle_{k+1} \).
The combination of the three outage modes changing $X_{k_0,k}(t)$ yields the node-space master equation describing the time evolution of the number of nodes with initial degree $k_0$ and actual degree $k$:

$$\frac{dX_{k_0,k}(t)}{dt} = -[\lambda_{k_0,k} + k \langle \lambda_{nn}(t) \rangle_k] X_{k_0,k}(t) + (k + 1) \langle \lambda_{nn}(t) \rangle_{k+1} X_{k_0,k+1}(t). \quad (4.20)$$

### 4.6.2 Link-space master equation

In order to estimate the expected degree of a neighboring node using the two node degree correlation, $P(k'|k,t)$, the number of links $L_{k,k'}(t)$ is needed. The derivation of the link-space master equation closely follows the above mentioned mechanisms of the node-space master equation. The expected number of links $L_{k,k'}(t)$, connecting nodes of actual degree $k$ with nodes of actual degree $k'$, can change in three ways being valid for both ends of a given link:

1. If a neighbor of a $k+1$ node fails and this node, in turn, is connected to a node with actual degree $k'$, a new link $k \leftrightarrow k'$ is formed. This is occurring at a rate $k$ times the expected failure rate of a neighbor of a node with actual degree $k+1$, $\langle \lambda_{nn}(t) \rangle_{k+1}$. Analogously, $L_{k,k'}(t)$ is increased if a neighbor of a node with actual degree $k'+1$ fails and this node is connected to a node with actual degree $k$.

2. One of the neighbors of the two nodes at each end of a link $k \leftrightarrow k'$ can fail with average rate $\langle \lambda_{nn}(t) \rangle_k$ and $\langle \lambda_{nn}(t) \rangle_{k'}$, respectively. This decreases $L_{k,k'}(t)$. As each of the neighbors may fail, the rates are multiplied by the actual number of connecting links, $k-1$ and $k'-1$ respectively.

3. One of the nodes at each end of a link $k \leftrightarrow k'$ can fail with average rate $\langle \lambda(t) \rangle_k$:

$$\langle \lambda(t) \rangle_k = \frac{1}{X_k(t)} \sum_{k_0 \geq 0} X_{k_0,k}(t) \lambda_{k_0,k} \quad (4.21)$$

and $\langle \lambda(t) \rangle_{k'}$ accordingly. This decreases $L_{k,k'}(t)$.
The combination of the three outage modes yields the link-space master equation describing the time evolution of the number of links connecting nodes of actual degree \( k \) with nodes of actual degree \( k' \):

\[
\frac{dL_{k,k'}(t)}{dt} = k L_{k+1,k'}(t) \langle \lambda_{nn}(t) \rangle_{k+1} + k' L_{k,k'+1}(t) \langle \lambda_{nn}(t) \rangle_{k'+1} - \left[(k-1) \langle \lambda_{nn}(t) \rangle_k + (k'-1) \langle \lambda_{nn}(t) \rangle_{k'} \right] + \langle \lambda(t) \rangle_k + \langle \lambda(t) \rangle_{k'} \right] L_{k,k'}(t). \tag{4.22}
\]

### 4.6.3 Comparison with Monte Carlo simulation

The accuracy of the master equation formalism has been tested by comparison with Monte Carlo simulations. Therefore, both a linear and a non-linear function for the stress has been used. The linear function implies that the failure rate of a node according to equation (4.16) increases linearly with the number of failed neighboring nodes:

\[
s_{k_0,k} = x_{k_0,k}. \tag{4.23}
\]

Regarding the non-linear increase in \( x_{k_0,k} \) the stress is assumed to follow a sigmoidal (logistic) function:

\[
s_{k_0,k} = \frac{1}{1 + \exp[-\beta(x_{k_0,k} - \theta)]} \tag{4.24}
\]

Choosing a sufficiently large value for the gain parameter of \( \beta=100 \) allows approximating a step function. The threshold \( \theta \) determines the level where the stress starts to significantly accelerate the wearout process. For generating the degree distribution \( P(k) \), again Erdős-Rényi random graphs [equation (4.13)] and scale-free networks [equation (4.14)] have been used. The system of coupled differential equations (4.20) and (4.22) has been integrated numerically using the standard ‘Matlab ode45 solver’ [57], whereas for both classes of networks the results are averaged over an ensemble of 20 realizations. The Monte Carlo simulations are averaged over 200 network realizations. Regarding the linear stress function, figure 4.5 a) displays the evolution of the total number of nodes, \( N(t) \), for Erdős-Rényi random graphs with an initial number of
Figure 4.5: Comparison of the master equation formalism (mfa) with Monte Carlo (MC) simulations. a) ER, $\langle k_0 \rangle \approx 4$, linear stress function with different values for $c$. b) ER and SFN, $\langle k_0 \rangle \approx 2$, linear stress function with $c=100$. c) ER and SFN, $\langle k_0 \rangle \approx 2$, logistic stress function with $c=100$, $\beta=100$ and $\theta=0.3$. The dashed-dotted lines represent the theoretical decay of an empty graph with $N_0=1000$, i.e. without any node interaction.

$N_0=1000$ nodes, an initial average degree of $\langle k_0 \rangle=4$ and different values for $c$. Figure 4.5 b) shows the results for both scale-free networks and Erdős-Rényi random graphs with an initial average degree of $\langle k_0 \rangle=2$. For the random graphs there is an excellent match between the numerical solution of the master equations and the results of the Monte Carlo simulations. At about 65% decomposition of the scale-free networks the result of the master equations starts to slightly deviate from the Monte Carlo simulation as the speed of the cascade becomes underestimated. With respect to the logistic stress function, figure 4.5 c) depicts the evolution of the total number of nodes for both Erdős-Rényi random graphs and scale-free networks with an initial average degree of $\langle k_0 \rangle=2$. The master equations again capture the evolution of the decaying random networks with a high accuracy, whereas regarding the scale-free networks they start to overestimate the speed of the failure propagation when about 30% of the nodes have failed. Hence, the approach seems to well describe heterogeneous networks, at least if they are characterized by a high or moderately low functionality of the nodes. Further model refinements are only necessary when it comes to the description of a total network collapse.
4.7 Conclusions

The present chapter introduced a generic model allowing to analyze the dynamics of failure propagation emerging from two coupled node outage mechanisms differing in their time scale. The results of the stochastic simulations performed on Erdős-Rényi random graphs and scale-free networks show a strong influence of the network connectivity on the failure time sequences. In weakly connected networks the degradation proceeds more smooth, whereas in highly connected networks the cascading node failures appear slightly delayed but at a significantly higher speed. The proposed master equation approach is capable to describe the evolution of the decaying networks, being in excellent agreement with Monte Carlo simulations, except for the later stages of the decaying scale-free networks. The model and the results from its application help to understand how changes in different parameters affect the overall component failure dynamics, so as to eventually optimize strategies for the prevention of cascading network breakdowns. The theoretical insights gained by the wearout analogy however remain to be further substantiated by the investigation on real-world systems.
Chapter 5

Decelerated Spreading Processes in Degree-correlated Networks

Parts of this chapter have been submitted for publication in

M. Schläpfer and L. Buzna, *Decelerated spreading processes in degree-correlated networks.*

5.1 Abstract

This chapter demonstrates that a broad spectrum of spreading processes taking place on complex networks can be categorized into two fundamentally different types, in which either positive or negative degree-degree correlation decelerates their propagation. This result can be explained by the role of the nodes with a high connectivity and their distribution in the network, while they act either as accelerators or delayers and effectively alter the spreading paths. This finding is relevant for controlling the velocity of many different spreading processes, ranging from cascading failures in electric power grids to dissemination of information and epidemic diseases.
Chapter 5. Decelerated Spreading Processes

5.2 Introduction

Understanding the mechanisms of spreading phenomena is a need shared across many scientific disciplines, with examples as seemingly diverse as reaction diffusion processes, pandemics or cascading failures in electric power grids. Substantial new insights have recently been gained through the application of statistical physics on the study of large-scale networked systems, the results of which show how the underlying connectivity pattern affects the dynamical interactions among the examined elements [11, 77, 78]. Going beyond characterizing the network topology by the essential degree distribution [74], extensive research has focused on degree-degree correlations [11]. A network with a positive degree-degree correlation is usually called assortative, and implies that nodes with a similarly small or large degree tend to be connected to each other [79]. If, by contrast, the nodes tend to be connected to nodes with a considerably different degree, the network is called disassortative, referring to a negative degree-degree correlation. Assortativity is typically found in social networks, disassortativity in biological and technical networks [79]. The impact of correlations on the dynamics appears to be non-trivial [80] and has so far been discussed by analyzing specific processes only. Interestingly, assortativity seems to be a suppressing factor for the spatio-temporal spreading of epidemic diseases [81], while at the same time disassortativity has been suggested to prevent the propagation of perturbations in protein networks [82] and to enhance the robustness of declining company networks [83].

5.3 Spreading Types and Stress Model

A spreading process is broadly defined here as the propagation of a certain nodal state along the links of a network. The two types are then given by the probability of a node to respond to the current states of its neighboring (i.e., directly connected) nodes. In type-I processes, the probability of adopting the propagating state is lower for nodes with a large degree than for nodes with a small degree, given that the same ratio of neighboring nodes already is in that state [see illustration in table 5.1]. In type-II processes, the probability is higher for nodes with a large degree. These response rules are inherent to present models for various phenomena. An example of type I is a model for a declining
Table 5.1: Examples of models for dynamic processes and their categorization based on the local response of a node to the states of its neighboring nodes. In the illustration node $A$ changes its state with probability $P_A$ and node $B$ with $P_B$ respectively, whereas the same ratio of neighboring nodes is already in the propagating state (dark color).

<table>
<thead>
<tr>
<th></th>
<th>Type I: $P_A &lt; P_B$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Declining company network [83]</td>
</tr>
<tr>
<td></td>
<td>Extinction of species [85]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Type II: $P_A &gt; P_B$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Reaction-diffusion processes [86]</td>
</tr>
<tr>
<td></td>
<td>Global epidemics [84]</td>
</tr>
<tr>
<td></td>
<td>Dissemination of information [87]</td>
</tr>
<tr>
<td></td>
<td>Cascading failures in electric power grids [65]</td>
</tr>
</tbody>
</table>

company network, whereas the probability for a company disappearing is inversely proportional to its degree [83]. After losing the same ratio of connected firms a company with a large (initial) degree has still more connections and thus a lower probability to disappear than a company with a small degree. An example of type II is the spreading of epidemics on the air-transportation network [84], since a highly linked city is more susceptible than a city with less links, given that the same ratio of connected cities has an equally infected population. Further examples are listed in table 5.1.

In order to focus on this local response mechanism and to analyze the type dependent effect of degree-degree correlations on the resulting spreading dynamics, the illustration of a decaying undirected network is used, which has been introduced and discussed in chapter 4. The spreading is again captured by a minimal two-state model, whereby the nodes are either fully functional or failed. A node becomes stressed when losing

---

1 Similarly to the declining company network (see text) an extinction sequence follows the deletions from the least to the most connected species (being a deterministic model in fact).

2 A node with a large degree has a higher probability to receive active particles than a node with a small degree, if the same ratio of neighboring nodes has the same density of active particles.

3 Large-degree nodes are likely to get the information at a lower ratio of neighboring nodes being already informed.

4 If a certain ratio of neighboring nodes fails, a node with a larger degree has a higher probability to become overloaded.
one of its neighbors and the stress, in turn, increases its probability to fail during the next time step. The stress $s_i$ on a node $i$ increases with the ratio $x_i = k_{f,i}/k_{0,i}$ of failed neighboring nodes $k_{f,i}$ to the initial node degree, $k_{0,i}$ [chapter 4, equations (4.1) and (4.2)]. Furthermore, $s_i = 0$ if $k_{f,i} = 0$ and the maximum stress $s_i = 1$ is independent of the initial degree and is reached when $k_{f,i} = k_{0,i}$. In order to vary the relative influence of the initial node degree the following formula for the stress $s_i(x_i)$ is introduced:

$$s_i(x_i) = \frac{x_i k_{0,i}^\theta}{1 + x_i (k_{0,i}^\theta - 1)},$$

(5.1)

where $\theta$ is the response parameter. The stress equals to $x_i$ if $\theta = 0$. As shown in figure 5.1 response type I is obtained for $\theta < 0$, as the absolute number of functional neighboring nodes becomes more important. For example, having lost 50% of the neighbors, a node $i$ initially with $k_{0,i} = 30$ neighbors is less stressed than a node $j$ with an initial degree of $k_{0,j} = 10$. Analogously, $\theta > 0$ corresponds to response type II, as the absolute number of failed neighbors becomes more important. For $\theta \gg 0$ the loss of just one neighbor causes maximum stress, independently of $k_{0,i}$. In the limit $\theta \ll 0$ the stress is $s_i \approx 0$ if $0 \leq x_i < 1$, but jumps to

Figure 5.1: Illustration of the stress function described by equation (5.1). For $\theta < 0$ the response of nodes to the loss of their neighboring nodes is according to type I and, inversely, $\theta > 0$, corresponds to response type II.
$s_i = 1$ when all neighbors have failed. Equation (5.1) can be replaced by other functions which qualitatively reproduce the two response types.

Starting at time $t = 0$ with $N_0$ nodes, the network decay (or degradation) is governed by both random (i.e., independent) and stress induced node failures, formulated by the time dependent failure rate of a node $i$:

$$\lambda_i(t) = \frac{f_i(t)}{1 - F_i(t)} = cs_i(x_i(t)) + \lambda_B, \quad (5.2)$$

where $F_i(t) = \int_0^t f_i(u)du$ is the life time distribution function. The parameters $\lambda_B \geq 0$ and $c \geq 0$ account for the random failures and for the stress respectively.

The relation to different classes of spreading models is given by the parameter range of equation (5.1) and equation (5.2). For instance, letting $\lambda_B \to 0$ in the limit $\theta \gg 0$ the susceptible-infected (SI) model is readily recovered as applied to study epidemics in which individuals are either susceptible or infected [88]. The proportional increase of $\lambda_i$, as $\theta = 0$, corresponds to the Bass model for the diffusion of innovations [89]. A further example are binary threshold models for social contagion processes [90, 91]; assuming a high (low) threshold these models can be approximated by setting a correspondingly low (high) value for $\theta$ at $\lambda_B \to 0$ and a high value of $c$.

5.4 Degree-correlated Networks

The spreading processes are studied on degree-correlated scale-free networks. Their degree distribution follows a power law, $P(k) \propto k^{-\gamma}$, which is found in many real-world systems, usually with $2 < \gamma \leq 3$ [77]. The global level of correlation is commonly quantified by the Pearson coefficient $r$ [79]

$$r = \frac{M^{-1} \sum_e j_e k_e - [0.5M^{-1} \sum_e (j_e + k_e)]^2}{0.5M^{-1} \sum_e (j_e^2 + k_e^2) - [0.5M^{-1} \sum_e (j_e + k_e)]^2}, \quad (5.3)$$

with $j_e$ and $k_e$ denoting the degrees of the two nodes being connected by edge $e$ and $M$ as the total number of edges. This measure is defined
in the interval $[-1, 1]$ whereas $r = 0$ corresponds to an uncorrelated network and a positive ($r_+$) (negative ($r_-$)) value denotes positive (negative) correlation. As a first step uncorrelated networks with a given scale-free exponent $\gamma$ have been constructed according to [92], restricting the degree $k_i$ of each node $i$ to $m \leq k_i \leq \sqrt{N_0}$ with $\sum_i k_i$ being even. The reshuffling method [93] subsequently is applied to impose the desired correlation in the range $-0.3 \leq r \leq 0.5$. The positive range includes higher absolute values in order to better account for real life networks. Their maximum degree is not restricted to $\sqrt{N_0}$, so that a higher positive correlation below this cut-off becomes concealed in the value of $r$ as nodes with a larger degree naturally induce negative correlations [92].

## 5.5 Numerical Results

The velocity of the stress propagation determines the nodal life expectancy, which represents the average time span until a randomly chosen node fails: the slower the spreading, the larger its value. Estimating this measure by extensive Monte Carlo simulations shows evidence for disassortativity decelerating spreading processes of type I and assortativity decelerating those of type II. This result is presented in figure 5.2, comparing the nodal life expectancies $\tau(r_+)$ in assortative and $\tau(r_-)$ in disassortative networks, being normalized with the corresponding value of the uncorrelated case, $\tau(0)$. The ensembles of scale-free networks have $N_0 = 10^4$ nodes, characteristic exponent $\gamma = 2.5$ and minimum degree $m = 2$. The parameters of equation (5.2) are set to $\lambda_B = 0.0001$ and $c = 1 - \lambda_B = 0.9999$.

Starting with response type I and a low susceptibility to failures, i.e. $\theta \ll 0$, the decay is dominated by random failures as the stress propagation is slow [see inset to figure 5.2]. Increasing $\theta$ strongly increases the influence of $s_i$ on the nodal life expectancy, with $\tau(r_+) < \tau(0) < \tau(r_-)$. These characteristics are more pronounced for larger values of $|r|$. After observing $\tau(r_+) \approx \tau(0) \approx \tau(r_-)$ in the intermediate range, the spreading sustains with high absolute velocity for type-II processes. Now the nodal life expectancy in assortative networks becomes larger, with $\tau(r_-) \approx \tau(0) < \tau(r_+)$. Finally, in the limit $\theta \gg 0$ the decelerating effect of the assortativity persists, being consistent with earlier findings on epidemic spreading [81].
The observed behavior can be attributed to the role of the nodes with a large degree and their distribution in the network. In type-I processes, they retain the propagation of the stress [see figure 5.1], acting as *delayers*, whereas in processes of type II they become *accelerators*. In assortative networks they become clustered in regions of high connectivity, which can be identified, for example, by the $k$-core decomposition [94]. A $k$-core is the maximum subgraph with all nodes having minimum degree $k$, and a $k$-shell contains the fraction of nodes belonging to the $k$-core but not to the $(k+1)$-core, see figure 5.3 (a) for illustration. Hence, either the delayers become bunched into higher-order $k$-shells, allowing for a fast stress propagation in the lower-order $k$-shells [see figure 5.3 (b)] or the accelerators become separated, effectively decelerating the spreading velocity within the lower-order $k$-shells [see figure 5.3 (c)]. The vast majority of the nodes remains in lower-order $k$-shells [see inset to figure 5.3 (c)] so that this opposite effect becomes directly reflected in the average nodal life expectancy. For type II, this general spreading progression from nodes with a large degree towards nodes with a small degree is again consistent with epidemic models [95].
In processes of type I the retention times of the stress in the delayers are long, so that the propagation, when possible, bypasses through nodes with a smaller degree. The opposite is true for the accelerators in type-II processes. Hence, the different encapsulation of the delayers (respectively accelerators) as captured by the $k$-core decomposition also changes the overall spreading paths through the entire network and thus the time needed until the stress hits a randomly chosen node.
To illuminate this effect, the ‘efficient paths’ through which the stress propagates most likely are examined. Its basic concept has been introduced in [96] for optimal routing in communication networks. The length of a path $P_{i,j}$, connecting node $i$ with node $j$ and containing the set of nodes $\mathcal{V}_P$, is calculated as

$$L^w(P_{i,j}; \nu) = \sum_{\ell \in \mathcal{V}_P, \ell \neq j} k_\ell^{-\nu}, \quad (5.4)$$

where $\nu$ is a parameter accounting for the degree dependent retention times. The efficient path length is then given by the minimum value of $L^w(P_{i,j}; \nu)$ for all possible paths between nodes $i$ and $j$. Averaging over the efficient path lengths between all possible pairs of nodes eventually leads to the average efficient path length $\langle l^w \rangle$. Setting $\nu = 0$ retrieves the geodesic shortest path. The resulting average efficient path lengths for different values of $\nu$ are depicted in figure 5.4. The values for $\langle l^w \rangle$ indeed show a qualitatively similar behavior as the result of the stress propagation [see figure 5.2]. For $\nu < 0$ disassortative networks exhibit a larger value for $\langle l^w \rangle$. The differences at $\nu \approx 0$ stem from the fact

![Figure 5.4: Average efficient path length of correlated networks, $\langle l^w(r) \rangle$, normalized with the uncorrelated reference case, $\langle l^w(0) \rangle$, versus the parameter $\nu$. Each curve corresponds to an average over 100 independent realizations of networks, having parameters as of figure 5.2.](image-url)
that assortative networks have a significantly larger (and disassortative networks a slightly larger) geodesic shortest path length than the uncorrelated reference networks. At \( \nu > 0 \) the average efficient path length becomes larger for assortative networks. Given this good agreement, the stress propagation seems to follow the efficient paths, suggesting that \( \langle l^w \rangle \) might be a robust indicator for the relative influence of degree-degree correlations on the spreading velocity. It further supports the categorization of spreading processes into the two proposed types.

5.6 Conclusions

In summary, this chapter has drawn a global picture on how degree-degree correlation affects the velocity of spreading processes. If the propagation follows response type I, disassortativity decelerates the velocity. Conversely, if the process is of type II, it is assortativity which prevents a fast spreading. The opposite influence can be explained with the underlying role of the nodes with a large degree, effectively altering the spreading paths. This finding eventually can be important for improving the robustness of networks against undesirable spreading processes such as cascading failures in large-scale technical networks or to accelerate the desired propagation of, for example, information in social networks.


Chapter 6

Network Splitting under Fluctuating Flows

Parts of this chapter are published in

6.1 Abstract

This chapter investigates the splitting of networks subject to fluctuating flows. Various flow patterns are produced by different groupings of the nodes, based on their random alternation between two possible states. Extensive Monte Carlo simulations reveal how the time span until the occurrence of a splitting depends on the flow patterns. Increasing the flow fluctuation frequency shortens this time span which reaches a minimum before rising again due to inertia effects incorporated in the model. The size of the largest connected component after the splitting is rather independent of the flow fluctuation frequency but slightly decreases with the link capacities. The findings carry relevant implications for real-world networks, such as electric power grids with a large share of fluctuating renewable energy sources.


6.2 Introduction

Assessing the robustness of networks against failures of nodes and links is an essential research topic across many scientific disciplines. Examples range from the extinction of species in food webs and malfunctions in protein networks to the vulnerability of the World Wide Web and cascading failures in electric power grids. In the last decade, substantial new insights have been gained through the application of methods from statistical physics [11, 77, 97, 98]. Random failures as well as targeted attacks have been addressed by first studying static properties such as different network topologies [70]. Later on, load redistribution models have been introduced to better represent networks supporting the flow of a physical quantity. For example, the load of a node has been defined by its betweenness centrality [67], by the total number of efficient paths passing through it [68], or enriched with stochastic flux fluctuations [99]. While these approaches model the failure propagation in a static manner, the dynamic flow properties have just recently been taken into account [69].

The contribution of this chapter is to investigate the impact of stochastic flow patterns on the potential occurrence of cascading link failures, eventually leading to a network splitting. Therefore, the model considers 2-dimensional lattices with different groups of nodes which randomly alternate between two possible states, i.e. they act as sources or sinks respectively. These state transitions induce time-varying stochastic flows on every link. Once reaching its capacity, a link fails with a time delay due to inertia effects.

The motivation for this dynamic flow model is the integration of fluctuating renewable energy sources (e.g., wind power, photovoltaic systems) into the electric power grid. This implies a higher ratio of non-dispatchable generation which, in turn, leads to less predictable and more fluctuating flows on the network [see chapter 3]. Consequently, the anticipation of undesired situations such as cascading transmission line overloads leading to a network breakdown becomes highly complicated [45]. In such a future infrastructure layout the network merely serves as a backbone for the redistribution of power from regions of energy surplus to regions with net power consumption. As detailed modeling and simulation approaches become limited due to the increased complexity of electric power systems with large share of renewables, a minimal approach is chosen in order to understand the fundamental
physics governing the dynamic behavior leading to a network splitting. Past experience has shown that such a splitting potentially results in a wide-area blackout with severe social and economic consequences [see chapter 1]. Questions to be tackled are: What is the relation between the stochastic behavior of the nodes and the emerging flow patterns on the network? How are these flow patterns affected by different groupings of the nodes? What, in turn, is the impact of these flow patterns on the probability of a network splitting? How do inertia effects influence the potential splitting process?

Although the definition of the model is based on the specific properties of future energy networks, it is expected to reflect basic features of other real-world networked systems, whose robustness is subject to fluctuating flows.

### 6.3 Dynamic Model and Simulation Procedure

The study system incorporates a model for the nodal state alternation, a flow model, a lattice layout model and a model for the cascading link outages.

#### 6.3.1 Stochastic nodal state alternation

The two possible states between which all nodes can alternate assume current injections of $P^+_i = 1$ (node state “up”) when the node acts as a source or $P^-_i = -1$ (node state “down”) when the node acts as a sink. This stochastic up-down-up cycle assumes for every node $i$ constant transition rates $\lambda_i$ and $\mu_i$ respectively. Hence, this alternating process is characterized by the cumulative distribution functions of the up-state and down-state times,

$$F_i(t_u) = 1 - e^{-\lambda_i t_u}, \quad F_i(t_d) = 1 - e^{-\mu_i t_d}, \quad (6.1)$$

where $t_u$ and $t_d$ are the time spans measured from the moment of entering the up-state and down-state respectively.\(^1\) The state transition

\(^1\)This model corresponds to the two-state model for individual generating units as introduced in section 3.5.1.
The frequency of every node is calculated by
\[ f_i = \frac{\lambda_i \mu_i}{\lambda_i + \mu_i} \] (6.2)
and corresponds to the average number of up-down-up cycles per time unit. For simplicity to every node \( i \) the same transition rates \( \lambda_i = \lambda \) and \( \mu_i = \mu \) is assigned, implying the same transition frequency \( f_i = f \). Moreover, the ratio is kept constant at \( \lambda/\mu = 1 \) in order to assign the same probabilities to both possible states.

### 6.3.2 Flow model

The flows on the network with \( N \) nodes are modeled by applying an electrical direct current model based on Ohm’s law. Thereby, the linear relation between the nodal current injections \( P_i \) and the voltages \( V_i \) can be put into matrix form
\[ P = BV. \] (6.3)

The conductance matrix \( B \) comprises the elements \( B_{ij} = -r_{ij}^{-1} \) and \( B_{ii} = \sum_{j \in \Omega_i} r_{ij}^{-1} \), where \( r_{ij} \) is the resistance of each link \((i,j)\) and \( \Omega_i \) is the set of all the directly connected nodes to \( i \). By assuming for simplicity that \( r_{ij} = 1 \) for all links, the flow on a link \((i,j)\) is given by
\[ P_{ij} = V_i - V_j. \] (6.4)

The sum of all the current injections at a given time instant is not necessarily equal to zero due to the stochastic nature of the up-down-up cycle. In order to satisfy the balance condition \( \sum_i P_i = 0 \) at all times, a lack or surplus of the total current injections within the network is compensated by an additional, equally distributed injection \( \pm|\sum_i P_i|/N \) at every node. Nevertheless, the satisfaction of the balance condition implies that the rows of \( B \) are linearly dependent. To make equation (6.3) uniquely solvable, one of the equations in the system is removed and the node associated with that row is chosen as the voltage reference \( V_{ref} = 0 \).

### 6.3.3 Lattice layout and node grouping

The model for the nodal behavior and the resulting flows is embedded into a regular square lattice of \( N \) nodes and \( L = 2N \) links with
periodic (or “wrap-around”) boundary conditions. In this way every node is directly connected to 4 neighbors, thus different conditions for boundary nodes are avoided. Furthermore, the lattice is partitioned into several square groups, each containing an equal number of nodes [figure 6.1 (a)]. All the nodes in a given group are in the same state at all times and alternate states simultaneously. The grouping factor $G$ corresponds to the number of groups in the network, thus $G = N$ represents total stochastic independence between all nodes. As depicted in figure 6.1 (b)-(e), an increased $f$ results in a higher fluctuation frequency of the flows. By further varying the grouping factor $G$, a broad spectrum of different stochastic flow patterns can be reproduced. A high value of $G$ is leading to more smooth flow time-series, while a small value implies a strong fluctuation around the mean value.

### 6.3.4 Link outage model

In order to incorporate inertia effects in the model, the link outage mechanism is based on the concept that the flow $P_{ij}(t)$ determines the “temperature” $T_{ij}(t)$ on the link $(i,j)$ according to

$$
\tau_{ij} \frac{dT_{ij}(t)}{dt} = q_{ij} P_{ij}(t) - T_{ij}(t).
$$

(6.5)

The link fails if $T_{ij}(t)$ reaches its capacity $T_{ij}^{c}$. In order to simplify equation (6.5) it is $q_{ij} = 1$. The parameter $\tau_{ij}$ represents the characteristic time (inertia) constant.

As an example, such an inertia is present in electric power grids where the power flows might heat the transmission lines up to a maximum allowable temperature [see chapter 3].

### 6.3.5 Simulation procedure

With respect to the implementation the discrete-event approach as introduced in chapter 4 has been further tailored to the needs of this study. This allows describing the time evolution of the nodal states and the resulting flows, as well as of the link outages and the resulting lattice status. By means of extensive Monte Carlo simulations the expected time until the splitting of the lattice is estimated. The simulation procedure comprises the following steps:
Figure 6.1: (a) Schematic plot of a $12 \times 12$ lattice ($N = 144$) with grouping factor $G = 9$ (left) and $G = 36$ (right). The coloring of the groups has been used only for illustrative purposes, all groups are stochastically independent with each other. The total flow $\sum_{i<j} |P_{ij}|$ versus time in the left lattice is depicted in (b) for $f = 0.01$ and in (d) for $f = 0.5$. Similar for the right lattice in (c) and (e) respectively.
1. Construct the $N \times N$ lattice adjacency matrix $A$ and the $N \times N$ conductance matrix $B$. For all the nodes $i$ in a single group determine their equal output states $P_i$ at $t = 0$ by a single Bernoulli trial with probability $p = 0.5$. Set the simulation step to $n = 0$, set $t(0) = 0$ and initialize the temperature of each link to $T_{ij}(t(0)) = 0$.

2. Calculate the flow $P_{ij}$ on each link $(i, j)$ by equation (6.4) after solving equation (6.3) for $V$. For all links $(i, j)$ determine the subsequent time step $\Delta t_{ij}^{\text{temp}}$ after which they fail. If $P_{ij}(t_n) \geq T_{ij}^{c}$, this time span is given by

$$\Delta t_{ij, (n+1)}^{\text{temp}} = -\tau_{ij} \ln \left( \frac{T_{ij}^{c} - P_{ij}(t_n)}{T_{ij}(t_n) - P_{ij}(t_n)} \right).$$

(6.6)

For every link calculate the point in time when it fails due to reaching $T_{ij}^{c}$ as $t_{ij}^{\text{temp}} = t(n) + \Delta t_{ij, (n+1)}^{\text{temp}}$ and build the vector $t^{\text{temp}}$ with elements $t_{ij}^{\text{temp}}$. Determine the time of the first link outage as $t^{\text{out, temp}} = \min [t^{\text{temp}}]$. Determine for every node $i$ of the network the point in time $t_{i}^{s}$ when it changes its state. Then, the time of the first state change is given by $t^{\text{change, s}} = \min [t^{s}]$ with $t^{s} = [t_{1}^{s} \; t_{2}^{s} \; \cdots \; t_{N}^{s}]$. Determine the time of the next simulation event as $t^{\text{next}} = \min [t^{\text{out, temp}}, t^{\text{change, s}}]$. Increment the simulation step to $n = 1$.

3. Proceed the simulation to $t(n) = t^{\text{next}}$. Remove the failed link (if any) from the lattice and update $A$ and $B$. Recalculate the output $P_{i}$ of each node $i$ based on equation (6.1). Recalculate the flow $P_{ij}$ and the temperature $T_{ij}$ on each link $(i, j)$. The flow $P_{ij}$ remains constant at least until the next event. The temperature $T_{ij}(t(n))$ is given by

$$T_{ij}(t(n)) = P_{ij}(t(n-1)) \left[ 1 - e^{-\frac{1}{\tau_{ij}} \Delta t(n)} \right]$$

$$+ T_{ij}(t(n-1)) e^{-\frac{1}{\tau_{ij}} \Delta t(n)},$$

(6.7)

where $\Delta t(n) = t(n) - t(n-1)$.

4. For each node and link recalculate $t_{i}^{s}$ and $t_{ij}^{\text{temp}}$ and update $t^{\text{change, s}}$ and $t^{\text{out, temp}}$ respectively as described in Step 2. Determine the time of the next event $t^{\text{next}}$. 

---

**6.3. Dynamic Model and Simulation Procedure**
5. Check the connectivity of the lattice. If it remains connected, increment the simulation step \( n \) and go back to Step 3. Otherwise, stop the simulation.

### 6.4 Numerical Results

#### 6.4.1 Average flows

To clarify the impact of different grouping factors \( G \) and lattice sizes \( N \) on the resulting flow patterns, the average flow per link \( \langle P_{ij} \rangle \) is estimated without considering the link outage model,

\[
\langle P_{ij} \rangle = \frac{1}{L} \lim_{t \to \infty} \left( \frac{1}{t} \int_0^t \sum_{i<j} |P_{ij}(t')| dt' \right) \\
\approx \frac{1}{t_{tot}} \sum_n \sum_{i<j} (|P_{ij}(t_{(n-1)})| \Delta t_{(n)}) ,
\]

where \( t_{tot} = \sum_n \Delta t_n \) is the sampled overall time span. The average flow is independent of \( f \), as by increasing the frequency the relative duration among all different nodal state combinations remains unchanged. Figure 6.2 shows the values of \( \langle P_{ij} \rangle \) versus the grouping factor \( G \) in lattices of different sizes \( N \). The average flows increase with the lattice size because for a given \( G \) the number of nodes in a group increases with \( N \) leading to a higher current exchange between the groups. Moreover, \( \langle P_{ij} \rangle \) decreases with the number of independently alternating groups. For a given lattice size \( N \), increasing \( G \) implies less nodes in the groups and thus less exchange among them. Seen from a different angle, a high value of \( G \) means that less nodes behave simultaneously in the same way, leading to a more local current exchange and less flows in the lattice. In contrast, a low \( G \) induces higher flows over longer distances. Interestingly, for a given \( N \) the decrease of the average flow with \( G \) follows a power law

\[
\langle P_{ij} \rangle \propto G^{-b}, G \geq 9 .
\]

The exponent \( b \) is rather small and slightly increasing with the size of the lattice.

As shown in figure 6.3 the data in figure 6.2 collapse onto a single curve, if the average flows are scaled with \( \sqrt{N} \). This result can be explained
Figure 6.2: Average flows $\langle P_{ij} \rangle$ versus the grouping factor $G$, which are well fitted by a power law with characteristic exponents $b$, slightly increasing with the lattice size $N$. The dotted lines serve as a guide to the eye.

Figure 6.3: Collapse of all the average flow data shown in figure 6.2 by scaling as $\langle P_{ij} \rangle / \sqrt{N}$, versus the grouping factor $G$. The collapsed data follow a power law with characteristic exponent $-0.27$. The dotted lines serve as a guide to the eye.

by the flow distribution on the lattice. The average flow is largely determined by the maximum flows which are encountered at the boundaries of the groups. Suppose two lattices with sizes $N_1$ and $N_2$ and same grouping factor $G$. Then the maximum possible flows induced by a (square) group on one of its boundary links, $P_{1}^{\text{max}}$ and $P_{2}^{\text{max}}$, are ap-
proximately proportional to $\sqrt{N_1/G}$ and $\sqrt{N_2/G}$ with the same factor respectively. Thus $P_1^{max}/P_2^{max} \approx \sqrt{N_1/N_2}$.

6.4.2 Lattice splitting

The robustness of a lattice is quantified by the expected time $\langle t_{split} \rangle$ when a splitting occurs and the lattice breaks into two parts. Note that the cascading link outages might propagate further, eventually leading to a lattice split in more than two components. Nevertheless, the simulation stops at this point as in real electric power grids a first splitting might already lead to a complete system breakdown (blackout) due to instability phenomena [see chapter 2].

The time span until the splitting can be interpreted as the life expectancy of the lattice and depends on the capacity $T^c_{ij}$ of each link $(i, j)$. To simplify matters, the same values $T^c_{ij} = T^c$ and $\tau_{ij} = \tau$ are assigned to all links. Figure 6.4 shows the behavior of $\langle t_{split} \rangle$ versus an increasing value of $T^c$ for two different grouping factors $G$ and state transition frequencies $f$.

![Figure 6.4: Relationship between the link capacity $T^c$ and the expected time until the splitting, $\langle t_{split} \rangle$, of a $24 \times 24$ lattice ($N = 576$) for two different grouping factors $G$ and state transition frequencies $f$. The inertia constant is set to $\tau = 1$. The error bars indicate the 95% confidence interval. The dotted lines serve as a guide to the eye.](image)
The expected time until the lattice splits increases exponentially with the link capacity. Hence, $\langle t_{\text{split}} \rangle$ is highly sensitive with respect to small changes of $T^c$. For a given value of $T^c$, a larger grouping factor $G$ leads to a significantly higher value of $\langle t_{\text{split}} \rangle$, as less flows are induced [figure 6.2]. The effect of varying the state transition frequency $f$ is similarly large and is examined in more detail in figure 6.5. Starting with a low value, an increase of $f$ leads to a shorter time span until the combined nodal states induce those minimum flows which are needed for the temperatures $T_{ij}$ to reach the capacities $T^c$ [figure 6.1]. Consequently, as depicted in figure 6.5 (a)-(d), the splitting times $\langle t_{\text{split}} \rangle$ are high at low values of $f$ and become significantly decreased as the value of $f$ is increasing. However, as $f$ is exceeding a certain value, the splitting times start to increase again, and the lattice becomes more robust. This result can be explained by the inertia effects according to equation (6.5).

While the flows on the lattice reach more often higher absolute values [figure 6.1], the average residence times of the underlying nodal states begin to fall below the minimum time needed to heat the links up to their capacity $T^c$.

With a small number of groups the combined output is more fluctuating between the extreme values [figure 6.1, left, compared to figure 6.1, right)]. This, in turn, is leading to a higher probability to encounter high flows and high link temperatures in a given time span. The splitting times $\langle t_{\text{split}} \rangle$ thus are significantly shorter, as depicted in figure 6.5 (a).

By considering the scaling behavior of the average flows [figure 6.3], the values of $\langle t_{\text{split}} \rangle$ collapse for lattices with different sizes, but equal model parameters otherwise, if the link capacities are set as $T^c \propto \sqrt{N}$. This result is demonstrated in figure 6.5 (b) for three different lattice sizes. Notice that for a given value of $f$ the average splitting times without adjusting $T^c$ differ by several decades [figure 6.5 (b), inset].

As shown in figure 6.5 (c) the link capacities $T^c$ can be adjusted in such a way, that the splitting times in lattices with equal $N$ but different grouping factors $G$ overlap for a wide range of the state transition frequency $f$. The adjusted capacities can be fitted by a power law with characteristic exponent $-0.23$ [figure 6.5 (c), inset], being remarkably close to the characteristic exponent $b$ of equation (6.9).

The effect of the inertia is shown in figure 6.5 (d) by varying the inertia constant $\tau$. Without any inertia, i.e. $\tau = 0$ implying $T_{ij}(t) = P_{ij}(t)$ [equation (6.5)], the splitting time declines with slope $1/f$.  

6.4. Numerical Results
Figure 6.5: Expected time $\langle t_{\text{split}} \rangle$ until the lattice splitting versus the state transition frequency $f$. If not stated otherwise, the lattices have size $N = 576$. The error bars indicate the 95% confidence interval. (a) Effect of different grouping factors $G$. The parameters of the dynamic model are set to $\tau = 1$ and $T^c = 4$. (b) Collapse of the splitting time data in lattices of different sizes $N$ with $G = 36$, by adjusting the link capacities according to $T^c = a\sqrt{N}$ with $a = 1/6$. The inertia constant is set to $\tau = 1$. The average splitting times without adjusting $T^c$ are depicted in the inset for $N = 576$ and $N = 900$. The corresponding values for $N = 144$ are omitted as the high splitting times induce prohibiting simulation run-times. (c) Collapse of $\langle t_{\text{split}} \rangle$ with $\tau = 1$ due to adjusting $T^c$ according to the grouping factor $G$. The inset shows the chosen value of $T^c$ for each value of $G$, being fitted by a power law. (d) Effect of the inertia constant $\tau$ on the splitting times with $G = 36$ and $T^c = 4$. 
On average, the number of state transition events increases linearly with \( f \) in a given time span. This, in turn, decreases the average time until a maximum allowable flow \( P_{ij} = T^c \) on a link \((i,j)\) is reached, in an inversely proportional manner. However, for \( \tau > 0 \) a minimum average splitting time arises for a roughly estimated value of \( f \approx 0.05/\tau \).

In order to quantify the damage after the splitting, the average relative size of the larger of the two remaining connected components is measured according to [66],

\[
\langle C \rangle = \langle N' \rangle / N,
\]

where \( \langle N' \rangle \) denotes the average number of nodes in the larger connected component. Figure 6.6 shows the average relative size of the larger connected component \( \langle C \rangle \) for different grouping factors \( G \) and link capacities \( T^c \) versus the state change frequency \( f \). The size of the larger connected component is rather independent of the flow fluctuation frequency as determined by \( f \). While keeping the same link capacities [figure 6.6, values for \( T^c = 4 \)] there is no clear indication with regard to the dependence of \( \langle C \rangle \) on the grouping factor \( G \). For a given \( G \), decreasing

![Figure 6.6: Average size of the larger connected component \( \langle C \rangle \) after the splitting into two parts versus the state transition frequency \( f \) for various grouping factors \( G \) and link capacities \( T^c \). The lattices have size \( N = 576 \). The error bars indicate the 95% confidence interval. The dotted lines serve as a guide for the eye.](image-url)
the link capacities $T^c$ slightly increases the value of $\langle C \rangle$. For a smaller value of $T^c$ lower flows are sufficient to overload the links and cascades may develop in smaller regions. Hence, the failing links envelop a lower number of nodes eventually breaking away from the lattice, implying a larger size of the remaining connected component after the splitting. However, for the chosen values of $T^c$ the average relative size remains approximately in the range $0.65 < \langle C \rangle < 0.75$.

### 6.5 Conclusions

To sum up, this chapter has introduced a parsimonious model for fluctuating flows on networks. These flows might induce cascading link overloads eventually leading to a network splitting into two parts. In order to better represent real-world systems an inertia is implied in such a way that a link does not fail immediately but rather delayed when it becomes overloaded. Extensive Monte Carlo simulations on regular square lattices unveiled how the time until such a splitting occurs depends on different flow patterns. With an increasing number of (stochastically) independent nodes the average flows decrease slowly, following a power law. With regard to the robustness of the lattices, a high sensitivity of the splitting time to the link capacities is observed. Increasing the flow fluctuation frequency (as determined by the nodal state alternation) decreases this time span until reaching a minimum, after which it rises again meaning a higher “life expectancy” of the lattice. Generally, both a higher stochastic independence among the nodes (i.e., more groups of simultaneously alternating nodes) and a smaller size of the lattice imply higher splitting times. However, these time spans seem to coincide by adjusting the link capacities according to a power law with respect to the node grouping, and according to the square-root of the lattice size respectively. Furthermore, it has been shown that the effect of the inertia is significant. Its absence implies a monotonic decrease of the splitting times, while introducing it results in remarkably higher values for higher inertia constants. As an indication of the damage after the splitting, the relative size of the larger connected component seems to be independent of the flow fluctuation frequency but slightly decreases with the link capacity.

The chapter concludes with some thoughts on the implications of these results for future energy networks, being characterized by a large share
of fluctuating renewable power sources. The more distributed the power sources are (being equivalent to more groups in the model), the lower the flows exchanged over the power grid [figure 6.2 and figure 6.5 (a)]. However, even in a highly distributed system, a considerable transmission capacity is still needed to keep the system at the desired level of security [figure 6.5 (c)]. Increasing the size of the grid can be expected leading to a disproportionally small increase of the flows [figure 6.3]. Restricting the capacities of the transmission lines or, equivalently, operating the system closer to its security margins might reduce the robustness of the network against cascading failures drastically [figure 6.4]. The inertia as induced by the heating of the transmission lines which might fail when reaching a maximum allowable temperature, potentially increases the robustness of power grids with large share of renewables [figure 6.5 (d)]. The same effect can be even exploited for increasing existing transmission line capacities, thus improving the economic performance of the system, as has been shown in chapter 3. If the grid breaks apart as a result of cascading line failures, the sizes of the two formed islands can be expected to be largely independent of the flow fluctuation frequencies. Nevertheless, they seem to slightly become more asymmetric with decreasing power transfer capacities thus leaving a larger remaining connected component of the network [figure 6.6].

The model provides insights into the underlying physics of networks subject to stochastic flows. Therefore, besides future energy systems, potential applications could be investigated on other large-scale technical networks such as transportation systems.
Chapter 7

Conclusions

This thesis contributes to the reliability analysis of large-scale technical networks by developing probabilistic modeling approaches, by introducing stochastic simulation techniques and by conducting investigations on two fundamentally different yet complementary levels of system abstraction. The electric power system serves as a particular example for a large-scale technical network and therefore is in focus throughout all the presented studies. According to the selected abstraction levels the thesis is divided into two main parts.

The first part aims at quantifying the reliability of specific systems based on replicating their real-world behavior in a detailed manner. Therefore, a probabilistic modeling and simulation framework has been developed combining object-oriented modeling methods with Monte Carlo techniques [chapter 2]. Its major advantage as elaborated in this work is the possibility to simulate the system explicitly in time and to include highly non-linear responses and non-technical factors such as human operators. This allows calculating expected frequencies of adverse events (e.g., power outages) versus their size (e.g., energy not supplied), thus representing the reliability in a differentiated way. Investigations on both a virtual electric power system and on a model of the Swiss high-voltage grid confirmed the functionality. For example, the simulation results demonstrated how already a small increase of the network loading or a delayed operator response to contingencies might carry a significantly higher risk of wide-area power outages. Beyond further model improvements such as the inclusion of voltage considerations or more intricate
operator behavior, the achieved technical realization might even be used as a viable starting point towards an application-oriented software tool for short- or mid-term power system planning purposes. The approach might further be deployed for revisiting current operational security criteria, such as the N-1 rule, or for elaborating novel probabilistic security targets. In addition, the high flexibility of object-oriented modeling suggests its application to other large-scale technical networks such as telecommunication or transportation systems, as well as to their interdependencies.

For accelerating such extensive stochastic simulations the thesis demonstrated the capability of the RESTART technique [chapter 3]. Computational time savings of several orders of magnitude were reached in comparison to crude Monte Carlo simulations. Due to its generality, RESTART can be applied to a broad spectrum of probabilistic models where the time evolution of the analyzed system needs to be explicitly considered. In this work, its algorithmic adaptation allowed for the investigation of fluctuating power flows as induced by renewable energy sources, and their impact on the occurrence of unfavorably high transmission line temperatures. A realistic study case, which discussed the integration of wind power into a transmission system, showed how the thermal inertia on the lines might effectively decrease the probability of reaching the maximum allowable temperature. This effect might be further exploited for increasing the economic performance of the existing transmission network.

The realistic simulation of large-scale technical networks requires the use of a high number of model parameters. Their evaluation is however limited by the availability of empirical data. Moreover, assessing their combined effects on the system response might become highly time-consuming and is further exacerbated by potentially non-trivial interactions. The problem can be tackled by gaining experience in applying the proposed approaches, in order to further clarify the importance of the individual parameters and to identify needs for the collection and provision of relevant system data.

The second part of the thesis uncovers basic underlying mechanisms being relevant for the reliability of large-scale technical networks. These mechanisms have been analyzed at a high level of system abstraction by using parsimonious models, which aim at being as generic as possible with only few embedded parameters as is often done in the field of statistical physics. In the first instance, a model for capturing the
basic dynamics of cascading failures has been introduced allowing to approximate the resulting time-dependent spreading process by a set of coupled differential equations [chapter 4]. Applying this model to degree-correlated networks revealed how a different encapsulation of the nodes with a high connectivity alters the spreading paths and effectively influences the propagation velocity [chapter 5]. The results suggest that many real-world spreading processes, ranging from cascading failures in electric power systems to epidemic diseases, can be categorized into two general types for which either positive or negative degree-degree correlations decelerate the spreading. This result contributes to the basic understanding of spreading phenomena, being relevant for solving a broad range of problems in a larger spectrum of scientific fields.

In the same line, the spreading model has been modified for analyzing cascading failures in networks subject to fluctuating flows [chapter 6]. The basic insights gained carry potential implications for future electric power systems with a large share of renewable distributed energy sources. For instance, the results suggest that even in a highly distributed system a considerable transmission capacity is still needed, or that the thermal inertia on the transmission lines increases the network robustness. This second result, in turn, confirms the generality of the corresponding finding on a specific transmission network using a highly detailed system model [chapter 3].

Such minimal modeling approaches provide generic phenomenological insights into the processes taking place on networks. The information about their actual relevance within the overall operation of specific real-world systems is however limited, as many particular features are highly simplified or are even neglected. Therefore, in a subsequent working step a thorough substantiation of the results is essential, making use of empirical system data or again of more detailed system models.

The elaborated advantages and limitations of the two opposed levels of system abstraction eventually lay the groundwork for defining their specific application areas as well as their conceptual interplay within a comprehensive reliability study of large-scale technical networks. Highly detailed modeling approaches with close adherence to reality may serve for quantifying the reliability of a specific system under given operational conditions. As a complementation, parsimonious models allow for a qualitative identification of basic underlying mechanisms and generic key factors, which are not restricted to specific systems. Such insights may again serve as valuable clues on where to put the focus with highly
detailed modeling approaches. On the other way round, quantitative results from specific systems may also be checked with regard to their generality by using parsimonious models.
Bibliography


