Probabilistic Transmission System Redundancy Optimization

Dissertation

Institute of Electrical Power Systems
Graz University of Technology

Author
DI Alexander Gaun

Supervisor
Ao. Univ.-Prof. DI Dr.techn. Herwig Renner

Reviewer
Ao. Univ.-Prof. DI Dr.techn. Herwig Renner
Graz University of Technology

Reviewer
Ao. Univ.-Prof. DI Dr.techn. Gerhard Theil
Vienna University of Technology

Head of Institute: Univ.-Prof. DI Dr.techn. Lothar Fickert
A - 8010 Graz, Inffeldgasse 18-I
Phone: (+43 316) 873 – 7551
Fax: (+43 316) 873 – 7553
http://www.ifea.tugraz.at
http://www.tugraz.at

Graz / October - 2010

In memory of my dear brother Peter

Acknowledgements

It is a challenging task to dedicate one page to mention and to give some proper recognition to all those people who supported me during my development of this doctoral thesis as a research associate at the Institute of Electrical Power System at Graz University of Technology.

Prof. Herwig Renner is at the forefront of those whose support and mentorship represents the major valuable input for completing this work. I want to thank him for all the useful advice he gave to me that will help me a lot in my future career and for his continuous support at every time.

I wish to acknowledge Prof. Lothar Fickert. His personal, leading and socials skills afforded me wonderful days at this university which I will always remember. I also want to thank Prof. Matti Lehtonen, to whom I'm deeply grateful for the wonderful research stay at Alto University Helsinki and the excellent support he devoted to me during this time.

I wish to express my most sincere gratitude to Georg Rechberger for the supportive discussions, constructive questions and encouragement. Without him, completing this doctoral thesis would have been even more rocky and challenging. Thank you for being such a friend. I also want to acknowledge all the colleagues, friends and staff members of the Institute of Electrical Power Systems, especially Dr. Ernst Schmautzer, for their useful help and assistance at every point of my work.

I would like to thank my parents for their patience and endless understanding and the family of my girlfriend Andrea for their friendship and support.

Andrea, I thank you for your valuable love, patience, motivation, continuous support and endless encouragement. I express my deep thank to you that you are with me, I don't know where I would be without you …

Statutory Declaration

I declare that I have authored this doctoral thesis independently, that I have not used other than the declared sources / resources, and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

Graz, October, 2010

……………………………………………………………………………………

(Alexander Gaun)

Abstract

Transmission system reinforcements, transmission restructuring or transmission network planning are nowadays increasingly determined by uncertainties, mainly those due to strong wind power supply, increased supply of distributed generation, unconfident load growth forecasts, new market requirements or new regulatory requirements. These uncertainties force a new, customized approach to solve the associated system design challenges as economically as possible and in compliance with planning conditions from the point of view of network operators. In recent years, reliability analysis within this framework has been increasingly used for a comparative assessment of power network (re-)design plans, for justification of additional lines, in a cost-benefit analysis, etc. Such an analysis utilizes not only economic criteria but also additional valuation parameters such as monetary valued outage energy for welfare-economic decision making.

Basically reliability estimation of a power transmission system is a time consuming process, where the computational burden increases nonlinearly with the number of transmission lines and with the number of substation equipment. This computation time can be reduced by applying special heuristics, for instance limited outage orders. In this doctoral thesis two minimal cutset based algorithms are proposed that provide a computationally efficient way to evaluate independent first and second order outages and common mode failures, even for real world sized utility driven transmission systems.

Based on these algorithms a novel reliability assessment program is developed that provides a fast probabilistic based computation of load point reliability indices, for instance expectation of energy not supplied with respect to uncertainties in input data. Since investigations of several power transmission systems clearly show the inadequateness of peak load and off-peak load cases, a probabilistic load flow approach considering mutual dependencies between input data is applied for reliability assessment. In this assessment, an in terms of transmission system planning pragmatical DC based load flow is applied. Moreover independent failures up to the second order, common mode failures, first order substation originated outages as well as first order dependent overloading of transmission

facilities, derived from a (n-1)-criterion, are investigated. A novel short-term congestion management probabilistic optimized power flow is used to simulate remedial actions in order to avoid power line overloading.

A novel hybrid genetic algorithm is used to incorporate expected outage costs in the redundancy optimization, which is also known as minimization of power system investment costs and reliability optimization. An advantage of such evolution based heuristic optimization approaches is that these algorithms can be easily utilized to perform nonlinear optimization based on expert knowledge planning heuristics. Since each individual in a generation is optimized by another genetic algorithm, applying different evolutionary strategies, this kind of optimization methodology is called hybrid genetic algorithm. The local optimization is done by a probabilistic load flow based binary coded genetic algorithm that optimizes power line types with respect to minimal investment costs and operational costs of a power system configuration based on a cyclic net present value approach.

The results of this doctoral thesis suggest that algorithms optimizing power system designs based on welfare-economic optimization of costs including substation outages and uncertainties should be part of every network design study in future in order to achieve an improved regulated electrical economy.

Kurzfassung

Titel: Probabilistische Übertragungsnetzzuverlässigkeitsoptimierung

Unsicherheiten, die im Zuge von geplanten Übertragungsnetzerweiterungen, -umstrukturierungen oder -planungen verstärkt auftreten können, wie zum Beispiel starke Windeinspeisungen, erhöhte Einspeisung von dezentralen Erzeugungsanlagen, unsichere Lastentwicklungsprognosen, neue Marktvorgaben oder neue behördliche Vorgaben, usw. forcieren ein neuartiges, angepasstes Konzept, um diese Herausforderungen im Sinne des Netzbetreibers möglichst wirtschaftlich, unter Einhaltung planerischer Randbedingungen, zu lösen. Zuverlässigkeitsanalysen werden in diesem Rahmen für eine vergleichende Bewertung der Aus- bzw. Umbaupläne, für die Rechtfertigung von zusätzlichen Leitungen, einer Kosten / Nutzen-Analysen der Restrukturierungspläne, etc. immer öfter herangezogen um neben wirtschaftlichen Kriterien zusätzliche Bewertungsparameter wie beispielsweise monetär bewertete Ausfallsenergie zur Entscheidungsfindung zur Verfügung zu haben.

Grundsätzlich ist die Zuverlässigkeitsberechnung eines elektrischen Energieübertragungs-systems rechenzeitaufwendig. Insbesondere steigt nichtlinear mit der Anzahl der zu betrachteten Netzelemente die Komplexität und Berechnungsdauer. Diese Berechnungszeit kann durch Anwenden bestimmter Heuristiken reduziert werden. Um eine weitere Rechenzeitreduzierung zu bewirken, wurden im Rahmen dieser Dissertation zwei auf minimalen Schnitten basierende Berechnungsalgorithmen entwickelt, welche die schnelle Analyse hinsichtlich unabhängiger Ausfälle erster und zweiter Ordnung sowie Common-Mode Ausfälle auch für reale, große Übertragungsnetze erlaubt.

Aufbauend auf diese Algorithmen ist im Rahmen dieser Dissertation ein neuartiges Zuverlässigkeitsberechnungsprogramm entwickelt worden, welches die schnelle wahrscheinlichkeitstheoretische Berechnung der Zuverlässigkeit von Übertragungsnetzen, z.B. Erwartungswert der nicht zeitgerecht zur Verfügung gestellten Energie, unter Einbezug von Unsicherheiten bei den Lasteingangsdaten erlaubt. Untersuchungen an mehreren Übertragungsnetzen zeigen, dass es nicht ausreicht diskrete Lastfälle für die Netzplanung zu

verwenden. Daher wird mittels analytischer probabilistischer Lastflussberechnung mit korrelierten Eingangsdaten, aufbauend auf die schnelle DC-Lastflussmethode, eine Zuverlässigkeitsanalyse durchgeführt. In der Zuverlässigkeitsanalyse werden unabhängige Ausfälle bis zur zweiten Ordnung, Umspannwerksausfälle erster Ordnung sowie aus dem (n-1)-Kriterium abgeleitete abhängige Überlastungen von Übertragungselementen berücksichtigt. Um die Behebung letzterer Ausfälle im Sinne von Kurzzeit-Engpassmanagement-Maßnahmen zu simulieren wird ein analytischer probabilistisch optimaler Lastfluss angewendete.

Damit in der kostenminimalen Optimierung des Übertragungsnetzes auch zuverlässigkeitsrelevante Indizes, wie erwartete Ausfallskosten, berücksichtigt werden können, wird ein neuartiger hybrid genetischer Algorithmus angewendet. Solche auf der biologischen Evolution basierende heuristische Optimierungsverfahren eignen sich hervorragend um auf Expertenwissen basierende Planungsheuristiken in der nichtlinearen Optimierung zu berücksichtigen. Hybrid genetisch wird der Algorithmus dann genannt, wenn während eines evolutionären Zyklus eine lokale Optimierung der einzelnen Lösungen durchgeführt wird. Dies wird in dieser Dissertation mittels eines auf probabilistischen Lastfluss basierenden zusätzlichen genetischen Leitungstypenauswahlalgorithmus bewirkt, welcher Systemkonfigurationen hinsichtlich Investitions- und Betriebskosten mittels zyklischer Barwertmethode minimiert.

Die Ergebnisse dieser Dissertation zeigen, dass Algorithmen zur Optimierung von elektrischen Übertragungsnetzen, z.B. basierend auf Investitions-, Betriebs und Ausfallskosten unter Berücksichtigung von Umspannwerken und Unsicherheiten, Bestandteil aller zukünftigen Netzplanungen sein sollten, um optimalere Rahmenbedingungen für die Elektrizitätswirtschaft zu erreichen.

I Table of Contents

II List of Abbreviations and Symbols

A	Branch incidence matrix
AC	Alternating current
B	Reduced system nodal susceptance matrix, where reduced means that the slack bus is eliminated in the matrix.
B^b	Reduced matrix with the branch susceptances of the power line reactances X where the "from value" of each branch is negative and the "to value" is positive.
C	Cycle in a graph
CIC	Customer interruption cost or customer damage function with an energy dependent part in €/MWh and with a load dependent part in €/MW.
COC	Customer outage cost, which are calculated based on CIC for each load point
DC	Direct current
DFS	Depth first search
DLF	Deterministic load flow
E	Edges of a graph, in power systems also known as power lines.
EIC	Expected interruption cost; sum of all load point COCs in a power system obtained via MCS.
Δf	Value used to obtain mutation, immigration and crossover probability based on a fuzzy logic approach.
G	Graph
GA	Genetic algorithm
GC	Graph connectivity matrix; a matrix that indicates whether nodes are connected or not, obtained by a simple DFS approach.
HGA	Hybrid genetic algorithm
IFS	Injection flow sensitivities: PTDF times A, used to deduce MC; it represents network connectivity between terminal nodes.

ℓ	A certain power line of a system
LODF	Line outage distribution factor
m	Number of uncertain RVs in a power system
MC	Minimal cutset
MCS	Monte Carlo Simulation
n	Number of bus injected powers
NPV	Net present value
OTDF	Outage transfer distribution factor
PE	Point estimate
PLF	Probabilistic load flow
PT	Power line type
POPF	Probabilistic optimized power flow
PTDF	Power transfer distribution factor
pu	Per unit
popSize	Population Size in the GA
RV	Random variable
R^2	Goodness of fit, how successful the fit is in explaining the variation of modeled PLF data
R_0	Minimal reliability constraint for improvement algorithm
S_F	Reciprocal of diagonal matrix with network connectivity information at its diagonal elements
SST	Substation type
t	Number of analyzed generation, where $t \in \{1 \dots \lvert Generations \rvert\}$
TEC	Transmission expansion costs
u	Fuzzy logic self control parameter
V	Vertices of a graph, in power systems also known as substations
X	Power line reactance in pu
X	Time series of nodal powers
$\lvert\cdot\rvert$	Number of $\cdot$
α	Fuzzy logic self control parameter
λ	Function used for estimation of mutation, immigration and crossover

III List of Publications

This dissertation consists of a short summary of following six publications, referred to as [P1]-[P6] in the text and further expansions of the proposed nonlinear optimization methodology.

[P1] A. Gaun, H. Renner, and G. Rechberger, "Fast Minimal Cutset Evaluation in Cyclic Undirected Graphs for Power Transmission Systems," *Proceedings of 2009 IEEE Bucharest PowerTech Conference*, pp. 1-8, Jul. 2009.

[P2] G. Rechberger, H. Renner, and A. Gaun, "Systematical Determination of Load Flow Cases for Power System Planning," *Proceedings of 2009 IEEE Bucharest PowerTech Conference*, pp. 1-6, Jul. 2009.

[P3] A. Gaun, G. Rechberger, H. Renner, and M. Lehtonen, "Probabilistic Load Flow Computation Via Enhanced Cumulant based Gram-Charlier Expansion Considering Mutual Dependencies," *Electrical Engineering*, in reviewing by the Journal, pp. 1-27, Apr. 2010

[P4] A. Gaun, G. Rechberger, and H. Renner, "Point Estimate Methods for Probabilistic Optimized Power Flow," *Proceedings CD of EnInnov2010*, Graz, Austria, pp. 1-8, Feb. 2010.

[P5] A. Gaun, G. Rechberger, H. Renner, and M. Lehtonen, "Enumeration Based Reliability Assessment Algorithm Considering Nodal Uncertainties," Proceedings of *Power & Energy Society General Meeting 2010 (IEEE PES '10)*, pp. 1-8, Jul. 2010.

[P6] A. Gaun, G. Rechberger, and H. Renner, "Probabilistic Reliability Optimization using Hybrid Genetic Algorithms," *Proceedings of the 7th International Conference 2010 Power Quality and Supply Reliability*, vol. 1, no. 7, pp. 151-158, Jun. 2010.

1 Introduction

1.1 Research Motivations

Power transmission systems consist of an interconnected set of overhead lines and cables, and related equipment, and are used for the transfer of electricity at high voltage levels between supply points and load points, such as customers and/or other electric systems. Power transmission systems are a major component of electric power utilities, and a fundamental backbone of each economy; they are characterized by high technical and organizational complexity, have developed over decades, are long-term investments and a natural monopole. Power systems have to be reliable and powerful under normal as well as under stress conditions and have to guarantee a certain level of supply security to their customers [1]. Supply reliability, which is beside power quality, commercial quality, energy policies and others a part of supply security can be quantified with either system index numbers or customer index numbers, for example customer outage cost (COC) functions [2].
The expansion planning of power transmission systems is mainly based on a heuristic criterion named (n-1)-criterion. In Austria this criterion is defined in [3]. Basically, applying this criterion may not lead to optimal expansion plans in terms of minimal costs satisfying a certain level of supply security or supply reliability. Thus by incorporation of reliability measures in the system planning a more justifiable approach can be developed, based on welfare-economic optimization.

The strategic planning of power systems always goes along with consideration of several uncertainties, for instance uncertainties that originate from high wind energy supply, increased supply by distributed generation like photovoltaic or other volatile renewable energy sources, uncertain load increase forecasts and new market and regulatory strategies, such as the incentive scheme discussed in [4]. These facts force new adapted approaches to solve the competitive challenges of power system expansion, restructuring or planning with optimized economic planning conditions. Therefore approaches such as probabilistic power flow method or even probabilistic optimized power flow approaches can be adapted, providing a tool for consideration of uncertainties in an optimization framework.

Reliability, which is *"… the probability of a device performing its purpose adequately for the period of time indented under the operating conditions encountered."* [5], is of particular interest in power transmission system expansion planning due to an evident need for predicting future performance of a power system.
The adequate purpose of power systems is generally considered to be the existence of sufficient facilities within the system in order to satisfy all system demand requirements and system constraints. This is associated with static conditions and present reliability evaluation methodologies generally related to adequacy assessment [6]. The definition of adequacy can be seen as formulation of an optimization problem with subjects to minimal number of facilities satisfying certain constraints such as reliability levels, heuristic criteria, economical criteria and system requirements. This type of power system optimization problem is well known as redundancy optimization [7] and is sometimes called transmission expansion problem in electrical power systems. The static transmission expansion planning problem, a synthesis transmission planning formulation for providing (quasi)-optimal transmission expansion plans for a certain time [8], [9] can be used to determine the optimal power line type for a certain power line connection. This type of optimization with respect to minimal investment, operation and renewal expenditure satisfying power flow constraints is referred to as modified transmission expansion problem in this doctoral thesis. Thus, a combination of both problems, redundancy optimization and modified transmission expansion problem, a novel powerful tool can be developed solving grid structure optimization efficiently based on load point reliability indices.

The motivation of this dissertation is to develop a tool, which can solve the outlined challenges (quasi)-optimally, in order to provide a methodology for future transmission network expansion projects applicable in every power system utility. Furthermore planning of power transmission systems is nowadays a field of high interest in the off shore wind industry with special focus on cost-minimized power systems with highest level of reliability and not only satisfying a certain reliability constraint. These types of off-shore grids are completely new, by now comprise a certain lack of unperformed investigations and consist of a suitable size for power system optimization algorithms.

The developed algorithm presented in this doctoral thesis is applied to a real 110-kV sub transmission network for demonstration purposes in order to show how weak spots in the power transmission system structure can be avoided.

1.2 Contribution of this Doctoral Thesis

The aim of this research work is to present an approach for developing optimal power transmission systems for network expansions, network restructuring, green field studies and network development based on the (quasi)-optimal structure of a power system incorporating load data uncertainties, reliability indices and a predefined set of different power line types.

Therefore a computational inexpensive approach for power transmission system reliability assessment with uncertain nodal powers and substation originated outages applying enumeration method, probabilistic load flow (PLF) method, probabilistic optimized power flow (POPF) method and Monte Carlo Simulation (MCS) method must be developed. The results of this algorithm have to be used in a heuristic hybrid redundancy optimization algorithm for network design optimization and reliability optimization of power transmission systems. The hybrid genetic algorithm (HGA) consists of a binary coded genetic algorithm for local optimization of power line types for a transmission expansion system and a genetic algorithm utilizing several expert knowledge based heuristic planning methodologies providing fast convergence of the algorithm. Furthermore, via consideration of outage costs based on the reliability assessment algorithm a welfare-economic approach will be developed, applicable in power system utilities obtaining (quasi)-optimal power system designs, on which, for instance, conclusions concerning planning strategies could be drawn and justified by the utility.

The computer program was completely developed in MATLAB by the author using libraries for graph calculations [10], [11] and the well known MATPOWER package [12]. The developed algorithm was never intended to be used as a powerful optimization tool for large power systems consisting of hundreds of substations and even more power lines, and therefore the coding was optimized only rudimentary with the aim to reduce overall computation time. However, using more powerful programming languages and optimized data structures the algorithm can be improved efficiently resulting in a broadened application field.

1.3 Doctoral Thesis Overview

This doctoral thesis is structured as follows. Section 2 provides a literature review on state of the art in power transmission system optimizing and a scope of research work with research questions and research hypothesis is given.

In the subsequent section the previously stated theorems are established by the use of six reviewed and presented papers [P1] - [P6]. A central theme for understanding each paper in the overall context, answers to upcoming questions from the papers are given and further enhancements of the algorithm are explained and confirmed. Moreover the main results of the publications are stated and the contributions of the authors conducted research work are discussed.

In chapter 4 the developed algorithm is applied to a real world 110-kV power sub transmission system. The obtained results of the (n-1)-criterion optimized power system are discussed and the main application area of the algorithm is highlighted.

Conclusions, dealing with the main results of the algorithm and the main scientific impacts as well as future research work, follow in section 5, and references are listed in section 6. The annexes in section 7 contain the framework papers of this doctoral thesis.

2 Transmission System Planning

2.1 Literature Review: State of the Art

The last decade has seen increased research activity in grid structure optimization (such as redundancy optimization / transmission expansion planning), mainly due to improved computational power and advanced software tools for solving optimization problems, as well as due to economic, environmental and political reasons that led to insufficient investments in transmission systems. The following section provides a comprehensive overview of relevant publications in the field of transmission system optimization, dealing with reliability and cost constrained transmission system optimization.

Optimizing a power transmission system subject to engineering and economic issues such as reliability and costs is a combinatorial optimization problem, where the applied optimization methodology is either analytically or modern heuristically. Considering analytical approaches for solving this non-convex and mixed integer nature of the optimization problems [13] papers are dealing with nonlinear mixed integer programming [14] [15] and with linear mixed integer programming [16], [17], [18],[19], [20], [21], [22], [23], [24], [25], [26], [27], [28], [29], [30], both solved via various techniques by any of the various available programs dealing with this specific problem. Successful solution approaches for non-convex mixed integer programming include decomposition technique or heuristics, where neither of them can guarantee optimality of the solution [31]. However, disjunctive mixed integer linear formulations applying penalty factors and used with analytical solution approaches are very sensitive to the correct value of the problem depending penalty factor and therefore can also lead to suboptimal solutions.

Linear problem formulations [32], [33], [34], as well as nonlinear problem formulations [13], [35], [36], [37], [38] of transmission expansion problems are used with modern heuristics, like genetic algorithms (GA), evolutionary programming, fuzzy-sets and game theoretic methods. In recent years also scenario technique for transmission expansion problem is used as optimization methodology [39], [40], [41] with the drawback of optimizing defined scenarios only. A detailed review of proposed methods for transmission expansion planning, including publications beyond the year 2000, can be found in [8]. Basically, GAs have the ability of optimizing linear and nonlinear models with a high degree of model accuracy. However, it's an evident fact that GA cannot ensure finding a global optimum of the formulated objective function. Contrariwise GAs are excellent in finding (quasi)-optimal solutions, where the solution is not the optimum but it is very close to optimality. They are empirically and theoretically proven to perform robust optimization in complex search spaces without getting trapped in local minima. GAs provide a high degree of confidence and these not necessarily optimal but (quasi)-optimal solutions satisfy complex system characteristics, constraints, and objectives of planning engineers [42].

The optimal solution of a bulk power system in terms of transmission and generation capacity expansion planning is carried out in [18], [29] and [34]. In deregulated power systems, where transmission planning is not coordinated with generation planning, the transmission planning usually has to be performed without exact knowledge of future generation capacities, hence, resulting in a random uncertain load and generation pattern [41]. These load and generation uncertainties have to be taken into account whenever power transmission system expansions are planned. Furthermore, conducting power systems planning based on peak load cases like [13], [39], [40], [36], [23], [26], [27], [28], [34], based on several load cases like [32], [20], [22], [15], [30] or even based on load duration curve like [14], [16], [17], [18], [19], [21], [25], [29] may lead to inconsideration of relevant loading conditions as it is shown in [P2] and [43]. Therefore a relevant technique should be developed in order to deal with this important planning issue.

Since power system facilities have physical limits, represented as power flow constraints in the (non-)linear optimization, a method of obtaining the flows in transmission expansion planning is incorporated in most publications. A maximum flow algorithm is used in [16], [17], [18], [21] and [25]. This algorithm has disadvantages in finding an optimal solution from a set of power line types since it doesn't model the physical laws of electrical power flows in an appropriate way. The AC-load flow method, applied by [24] and [34], is usually used to compute power flows on daily routine. However, this method is computational expensive, compared to DC methods, is nonlinear and computes power system results such as reactive power or voltage profiles / levels in the long term planning process which could be neglected by some assumptions. Therefore DC load flow, which is an efficient modified linearized version of an AC load flow, is used in many transmission expansion planning publications [13] - [15], [19], [20], [22], [23], [26] - [33], [36] - [41].

Redundancy optimization including reliability optimization has been taken into account by the transmission expansion planning only in a limited number of publications like [16], [18], [34] or [44] until now. Basically, reliability can be considered as a certain constraint in the transmission system expansion planning. This constrained optimization is based on contingency analysis and deterministic (n-1)-criterion, or even higher, respectively in [13], [14], [15], [25], [26], [29], [33] and [36]. In recent years nodal prices, locational marginal prices or spot prices based constrained optimization was conducted by [19], [20], [39] and [41]. Nodal reliability indices, such as loss of load expectation, expected energy not served, loss of load costs, outage costs, load curtailment costs or useful combinations of these indices, are used in [17], [18], [21], [22], [23], [32] and [40]. The nodal reliability indices are obtained for instance via state enumeration methods [45], via minimal cutsets (MC) [18], via MCS [32] or via genetic algorithms [46]. By using nodal power system reliability cost functions the reliability optimization could be included in the transmission expansion planning, thus minimizing investment costs and for instance customer interruption cost (CIC) based outage costs, like in [15], [32] and [34]. In order to use reliability load point indices in the optimization function a suitable and fast way calculating these indices should be developed.

Redundancy optimization includes per definition investment costs and, depending on the model formulation, reliability costs like load curtailment costs, outage costs and congestion costs. Firstly, real world utilities have to consider additional costs like operational costs, for instance maintenance, losses, taxes, insurance rates, etc. Secondly, facilities of electrical power systems are very durable. Lifetimes exceeding forty years and more are common practice, but there are differences in the expected life time of power system components like cables and overhead transmission lines. These differing expected lifetimes influence the optimal expansion plan due to relevant renewal of power system devices. Hence it is strongly recommended to consider these impacts in the optimization model. There are several dynamic investment planning methods like net present value (NPV) method, internal rate of return method, amount of annuity method and others [47]. In this doctoral thesis the commonly used NPV method [29], [32], [34], [37] is applied.

Congestion, where a particular power line cannot support a particular power flow, could emerge due to manifold reasons like, trading patterns (too high supply-demand imbalances), different nodal prices, serious reliability problems (unavailability of power lines or generation), etc. [19]. Special focus has to be put on the calculation of congestion revenues in the case of random load uncertainties, for instance whenever long term load forecasts are used. In [32] a MCS is used to obtain congestion costs including load forecast uncertainties based on load duration curve. It is an accepted fact that MCS can deal with uncertainties but it is also commonly known that MCS is computationally expensive and hence techniques for a fast probabilistic congestion cost computation have to be evaluated and implemented.

Power transmission systems consist of power lines, substations, circuit breakers, bus bars, transformer, etc. All these facilities have a certain reliability but also disturbances, malfunctions, (deliberate) outages and so can have influence on the operating conditions of a power transmission system. Furthermore some of the substation elements have a certain nominal rating, like transformers and these elements have to be considered in the optimization model. Since substations have an evident and significant impact on reliability indices of power systems [48], these power system elements should be also considered in the optimization objective. By the best knowledge of the author only in [16] substation elements are modeled with their capacity limits and no publication has dealt with substations in the power transmission system reliability optimization. Therefore suitable methodologies should be developed in order to efficiently optimize the whole power system consisting of power line facilities and substation facilities.

Transmission expansion planning can either be a static process, where the optimal power system for a certain point in the future, normally between five and 20 years, is designed [16]

-[19], [21]- [28], [30] - [41] or it is a dynamic process with a multiyear dynamic nature and with mutual dependencies between each year [20], [29], [49] and [47]. It is evident that dynamic transmission planning can be computationally very expensive due to its complex modeling. Moreover each static model can be extended to a dynamic model, for instance via obtaining the optimal path to a predefined future power system configuration for each time step. Therefore and since the aim of this doctoral thesis is to find an algorithm for optimization of power system configurations with respect to minimal overall utility costs (welfare-economic optimal) static modeling is considered in this doctoral thesis.

2.2 Scope of Research Work

2.2.1 Research Questions

Based on the comprehensive literature review of section 2.1 following research questions are going to be stated in a logical order. These research questions are then applied to obtain research theorems which are answered and justified in the following chapters.

Question	Theorem
How can MC of a power transmission network be evaluated efficiently?	1
How can interdependent load flow data representing a certain measured duration be modeled with probabilistic approaches and how can the accuracy of the obtained results, based on a set of measured nodal powers including random load forecast uncertainties, be improved?	2
How can uncertainties be incorporated in congestion management by a fast approach?	2
How can load data uncertainties be incorporated in short-term congestion management?	3
How accurate is the calculation of load point reliability indices with certain model simplifications compared to a commercial available algorithm?	4
Is it worth to use MC computation for reliability estimation?	4
How can nonlinearities in the optimization of a power system considering investment costs, costs of operation and renewal expenditure of different power line types and substation types be considered efficiently?	5
What kind of algorithm can be utilized for reliability based network structure optimization of a power system?	5
What kind of reliability measure can be used in the optimization algorithm in order to guarantee an effective objective function?	6

Table 2.1: Research questions based on literature review

2.2.2 Research Hypothesis

In the contents with the recent work, primarily conducted by the author of this doctoral thesis, following research hypotheses are worded.

Hypothesis	Description	Section	Reference
1	By applying specialized graph reduction methods or network structure representation methods it is possible to develop a fast MC calculation algorithm.	3.2	[P1]
2	Via probabilistic load flow methods, applying Gram-Charlier Expansion, power system planning is based on significant data, and via the method of joint cumulants mutual dependencies of nodal powers can be incorporated in PLF computation.	3.3	[P2], [P3]
3	It is possible to use a POPF approach for short-term congestion management.	3.4	[P4]
4	It is possible to develop an accurate fast power transmission system reliability assessment algorithm based on theorem 1-3.	3.5	[P5]
5	It is possible to minimize the investment costs of power transmission systems with respect to different power line types and substation types by the use of genetic algorithms.	3.6	[P6]
6	Via assessing expected customer outage costs it is possible to perform a redundancy optimization of power transmission systems.	3.7	[P6]

Table 2.2: Research hypotheses based on research questions

3 Conducted Research Work

3.1 Research Work Overview

This section gives an overview of the conducted research work to provide a central theme for the reader of this doctoral thesis. This overview should help to understand each individual paper in the annex, representing essential parts of this doctoral thesis, in the overall context of the developed optimization methodology.
In this doctoral thesis three computation and optimization algorithms have been developed. The first one performs a fast load point based reliability calculation of power systems (see "Reliability Assessment" in Fig. 3-1) and the second one allows optimizing power systems with respect to different power line types (PT) connecting predefined load points, satisfying certain constraints like (n-1)-criterion (see "Modified Transmission Expansion" in Fig. 3-1). The third algorithm is a powerful optimization tool using heuristic evolutionary optimization techniques in a genetic algorithm and applying both previously quoted algorithms (see "Hybrid Redundancy Optimization" in Fig. 3-2).

More detailed, the reliability computation algorithm, which is the essential algorithm for the redundancy power system optimization has been developed with the focus on a fast but accurate load point indices calculation. This is necessary, since reliability assessment is known to be cumbersome. Therefore two efficient minimal cutset algorithms are proposed in the next chapter, allowing a fast computation of independent first and second order outages as well as common mode outages.

Dependent outages, for instance forced by overloading of power lines, occur regarding to different load cases. Hence a method was developed in order to account for different load cases and uncertainties. In section 3.3 an enhanced method for probabilistic load flow is presented, where mutual dependencies are considered via joint cumulants — a quantitative measure for distributions that takes correlations into account. The computation of the probabilistic load flow is done by the help of PTDFs and the distributions are obtained via the well known Gram-Charlier Expansion.

Based on this probabilistic load flow computations the power system is also analyzed during the reliability calculation with respect to (n-1)-contingencies. Each secondary failure due to overloading is avoided via remedial actions obtained by a probabilistic optimized power flow, minimizing the effect of necessary generation dispatch and / or load shedding. Due to computation reasons like possible nonlinearity of optimized power flows a technique named Point Estimate method is used to estimate the short term congestion costs distributions. This algorithm is explained in section 3.4.

Substation outages, which are essential for a significant reliability estimation of power systems due to the large impact probability of such an outage, can be incorporated in a fast reliability calculation via a heuristic approach.
All previously mentioned procedures are merged together into the powerful frequency and duration based power system reliability assessment algorithm calculating the net present value of expected interruption costs (EIC_{NPV}). Details, further explanations of the algorithm and its accurateness are given or at least referenced in section 3.5.

In an intermediate step a simple and fast genetic algorithm is used to determine an optimized set of PTs and substation types (SSTs), thus finding minimized cyclic net present value of transmission expansion costs (TEC_{NPV}) including operation costs. Whenever power systems are optimized, PTs are a possible degree of freedom. In this doctoral thesis the right power line type is obtained via a genetic algorithm. The nonlinear transmission expansion problem minimizing investment costs, renewal costs, operation costs for power lines SSTs is solved by the help of the proposed linear probabilistic load flow calculation technique. Further details are given in section 3.6.

The redundancy optimization, where power system configurations are minimized based on the sum of TEC_{NPV} and EIC_{NPV} is explained in section 3.7. A certain number of optimized power system topologies of each generation, satisfying a certain minimum reliability constraint R_0, are used to assess the load point reliability including substation outages. Moreover heuristic evolutionary strategies are implemented to model expert knowledge based power system planning. All together represents one novel hybrid methodology, displayed in Fig. 3-2, to optimize power systems with respect to costs such as reliability costs, operation costs and investment costs.

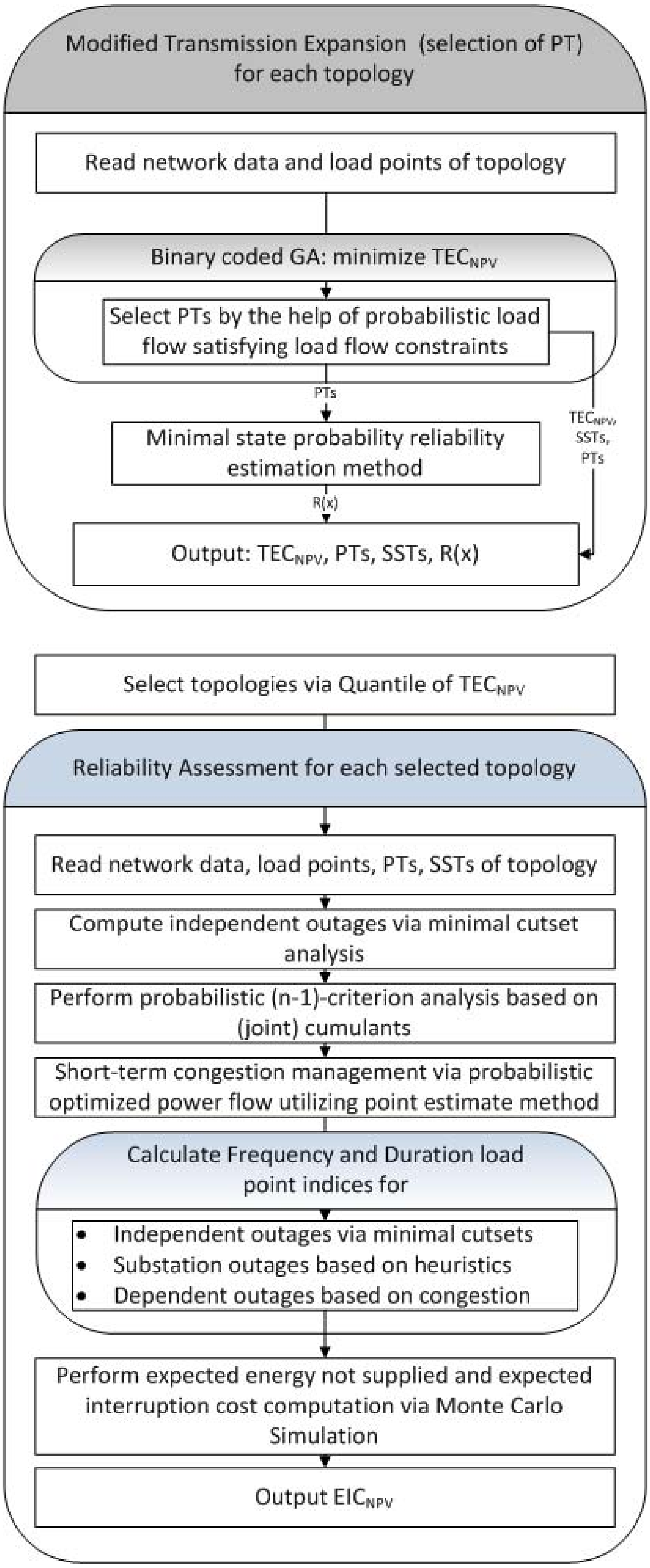

Fig. 3-1 Modified Transmission Expansion algorithm and Fast Reliability Assessment algorithm

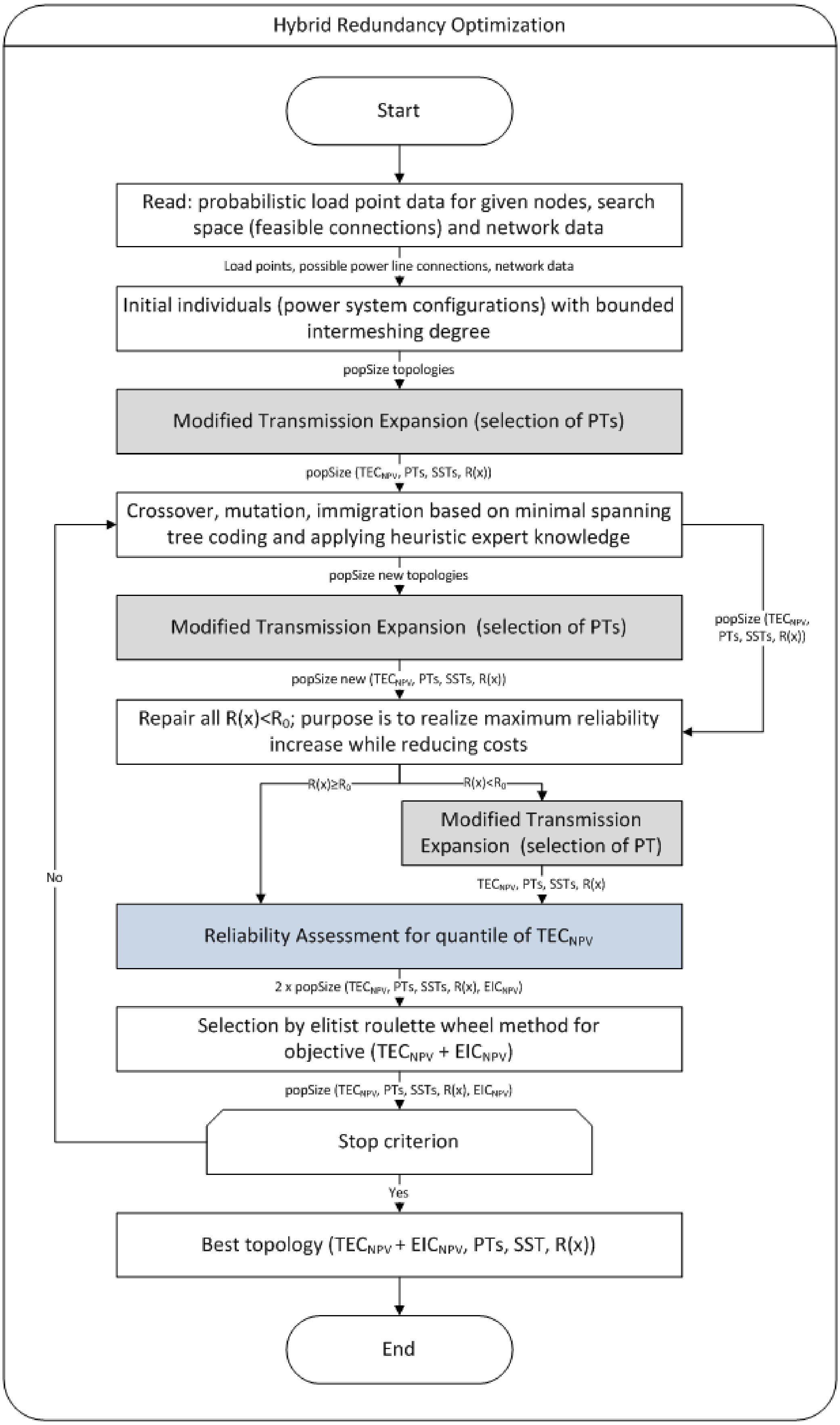

Fig. 3-2 Flowchart of the developed algorithm in this doctoral thesis where popSize is the population size

3.2 Fast Minimal Cutset Calculation Algorithm

The evaluation of network reliability is an important topic in the planning, design, and operation of power systems. The aim of [P1] was to investigate and tune up the speed of calculating 2-terminal minimal cutsets (MC). Following short summary of [P1] gives necessary information of the proposed algorithm.

Basically, MC approach is one of the most popular used techniques to evaluate 2-terminal network reliability of a graph G = (V, E), with V being a set of nodes and E being a set of edges, since MC-technique provides a complete list of independent events that cause network failures. As a matter of principle the complexity of calculating all MCs in a network is proven to be NP-hard, which means in the worst case that the computational difficulty grows exponentially with |V| and |E| and therefore methodologies and techniques have to be developed to allow fast MC computation. MC can be obtained via minimal path computation, which requires high memory demand, and via intuitive theorems. By applying heuristic knowledge of transmission systems it is possible to develop a novel intuitive strategy for computation of MC.

Most power transmission networks can be modeled via a planar graph and simultaneous outages of order greater than two are very unlikely in transmission systems [P1]. The basic concept of the intuitive algorithm is based on chordless cycles and their computation. A chordless cycle C in G is a cycle in G forming an induced sub graph that has no chords, with a chord of C being an edge that joins two nodes of C but is not itself an edge of C [P1]. Using the knowledge of induced cycles a graph reduction scheme, based on merging vertices of chordless Cs together, can be applied. This reduced graph is then analyzed by a fast extended recursive node merge algorithm. For a fixed number of nodes, the computation time of the MC algorithm depends linearly on the graph density of the network. In [P1] three different algorithms have been benchmarked by the help of eleven benchmark networks and the new developed algorithm is the best in terms of computation time. This type of algorithm can compute all MC based on a simple adjacency matrix without requiring any further graph data. A detailed description, comparisons and benchmarks details with other algorithms, as well as mathematical proven Lemmas, verifying that the algorithm finds all MC up to and including 2^{nd} order, can be found in [P1].

The author of this doctoral thesis was fully responsible for carrying out the heuristic graph reduction concept, the mathematical proofs, the algorithm design and the benchmark computations in [P1]. The co-authors of the paper supported the development of the algorithm and acted as valuable discussion opponent.

Although the algorithm developed by the author is already very fast in terms of computation time and has its eligibility for power systems represented as graphs another even more efficient methodology which is specially designed for electric power systems exists. Contrary to the new algorithm in [P1] this algorithm requires more information about the power system, for instance power line data, since it utilizes power system matrixes. Based on DC load flow, the so-called power transfer distribution factors (PTDF), which are sensitivities of bus injection changes to power line flows, the IFS-matrix, a matrix of injection pairs to flow sensitivities, can be derived. This IFS-matrix has a one-to-one relationship with first order outages, causing island formations, which has been proven in [50]. First order network separations are single line outages that lead to a disconnected power system and can be referred to as multiple MC. Note that IFS is also sometimes called PTDF; however, since PTDFs are related to bus injections the name IFS was introduced to distinguish it from PTDFs. PTDF-matrix is constructed by eqn. (3.1), where B is the reduced nodal susceptance matrix and B^{b} is a reduced matrix with the branch susceptances of the power line reactances X [P3], where the “from-value” of each branch is negative and the “to-value” is positive.

$$\mathbf{PTDF} = \mathbf{B^b} \times \mathbf{B}^{-1} \tag{3.1}$$

Eqn. (3.1) can be extended to the IFS-matrix by applying multiplication with the transposed branch incidence matrix $\mathbf{A}$ (see eqn. (3.2)) [50].

$$\mathbf{IFS} = \mathbf{PTDF} \times \mathbf{A}^{\mathbf{T}} \tag{3.2}$$

Every power line ℓ_k with $\ell_k \in$ E that is an inter-sub network connection is now characterized in the IFS-matrix by the following relationship

$$\mathbf{IFS}_{\ell_m}^{\ell_k} = \begin{cases} 1, & \ell_m = \ell_k \\ 0, & \ell_m \neq \ell_k \end{cases} \tag{3.3}$$

where ℓ_m is one of 1…|E| power lines of the power system and ℓ_k is the supposed power line that causes a network separation into two sub networks. Since a second order network separation is a simultaneous combination of two first order cuts all 2^{nd} order multiple MC can be estimated in worst case via |E| matrix operations, where every power line is assumed to be outaged once and the respective first order inter-sub network connection is combined with the outaged power lines to find all 2^{nd} order multiple MC. The PTDF and IFS-matrix has to be calculated in every single outage case because a change in the network structure changes the flow sensitivities. Via simple network connectivity analysis the graph connectivity GC-matrix is calculated in order to obtain 2-terminal MC. This computation is for instance based on a depth first search (DFS) of G’, where G’ = (V, E’) is the network graph without the considered first or second order outage, causing island formation. Every sink and source combination that is not connected in the outage case are then terminal nodes of the 2-terminal MC. This approach can be implemented in MATLAB elegantly, resulting in a very efficient method for computing all MC.

The pseudo code of the proposed and in the optimization algorithm of this doctoral thesis used method is listed below.

procedure: MC estimation

input: $\mathbf{G}=\left(\mathbf{V},\mathbf{E}^{\mathbf{I}},\mathbf{X}^{\mathbf{I}}\right)$

output: **all 2-terminal MC**

$\mathbf{IFS} \leftarrow eval\ \mathbf{IFS}\left(\mathbf{V},\mathbf{E}^{\mathbf{I}},\mathbf{X}^{\mathbf{I}}\right)$

if $any\left(\left(IFS_{\ell_m}^{\ell_k}=1\right)\&\left(IFS_{\ell_m}^{\ell_j}=0\middle|\ell_j=\left\{1...\left|\mathbf{E}^{\mathbf{I}}\right|\backslash\ell_k\right\}\right)\right)$ **then**

  for k = $1...\left|\ell_m\right|$ **do**

    $\mathbf{E}^{\mathbf{II}}=\mathbf{E}^{\mathbf{I}}\backslash\ell_k$

    $GC\leftarrow eval\ connectivity\, G=\left(\mathbf{V},\mathbf{E}^{\mathbf{II}}\right)$

    $MC\leftarrow eval\, GC$

  end

end

$\mathbf{E}^{\mathbf{III}}\leftarrow\mathbf{E}^{\mathbf{I}}\backslash\ell_m\in MC$

for z = $1...\left|\mathbf{E}^{\mathbf{III}}\right|$ **do**

  $\mathbf{E}^{\mathbf{IV}}\leftarrow\mathbf{E}^{\mathbf{III}}\backslash\mathbf{E}_{\mathbf{z}}^{\mathbf{III}};\ \mathbf{X}^{\mathbf{IV}}\leftarrow\mathbf{X}^{\mathbf{I}}\backslash\mathbf{X}_{\mathbf{z}}^{\mathbf{I}}$

  $\mathbf{IFS} \leftarrow eval\ \mathbf{IFS}\left(\mathbf{V},\mathbf{E}^{\mathbf{IV}},\mathbf{X}^{\mathbf{IV}}\right)$

  if $any\left(\left(IFS_{\ell_m}^{\ell_k}=1\right)\&\left(IFS_{\ell_m}^{\ell_j}=0\middle|\ell_j=\left\{1...\left|\mathbf{E}^{\mathbf{IV}}\right|\backslash\ell_k\right\}\right)\right)$ **then**

    for k = $1...\left|\ell_m\right|$ **do**

      $\mathbf{E}^{\mathbf{V}}=\mathbf{E}^{\mathbf{IV}}\backslash\ell_k$

      $GC\leftarrow eval\ connectivity\, G=\left(\mathbf{V},\mathbf{E}^{\mathbf{V}}\right)$

      $MC\leftarrow eval\, GC$

    end

  end

End

Table 3.1: **Pseudo code of the MC-algorithm, based on IFS-matrix and graph connectivity**

3.3 Enhanced Probabilistic Load Flow Method

Reference [P3] of this doctoral thesis works out the different methodologies for Probabilistic Load flow, like Monte Carlo Simulation, Point Estimate (PE) Method, cumulant based evaluation, Convolution, etc. In order to allow a fast optimization of power systems, a fast reliability assessment is necessary. Hence, a computationally cheap method considering mutual dependencies, has to be utilized which is presented in [P3].

The PLF method based on the PE method has become popular during the past years. In [P3] it is shown how mutual dependencies between real measured sets of input random variables can be numerically implemented in the calculation by the means of joint cumulants. The modified method can be referred to as an "Enhanced Cumulant based Gram-Charlier Expansion". First, it is proven by means of mutual information that the measured input data is non independent. Then, the accuracy of the proposed analytical method is demonstrated via comparing the results with discrete load flow simulations and a Point Estimate Method, by use of a synthetic example and measured load data of a real world 110 kV sub transmission power system. Lastly, a suitable framework to use the proposed method in fast probabilistic load flow studies is provided while maintaining a high degree of accuracy.

Electrical deterministic load flow (DLF) calculation is used to analyze and evaluate the planning, calculating and operating of power systems in almost every utility. To determine the system states, DLF uses fixed values of bus injected power for generation and load demands of a selected network configuration. The DLF approach for power system expansion planning is usually performed with one or two worst case scenarios, winter peak load case and summer off-peak load case, for instance like in [26] or [39]. It can be shown [P2], [43], that this methodology does not necessarily contain all relevant loading conditions of power lines and is therefore insufficient for power system planning. The PLF approach is a method to consider uncertainties in nodal powers and can be efficiently used in power system expansion planning [P3], avoiding different discrete load cases.

A set of measured nodal powers can be expressed as a probability density function (PDF) by using a histogram. If this set of nodal powers is seen as an input probability distribution for

the PLF computation, a whole year, or whatever duration the power data is measured, can be calculated by a single power flow calculation. As already explained in Sec. 2.1 on transmission system expansion planning it is reasonable to use DC load flow based approach.
There exist various reasons for dependencies between nodal powers. Mutually dependent loads are caused by common environmental factors, such as temperature, sunset, rainfall, season, etc., and by social factors, such as sporting events, television programs, meal times, working habits, wealth, educational background etc. Some of these factors affect all loads of a similar nature in a related manner and therefore a certain degree of linear dependence exists [P3].
This load correlation is assumed to be constant over a certain period of years, [51] and so the same correlation exists between loads in the basis year and between loads in the forecast year. By applying a joint cumulant approach, as described in [P3], this mutual correlation can be incorporated into the PLF computation, since it is known that load correlation can have a great impact on power system reliability evaluation [51].

Based on the Gram-Charlier Series Expansion [52], which is a series that approximates a probability distribution with the help of its cumulants, the author of this work has developed enhanced formulae to consider mutual dependencies up to the fourth order. Therefore the traditional formulae of cumulants [53] have been extended to joint cumulants, where mutual dependencies are included. In addition the formulae of joint cumulants have been generalized to equations where n random variables (RVs), with n being the number of bus injected powers, are involved to calculate the probability distribution of active power flows on power lines.
An investigated PE scheme, where correlation between RV is directly accounted for with a rotational transformation of the input RVs, achieves in an overall view an inferior accuracy than the proposed joint cumulant method and is computational more expensive, compared with the method where 2^{nd} order joint cumulants are computed [P3].
The computation time of higher order joint cumulants including all required input RVs is nearly growing exponentially. By taking into account an insignificant error the consideration of 2^{nd} order joint cumulants is sufficient to achieve an equivalent level of accuracy while reducing the calculation time [P3].

Furthermore, by considering the dependencies between RVs in formulae of joint cumulants the proposed approach can be applied to other expansions, such as the Cornish-Fisher and the Edgeworth Expansion [P3] or moment based Maximum-Entropy method.

The author of this doctoral thesis enhanced the rudimentary approach introduced in [54] to a general formulation for joint cumulants up to the fourth order. Moreover the author was fully responsible for the conducted computations, implementation of the algorithm and development of suitable formulations for a certain number of input variables. The co-authors offered advice, valuable comments and encouragement during the research process. The author of this doctoral thesis acted as co-author in [P2] and was there mainly participating in the development of the applied genetic optimization algorithm for the load case estimation.

To explain how uncertainties can be incorporated in the planning process in order to obtain input distributions for the proposed enhanced PLF method, following example is given.

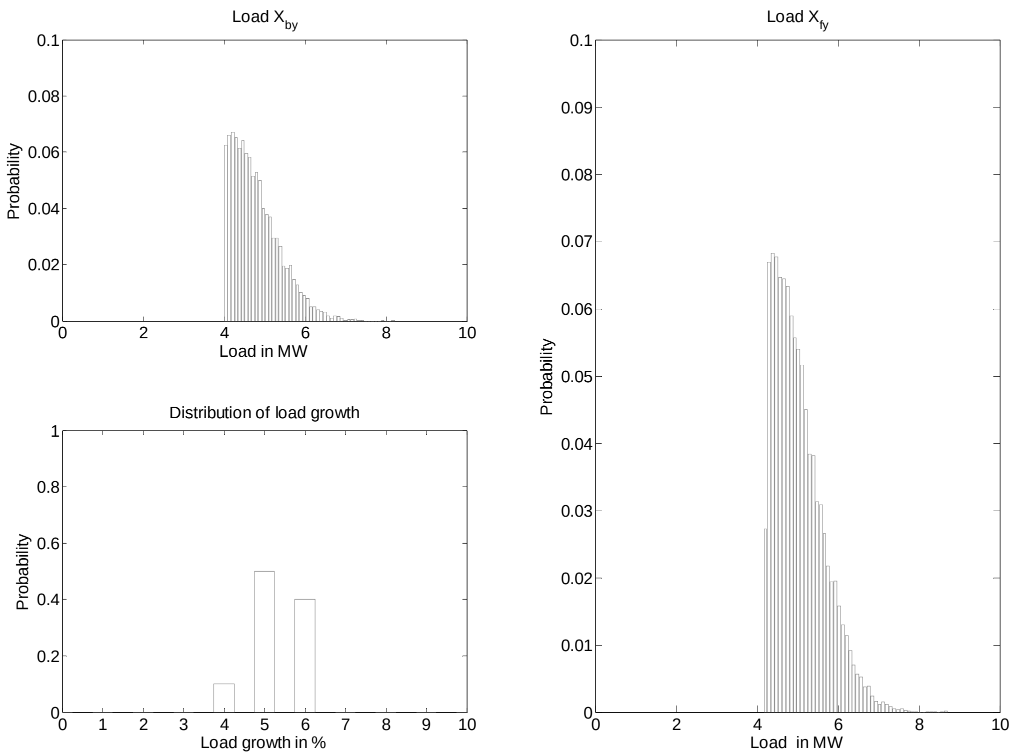

Fig. 3-3 PDF of a load (left upper part), forecasted load growth (left lower part) and estimated load distribution in the forecast year, which can be used as an input probability distribution for PLF calculations.

Let us consider a time series of nodal powers $X_{by} = 4 + |\varepsilon|$ in the basis year, where ε is a normally distributed random variable with zero mean and unit variance, and a load growth forecast for a certain year, where the forecast value is not a single number but a probability distribution (stochastic forecast) based on scenario technique or any other data collecting method. Let us assume that there is a likelihood of 10 % that the load growth is 4 %, a 40 % chance that the growth is equal to 6 % and a chance of 50 % that the load increases by 5 %. Applying a simple MCS, the resulting probability distribution of the forecasted load X_{fy} in the forecast year can be estimated (see Fig. 3-3).

As described in [P3] the joint cumulant based PLF approach is characterized by a great accuracy if normally or almost normally distributed input nodal powers are considered. However, not all nodal powers are normally distributed, as illustrated by a load duration curve of a thermal power plant that can be modeled by two or more discrete values, own consumption load in the summer and a few certain load levels in winter.

This challenging problem can be solved by dividing the observation period into, for instance, two sub durations, each of them having almost normally distributed input nodal powers. Via duration weighted summing up of the obtained probability distributions the final load flow on each power line can be calculated. Since this method is computationally more expensive, its use has to be justified by a great improvement of the calculation results which is a problem depending matter.

Simulation results show that, based on the two goodness measures introduced in [P3] and based on the given load flow data [P5] of a real world power sub transmission system, for instance shown in [P3], [P5], an improvement in the system R^2 measure of 0.5 % (from 39.66 to 39.86) and an improvement in the overall Average Root Mean Square measure of 2.6 % can be achieved. These results were obtained with a visually optimized division of the time intervals and could further be optimized for instance with genetic algorithms. The computation time increased by 40 % which leads to two conclusions:

Firstly the proposed method results in a greater accuracy of the computation results and secondly for the considered power system the use of this approach is not justified, since the accuracy is already suitable for significant results in power system planning and the increased computation time is a major disadvantage.

3.4 Probabilistic Short-term Congestion Management Approach

Whenever adverse load flow situations during a reliability assessment have to be solved, i.e., when a first order power line outage leads to power line overloadings, a suitable algorithm has to perform congestion management. This can either be done via a simple load shedding algorithm or cleverer via optimized power flow methods minimizing outage costs. In [P4], Point Estimate (PE) methods are benchmarked to find the best probabilistic optimized power flow method for the requirements in the proposed reliability assessment algorithm.

Congestion management is an important issue whenever a particular power line cannot support a particular power flow. In this doctoral thesis, congestion within an optimized power flow framework means any adverse situation that leads to system states where power lines cannot serve its calculated power flows due to limited rating. In these situations either cost-optimal redispatch of generation or cost optimized load curtailment of consumer loads or a cost-optimal combination of both has to assure that all power lines and generators are within their limit values.

A benchmark of four different PE methods is presented in this paper, in order to identify the most suitable method for a load curtailment and / or generation dispatch probabilistic optimized power flow (POPF) with correlated input RVs. PE is used to calculate moments, characteristic values of RVs, of a random quantity that is a function of RVs. Several PE methods have been developed over the past years, and all methods are characterized by an input value depending weighting of the DLF results in order to obtain a load flow probability distribution. Taking advantage out of this fact, the PE method can be combined with a deterministic linear optimization algorithm, calculating POPF distributions. This optimization algorithm, minimizing congestion costs subject to load flow constraints and generation limits is also stated in [P4].
Four different PE methods, which are mathematically explained in [P4], are benchmarked against a sequential Monte Carlo Simulation. The Benchmark is based on calculation time, calculation accuracy and problem applicability. By the use of a real world sub transmission power network with measured probability density distributions of nodal input active powers

and significant congestion cost functions the calculation accuracy of the proposed PE methods is determined. Therefore a R^2 measure and the expectation value of via Gram-Charlier Expansion obtained expected energy not supplied distributions for several selected contingencies are used.

Furthermore it is demonstrated that the rotational transformation, used to take into account the correlated nodal input RVs, is essential for the proposed methodologies in order to achieve sufficient calculation results compared with the MCS. Summarizing the results of the benchmark, Harr's algorithm [55] is the most adequate method for the proposed short-term congestion management POPF problem. This algorithm has the lowest computational burden, yields in reasonable results for all test contingencies and provides the best approximation (highest R^2 and best E(X)) for distributions with high probability of zero expected energy not served and low probability of positive nonzero expected energy not served, which is the most likely case for the proposed POPF application in real world power systems

The obtained congestion cost, which is a function of the curtailed load or dispatched power times a cost function, is also congestion duration dependent. Harr's method, combined with the DC optimized power flow framework, results in a probability cost distribution. For instance, with 8760 measured values per year, the outage cost is based on one hour outage duration. By combining Monte Carlo Methods and different POPF calculations it is possible to consider this fact in the congestion cost calculation but this is very time consuming [P4].

The proposed method in [P4] computes the distribution of minimal system congestion costs. Adverse situations, lasting longer than one hour, force higher congestion revenues compared to the obtained congestion revenues. In other words, via the proposed approach and based on the obtained nodal power distributions a minimum value, a lower bound of congestion costs, is calculated by the algorithm since overloading cases lasting longer than one hour lead to additional congestion costs. Furthermore this congestion costs are one part of the objective in the power system reliability optimization, and therefore, by the use of POPF results, the congestion costs are minimized within the genetic algorithm framework. Thus, less adverse power system states also lead to a minimized impact of long-lasting congestion events.

The author of this doctoral thesis had the idea for short-term congestion management based on probabilistic load flow and was fully responsible for implementing the approach and performing the benchmark calculations. Furthermore the author of this doctoral thesis highlighted an approach to enhance the calculation results of the method for long term outages. The co-authors offered advice, discussion and additionally contributed as conclusion controllers.

3.5 Accurate Fast Power Transmission System Reliability Assessment Algorithm

In [P5] the previous three algorithms, the fast minimal cutset algorithm, the enhanced probabilistic load flow algorithm and the probabilistic short-term congestion management algorithm are merged together into a novel fast power system reliability assessment algorithm. Additionally to common reliability assessment algorithms, this algorithm which is based on a probabilistic load flow approach considers nodal uncertainties. The Frequency and Duration indices for each load point and the expected energy not supplied as well as the expected interruption costs (EIC) are derived with the proposed methodology. By reducing the number of necessary system states a huge decrease of computation time is achieved while maintaining an adequate degree of accuracy. In order to verify the accuracy and the computational expensiveness of the proposed algorithm, a commercial available program was used to compare the calculation results of a utility driven sub transmission power system. A DC load flow based reliability calculation including substations modeled with circuit breakers, bus bars and load switches according to Fig. 2 [P5] was implemented in NEPLAN (see [56]). Common mode outages as well as single and double outages were considered during the computation. Each load was modeled as a five point load duration curve. It was demonstrated that a fast calculation of the load indices, including station originated outages, can be performed based on uncertain nodal powers, for instance nodal injections of stochastic power system loads, with a high degree of accuracy.

In order to include reliability optimization in the expansion planning algorithm, reliability costs are included in the objective of the optimization algorithm. Therefore significant power system reliability costs have to be used. These costs are for instance customer outage costs (COC) including first and second order power line outages as common mode and independent outages and first order substation originated outages and reliability costs based on congestion costs due to secondary failures (dependent outage). The proposed algorithm in [P5] can calculate the described costs accurately and fast.

By applying the MC based algorithm developed in chapter 3.2 an efficient way for the proposed problem of analyzing independent power line outages up to the 2^{nd} order is used.

In a system state enumeration a total number of $|E|+(|E|\times(|E|-1))/2$ first and 2nd order system states have to be analyzed by a system state analysis. As described in [45], power system steady-state analysis tools, such as network flow and AC / DC load flow methods, are the most common used methods. In order to evaluate the reliability, all systems states have to be analyzed by one of these methods, for instance a network flow method based on a simple DFS algorithm.

The evaluation of all MC of the proposed MC enumeration algorithm provides the same information about independent network failure as the system state enumeration method. Since the number of network connectivity analysis depends strongly on the type of power system, a worst case estimation can give an idea about the effectiveness of the algorithm. More specifically, a cycle graph consisting exclusively of 2nd order MC, requires $|E|$ IFS-matrix calculations and $|E|\times(|E|-1)$ GC analysis, whereas a graph representing a spanning tree (which consist only of 1st order MC) requires one IFS-matrix calculation and $|E|$ GC analysis. Since real world power systems are a connected combination of these two types the total number of necessary IFS-matrix and GC computations is somewhere between the two worst case values, which is, if it is assumed that a IFS-matrix computation is equal to a GC computation in terms of computational efforts, better than the system state enumeration. In other words, since the proposed MC enumeration algorithm applies a graph topological reduction of possible 1st and 2nd order system states, it is smart to perform reliability calculations of transmission systems based on MC.

As shown in Fig. 3-6, substation outages represent the main part of outage costs in (n-1)-optimized power systems. Therefore substation outages should to be considered in a reliability analysis, although this process increases the computational effort dramatically, if all substations are implemented in the analysis as a configuration consisting of edges, for instance circuit breaker, load switches, and bus bars represented as nodes. This circumstance can be prevented when the number of possible substation types (SST) is limited to a fixed number of configurations where advanced SST are flexible in the number of connected feeders, as for instance the four substation types, Block-connection, H-connection, single bus bar and double bus bar substation, all described in [47], [P5]. In the fast reliability estimation algorithm these substations are included as nodes with a certain outage probability and certain outage duration, both factors mainly depending on the number of feeders and possible remedial switching actions in the case of power line -, bus bar -, circuit breaker -, transformer - and load switches failure. These reliability values are implemented in the algorithm as a function of connected loads respectively transformers, connected sources and connected power lines. The used formulas include all 1st order outages obtained by a substation fault tree analysis. For further information see reference [P5].

Dependent outages due to power line overloading are repaired via short-term congestion management described in the previous section. Using the IFS-matrix in eqn. (3.2), line outage distribution factors (LODF) can be derived, where eqn. (3.4) gives the relationship between IFS and S_F , a matrix with $1/\left(1-diag\left(\mathbf{IFS}\right)\right)$ as its diagonal elements and all other elements being zero. Note that, IFS-matrix is referred to as F-matrix in [P5].

$$\mathbf{LODF} = \mathbf{IFS} \times \mathbf{S_F} \tag{3.4}$$

LODF-matrix can then be used to evaluate those first order power line outages that result in any overloading on any power line, thus requiring congestion management to perform remedial actions to clear adverse situations [P5]. Since the conventional transmission system planning considers only the so-called (n-1)-criterion, 2nd order outages, which would require a much higher computational effort, are neglected in the algorithm, although the evaluation and analysis of these situations could be implemented in the algorithm in the same way. As demonstrated in [P4], Harr's algorithm is the most suitable one for short-term congestion management. Therefore this algorithm is applied to calculate the impact of customer outage costs obtained with a probabilistic optimized power flow utilizing load curtailment and / or generation dispatch.

The outage transfer distribution factor method (OTDF) can be utilized for estimating those power lines that are secondary outages, for instance overloaded power lines due to a 1st order outage. OTDF, a factor that calculates a flow on a branch after a contingency occurs, can be determined via eqn. (3.5). By doing this all first order power line outages are analyzed via PLF computation considering the pre-outage power flow direction.

$$\mathbf{OTDF} = \mathbf{PTDF} + \mathbf{LODF} \times \mathbf{PTDF} \tag{3.5}$$

Compared to a reliability analysis where all 1^{st} and 2^{nd} order outages are considered the developed fast reliability assessment algorithm neglects 2^{nd} order dependent outages and 2^{nd} order substation outages. Since the probability of two simultaneous outages, compared to a single outage, is the product of the two outage probabilities, which is fairly small, the resulting difference is evident but in the author's opinion not worth being considered in the computation, especially with the disadvantage of huge computation time increase. The algorithm in [P5] cannot perform a reliability assessment on isolated power networks. This circumstance is avoided by the implemented repair algorithm and start solution algorithm described at section 3.7.

One major advantage of the proposed methodology can be seen in the fast computation time. Compared with the commercial program an impressive reduction in computation time follows. In order to emphasize this advantage it is now possible to use the proposed algorithm in a genetic optimization tool, to perform reliability optimized power system planning in reasonable time.

The co-authors of this paper provided useful assistance during the development of the algorithm and worked as discussion opponents as well as supervisors. The author was fully responsible for the whole algorithm, including algorithm and programming optimization, testing and implementation. The author also had the idea of using fast MC computation for enumerative reliability computation and nodal reliability results for incorporating fixed SST with variable number of touching power lines in the reliability estimation.

3.6 Genetic Algorithms Minimizing Power Line Investment Costs

Whenever power systems are optimized, power line types are a possible degree of freedom. In a local search step the power system is optimized to compute the right PT. The nonlinear transmission expansion problem minimizing investment costs, renewal costs, operation costs for power lines and substation is solved by the help of the proposed linear probabilistic load flow calculation technique implemented in a the GA. The motivation of this intermediate optimization is to narrow the search space in order to provide a fast convergence of the overall optimization algorithm.

Power transmission system planning, i.e., optimizing the type of power lines out of a given set of available power line types (PTs) subject to power flow constraints, can be formulated as a linear mixed integer problem utilizing the method of a disjunctive formulation. This approach is very sensitive to the used penalty factor. Basically considering different SST in the objective function based on a heuristic criterion as well as considering renewal expenditure and equipment lifetime via cyclic NPV lead to a nonlinear formulation. Furthermore taking uncertainties of loads into consideration heuristic optimization methods are an adequate approach for finding results for an optimization objective function. The author of this doctoral thesis conducted a comparison research work dealing with linear mixed integer problems and GA, concluding that it is more suitable to use GA for this kind of nonlinear problem.

Basics about GA can be found in the literature. Special features of the applied GA are crossover and mutation. The former operator creates a new child, depending on the binary coded parents and returns the best two of three possible individuals. The latter changes randomly positions in the binary coded chromosome. A roulette wheel reproduction operator is used to select new individuals and a penalty factor is used to punish the NPV of individuals not satisfying load flow constraints.

The following gives further details about the algorithm. PTs are represented in each chromosome as an integer value, ranging from one to the number of different power line types $|PT|$. In this case binary coding is an efficient way to encode the problem. Since different transmission system regions can have a limitation to a specified PT, for instance cables are the only realistic option in the inner-city area. This circumstance is also implemented in the algorithm via considering linear constraints in the coding of the GA, resulting in feasible solutions in every generation in terms of investigated power line cable and overhead power line connections.

The transmission expansion cost (TEC) objective is calculated as the cyclic NPV of investment costs, renewal costs and operational costs of all power system facilities like overhead transmission lines, power lines, circuit breakers, load switches, transformers and bus bars for certain duration. Costs of losses is currently not implemented but can be incorporated into the algorithm for instance by the approaches presented in [27] and [57]. The realized SST is defined by a heuristic criterion given in [P6].

The convergence of GA is strongly dependent on the goodness of the start solution. Furthermore, trying to minimize the computation time requires also a well chosen start population. PLF computations, based on the reactance of power lines with highest nominal ratings, yield to a worst case approximation of line loadings. Those power lines, where the difference between nominal rating and estimated worst case power line loading exceeds a certain value, are realized in the start solution with the next lower rated PT. Combining this start solution with the start solutions stated in [P6] yields in a reduced computation time.

Constraints, such as nominal ratings of power lines, are considered via a PLF calculation including possible uncertainties. These power flow constraints could include the well known (n-1)-criterion, where via OTDF (eqn. (3.5)) all first order power line outages are analyzed. The (n-1)-criterion is a heuristic planning method to provide reliability to systems. The system is sufficiently reliable if it is able to operate under any unplanned outage of one component. In an Austrian power system, for instance, the criterion implies that the failure, not considering common mode outages, of any line or transformer after a single power system outage, will not cause any violation of normal operation conditions.

Hence, it is possible to compare an optimized power transmission system satisfying (n-1)-criterion and PLF based power flow constraints with the same optimized power transmission system satisfying the PLF based power flow constraints. The necessary power system data of the analyzed real world transmission system, the different possible PTs, the predefined substation types, the cost data and fixed locations of load nodes and generation nodes are given in [P6]. Fig. 3-4 and Fig. 3-5 represent the best obtained solution for (n-0)-criterion optimized power system and (n-1)-criterion optimized power system each out of $|PT|^{|E|}$ possible solutions, which is in the present power system example 4^{32}.

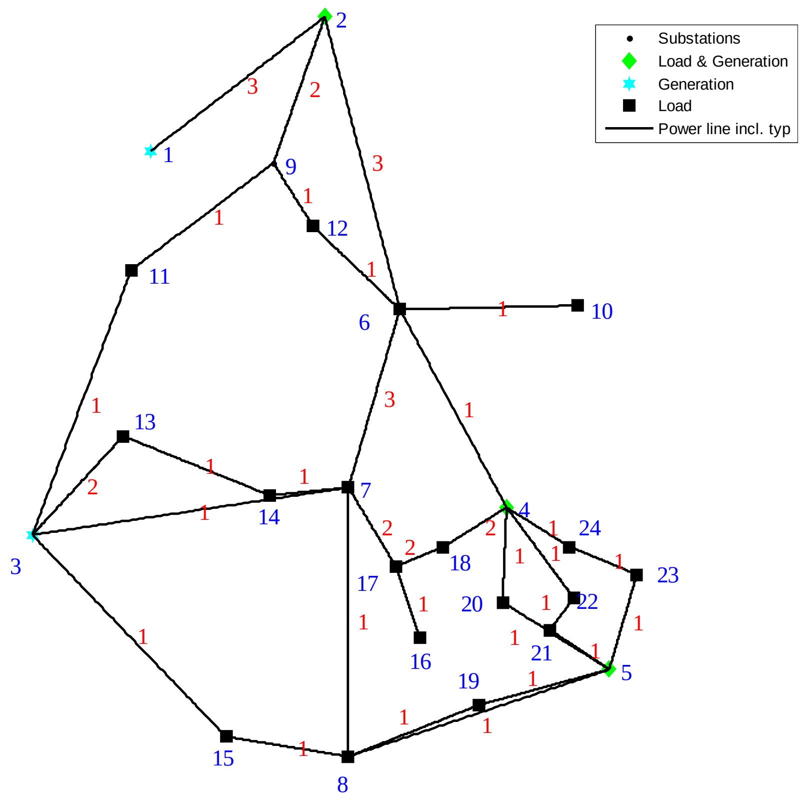

Fig. 3-4 (n-0)-criterion optimized power sub transmission system. Red numbers in the center of the lines designate the power line type, bus numbers are blue.

Comparing realized power line types in the power system shown in Fig. 3-4 and the power system configuration shown in Fig. 2 in [P6] five power lines are realized with different types in this demonstration power system which can be returned to more detailed input data in terms of higher order moments for PLF computations in this doctoral thesis.

The (n-1)-criterion optimized power sub transmission system satisfies the (n-1)-criterion for stub lines and also for meshed power lines. The differences in investment costs for substations and power lines, calculated for 50 years with an interest rate of 5 %, are given in Table 3.2. It is obvious that realizing an (n-1)-secure power system, where one outage of a power line does not lead to an adverse situation, forces higher investments (11.4 %) than a less restrict optimized power system; thus the question if it is worth to build (n-1)-secure power systems is at least justified (see also 3.7). Power line NPV is increased by 22.9 % and NPV of substations is raised by 6.5 % where both power systems satisfy normal operation under real world conditions. This increase is evidence that substations have to be considered in power system expansion planning explicitly.

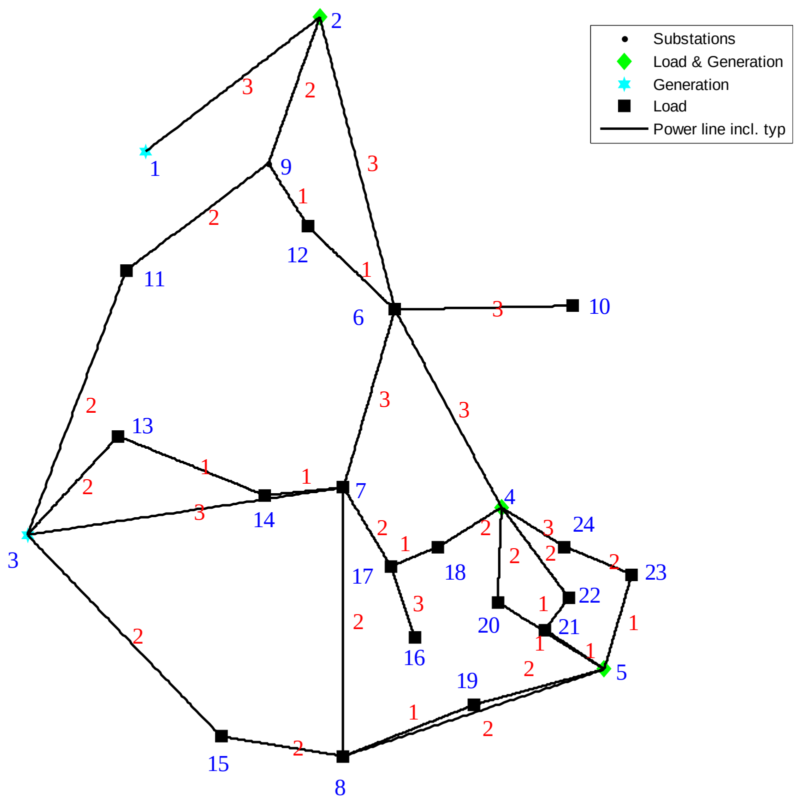

Fig. 3-5 (n-1)-criterion optimized power sub transmission system. Power line types from one to two are single power lines and from three to four are double power lines.

Criterion	NPV of power lines in %	NPV of substations in %
n-1	33.1	66.9
n-0	30.0	70.0

Table 3.2: Comparison of NPV of (n-1)-secure power system and (n-0)-secure power system subdivided into pu-NPV for substations and power lines.

3.7 Expected Customer Outage Costs Objective for Redundancy Optimization

In order to optimize reliability in a redundancy optimization algorithm, which is known as minimization of power system investment costs and reliability optimization, a monetary value of reliability has to be used to formulate an objective function including local optimized transmission expansion costs. Basically there are several methods to estimate the reliability worth of a power system as a monetary value for a comprehensive cost function for redundancy optimization. Beside the methods mentioned in the literature review ([15], [32] and [34]) commonly agreed sector customer damage functions based outage costs, here referred to as CIC, for instance collected by the Finnish Energy Market Authority [58], are a valid opportunity to estimate such reliability data. In this doctoral thesis, CIC is applied to loads connected to transformers at power transmission system substations in order to obtain expected interruption costs (EIC) via summing up all COC of power system load points. This approach does not take distribution facilities into consideration, like potential remedial switching actions in distribution substations, since this problem increases the computational effort dramatically because of the scale of the problem, and the aim of this doctoral thesis is finding an algorithm for optimal power transmission systems which are so far not affected by distribution substations [59]. Therefore the redundancy optimization is limited to transmission facilities [60]. Maintenance and normal non operation durations of generators, for instance thermal power plants in summer period, can be included in the probabilistic load data. Generation outages can be incorporated in the calculation via modifying PTDFs. However, in this doctoral thesis, ideal generation facilities are assumed, due to the main objectives of this doctoral thesis.

A predefined reliability constraint R_0, evaluated by a minimal state probability reliability estimation algorithm, is used to optimize power systems satisfying this constraint. A developed repair / improvement algorithm is used for individuals not satisfying this reliability constraint. Since EIC is used as objective function the algorithm optimizes all power systems to a minimum sum of TEC and EIC. This means, in other words, that the algorithm converges to a power system satisfying the reliability constraint and realized with minimized EIC which is a more restrict optimized power system with respect to reliability then a power system just achieving a certain reliability limit.

Hybrid algorithms are a combination of at least two different techniques. In this case expert knowledge based heuristics for mutation and repair and random crossover of a minimal spanning tree coded GA is combined with a random mutation and crossover based binary coded GA of section 3.6 to the hybrid GA. This is done in order to avoid an extension of search space to additionally $|PT|^{|E|}$ power line types. The redundancy optimization algorithm based on [61] with modified selection and evolutionary operators has the fastest convergence of the compared algorithms in [61] and is therefore chosen as main optimization algorithm in this doctoral thesis. The algorithm to estimate COC is described in [P5], and the redundancy optimization algorithm is described in detail in [P6].

A heuristic highest node degree, lowest cost mutation operator deletes the power line with highest cost at the substation with the highest node degree. This power line is, depending on a random variable, replaced at a node with minimum node degree by a power line with lower costs. This mutation simulates expert knowledge that H-connection substations with node degree 2 are privileged in terms of costs and reliability. The repair / improvement algorithm utilizes a cost weighted shortest path algorithm in order to repair the network not satisfying the reliability constraint.

The formulae Δf, used to obtain mutation, immigration and crossover probability based on a fuzzy logic approach is given in (3.6) where $\lambda(\Delta f_{avg})$ is one if it is greater or equal zero and zero otherwise and t is the number of current generation. More details about the implemented fuzzy logic self-controlled GA and the used functions can be found in [P6] and [61].

$$\Delta f=\sum_{i=1}^{u} 2^{t-i}\times\lambda\left(\Delta f_{avg}\left(t-i\right)\right) \qquad t\geq u \tag{3.6}$$

As stated in [P6], not all possible candidate networks, optimized by the optimal branch type determination algorithm described in the previous section, are analyzed with the fast enumeration based reliability calculation algorithm [P5]. Only a predefined quantile of the transmission expansion objective function is used to estimate the candidate networks where COC and EIC are assessed. Other networks are punished via a penalty factor. By doing this, it is assumed that candidate networks having small TEC are also more likely to have small EIC, resulting in an overall minimized cost function.

In order to verify the assumption that candidate networks having small TEC are more likely to have small EIC, a deeper understanding of the contribution of different outages to total COC is necessary. Let us consider the two optimized power systems in Fig. 3-4 and Fig. 3-5 to figure out the different outage costs of single and double line independent outages, substation outages and first order dependent power line outages applying short time congestion management via POPF. The following Fig. 3-6 gives an overview of the different outages and their percentage on the overall EIC per year.

The substation originated outages represent the major part of the overall EIC when (n-1) - criterion optimized power system is analyzed, since substations are critical facilities in a power system and already a single bus bar outage can lead to power supply interruption at load points.

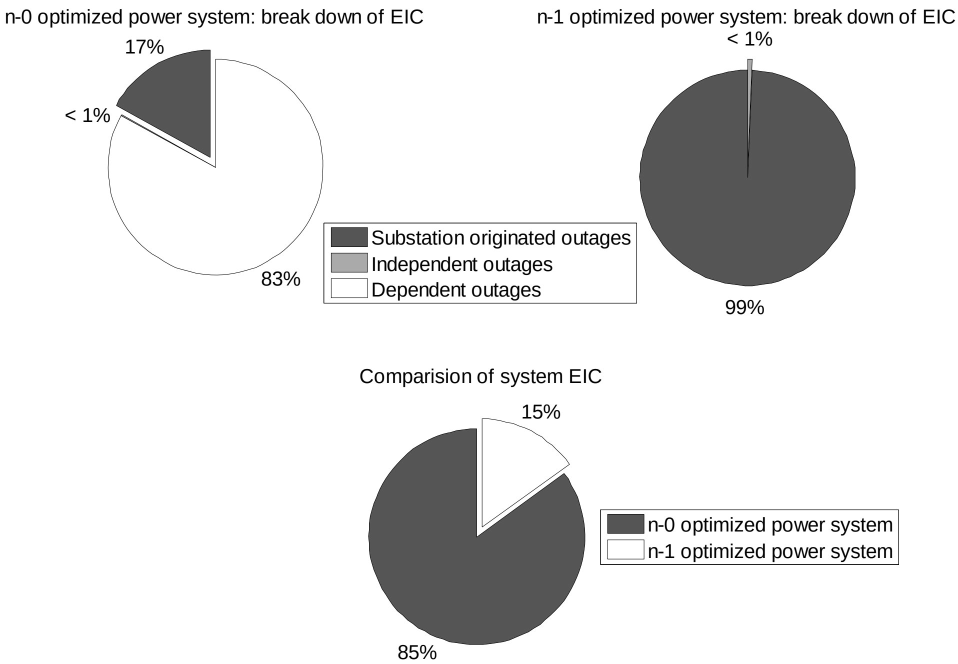

Fig. 3-6 Contribution of considered different outage types to total EIC per year of the (n-1)-criterion and the (n-0)-criterion satisfying power system and proportion of total EIC

As shown with Fig. 3-6 EIC depend mainly on the different SST and on dependent outages. Independent outages, like 1st and 2nd order outages including common mode outages, occur in both power systems, but their contribution to overall EIC is insignificantly small, in particular smaller than 1 %. However, since the enumeration based fast reliability assessment algorithm [P5] performs an expected outage calculation for 1st order dependent outages only, dependent outages are the most important factor in the power system satisfying (n-0)-criterion for realistic EIC.

This all leads to the conclusion for (n-1)-criterion optimized power systems that the number of power lines affects the overall expected interruption costs, but substation originated outages have a much higher impact. Since investment costs for substations are more or less twice as high as investment costs for power lines in the demonstrative real world example and the number of power lines is a direct impact factor on TEC and therefore also on EIC the applied approach, investigating a predefined quantile of individuals having small TEC and punishing all other individuals, can be justified.

Considering (n-0)-criterion optimized power systems, where dependent outages also affect overall EIC a different conclusion can be made. The repair algorithm of the developed genetic algorithm [P6] generates power systems having node degrees greater than or equal to two, which means no stub lines for load branches and having the ability to provide an alternative path for power flow in the case of an outage. Hence, the curtailed load by the applied POPF [P4] at a specific load point is always a proportion of the total load at this load point or equal to this load. Moreover the probability of such a state is always smaller than the related first order outage or equal but can affect more load points so that the EIC caused by dependent outages could result in individuals having a certain TEC exceeding the predefined quantile and hence resulting in a smaller overall cost function.
Therefore, the selection of reliability estimated individuals is changed in this doctoral thesis to additionally choosing a certain number (for example 10% of all individuals) of randomly selected individuals that are additionally investigated by [P5] in order to model the stochastic nature of genetic algorithms during the redundancy optimization.

As shown in [59] and [62] H-connections [P5] are better than single bus bar substations in terms of investment costs and expected outage frequency. Double bus bar substations are the most expensive ones and similar to H-connections in terms of reliability. Block connections are the worst in terms of reliability. The described heuristics in [P6], based on planning expert knowledge, is used to determine the substation type, depending on the number of touching power lines and generators. Analyzing the obtained (quasi)-optimal power system in [P6] with respect to power systems substations and compared to the modified utility driven power sub transmission system in [P6] one can obtain that the number of H-connection in power system increased (from 12 to 14) while the number of double bus bar substations remained constant and all block connections where removed by H-connections. This is evidence that the redundancy optimization also optimizes substations within expert knowledge based constraints.

Moreover, based on the investigations in [59] and [62] the genetic operator mutation and the repair algorithm modify each individual so that the node degree of the whole power system is minimized while satisfying a minimal node degree of 2 which results in a maximized number of H-connections.

The co-authors acted as discussion opponent and provided valuable inputs for expert knowledge based power system planning heuristics. The author of this doctoral thesis was fully responsible for development and formulation of the used planning heuristics in the crossover procedure and mutation procedure as well as for implementing the whole algorithm, testing and tuning. Moreover the author of this research work conducted all calculations, designed the repair / improvement algorithm and was fully responsible for drawn conclusions and results.

4 Optimal Power System for a Real World Sub Transmission Network

4.1 Real World Sub Transmission Network

4.1.1 Schematic View of the Power System

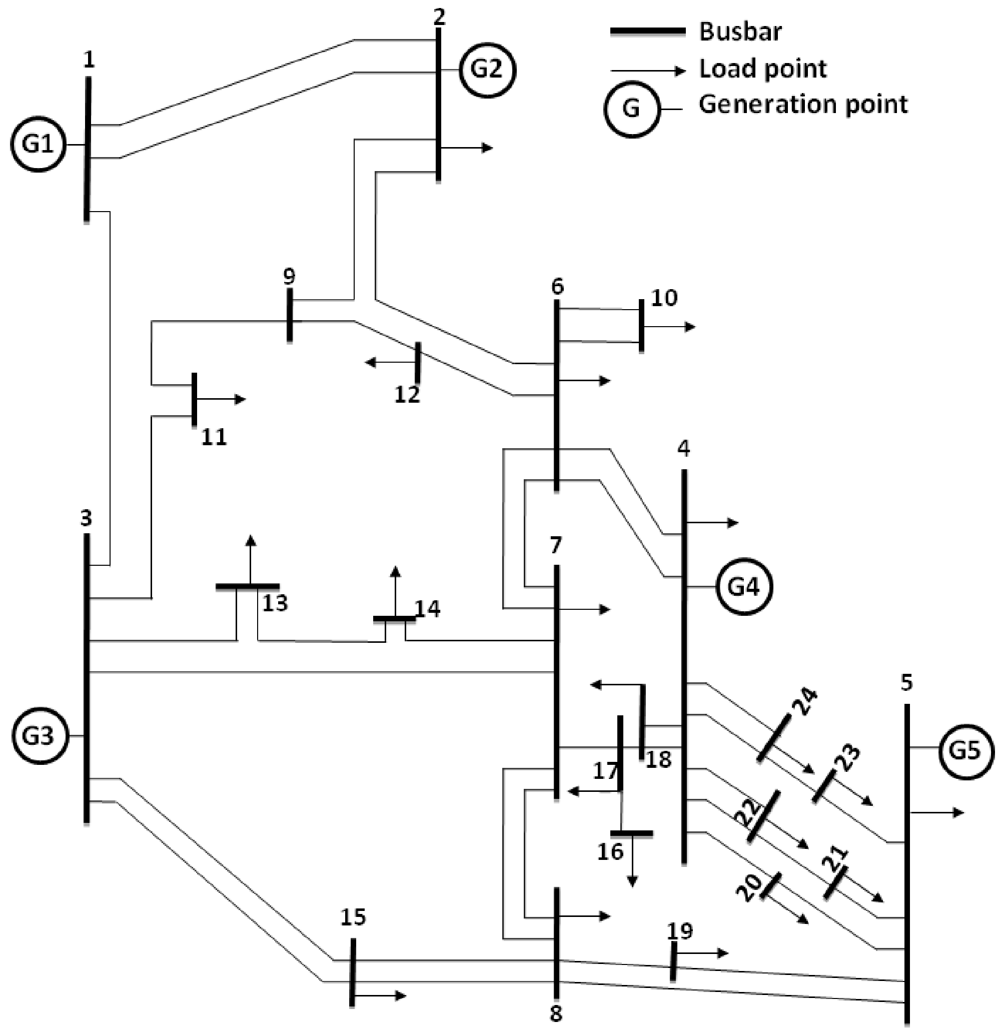

Fig. 4-1 Schematic view of the used real world 110-kV power sub transmission system consisting of 5 generation points, 21 load points, 24 substations and 43 power lines; power line from 1 to 3 is used as reserve connection.

To verify the enhancements of the developed redundancy optimization algorithm a real world 110-kV power sub transmission system, shown in Fig. 4-1, is optimized with respect to (n-1)-criterion. The optimal power system is calculated based on 84 identified possible power line connections and on necessary cost data [P6]. For ease of reference, Table 4.1 to Table 4.3 summarize the settings used in the optimization. The relevant power system data can be also found in the corresponding papers in section 7.

4.1.2 GA Parameter

Genetic algorithm parameters such as population size, number of generations, crossover, mutation and immigration are sensitive to the quality of the obtained calculation results. Crossover probability, mutation probability and immigration probability are adjusted by a fuzzy logic control approach after each generation providing a balanced exploitation and exploration of the search space.

Because the overall computation time depends on the population size and the number of generations, their impact has to be analyzed to obtain optimal GA settings for the algorithm. Genetic operator probabilities are adjusted after each generation; hence sensitivity analysis like the one performed in [47] are not significant since the exploration and exploitation of the search space and the punishing factor used for each uninvestigated individual go along with diversified standard deviation. However, simulation results show that a minimal population size of 40 and a minimum generation size of 40 are crucial for the performance of the developed GA for the proposed real world power sub transmission system. Obviously for other power system configurations with changed number of load points and / or generation points these numbers have to be adjusted in order to obtain meaningful results.

4.1.3 HGA Settings

In Table 4.1 the used GA settings for the modified transmission expansion problem are given and in Table 4.2 the used GA settings for the redundancy optimization are stated. Table 4.3 contains settings of the optimization algorithm framework.

Moreover mutual dependencies between different load points are considered during the optimization by the method presented in [P3].

Value	Comment
40	Population size
80	Maximum number of generations
0.9	Crossover probability
0.05	Mutation probability
0.30	If the difference in pu. between nominal rating and estimated worst case power line loading exceeds this value the next higher rated power line is chosen; see section 3.6.
10000	Multiplication factor for penalty factor for those power lines exceeding nominal rating

Table 4.1: Settings for modified transmission expansion problem genetic algorithm

Value	Comment
40	Minimal Population size (popSize)
40	Minimal number of generations
0.45	Initial crossover probability
0.45	Initial mutation probability
0.1	Initial immigration probability
0.4	Quantile for individual selection for EIC estimation
10	Percentage of additional randomly chosen individuals used for EIC estimation

Table 4.2: Settings for redundancy optimization genetic algorithm

Value	Comment
1.15	Minimum intermeshing degree
1.5	Maximum intermeshing degree
3	Fuzzy logic self control parameter u [61], [63]
2	Fuzzy logic self control parameter α [61], [63]
5	Interest rate in %
50	Period for NPV calculation
0.8	Quantile for minimum state probability reliability estimation
1	Bus one is slack bus
7	Number of moments used for PLF computation
0.001	Limit value at both ends of the cumulative distribution curve to estimate maximum load flow

Table 4.3: Settings for algorithm combining modified transmission expansion problem and redundancy optimization

4.2 (n-1)-Criterion Optimized Power System

Fig. 4-2 displays the (n-1)-criterion optimized power system. In this power system the NPV of TEC contributes 52.6 % to the overall NPV, and the NPV of the EIC consists the residual 47.4 %. This power system consist of 15 H-connection substations, which seems to be the maximum realizable number of H-connection substations, since the power system has 24 buses and five of them, the generation buses, are fixed to double bus bar substations. Maximizing the number of H-connection substation leads to minimizing the number of double circuit power lines, and therefore small loops of connected load points are mandatory. In Fig. 4-2 at most three load points with H-connection substation are connected to generation buses in one power line loop.

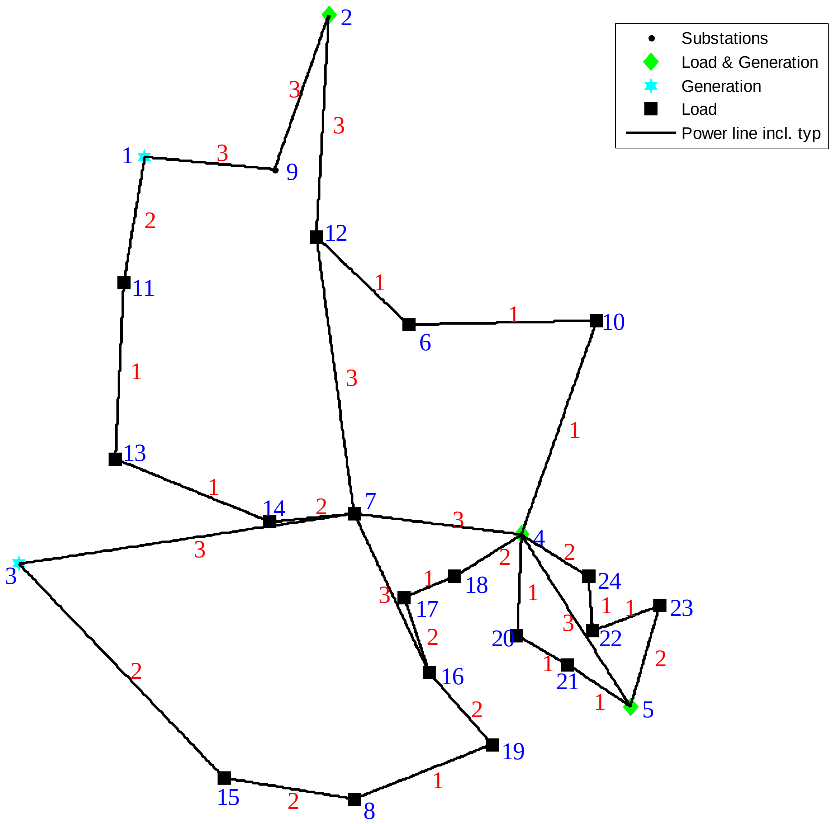

Fig. 4-2 The (n-1)-criterion optimized power system consists of 37 power lines, 15 H-connection substations, 6 double bus bar substations and 3 single bus bar substations; Best solution of 15 runs;

In the Fig. 4-3 the improvement by the optimization is depicted related to the real world power system shown in Fig. 4-1, where the NPV of EIC and NPV of TEC has been estimated based on the data used in this doctoral thesis. The HGA results in a network structure with a

22.3 % decreased objective function compared to the real world power system in Fig. 4-1. Although the real world power system is realized with more power lines the NPV of the system expected interruption costs (EIC) is higher than in the redundancy optimized power system. This is mainly related to the fact that the real world power system is not (n-1)-secure and that the implemented SSTs are not optimal in terms of the used reliability and cost data.

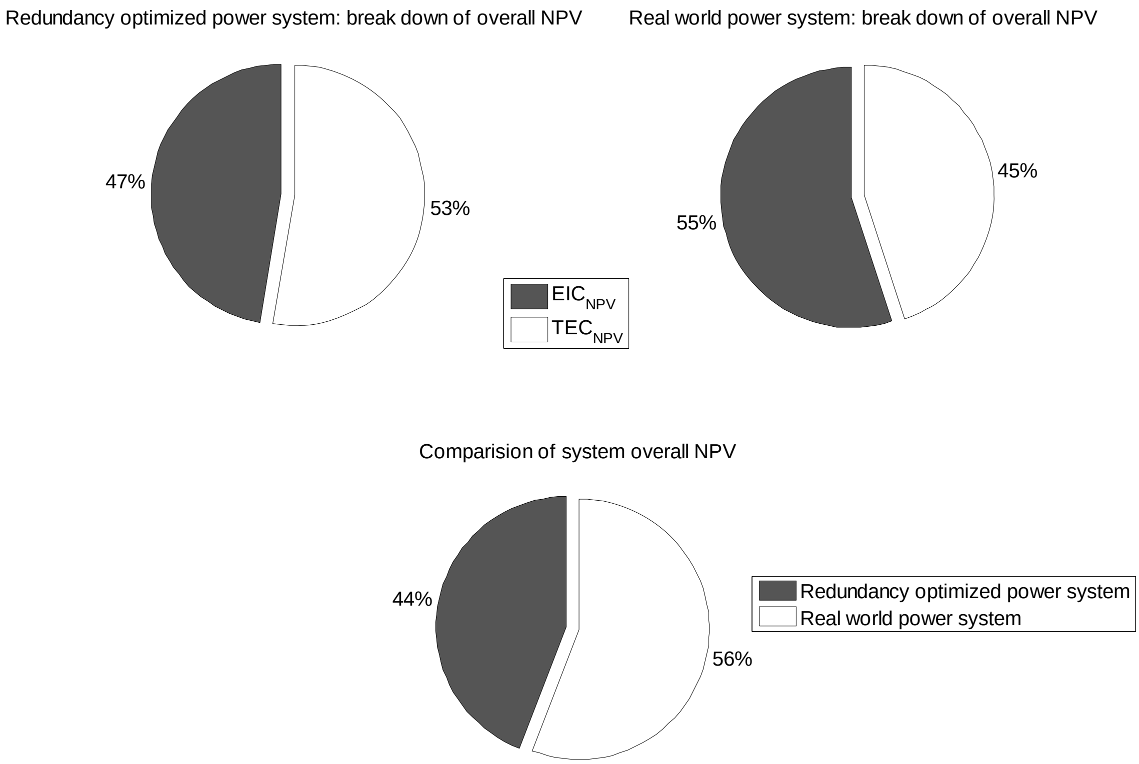

Fig. 4-3 Contribution of TEC and EIC to total NPV of the redundancy optimized power system and the real world power system in Fig. 4-1 and proportion of total NPV

4.3 Discussion of Obtained Results

Comparing the obtained (quasi)-optimal network structure in Fig. 4-2 with the real world 110-kV power sub transmission system in Fig. 4-1, shows that the number of power lines has decreased by 14 % while the number of H-connections has increased from 9 to 15. The power system in Fig. 4-2 satisfies the (n-1)-criterion while the intermeshing degree [P1] decreased from 1.79 to 1.54. The results of Fig. 4-3 clearly show that there is still room for improvement in the real world power system. For instance, a possible improvement is to build a power line from bus 7 to bus 16 and to deconstruct the power line from bus 7 to 17. This reduces the gap between the optimized power system and the real world power system considerably.

As shown in Fig. 3-6, the EIC net present value of each (n-0)-criterion optimized power system is the dominant factor in this optimization problem objective function. Since dependent first order outages represent the major contribution to total EIC, those power systems satisfying less (n-1)-contingencies are privileged during the optimization process. However, due to the high EIC costs (see Fig. 3-6), even if only best case values in terms of considered one hour outage duration are taken into account, this kind of optimized power system is not at all competitive in terms of the optimization objective value compared to (n-1)-criterion optimized power systems. Using the two (quasi)-optimal power system configurations (base case) obtained in section 3.6, the impact of power line rating increase and decrease is shown to strengthen the influence of load curtailment and power system expansion. All PT of the power system in Fig. 3-4 are sequentially increased to the next higher PT, and the NPV differences of the base case objective function and the NPV of each alternative are depicted in the box-and-whisker diagram in Fig. 4-4 (improved n-0). Similar results are obtained for the sequential decrease to the next lower PT of the power system in Fig. 3-5 (degraded n-1). These alternatives include either power line rating upgrade of a single circuit power line or the change from a single circuit to a double circuit power line including power line rating upgrade.

The investigated power line extensions of the power system in Fig. 3-4 cause a maximal increase of TEC NPV of 2 %. These improvements reduce the NPV of EIC in terms of the interquartile range between zero and 8 %. The optimization objective function, the net present value of TEC and EIC, vary between +3 % and -13 %. The change of a power line in Fig. 3-5 to the next lower rated PT leads to a decrease of TEC NPV of at most 2 %. In terms of EIC NPV the degraded power systems causes an increase of approximately 30 % as median and 94 % as maximum. Summing up both values the variation is still between -1 % and +26 %, neglecting possible outliers.

Moreover, the variation of TEC NPV for exchanging power line types is quite similar, as shown in Fig. 4-4. However, since each (n-1)-limit violation is eliminated by congestion via load curtailment and generation dispatch, these (n-1)-cases result in high changes for the EIC NPV. As depicted in Fig. 4-4 by the interquartile range of the (EIC + TEC) NPV box-and-whisker, each additional (n-1)-contingency can lead to an increase of the objective function between 8 % and 19 % with a median value of 15 %. This is by far more than all independent outages cause in the power system of Fig. 3-5 (see Fig. 3-6). Since the NPV TEC of the base case for the (n-1)-optimized power system is 48 % of the total NPV and the NPV of EIC is 52 % a median increase of EIC widens the gap between TEC and EIC resulting in an higher overall objective value compared to (n-1)-optimized power systems.

Based on the conducted optimization calculations and the developed optimization algorithm the following conclusions may be drawn. Even if a power system utility is able to implement a fast, powerful and effective short-term congestion management system it will be less efficient in terms of outage costs and investment costs compared to (n-1)-optimized power system. Therefore it is still mandatory to consider (n-1)-criterion when planning reinforcements or restructuring of power systems or when designing new power transmission systems. More importantly, the question if it is worth to build (n-1)-secure power systems, posed in section 3.6, can be definitely answered with yes.

As for instance discussed in [64], (n-1)-criterion based power systems, designed and based on heuristic planning strategies, may lead to welfare-economic overdesigned power systems. Also the famous "Zollenkopf"-criterion [64], a customer outage based power system planning approach, is not sufficient to ensure a welfare-economic ideal power system. Several other planning criteria, like those discussed in [19], [25] and [65], have been developed over the last years. Some of them have been derived by optimization algorithms.

However, they are all indented to be used by planning engineers and therefore do not guarantee (quasi)-optimal network designs.

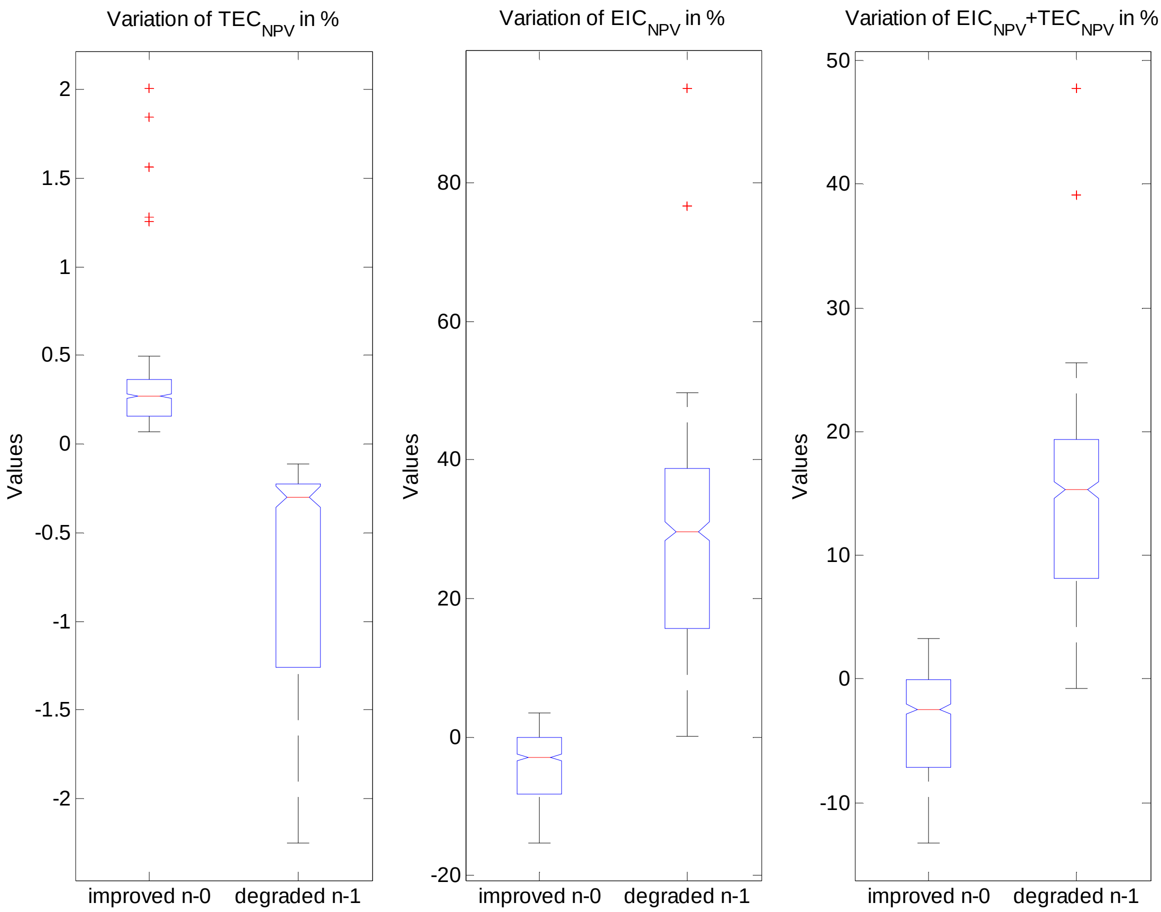

Fig. 4-4 NPV variation of TEC, EIC and (EIC+TEC), related to the corresponding base case value (see Fig. 3-4 and Fig. 3-5) to justify (n-1)-optimized power system expansion. The box-and-whisker diagram shows the lower quartile, upper quartile (blue box), median (red line) as well as whiskers to the adjacent values of data base. Outliers are displayed with a red + sign.

The results of this dissertation show that an efficient and fast optimization of power systems with respect to heuristic criteria can be conducted via the help of computer algorithms, resulting in (quasi)-optimal power system designs considering substations and uncertainties. Thus, planning of power systems based exclusively on engineer's skills and knowledge may become unnecessary. The results also suggest that the valuable assistance of computationally tractable power system optimization algorithms should be part of every network design study and planning problem, irrespective of the type of the conducted study and of the problem, in future.

5 Conclusion

5.1 Doctoral Thesis's Main Results

In this work it has been shown that the computation of MC up to the 2nd order is an appropriate method for fast reliability estimation of power transmission systems. The two proposed algorithms are based on different methodologies — a novel graph reduction methodology and a PTDF based path estimation methodology via matrix operations – are fast and efficient in deducing power system failure states. It has been theoretically shown that this method is computationally at least as fast, and in most cases even faster, than for instance analyzing all possible failure states up to the 2nd order in a power system.

The classical planning scenarios, such as peak load, peak generation and off-peak load, usually only cover a limited number of critical loading situations, and are therefore not adequate to give worst case power flow scenarios of a power system, as demonstrated in this doctoral thesis. Hence new approaches have to be developed. One approach is to determine a characteristic set of load flow cases as a subset of measured annual load flow data with a combination of a heuristic search algorithm and a genetic optimization algorithm. The other one is to use PLF also based on measured annual load flow data and measured nodal powers respectively.

By means of a conservative mutual information measure, real measured load data of a 110-kV power sub transmission grid has been proven to contain mutual dependencies. Therefore, using a set of RVs as input data for PLF calculation, these mutual dependencies have to be taken into account to obtain more realistic and significant calculation results.
Based on Gram-Charlier Expansion the doctoral thesis author has developed enhanced formulae to consider mutual dependencies up to 4th order. The traditional formulae of

cumulants have been extended to joint cumulants to include these mutual dependencies; additionally, the formulae of joint cumulants have been generalized to equations where n RVs of bus injected powers are involved to calculate the distribution of active power flows on power lines.

A synthetic example clarifies the influence of different orders of joint cumulants on the goodness of the distribution. The real world calculation results are also verifying that, the accuracy of the obtained results enhances with increasing order of considered joint cumulants. A 2m+1 PE scheme, where correlations between RVs are directly accounted for with a rotational transformation of the input RVs, is less suitable for this kind of problem compared with the method where 2nd order joint cumulants are computed. Hence, in the author’s view the 2nd order joint cumulants method is the best methodology for the proposed application with respect to computation accuracy and computation time.

A benchmark of four different PE methods — Harr’s PE method, 2m, 2m+1 and 4m+1 scheme — with benchmark criteria computation time, computation accuracy and problem applicability verifies that Harr’s algorithm is the most suitable short-term congestion management algorithm utilizing POPF from the point of view of benchmark criteria. Although Harr’s algorithm has been originally designed for Gaussian RVs, it has also reasonable advantages when Non-Gaussian RVs are applied. Basically, the results of the POPF calculation depend on the input distributions and on the amount of the curtailed load or dispatched generation. Power overloadings, occurring when single power line outages lead to an increased power flow, are optimally cleared via remedial actions based on this algorithm.

A novel approach for assessing reliability indices, including substation originated outages, with uncertain nodal powers has been developed in this doctoral thesis. The methodology is based on PLF and on POPF. Comparing results of the proposed methodology with the indices obtained with commercial available software an adequate degree of accuracy can be observed. Furthermore, compared to commercial software, a major advantage of the proposed methodology is the significantly reduced computation time. The implemented POPF used for optimal load curtailment calculation and generation dispatch (short-time congestion management) in any contingency case is a computationally efficient way to consider uncertainties in the load curtailment process. Furthermore, the obtained distributions show, compared to MCS, a high goodness of fit. As demonstrated, the proposed methodology allows including annual distributions of nodal powers in a reliability enumeration algorithm while maintaining an adequate accuracy for all used reliability indices.

A fuzzy logic self-controlled HGA is used for redundancy optimization of a power transmission system with uncertain load data. The modified transmission network expansion problem, i.e., finding optimal power line types for a certain PLF, is solved with respect to minimal cyclic NPV by a binary coded GA, including different substation types depending on the number of touching power lines. This GA can be used for this nonlinear optimization task, resulting in an optimal or (quasi)-optimal grid structure satisfying, if wanted, (n-19-criterion based power flow constraints. The redundancy optimization is done by a GA based on a reliability improvement algorithm that satisfies a minimum node degree of two while reducing the power system investment cost. By applying a shortest path algorithm and a heuristic modification, each individual, having a non availability greater a certain limit, is repaired / improved. This methodology, combined with a special mutation and crossover routine, provides an improvement of the overall investment, operating, renewal and expected outage costs based on CIC and NPV calculation. The proposed powerful HGA can be used for green field studies, power system expansion planning, benchmark studies, restructuring studies, reinforcement studies, planning studies and grid design studies, all including uncertainties and transmission network reliability.

The conducted research suggests that powerful optimization algorithms, like the one developed in this dissertation, should be part of every power system design study in future in order to ensure (quasi)-optimality in power system designs, thus resulting in minimal welfare - economic costs.

5.2 Scientific Importance of Conducted Work

The scientific importance of the conducted work proposed in this dissertation is based on following five factors.

- To the best of the author's knowledge, substation related outages have so far not been considered in redundancy optimization and transmission expansion planning. By the incorporation of these substations related outages a more detailed and more realistic reliability measure is used as optimization objective.

- Via short-term congestion management utilizing probabilistic optimized power flow dependent outages are considered in an enumerative reliability assessment algorithm. It is not completely new to consider dependent outages during a reliability assessment but the applied probabilistic congestion management approach is by the best of the author's knowledge completely novel and has not been proposed before.

- Minimal cutsets are a well known method for reliability estimation of engineering systems but have not been used extensively for power systems due to high computational efforts. MC estimation via the two proposed approaches in this doctoral thesis eliminates the drawback of cumbersome computation and provides therefore a powerful method to deduce network states resulting in failure states.

- The concept of using joint cumulant based probabilistic load flow, considering mutual dependencies of input random variables, has been deduced by the author and it was shown that consideration of joint cumulants up to the second order is sufficient for reasonable accuracy of the obtained calculation results. By the best of the author's knowledge, this point has not been investigated by anyone else in this context before.

- All the previous listed facts are merged into a novel nonlinear algorithm performing power system optimization based on an objective function considering investment costs, operational costs, expenditure for equipment renewal and expected interruption costs of load points. Due to the nonlinearity of the problem, finding the global optimum cannot be guaranteed. However, the method presented in this doctoral thesis allows to estimate (quasi)-optimal solutions, enabling the optimization of power systems (including substations) to not only satisfy a certain reliability constraint but also to minimize the even more restricted objective of expected interruption costs.

5.3 Future Research and Development

Future research can be carried out on the topic of fast PLF computation based on AC load flow, including mutual dependencies consideration. This is necessary since voltage profiles are assumed to be constant in DC power flow but with increasing network size or increasing uncertain power supply, for instance due to wind power plants, large photovoltaic power plants, etc. voltage level considerations can get more importance in transmission system planning. In the authors view this could either be based on fast, problem optimized MCSs or on enhanced convolution methods.

Due to increasing uncertainties in power flows, probabilistic congestion management utilizing POPF will gain importance in future transmission planning processes. Hence, fast improved methods for long term congestion management cost estimation are mandatory. Possible solution approaches and research areas should, in the author's view, be based on computation speed optimized Monte Carlo Simulations.

Beside load data uncertainties other uncertainties could become worth considering. Forecast values, outage durations, outage frequencies, etc. are in real world problems not one single number but probability distributions. These uncertainties can be included in the reliability assessment algorithm with one of the proposed PE methods.

The proposed optimization algorithm is a tool for static planning purposes. In some real world problems it is necessary to develop optimal expansion plans or restructuring plans in a dynamic sense, which means not all power system modification arrangements have to be established at the same time. This task can be solved for instance by a step wise optimization of each particular planning year. However, if power system restructuring or reinforcement is considered as a way of getting optimal from one planning year to another additional optimization processes have to be carried out for these transition states. Based on [47], all these duties can be solved via the proposed hybrid genetic algorithm with arguable effort on modifying the algorithm structure and objective function.

A further topic for future research could include the development and investigation of methodologies to consider flexible alternating current transmission system (FACTS) in the optimization function. FACTS are static equipment mend to enhance controllability and increase power transfer capability and are an interesting research topic. Therefore it could be worthwhile taking this kind of equipment into account in an optimization algorithm as it has already been demonstrated by the author of this doctoral thesis in [66].

The above-mentioned implementation of transmission losses and generation outages in the algorithm is another development point for the future. Additionally, modeling of adverse weather situation, lightning and other reliability affecting and monetary ratable impacts could be further seen as interesting developments for the proposed algorithm in future. Last but not least some effort should also be put on computation time decrease and algorithm convergence increase for instance with different programming languages, different genetic strategies or more suitable and sophisticated data types.

6 References

[1] B. Österreich. (2009, Apr.) Bundesgesetz, mit dem die Organisation auf dem Gebiet der Elektrizitätswirtschaft neu geregelt wird (Elektrizitätswirtschafts- und -organisationsgesetz – ElWOG), idF BGBl. I Nr. 112/2008. Internet, Statute in German.

[2] A. Haber, *Entwicklung und Analyse eines Qualitätsregulierungsmodells für die österreichischen Mittelspannungsnetze*. University of Technology, Graz, Austria, 2005, Dissertation in German.

[3] E.-C. GmbH. (2008) Technische und organisatorische Regeln für Betreiber und Benutzer von Netzen; Teil B: Technische Regeln für Netze mit Nennspannung >= 110 kV. Document in German.

[4] T.-O. Léautier and V. Thelen, "Optimal expansion of the power transmission grid: why not?," *Journal of Regulatory Economics*, vol. 36, pp. 127-153, Jan. 2009.

[5] R. Billinton and R. N. Allan, *Reliability Evaluation of Engineering Systems, Concepts and Techniques*, 2nd ed. New York, N.Y. 10013: Plenum Press, 1992, Page 7.

[6] R. Billinton and R. N. Allan, *Reliability Evaluation of Power Systems*, 2nd ed. New York, N.Y. 10013: Plenum Press, 1996, Page 8-9.

[7] A. Lisnianski, G. Levitin, H. Ben-Haim, and D. Elmakis, "Power System structure optimization subject to reliability constraints," *Electric Power System Reserach*, no. 39, pp. 145-152, 1996.

[8] G. Latorre, R. D. Cruz, J. M. Areiza, and A. Villegas, "Classification of Publications and Models on Transmission Expansion Planning," *IEEE Transactions on Power Systems*, vol. 18, no. 2, pp. 938-946, May 2003.

[9] E. Luiz da Silva, H. A. Gil, and J. M. Areiza, "Transmission Network Expansion Planning Under an Improved Genetic Algorithm," *IEEE Transactions on Power Systems*, vol. 15, no. 3, pp. 1168-1175, Aug. 2000.

[10] K. P. Murphy. (2001, Oct.) The Bayes Net Toolbox for Matlab. http://HTTP.CS.Berkeley.EDU/~murphyk/Papers/bnt.ps.gz.

[11] D. Gleich. (2008, Oct.) MatlabBGL A Matlab Graph Library. http://www.stanford.edu/~dgleich/programs/matlab_bgl/.

[12] R. D. Zimmerman and C. E. Murillo-Sánchesz. (2007, Sep.) MATPOWER, A MATLAB™ Power System Simulation Package. http://www.pserc.cornell.edu/matpower.

[13] P. Maghouli, S. H. Hosseini, M. O. Buygi, and M. Shahidehpour, "A Multi-Objective Framework for Transmission Expansion Planning in Deregulated Environments," *IEEE Transactions on Power Systems*, vol. 24, no. 2, pp. 1051-1061, May 2009.

[14] P. S. Georgilakis, "Market-based transmission expansion planning by improved differential evolution," *Electrical Power and Energy Systems*, vol. 32, pp. 450-456, 2010.

[15] G. C. Oliveira, S. Binato, and M. V. F. Pereira, "Value-Based Transmission Expansion Planning of Hydrothermal Systems Under Uncertainty," *IEEE Transactions on Power Systems*, vol. 22, no. 4, pp. 1429-1435, Nov. 2007.

[16] J. Choi, A. R. A. El-Keib, and T. Tran, "A Fuzzy Branch and Bound-Based Transmission System Expansion Planning for the Highest Satisfaction Level of the Decision Maker," *IEEE Transactions on Power Systems*, vol. 20, no. 1, pp. 476-484, Feb. 2005.

[17] J. Choi, et al., "A Method for Transmission System Expansion Planning Considering Probabilistic Reliability Criteria," *IEEE Transactions on Power Systems*, vol. 20, no. 3, pp. 1606-1615, Aug. 2005.

[18] J. Choi, T. Tran, T. D. Mount, and R. Thomas, "Composite Power System Expansion Planning Considering Outage Cost," *Proceedings of the 2007 IEEE Power Engineering Society General Meeting* , pp. 1-5, Jun. 2007.

[19] G. B. Shrestha and P. A. J. Fonseka, "Congestion-Driven Transmission Expansion in Competitive Power Markets," *IEEE Transactions on Power Systems*, vol. 19, no. 3, pp. 1658-1665, Aug. 2004.

[20] H. A. Gil, E. L. da Silva, and F. D. Galiana, "Modeling Competition in Transmission Expansion," *IEEE Transactions on Power Systems*, vol. 17, no. 4, pp. 1043-1049, Nov. 2002.

[21] J. Choi, T. D. Mount, R. J. Thomas, and R. Billinton, "Probabilistic reliability criterion for planning transmission system expansions," *IEE Proceedings Generation, Transmission and Distribution*, vol. 153, no. 6, pp. 719-727, Nov. 2006.

[22] P. Jirutitijaroen and C. Singh, "Reliability Constrained Multi-Area Adequacy Planning Using Stochastic Programming With Sample-Average Approximations," *IEEE Transactions on Power Systems*, vol. 23, no. 2, pp. 504-513, May 2008.

[23] P. Attaviriyanupap and A. Yokoyama, "Transmission Expansion in the Deregulated Power System Considering Social Welfare and Reliability Criteria," *Transmission and Distribution Conference and Exhibition: Asia and Pacific, 2005 IEEE/PES*, pp. 1-6, Dec. 2005.

[24] S. de la Torre, A. J. Conejo, and J. Contreras, "Transmission Expansion Planning in Electricity Markets," *IEEE Transactions on Power Systems*, vol. 23, no. 1, pp. 238-248, Feb. 2008.

[25] J. Choi, T. D. Mount, and R. J. Thomas, "Transmission Expansion Planning Using Contingency Criteria," *IEEE Transactions on Power Systems*, vol. 22, no. 4, pp. 2249-2261, Nov. 2007.

[26] L. S. Moulin, M. Poss, and C. Sagastizábal, "Transmission expansion planning with re-design," *Energy Systems*, vol. 1, no. 2, pp. 113-139, May 2010.

[27] N. Alguacil, A. L. Motto, and A. J. Conejo, "Transmission Expansion Planning: A Mixed-Integer LP Approach," *IEEE Transactions on Power Systems*, vol. 18, no. 3, pp. 1070-1077, Aug. 2003.

[28] N. Alguacil, M. Carrión, and J. M. Arroyo, "Transmission network expansion planning under deliberate outages," *Electrical Power and Energy Systems*, vol. 31, pp. 553-561, 2009.

[29] A. Khodaei, M. Shahidehpour, and S. Kamalinia, "Transmission Switching in Expansion Planning," *IEEE Transactions on Power Systems*, pp. 1-12, Feb. 2010, has been accepted for inclusion in a future issue of this journal.

[30] M. Carrión, J. M. Arroyo, and N. Alguacil, "Vulnerability-Constrained Transmission Expansion Planning: A Stochastic Programming Approach," *IEEE Transactions on Power Systems*, vol. 22, no. 4, pp. 1436-1445, Nov. 2007.

[31] L. Bahiense, G. C. Oliveira, M. Pereira, and S. Granville, "A Mixed Integer Disjunctive Model for Transmission Network Expansion," *IEEE Transactions on Power Systems*, vol. 16, no. 3, pp. 560-565, Aug. 2001.

[32] M. Lu, Z. Y. Dong, and T. K. Saha, "A Hybrid Probabilistic Criterion for Market-based Transmission Expansion Planning," *Proceedings of the 2006 IEEE Power Engineering Society General Meeting*, pp. 1-7, Jun. 2006.

[33] I. de J. Silva, M. J. Rider, R. Romero, A. V. Garcia, and C. A. Murari, "Transmission network expansion planning with security constraints," *IEE Proceedings Generation, Transmission and Distribution*, vol. 152, no. 6, pp. 828-836, Nov. 2005.

[34] J. A. Gómez-Hernández, J. Robles-García, and D. Romero-Romero, "Transmission project profitability including reliability optimization of bulk power systems," *Electrical Power and Energy Systems*, vol. 28, pp. 421-428, 2006.

[35] F. Cadini, E. Zio, and C. A. Petrescu, "Optimal expansion of an existing electrical power transmission network by multi-objective genetic algorithms," *Reliability Engineering and System Safety*, vol. 95, pp. 173-181, 2010.

[36] Y. Wang, et al., "Pareto optimality-based multi-objective transmission planning considering transmission congestion," *Electric Power Systems Research*, vol. 78, pp. 1619-1626, 2008.

[37] Z. Xu, Z. Y. Dong, and K. P. Wong, "Transmission planning in a deregulated environment," *IEE Proceedings Generation, Transmission and Distribution*, vol. 153, no. 3, pp. 326-334, May 2006.

[38] H. A. Gil and E. L. da Silva, "A reliable approach for solving the transmission network expansion planning problem using genetic algorithms," *Electric Power Systems Research*, vol. 58, pp. 45-51, 2001.

[39] M. O. Buygi, G. Balzer, H. M. Shanechi, and M. Shahidehpour, "Market-Based Transmission Expansion Planning," *IEEE Transactions on Power Systems*, vol. 19, no. 4, pp. 2060-2067, Nov. 2004.

[40] M. O. Buygi, H. M. Shanechi, G. Balzer, M. Shahidehpour, and N. Pariz, "Network Planning in Unbundled Power Systems," *IEEE Transactions on Power Systems*, vol. 21, no. 3, pp. 1379-1387, Aug. 2006.

[41] M. O. Buygi, *Transmission Expansion Planning in Deregulated Power Systems*. University of Technology Darmstadt, Germany, Sep. 2004, Dissertation.

[42] K. Y. Lee and M. A. El-Sharkawi, *Modern Heuristic Optimization Techniques*, 1st ed., J. W. &. Sons, Ed. Hoboken, New Jersey, 2008.

[43] G. Rechberger, A. Gaun, and H. Renner, "Determination and Analysis of Load Flow Cases for Transmission System Planning," *Proceeding of Operation and Development of Power Systems in the New Context*, pp. 1-10, 2009, Symposium Guilin, China.

[44] D. Romero-Romero, J. A. Gómez-Hernández, and J. Robles-García, "Reliability optimisation of bulk power systems including voltage stability," *IEE Proceedings Generation, Transmission and Distribution*, vol. 150, no. 5, pp. 561-566, Sep. 2003.

[45] W. Zhang, *Reliability Evaluation of Bulk Power Systems using Analytical and Equivalent Approaches*. University o f Saskatchewan, Canada, 1999, Dissertation.

[46] N. A. A. Samaan, *Reliability Assessment of Electric Power Systems using Genetic Algorithms*. Texas A&M University, United States of Amerika, 2004, Dissertation.

[47] H.-C. G. Maurer, *Integrierte Grundsatz- und Ausbauplanung für Hochspannungsnetze*. RWTH Aachen, Germany, 2004, Dissertation in German.

[48] R. Nighot and R. Billinton, "Reliability evaluation of the IEEE-RTS incorporating station related outages," *Proceedings of IEEE Power Engineering Society General Meeting 2004*, vol. 1, pp. 118-123, Jun. 2004.

[49] A. S. D. Braga and J. T. Saraiva, "A Multiyear Dynamic Approach for Transmission Expansion Planning and Long-Term Marginal Costs Computation," *IEEE Transactions on Power Systems*, vol. 20, no. 3, pp. 1631-1639, Aug. 2005.

[50] T. Güler and G. Gross, "Detection of Island Formation and Identification of Causal Factors Under Multiple Line Outages," *IEEE Transactions on Power Systems*, vol. 22, no. 2, pp. 505-513, May 2007.

[51] R. Billington and W. Li, *Reliability Assessment of Electric Power Systems Using Monte Carlo Methods*, 1st ed. New York, United States of Amerika: Plenum Press, 1994.

[52] A. Hald, "The Early History of the Cumulants and the Gram-Charlier Series," *International Statistical Review*, vol. 68, no. 2, pp. 137-153, Aug. 2000.

[53] P. Zhang and S. T. Lee, "Probabilistic Load Flow Computation Using the Method of Combined Cumulants and Gram-Charlier Expansion," *IEEE Transactions on Power Systems*, vol. 19, no. 1, pp. 676-682, Feb. 2004.

[54] L. A. Sanabria and T. S. Dillon, "Power system reliability assessment suitable for a deregulated system via the method of cumulants," *Electrical Power & Energy Systems*, vol. 20, no. 3, pp. 203-211, 1998.

[55] M. E. Harr, "Probabilistic estimates for multivariate analyses," *Applied Mathematical Modelling*, vol. 13, no. 5, pp. 313-318, May 1989.

[56] BCP Busarello, Cott and Partner AG. (CH-8703 Erlenbach, Switzerland) NEPLAN Version 5.3.51. [Online]. http://www.neplan.ch

[57] A. Ozdemir, R. Caglar, and C. Singh, "Transmission losses based MW contingency ranking," *Electrical Engineering*, vol. 88, no. 4, pp. 275-280, 2006.

[58] M. Hyvärinen, *Electrical networks and economies of load density*. Helsinki University of Technology, Espoo/Finnland, 2008, Doctoral Thesis.

[59] P. Wolffram and C. Maurer, "Ermittlung langfristiger optimaler Strukturen für 110-KV-Netze unter Berücksichtigung der Stationskonzepte," *A. B. z. Energieversorgung (IAEW-FGE-Jahresbericht 2003)*, pp. 93-96, 2003, in German.

[60] R. Billinton and R. N. Allan, *Reliability Evaluation of Power Systems*, 2nd ed. New York, N.Y. 10013: Plenum Press, 1996, Page 10-11.

[61] L. Lin and G. Mitsuo, "A Self-controlled Genetic Algorithm for Reliable Communication Network Design," *Proceedings of 2006 IEEE Congress on Evolutionary Computation*, pp. 640-647, Jul. 2006.

[62] A. Braun, *Anlagen und Strukturoptimierung von 110-kV-Netzen*. RWTH Aachen, Germany, 2001, Dissertation in German.

[63] M. Gen, R. Cheng, and L. Lin, *Network Models and Optimization; Mutliobjective Genetic Algorithm Approach*, 1st ed. London, UK: Springer-Verlag London Limited, 2008.

[64] L. Fickert, et al., *110-kV-Kabel / -Freileitung, Eine technische Gegenüberstellung*, 2nd ed., IFEA/IHS, Ed. Graz, Austria: Verlag der Technischen Universität Graz, 2005, in German.

[65] T. Tran, et al., "A Study on Optimal Reliability Criterion Determination for Transmission System Expansion Planning," *KIEE International Transactions on Power Engineering*, vol. 5-A, no. 1, pp. 62-69, Jan. 2005.

[66] A. M. Othman, A. Gaun, M. Lehtonen, and M. El-Arini, "Real World Optimal UPFC Placement and its Impact on Reliability," *IASME/WSEAS International Conference on Energy and Environment*, pp. 90-95, Feb. 2010.

7 Annexes

7.1 Annex A: Reference P1

Paper printed in conference proceedings and presented at 2009 IEEE Bucharest Power Tech Conference, June 28th - July 2nd, 2009, Bucharest, Romania

Fast Minimal Cutset Evaluation in Cyclic Undirected Graphs for Power Transmission Systems

A. Gaun, H. Renner, G. Rechberger

Abstract-- **The evaluation of network reliability is an important topic in the planning, design, and operation of power systems. The aim of this publication is to investigate and tune up the speed of calculating 2-terminal minimal cutsets (MC). Two fast and well known algorithms are compared with a new developed MC algorithm concerning the speed of computation. The new designed algorithm is based on a novel graph reduction method, and on an adapted recursive merge method. Eleven benchmark-networks are used to analyze all three MC algorithms. Experimental results show that the new developed MC algorithm has a linear dependency between the computation time and the graph density of a network for a fixed number of nodes. Furthermore it is shown that the proposed algorithm is faster than the minimal path based algorithms and the currently best available MC algorithm for 2-terminal reliability in complex power transmission networks. A representative 57 node power transmission network demonstrates that the new proposed algorithm reduces the computation time for all relevant MC by 96.2 %.**

Index Terms-- **2-terminal network reliability, computation time, graph reduction, induced cycle, minimal cutset, power transmission system, recursive merge;**

I. Acronym, Notations, Nomenclature and Assumptions

A. Acronym

CI	cycle-incidence matrix
MC	minimal cut(s) / minimal cutset(s)
MP	minimal path(s) / minimal path set(s)
s	specified source node
t	specified sink node

B. Notations

G (V, E)	A connected network (graph) G with the node set $V = \{s, t, x_0, x_1, \ldots, x_{k-2}\}$ and the edge set E [1]. For example, Fig. 1 center part is a network. This paper treats only planar cyclic undirected networks.
$e_{x_1x_2}$	$e_{x_1x_2} \in E$ is an arc between nodes x_1 and x_2.
$\lvert\cdot\rvert$	The number of elements of ·, e.g., $\lvert V\rvert$ is the number of nodes in V.
P	A non-empty graph $P = (V, E)$ with $V = \{x_0, x_1, \ldots, x_k\}$ and $E = \{e_{x_1x_2}, e_{x_2x_3}, \ldots, e_{x_{k-1}x_k}\}$ [2]. P is called a path.
C	If $P = x_0 \ldots x_{k-1}$ is a path and $k \geq 3$, then the graph $C = P + e_{x_{k-1}x_0}$ is a cycle [2].
SS	A connected source set of nodes with $s \in SS$ [3].
G * x_1	Node x_1 is merged into SS of G by eliminating any edge connecting x_1 and SS [3].

C. Nomenclature

unavailability	Is defined as $\lambda_{MC} / (\lambda_{MC} + \mu_{MC})$ in h / year with the failure rate λ_{MC} and the repair rate μ_{MC} for each minimal cut of the MC [4], [5].
cut order	The cut order is defined as the number of elements that must fail, to disconnect the sink t and the source s. In this paper written as o(MC).
node degree	The node degree is the number of all edges that touch the node, written as $d(x_1)$; $\delta(V)$ is the minimum degree of V with $\delta(V) := \min\{d(x_i) \mid \forall x_i \in V\}$ [2].
adjacent node	A node x_1 is said to be adjacent to a subgraph of G if there exists an undirected edge $e_{x_1x_2}$ leading from this subgraph to the node x_1 [3]. If G contains G' (G' ⊆ G), G' is a subgraph of G.
redundant node	A redundant node is a node which is adjacent to SS and has no P to t without going through any node in SS [3]. For example see node 10 in Fig. 4 Step 4. A redundant loop contains only redundant nodes (see Fig. 4 Step 4 loop 18-8-13-18).
MC	A 2-terminal cutset is a set of arcs such that, by removing these arcs, there is no P in G from node s to node t. A MC is a cutset with no subset of it is a cut. For example, see Fig. 4 Step 4 MC1. MC are also expressed in terms of node sets [6].

A. Gaun is with the Institute of Electrical Power Systems, Graz University of Technology, A-8010 Graz (e-mail: alexander.gaun@tugraz.at).
H. Renner is with the Institute of Electrical Power Systems, Graz University of Technology, A-8010 Graz (e-mail: herwig.renner@tugraz.at).
G. Rechberger is with the Institute of Electrical Power Systems, Graz, University of Technology, A-8010 Graz (e-mail: georg.rechberger@tugraz.at).

MP	A path P between s and t is minimal, if no node or branch is traversed more than once in this path.
induced cycles	An induced cycle C in G is a cycle in G forming an induced subgraph that has no chords. A chord of a C is an edge that joins two nodes of C but is not itself an edge of C [2]. For example see chordless cycles I-VI Fig. 4 Step 2. A graph consists of |E| - |V| + 1 induced (chordless) cycles.

A. Assumption

The networks satisfy following assumptions [6], [7]:

1) Perfectly reliable nodes.
2) Connected planar cyclic undirected graph with no parallel branches.
3) Each edge is either in working or failed state with known probability.
4) All power flows in the network obey the conservation law.
5) All nodes have to be numbered in increasing order starting with s and ending with t.

II. Introduction

POWER transmission network reliability is an important topic in the planning, design, and operation of power systems. Several algorithms for evaluation of terminal reliability evaluation are proposed and classified in the literature [8]. Among them the MC approach is one of the most popular used techniques to evaluate 2-terminal network reliability [1], [9]. MC provide a list of events that cause network failures, a disconnection between the 2-terminals s and t, and therefore MC are preferred to calculate the network reliability. The determination of MC or MP is necessary to reduce the sum of disjoint product terms and, hence, the overall reliability computation time. For some networks it is simply impractical to enumerate all MP [3]. For instance, a 2x49 lattice network (see Fig. 7 (11)) contains 2^{49} paths. Although it has a huge number of MP, it only contains 2500 MC. Hence, the computation of all MC and the reliability of a network is a time intensive operation and can grow exponentially with the number of nodes (NP-hard), even if no MP are calculated [6], [10], [11]. The aim of this publication is to investigate and tune up the speed of calculating MC for power transmission networks.

Usually elements of power transmission networks have a very small unavailability. It is proven, that in this case MC up to order plus one of the lowest cut order of the investigated network have to be considered [5], [12]. In power transmission networks in the majority of cases the lowest MC order of a network is equal to one, if substations are considered [12], or two. Thus it is sufficient in the majority of the cases to calculate MC up to the 2nd order for reliability analysis, especially in the context with transmission grid reliability optimisation, e.g. with Genetic Algorithms [13] where hundreds and thousands of these operations have to be done in large networks. In this paper three 2-terminal reliability algorithms are benchmarked with eleven planar cyclic undirected benchmark-networks [3], [14] (see Fig. 7). Two of the algorithms are well known from the literature [6], [15] and one is a new proposed algorithm that is based on a novel network reduction method and on an adapted recursive merge algorithm.

This paper is organized as follows: in Section III a short overview of related researches is given and a short description of the used algorithms is provided. Section IV contains a detailed report about the network reduction, including the relevant theorem to show the accurateness of the Algorithm and the extended Algorithm itself. The third main part in Section IV deals with a representative example. Section V contains the benchmark of the three different networks with remarks to the computation time and the performance with different sizes of networks. With a 57 node sample network the effectiveness of the new proposed algorithm is demonstrated in computation all MC. Concluding remarks are presented in Section VI.

III. Description of the used Algorithms

An overview of related researches in determining (all) MC in graphs up to the year 2003 can be found in [6]. Within the last five years a new approach to calculate MP has been published in [7]. This algorithm requires fewer calculations to generate MP and is more effective in generating MP without duplicates and unfeasible MP [7]. Although this algorithm provides very good calculation results it is not used in this paper for the MP estimation, due to a higher calculation time for MP compared to the implemented MP-algorithm. This disadvantage is caused by the implementation in MATLAB and not by the algorithm itself. Further work has been done on the improved search for all MC in modified networks [16], which is useful if modifications on networks are performed in planning processes, reinforcement evaluation and network expansion.

Reference [6] deals with the authors current best known algorithm for the MC problem between all node pairs and between two special nodes. This algorithm has a time complexity $O(|V|\cdot 2^{|V|})$ for the MC problem between two special nodes and also all node pairs [6]. It is based on some simple intuitive theorems that characterize the structure of the MC using a node set and it can only be used for undirected graphs. This algorithm is easy to understand and to implement and it has the advantage that it can calculate all MC within reasonable time [6]. This algorithm has the disadvantage that it cannot deal with graphs, where the minimal node degree $\delta(V)$, except for the source node s and the sink node t, is equal to one. The edge adjacent to these nodes with node degree $d(x_1) = 1$ can be removed from the network in a preprocessing step, without losing any MC, before starting the MC estimation. Another disadvantage of the algorithm can be seen in the fact that it can not deal with self loops [6]. For the further investigations in this paper this algorithm is called Alg. A. Alg. A can also be used to check the accuracy of the other two implemented algorithms.

Alg. B [15] is based on MP and can therefore be utilized with directed and undirected graphs. It has an exponential worst time complexity in the number of minimal paths [10].

memory demand. The MP-algorithm in this paper is very simple. Firstly the algorithm estimates all paths from the source with a specified path length with a breadth first search approach. Since every connected graph G (V, E) contains at least one path of length min{2δ(V), |E|-1} [2], the specified path length is |E|-1 in the first step so that at least all edges are explored once. Secondly all paths with the sink as endpoint are chosen as MP. Once all MP are deduced an incidence matrix is constructed (|MP|x|E|). Columns that have all non-zero entries are MC of order one and must be eliminated from the incidence matrix. In the 2^{nd} step all combinations of two columns are compared to find columns with non-zero entries. This MC are cuts of the 2^{nd} order. To find higher order cuts the 2^{nd} step has to be repeated with combinations of 3, 4, 5, … columns and any cuts of lower order are eliminated [4]. Alg. B has the advantage that it can find MC up to a specified order. This reduces the computation time and Alg. B can deal with graphs having nodes with degree equal to one, due to the fact that they are ignored by the MP. Alg. B has the disadvantage that calculating all MP is simply impractical for special types of networks (e.g. Fig. 7 (11)), although a wide range of transmission power grids can be evaluated with this method [10], [15].

Alg. C, which is based on [3], is a new recursive algorithm that uses a merge approach and a novel intuitive graph reduction in a special preprocessing step to evaluate MC up to the 2^{nd} order. This algorithm combines the major advantages of Alg. A and Alg. B. Experimental results in [3] show that, this algorithm has a linear running time for different graph densities with a given number of nodes and an exponential running time with different numbers of nodes. The MC can be evaluated up to the 2^{nd} order without calculating the MP. The new designed algorithm impresses by a very high computation speed, even higher than Alg. B and it is not limited by the network size and structure, as it would be, if a MP approach would be used for the MC estimation.

IV. The New Proposed Algorithm Alg. C

A. *Network reduction*

To calculate the reliability of power transmission networks, due to the low probability of occurrence, MC of $o(MC) \geq 3$ are not incorporated. Thus a novel intuitive method is proposed to reduce the network graph to estimate first and second order MC. This method is fundamentally based on the following lemmas and theorem.

Lemma 1: The order of a MC is at least equal or greater three, provided that an edge $e_{x_1x_2}$ of an induced cycle is a member of the MC.

Proof: Consider a simple cycle appears with three nodes x_1, x_2, and x_3 and three edges $e_{x_1x_2}$, $e_{x_2x_3}$ and $e_{x_3x_1}$. Every MC between e.g. $s = x_1$ and $t = x_3$ in this simple cycle has $o(MC) = 2$ (see Fig. 1, left part). Consider a second simple cycle with three nodes x_4, x_2, and x_3 and three edges $e_{x_4x_2}$, $e_{x_2x_3}$ and $e_{x_3x_4}$. Now, with $s = x_1$ and $t = x_4$ those two MC ({ $e_{x_4x_2}$, $e_{x_1x_3}$ and $e_{x_2x_3}$ } and { $e_{x_1x_2}$, $e_{x_3x_4}$ and $e_{x_2x_3}$ }), that include the edge $e_{x_2x_3}$ have $o(MC) = 3$ and the other two MC have $o(MC) = 2$ (see Fig. 1, center part). Since every other connected cyclic network can be reduced to such kind of network with a merging approach (see Algorithm in [17]), it is obvious that the degree is equal or greater three. In other words this means every MC that cuts an edge of two chordless cycles has at least twice the minimal degree of a cut of a simple cycle minus one, which is equal or greater three. □

To consider the special case if s or t is a member of two or more induced cycles following Lemma 2 is defined.

Lemma 2: All edges touching 2-connected ($d(x_i) \leq 2$) nodes x_i, with $x_i \neq s \vee t$ and x_i is exclusively connected over 2-connected nodes with $s \vee t$ or adjacent to $s \vee t$, of a chordless cycle C, with $s \vee t \in C$, $x_i \in C$ and $d(s \vee t) = 2$, are excepted from the network reduction subroutine. If $d(s \vee t) = 1$ and there exists $x_{s,t}$, with $d(x_{s,t}) = 3$, where s or t and $x_{s,t}$ is connected over exclusively one P, then all edges touching 2-connected x_i or $x_{s,t}$, with x_i is exclusively connected over 2-connected nodes with $x_{s,t}$ or adjacent to $x_{s,t}$, of an induced cycle C, with $x_i \wedge x_{s,t} \in C$, are also excepted from the network reduction. For example see the single-connected node t in Fig. 7 (9), where all edges between to the two dashed line nodes and the before mentioned t have to be considered for MC with $o(MC) \leq 2$.

Since single-connected nodes are never a member of an induced cycle and therefore are not considered by the network reduction, the following important theorem derives from Lemma 1 and 2.

Theorem 1: Each edge $e_{x_2x_3}$ that is an edge of two different induced cycles C and C', and not an edge of Lemma 2, can be eliminated by merging node x_2 and x_3 into node x_3 without deleting any MC with $o(MC) \leq 2$ (see Fig.1 right part).

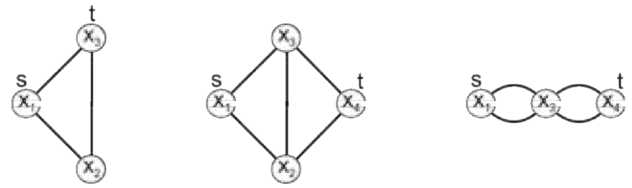

Fig. 1. An example network to explain the network reduction algorithm with two simple (chordless) cycles.

B. *Detailed description of the algorithm*

The algorithm consists of four main subroutines. The first subroutine deletes all nodes x_i from the graph with $d(x_i) < 3$ excluding node s and t and stores the information of deleted edges $e_{x_ix_k}$ and the belonging new edges. The second subroutine calculates all chordless cycles of the reduced graph. This is done with a shortest path approach based on the following proposition [2].

Proposition 1: Every G contains a P of length δ(G) and C of length at least δ(G) + 1 (provided that δ(G) ≥ 2)

Proof: The longest P in G is $x_0 \ldots x_k$. Then all the neighbours of x_k lie on P. Hence $k \geq d(x_k) \geq \delta(G)$. If $i < k$ is minimal with $e_{x_ix_k} \in E$, then $x_i \ldots x_kx_i$ is a C with length

$C \geq \delta(G) + 1$ [2]. □

The shortest paths are calculated with the well known algorithm of Dijkstra [18]. Afterwards the cycles are extended with the information of the deleted edges of subroutine one. Hence the original network is reconstructed and the gained information of chordless cycles can be used to generate a cycle-incidence matrix CI = (|E|x|C|).

In subroutine three the network is reduced with Theorem 1. Each node of an edge, that has a least two non-zero entries in the edge-row of CI, is merged into the graph in that way, that one of two nodes is replaced by the other one and the touching edge is deleted. Thus the original network is reduced to a graph that contains nodes with $d(x_i) > 3$ and nodes with $d(x_i) \leq 3$. The earlier group is also connected to one or more loops that may contain nodes or be a self-loop, see Fig. 4 Step 4. This reduced graph is thus analyzed with the extended algorithm of [3] in subroutine four.

The extended recursive merge algorithm is based on the obvious idea that s is prevented from arriving at t, if all edges emitting from s are deleted. If one of these edges is not deleted, then s has a MP to t since G is connected and these edges are a MC of G. The basic idea is to build a SS – adjacent nodes to s are merged one by one into this set –, where all edges emitting from this SS are a MC. Furthermore redundant nodes are also considered by merging this kind of nodes into SS before a MC is calculated.

Due to the fact, that the network reduction creates also redundant nodes (loops) that are not relevant for the MC estimation, all self loops are eliminated. To verify that this algorithm enumerates all first and second order cuts correct, consider Lemma 3 and 4.

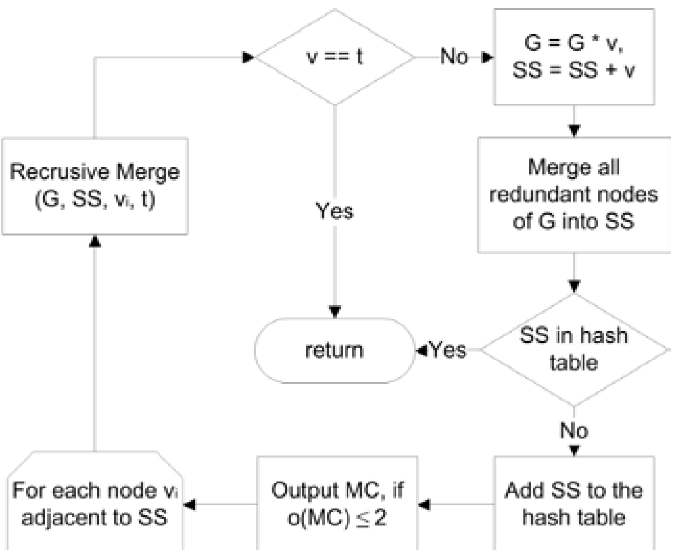

Fig. 2. Flowchart of subroutine four; the extended recursive merge algorithm based on [3].

Lemma 3: Both, the network reduction and the extended Alg. C do not produce isolated nodes in any subroutine.

Proof: In the network reduction subroutine, which is the only subroutine in the pre-processing steps that eliminates edges and nodes, only nodes are merged together. Since all nodes are connected and the merging-process does not disconnect nodes, no isolated nodes occur (see Lemma 1 in [3]). □

Lemma 4: Any first and second order MC generated by Alg. C can also be produced by the Algorithm of [3].

Proof: Since the Algorithm of [3], which itself is verified on accuracy by the algorithm in [17], is except two special requests concerning the order of MC and redundant loops, exactly the same as Alg. C, and since edges of a MC with o(MC) = 1 are not eliminated by the network reduction, Alg. C evaluates all MC correct up to the 2^{nd} order. □

In Fig. 2 the flowchart of the extended recursive merge Alg. C is depicted.

C. An Example

A moderate size network is chosen to demonstrate the methodology of the new proposed algorithm. This simple network is evaluated with the algorithm to find the MC between s and t. Not all steps are depicted detailed. Those steps that eliminate more than one edge are a series of single edge deleting and node replacing processes.

Consider the initial network G in Fig. 4 Step 0. After deleting all nodes x_i with degree $d(x_i) < 3$, the graph can be displayed with Fig. 4 Step 1. The edge-deleting-subroutine is necessary to estimate all induced cycles with the shortest path approach. This approach is based on the idea (see Proposition 1), that the shortest path between node e.g. 1 and 2, if edge (1) is deleted, includes edge (2) and (3) (see Fig. 7 (7)). This operation is repeated until all |E| - |V| + 1 induced cycles are found. CI of the example is shown in Fig. 3.

$CI|_{|E|\times|C|} =$

	I	II	III	IV	V	VI
(1)	1	0	0	0	0	0
(2)	1	0	0	0	0	0
(3)	1	1	0	0	0	0
(4)	0	1	1	0	0	0
(5)	0	1	0	0	0	0
(6)	0	1	1	0	0	0
(7)	0	0	1	1	0	0
(8)	0	0	1	1	0	0
(9)	0	0	1	0	0	0
(10)	0	0	0	1	1	0
(11)	0	0	0	1	0	0
(12)	0	0	0	0	1	0
(13)	0	0	0	1	0	0
(14)	0	0	0	1	1	0
(15)	0	0	0	1	0	1
(16)	0	0	0	0	1	0
(17)	0	0	0	0	1	1
(18)	0	0	0	0	0	1
(19)	0	0	0	0	1	1
(20)	0	0	0	0	0	1
(21)	0	0	0	0	1	0
(22)	0	0	0	0	1	1
(23)	0	0	0	0	0	1
(24)	0	0	0	0	0	1
(25)	0	0	0	0	0	1
(26)	0	0	0	0	0	1

Fig. 3. The CI of the example with 26 nodes (first column) and 6 induced cycles (upper row). The deleted edges in subroutine three are labeled bold.

In the next subroutine the graph is reduced as schematically depicted in Fig. 4 Step 3.1 to Step 3.6. Step 3.0 in Fig. 4 is the initial state. The algorithm in subroutine three firstly replaces node 2 by node 3 and eliminates the edge (3). Next node 3 is replaced by node 5 and the edges (4) and (6) are deleted as well as node 4, which is not depicted in Fig. 4. This node is the offspring of edge (4) and (6) and is therefore replaced by node 5 (Theorem 1).

With this operation, a self-loop of node 5 with edge (5) is created. It is obvious, that this self-loop is never a member of a MC due to the definition of MC. In Step 3.3 node 5 is replaced by node 7 and edges (7) and (8) are eliminated. In this step the second self-loop at node 7 with edge (9) appears. Step 3.4 shows how node 7 is replaced by the adjacent node 11 and edges (10) and (14) are deleted. The following Step 3.5 merges node 12 into node 11. This creates a new graph without edge (15) and a third loop with node 10 (see Step 4 in Fig 2.) and edges (11) and (13). As in Step 3.2 it is obvious that the edges of this loop are never a member of a MC because they are not on a path from s to t. Step 3.6 finishes the network reduction subroutine three by merging node 18 into node 12 and eliminating edges (17), (19) and (22). Subroutine four (for results see Step 4 in Fig. 4) operates recursive (see Fig. 2 and [3]). The main difference between the algorithm in [3] and its extended version used here is that the extended version in subroutine four can deal with self-loops. Redundant loops, as the loop depicted in Fig. 4 Step 4 with nodes 18-8-13-18, are merged into SS in one step and not one by one as in [3] before creating a new MC. For example consider node 8 in Fig. 4 Step 4. If SS = {1, 18}, node 8 is adjacent to SS and on a redundant loop. The touching edges are never a MC for the connection between s and t, respectively node 1 and 21. Therefore G * 8 and G * 13 in one step, without creating a MC. Both Alg. B and Alg. C provide the same results for the graph example: six 2nd order MC MC 1 … MC 6 with the edges (see Fig. 4 Step 4) {(1), (2)}, {(25), (18)}, {(25), (20)}, {(25), (23)}, {(25), (24)} and {(25), (26)}.

V. Benchmark of Three Different Algorithms

A. Processing time of Alg. C.

The new proposed algorithm was tested with randomly generated planar cyclic undirected grid-graphs G with fixed numbers of nodes. The time to enumerate all MC was measured and Fig. 5 depicts the results of running time versus graph density, which is defined as $2\cdot|E| / (|V|\cdot(|V|-1))$. The number of nodes ranges from 24 to 40 at a step size of 4 and the graph density ranges from 0.1 to the maximal size of planarity at a step size of 0.005. The maximal size of the planar graphs was calculated with [20]. The minimal size results from the constraint that the graph has to be connected

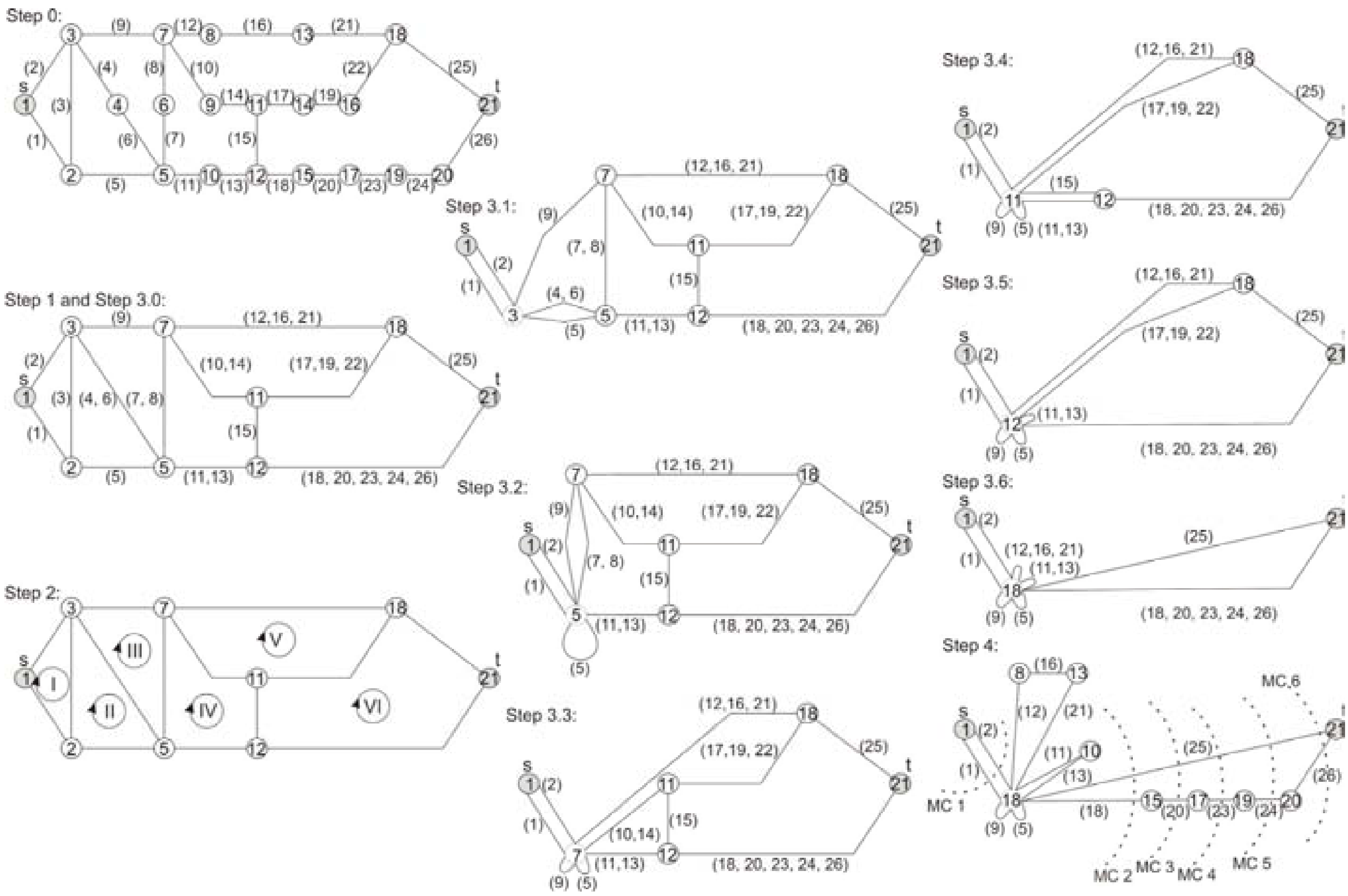

Fig. 4. Example network with source s und sink t from [3]; Step 0: Initial network; Step 1: result after subroutine 1; Step 2: result after subroutine 3; Step 3 and Step 4: The different steps of the network reduction subroutine three and the six second order MC (see doted curves in the figure with title Step 4), when the extended Algorithm of subroutine four (Alg. C) is applied.

without any intersections. This is guaranteed for the considered range with a graph density exceeding 0.1. To test either the randomly generated graph is planar or not algorithm form [20], [21] was used.

Due to the stochastic behavior concerning the number of, the induced cycles of the random graphs, the induced cycle depended computation time versus the graph density was fitted with a robust fit regression. This regression mode is used, if outliers have to be considered with lower importance, which is appropriate in this case. The results of this regression for the considered random networks are also depicted in Fig. 5. It is evident that there is a linear relationship between the graph density and the computation time for the 2-terminal MC. Furthermore the computation time for the 2-terminal MC per fixed number of nodes is increasing with the graph density. This gradient of the linear function is unequal for different number of nodes which is already pointed out in [3].

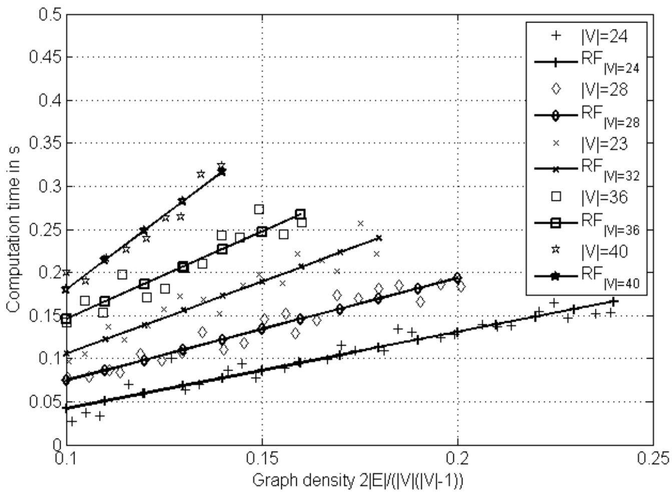

Fig. 5. The computation time versus graph density 2·|E| / (|V|·(|V|-1)) with fixed number of nodes. RF … Robust Fit regression for the relevant range of planar cyclic undirected networks

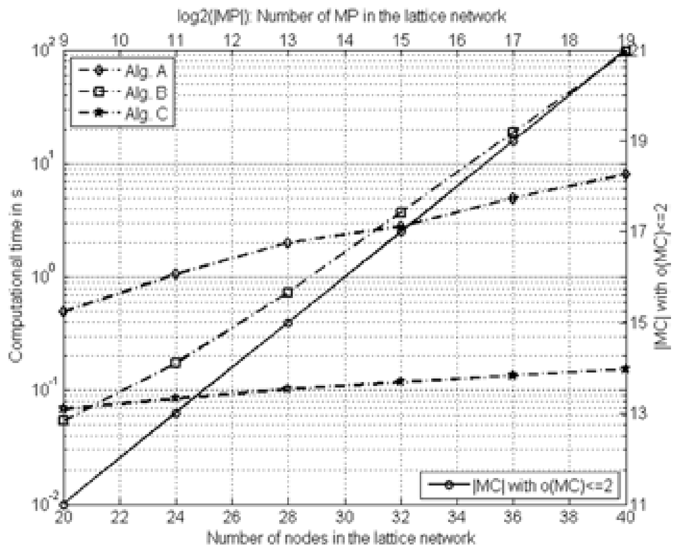

Fig. 6. The computation time for 2-terminal MC versus the increasing node number of lattice networks is shown for all three used algorithms. Additional, in the figure the number of 2nd order MC and the number of MP is depicted in the left y-axis and upper y-axis respectively.

In Fig. 6, six lattice networks (for example see Fig. 7 (11)), ranging from 20 nodes to 40 nodes in steps of 4 nodes are compared concerning the computation time. Alg. A, Alg. B and Alg. C are used to show the exponential computation time behaviour of the problem. All three algorithms were implemented in MATLAB on an Intel® Core™2 Duo CPU with 3.33 GHz.

It's worthless noting, that Alg. C is the fastest algorithm of all three presented ones, if the network exceeds a certain size. Alg. A, which enumerates all MC is the second best algorithm in terms of computation time, if lattice networks with more than 32 nodes are considered. Due to the huge computational effort to calculate the MP, e.g. the lattice network with 32 nodes has $2^{15} = 32768$ MP, Alg. B is the slowest for lattice networks with more than 32 nodes, even if MC up to the 2nd order are estimated.

B. Benchmark with test grids

The intermeshing degree ν in Table I is defined as |E| / |V|. In power networks the intermeshing degree is normally between 1 and 2, in power transmission networks, which are focussed in this paper, it is between 1 and 1.5 [19]. Depending on the number of nodes, for example the before mentioned range for power transmission network ranges around a graph density of 0.1 for a network with 30 nodes.

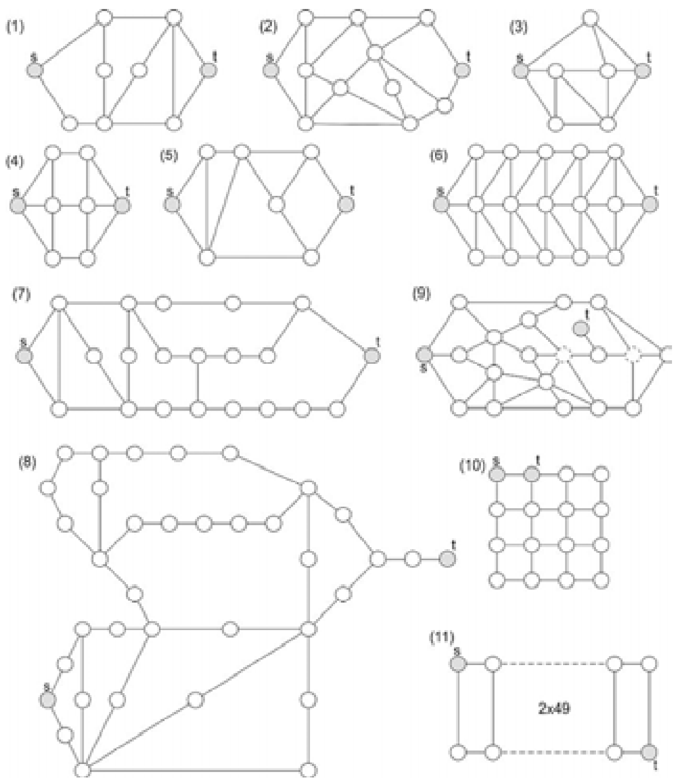

Fig. 7. Eleven planar Benchmark-networks with source node s and sink node t based on [3], [14]; Network (7) is used for the example to explain Alg. C. Network (9) is used for the example in Lemma 2

In Table I the normalized computation time, related to the respective fastest algorithm based on the best value out of ten calculations, for each benchmark-network in Fig. 7 is presented. The faster one of the two compared has a 1.00

TABLE I
NETWORK INFORMATION TABLE AND RESULTS OF THE BENCHMARK

Network	Intermeshing degree ν	1^{st} order Cuts	2^{nd} order Cuts	3^{rd} order Cuts	Unavail-ability [a]	Normalized computation time Alg. A vs. Alg. B [b]		Normalized computation time Alg. B vs. Alg. C	
${}^{12}_{9}G(1)^{13}_{28}$	1.33	0	3	7	$3.43 \cdot 10^{-8}$	1.98	1.00	1.00	4.24
${}^{20}_{12}G(2)^{150}_{128}$	1.66	0	2	3	$2.28 \cdot 10^{-8}$	2.42	1.00	1.00	2.18
${}^{12}_{7}G(3)^{25}_{20}$	1.74	0	0	2	$2.61 \cdot 10^{-14}$	1.02	1.00	1.00	4.14
${}^{13}_{8}G(4)^{29}_{29}$	1.63	0	0	3	$3.91 \cdot 10^{-14}$	1.43	1.00	1.00	3.35
${}^{12}_{8}G(5)^{24}_{19}$	1.50	0	2	5	$2.28 \cdot 10^{-8}$	1.13	1.00	1.00	3.62
${}^{36}_{17}G(6)^{22401}_{560}$	2.12	0	0	2	$2.61 \cdot 10^{-14}$	1.00	3.11	19.44	1.00
${}^{26}_{21}G(7)^{44}_{528}$	1.24	0	6	66	$6.85 \cdot 10^{-8}$	13.50	1.00	1.00	1.76
${}^{43}_{36}G(8)^{56}_{2666}$	1.19	2	8	16	0.02	39.77	1.00	1.00	1.08
${}^{35}_{20}G(9)^{4008}_{2545}$	1.75	1	1	2	0.01	8.80	1.00	4.66	1.00
${}^{24}_{16}G(10)^{98}_{105}$	1.50	0	1	2	$1.14 \cdot 10^{-8}$	1.00	3.11	1.00	2.28
${}^{148}_{100}G(11)^{2^{49}}_{2500}$	1.48	0	51	98	$5.82 \cdot 10^{-7}$	1.00	*	*	1.00

${}^{e}_{v}G(z)^{p}_{c}$ Benchmark-network number z with v nodes, e edges, p MP and c MC.

* Due to memory problems no calculation could be performed.

[a] In h / year; it is the same if all MC up to the highest order are considered or if MC up to the 2^{nd} order are considered, $\lambda = 0.01$ 1 / year and $\mu = 1$ h.

[b] The normalized computation time, if Alg. B evaluates all MC up to the 3^{rd} order and Alg. A evaluates all MC.

entry in the table. Alg. B considers MC up to the 3^{rd} order for comparison with Alg. A and up to the 2^{nd} order for comparison with Alg. C (see Table I). Alg. C determines all MC up to the 2^{nd} order. As a result of the investigation it can be pointed out, that in the case of an intermeshing degree v exceeding 2, the Alg. A is better for the MC calculation from the point of view of computation time compared to Alg. B if 3^{rd} order cuts or higher are analyzed. If Alg. B for 1^{st} and 2^{nd} order MC is compared to Alg. C, it can be seen, that for all sample graphs except graph (6), (9) and (11), Alg. B has lower computational time. Alg. C is the fastest for these graphs that have a high number of MP, which is increasing with the number of nodes and branches, respectively with the graph density. The system unavailability in Table I is calculated with the frequency and duration approach [4], [5]. The failure rate λ of all components is 0.01 1 / year and the repair rate μ of all components is 1 h. The results in Table I show that only 1^{st} order cuts have a major impact in the reliability evaluation if λ and μ are considered.

C. Computation of all MC

With the network in Fig. 8, that is more representative for a real world power transmission network, the benefit of the new proposed algorithm is shown when all MC up to the 2^{nd} order of a network have to be calculated.

The network in Fig. 8 is from [14] and has 57 nodes and 78 edges. It consists of five sources, 52 loads, 710 2^{nd} order MC and five 1^{st} order MC. The network has an intermeshing degree of 1.37, a graph density of 0.049 and 11497979 MP. All MC up to the 2^{nd} order are evaluated with Alg. B and Alg. C. Alg. B needs 5242.1 seconds and Alg. C needs 198.7 seconds. This is a saving in computation time of more than 5043.3 seconds or in other terms, Alg. C needs only 3.8 % of the computation time of Alg. B.

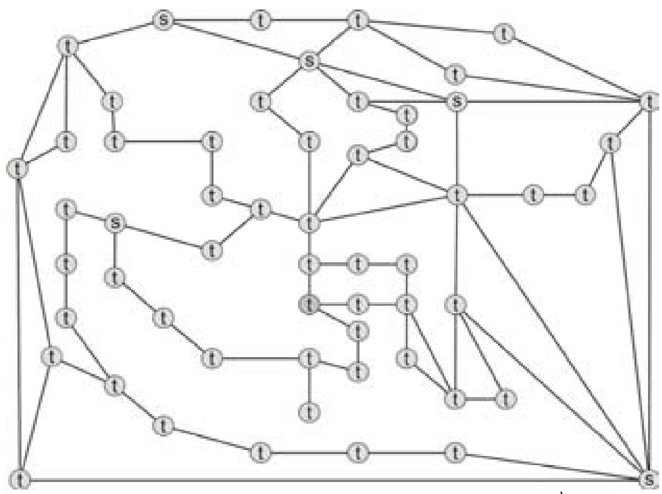

Fig. 8. Test network for the computation of all MC up to the 2^{nd} order based on [14] with 57 nodes (five sources and 52 loads).

VI. CONCLUSION

It is sufficient to calculate MC up to the 2^{nd} order for 2-terminal power system transmission reliability analyses, especially in the context with transmission grid reliability optimisation with Genetic Algorithms, where hundreds and thousands of these operations have to be done. Since the estimation of reliability in transmission grids is NP-hard, research interest focuses on fast MC algorithms to reduce the computation time to a minimum. The proposed algorithm, which is based on a novel intuitive network reduction and on a recursive merge approach, calculates all MC up to the 2^{nd} order in satisfying time and even faster as the currently best know algorithm. Depending on the number of nodes respectively on the graph density, it is also faster as a simple, but for small networks powerful, MP algorithm that generates all MC up to the 2^{nd} order. Furthermore with a 57 node test network it is demonstrated that the new proposed algorithm can calculate all MC up to the 2nd order. In terms of computation time this is a reduction by 96.2 % compared to the MP algorithm.

VII. References

[1] W.-C. Yeh, "Search for all MCs in networks with unreliable nodes and arcs," *Journal of Reliability Engineering and System Safety 79* (2003), pp. 95-101.

[2] R. Diestel, "*Graph Theory*," Electronic Edition 2000. New York: Springer, 2000, pp. 4-166.

[3] H.-Y. Lin, S.-Y. Kuo, and F.-M. Yeh, "Minimal Cutset Enumeration and Network Reliability Evaluation by Recursive Merge and BDD," 2003, in *Proc. of the Eighth IEEE International Symposium on Computers and Communication (ISCC'03)*, 6 pages.

[4] R. Billinton, and R. N. Allan, "*Reliability Evaluation of Engineering Systems, Concepts and Techniques*," 2nd ed., New York: Plenum Press, 1992.

[5] R. Billinton, and R. N. Allan, "*Reliability Evaluation of Power Systems*," 2nd ed., New York: Plenum Press, 1996.

[6] W.-C. Yeh, "A simple algorithm to search for all MCs in networks," *European Journal of operational Research* 174 (2006), pp. 1694-1705

[7] W.-C. Yeh, "A Simple Heuristic Algorithm for Generating All Minimal Paths," *IEEE Trans. On Reliability*, vol. 56, No. 3, pp. 488-494, September 2007.

[8] S. Soh, and S. Rai, "Experimental Results on Preprocessing of Path/Cut Terms in Sum of Disjoint Products Technique," *IEEE Trans. On Reliability*, vol. 42, No. 1, pp. 488-494, March 1993.

[9] M. Fotuhi-Firuzabad, R. Billinton, T.S. Munian, and B. Vinayagam, "A Novel Approach to Determine Minimal Tie-Sets of Complex Network," *IEEE Trans. On Reliability*, vol. 53, No. 1, pp. 61-70, March 2004.

[10] D. R. Shier and D. E. Whited, "Iterative Algorithms for Generating Minimal Cutsets in Directed Graphs," *Networks*, Vol. 16, (1986), pp.133-147.

[11] M. O. Ball, "Computational Complexity of Network Reliability Analysis: An Overview," *IEEE Trans. On Reliability*, vol. R-35, No. 3, pp. 230-239, August 1986.

[12] H.-J. Haubrich, "*Zuverlässigkeitsberechnung von Verteilungsnetzen, Grundlagen – Verfahren – Anwendungen*, " ABEV Band 36, Verlag der Augustinus Buchhandlung, 1996.

[13] A. Gaun, H. Renner, and G. Rechberger, "A reliability and environmental based Multi-Criteria Decision Making genetic algorithm (MCDMGA)," in 6th *Power Quality and Supply Reliability Conference, 2008. PQ 2008,* pp. 217-223, Aug 2008.

[14] A. R. Abdelaziz, "A New Approach for Enumeration Minimal Cut-sets in a Network," 2000, in *Proc. of 7th IEEE International Conference on Electronics, Circiuts and Systems, (ICECS 2000)*, pp. 693-696.

[15] R. N. Allan, R. Billinton, and M. F. De Oliveira, "An Efficient Algorithm for Deducing the Minimal Cuts and Reliability Indices of a General Network Configuration," *IEEE Trans. On Reliability*, vol. R-25, No. 4, pp. 226-233, October 1976.

[16] W.-C. Yeh, " An improved algorithm for searching all minimal cuts in modified networks," *Reliability Engineering & System Safety*, vol. 93, Issue. 7, pp.1018-1024, July 2008.

[17] S. Tsukiyama, I. Shirakawa, and H. Ozaki, "An Algorithm to Enumerate All Cutsets of a Graph in Linear Time Per Cutset," *Journal of A.C.M., vol. 27,* No. 4, pp. 619-632, October 1980.

[18] K. P. Murphy, "The Bayes Net Toolbox for Matlab," Department of Computer Science, University of California, Berkeley. Available: http://HTTP.CS.Berkeley.EDU/~murphyk/Papers/bnt.ps.gz, October 2001.

[19] D. Oeding, and B. R. Oswald, "*Elektrische Kraftwerke und Netze*," 6. Auflage, Springer-Verlag Berlin Heidelberg New York, 2004.

[20] D. Gleich, "MatlabBGL A Matlab Graph Library", Institute for Computational and Mathematical Engineering, Stanford University. Available: http://www.stanford.edu/~dgleich/programs/matlab_bgl/, October 2008.

[21] J. M. Boyer, and W: J: Myrvold, "On the Cutting Edge: Simplified O(n) Planarity by Edge Addition", *Journal of Graph Algorithms and Applications, vol. 8,* No. 3, pp. 241-273, October 2004.

VIII. Biographies

Alexander Gaun was born in Kufstein, Austria, in 1979. He received his diploma degree in 2005 at Graz University of Technology, Graz. He is currently a scientific assistant at the Institute of Electrical Power Systems. His research interests include electrical power system planning, reliability computation, genetic algorithms and electromagnetic compatibility.

Herwig Renner was born in Graz, Austria, in 1965. He received his doctoral degree in 1995 at Graz University of Technology, Graz, where he currently holds a position as associate professor at the Institute of Electrical Power Systems. His main work in research and teaching is in the field of electrical power system analyses with special emphasis on quality, supply reliability and power system control and stability. In 2007 he was employed as visiting professor at Helsinki University of Technology.

Georg Rechberger was born in Linz, Austria, in 1978. He received his diploma degree in 2005 at Graz University of Technology, Graz, where he currently holds a position as scientific assistant at the Institute of Electrical Power Systems. His main work in research and teaching is in the field of electrical power system planning.

7.2 Annex B: Reference P2

Paper printed in conference proceedings for 2009 IEEE Bucharest Power Tech Conference, June 28th - July 2nd, 2009, Bucharest, Romania

Systematical Determination of Load Flow Cases for Power System Planning

G. Rechberger, H. Renner, A. Gaun

***Abstract*-- This paper presents a novel method to determine a minimum number of representative load flow cases with load and generation data in a meshed transmission grid, which cover all critical power line loading situations in a given time frame. Measured active and reactive line loading data obtained from SCADA are the basic data input for this method. In order to find the optimum number of load flow cases a heuristic search algorithm in combination with a genetic optimization algorithm is used. The comparison of the required load flow cases with conventional load situations e.g. peak load or peak generation show, that conventional load flow cases do not cover all critical power line loading situations. Load flow cases determined with the presented method give an objective and representative picture of the expected situations and can be used for power system planning in the view of increasing demand and new generation capacity installations.**

***Index Terms*-- load flow case optimization – power flow measurement – power system planning – genetic algorithm – set of load flow cases**

I. Introduction

THE measured time series of reactive and active power, usually available in 15 min resolution, concerning consumption, generation and loading of lines provide a huge amount of data representing the state of an electrical power system. Based on this data set, a single load flow case for each time interval can be extracted. These load flow cases, consisting of load and generation values, represent a base for load flow analyses considering future demand increase, installation of new power plants and transmission network expansions for future situations.

In [2] peak load and peak generation cases for exemplary summer and winter snapshots are used as the planning bases. Other methods apply typical snap shot data sets for the planning purposes [6]. These situations are not necessarily representative for the maximal loading of each power line, especially in case of systems with high wind penetration, pump storage power plants or extensive cross boarder energy trading. In [7] a random set of load flow cases from a one year set of measurements with hourly resolution is recommended. The number of load flow cases in the dataset makes this method highly computationally inefficient, requiring considerable effort concerning handling and utilization. The usage of selected load flow cases, with respect to maximal loading of a specific power line, provides a solution for this problem.

In this paper, a novel methodology for finding a minimum set of load flow cases based on active and reactive power flow measurements by the SCADA system is presented. The most straightforward method determines a set of load flow cases containing specific load flow cases for every worst-case bi-directional power flow scenario of a single power line. This set is minimized by a combination of a heuristic search algorithm and an optimization algorithm and covers all relevant critical loading situations of power lines.

G. Rechberger is with the Institute of Electrical Power Systems, Graz, University of Technology, A-8010 Graz (e-mail: georg.rechberger@tugraz.at)

H. Renner is with the Institute of Electrical Power Systems, Graz University of Technology, A-8010 Graz (e-mail: herwig.renner@tugraz.at)

A. Gaun is with Institute of Electrical Power Systems, Graz University of Technology, A-8010 Graz (e-mail: alexander.gaun@tugraz.at)

II. Method

A. Data Bases

The introduced method is based on the measured apparent power $\mathbf{S}_{meas}$ with the elements $S_{meas,t,i}$ with the columns i, depending on the number of lines L in the investigated grid area and T rows for each single time step (index t) of the interested period T.

With respect to a worst case approach to the bi-directional power flows on the power lines the information about the load flow direction in the matrix S is necessary. For example, a planned power plant can cause additional line power flows in various directions, which have to be added or subtracted from the actual values of a load flow case. Since the apparent power $\boldsymbol{S}_{meas}$ has no information about the load flow direction, a direction (sign) of the active power is assigned to the matrix $\mathbf{S}_{meas}$ which results in the matrix $\boldsymbol{S}$ as follows:

$$S_{t,i} = S_{meas,t,i} \cdot sign\left(P_{t,i}\right) \quad (1)$$

where $P_{t,i}$ are the measured active power flows on the power lines.

With the given maximal allowed currents by the vector $\mathbf{I}_{lim}$ and the elements $I_{lim,i}$ of each power line, the conventional thermal transmission capacities $S_{lim,i}$ and the loadings $s_{t,i}$ in p.u. can be calculated (2).

Here is assumed that the actual voltage does not considerably deviate from the rated voltages U_i in the vector $\mathbf{U}$ where it is also possible to use the actual voltage at the maximal measured apparent power.

In the following line loadings are expressed in p.u. according to (2) and named with lower case characters.

$$S_{\lim,i} = \sqrt{3} \cdot U_i \cdot \mathrm{I}_{\lim,i} \quad \text{and} \quad s_{t,i} = \frac{S_{t,i}}{S_{\lim,i}} \tag{2}$$

The separate measurement for each circuit of double circuit lines with the same start and end node for each circuit can be reduced to a single value.

The data corresponding to any irregular states of the grid e.g. failure, planned outage or special grid configurations due to congestion management, have to be considered separately. These irregular states require separate corresponding grid states for further load flow calculations.

B. Defined minimum percentage of analyzed load flows

When analyzing the network, it is useful to take into account only power lines with a maximal loading above a specified loading limit s_{base}. With s_{base} = 20 % of their installed capacity all lines in the grid with peak loading values exceeding 20 % are taken into account.

With s_{base} = 60 % only lines with peak loading over 60 % of the maximal transmission capacity of this line are evaluated. If an increase of loading, caused by load increase or new power plants, a s_{base} = 60 % is not useful because in meshed grids loading over 60 % are often not within the (n-1) security criteria [6]. Choosing s_{base} should regard the expected loading increase within the actual planning period. To estimate the loading increase of single power lines load flow calculations or the calculation of power distribution transfer factors PDTF [8] have been proposed.

The matrix $S_{t,i}$ (1) can be thus reduced by eliminating the 'non critical' power lines with maximal loading below the predefined percentage s_{base} to the vectors $\mathbf{s}_{\max}$ and $\mathbf{s}_{\min}$ with the elements $s_{\max,i}$ and $s_{\min,i}$ by the following equations:

$$\begin{aligned} s_{\max,i} &= s_{t,i} \forall \max_t(s_{t,i}) \geq s_{base} \\ s_{\min,i} &= s_{t,i} \forall \min_t(s_{t,i}) \leq -s_{base} \end{aligned} \tag{3}$$

Every line with loading situation exceeding s_{base} or below $-s_{base}$ is classified in further analyses as critical situation.

C. Critical range

Depending on the maximum in both load flow directions of the apparent power ($s_{\max,i}$ and $s_{\min,i}$) on every power line a critical range ε expressed as percentage of the loading limit in each direction can be defined (see (4)). Each load flow case with a line loading within the critical range ε is assumed to a representative peak loading case for this line. The critical range is described with the lower and the upper apparent power $s_{lower,i}$ and $s_{upper,i}$.

With ε = 1 % nearly all peak loading situations within the time period of measurement are obtained, with ε = 20 % deviations up to 20 % from the selected load flow cases to the peak load situations are expected. Another possibility is to derive the critical range from the duration curve for the loading of each line. For further proceeding the definition for the critical range in (4) is used.

In Fig. 1 an example of load and duration line of the load flow on a high voltage power line illustrates the critical range. In this example the critical range ε is defined with ε = 20 %.

$$\begin{aligned} s_{upper,i} &= s_{\max,i} - s_{\lim,i} \cdot \frac{\varepsilon}{100} \\ s_{lower,i} &= s_{\min,i} + s_{\lim,i} \cdot \frac{\varepsilon}{100} \end{aligned} \tag{4}$$

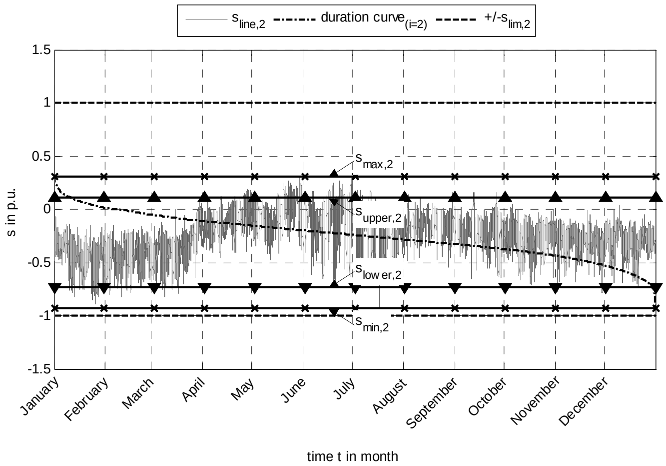

Fig. 1 Power transmission line (i = 2) apparent power $s_{line,2}$ in p.u., duration curve, lower and upper critical area with ε = 20 %.

A general survey of the general load flow situation for 10 typical lines chosen from the studied dataset in the interesting period is given in Fig. 2. This kind of diagram of the load flow situation used also in the following shows only the first 10 of the 38 considered power lines for the sake of a clear arrangement. This figure shows the measured line loadings $s_{line,t,i}$ including the maximal and minimal values $s_{line,i,max}$ and $s_{line,i,min}$. The critical range is between $s_{upper,i}$ and $s_{line,i,max}$ for one load flow direction and between $s_{lower,i}$ and $s_{line,i,min}$ for the opposite load flow direction. The dotted lines mark the s_{base} percentage.

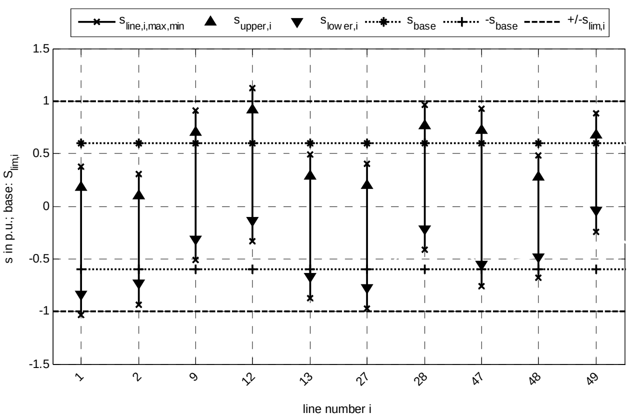

Fig. 2 Load flow situation of 10 power lines in a meshed grid (critical range for ε = 20 %, s_{base} = 60 %).

All load flow values $S_{\max,t,i}$ and $S_{\min,t,i}$ of the remaining lines are tested whether they represent a peak loading situation in accordance to the above definition of the critical range for one or more lines. The results of the mentioned comparison are the logic suitability matrices $\mathbf{C}_{max}$ and $\mathbf{C}_{min}$ with entries for every time step and every line.

$$C_{\min,t,i} = \begin{cases} 1, & \forall s_{\min,t,i} < s_{i,lower} \\ 0, & \forall s_{\min,t,i} < s_{i,lower} \end{cases} \tag{5}$$

$$C_{\max,t,i} = \begin{cases} 1, & \forall s_{\max,t,i} > s_{i,upper} \\ 0, & \forall s_{\max,t,i} > s_{i,upper} \end{cases} \tag{6}$$

The matrix $\mathbf{C}_{max}$ for one load flow direction merged with $\mathbf{C}_{min}$ for the other load flow direction merged to the matrix $\mathbf{C}_0$ as input data for further optimization.

$$\mathbf{C}_0 = \left[\mathbf{C}_{\max} \quad \mathbf{C}_{\min}\right] \tag{7}$$

Each row in matrix $\mathbf{C}_0$ represents a single load flow case and each column a load flow direction of a single line with peak loadings exceeding s_{base}. A logical 1 as matrix element declares the according load flow case as peak loading case for this line. The optimization problem is to find a minimum set of load flow cases, covering all peak loading situations in each direction. This selection can be done as described below.

D. Minimizing the data set

To reduce computational time, the matrix $\mathbf{C}_0$ is minimized before further steps to the matrix $\mathbf{C}$. Firstly all rows of the matrix $\mathbf{C}_0$, with the same characteristic, identified by all none zeros entries at the same position, can be reduced to a single row. The row with the maximal mean loading of this set with same characteristic will be selected for further calculations. Secondly all rows with only zero entries representing non critical load flow cases also have to be eliminated to reduce unnecessary calculations. This remaining set of load flow cases represented by the matrix $\mathbf{C}$ has at least one critical loaded line for each remaining time interval.

E. Selection of load flow cases

The trivial solution is to use two separated load flow cases for each line selected with the criteria e.g. maximal loading in both directions. In this case the number of required load flow cases N_{LC} is equal to $N_{LC} = 2 \cdot L$. This builds a complete worst case set of load flow cases. The following method allows the reduction of required load flow cases to minimize computation time on the one hand and improve the handling of complex planning calculations on the other hand.

In step 1 the sum of the non-zero entries of each column (sc) of $\mathbf{C}$ is calculated. If $\mathbf{C}$ includes a column with only a single non-zero entry, this row respectively this load flow case is selected because this peak loading situation for this line is not contained in any other situation. This row represents the first selected load flow case with the time $t_{LC,k=1} = t_1$ (see (8)) and the first generated solution vector $\mathbf{C}_{comb} = \mathbf{C}_{t1}$. In step 2 the matrix $\mathbf{C}$ will be reduced by excluding the row $\mathbf{C}_{tk}$ and all columns with none zero entries of $\mathbf{C}_{tk}$.

Step 1 and step 2 will be repeated as long as there are single none zeros entries in the columns of the new reduced matrix $\mathbf{C}$.

$$t_k = t\left[first\left(\underset{t=1\ldots T,\,i}{sum}(\mathbf{C}) == 1 \right) \right] \tag{8}$$

If no more single critical situation is found, in step 3 the sum of the non-zero entries of each row (sr) of $\mathbf{C}$ is calculated. The next selected load flow case with the time t_k (see (9)) is the row with the maximal sum over the rows. Now the matrix C will be reduced again as pointed out in step 2.

$$t_k = t\left[\underset{first}{\max}\left(\underset{t,\,i=1\ldots N}{sum}(\mathbf{C}) \right) \right] \tag{9}$$

Thus it is possible to find a new time t_k and the corresponding load flow case $\mathbf{C}_{tk}$ with step 1 or step 3. The new solution vector $\mathbf{C}_{comb}$ is the result of the combination of the old solution vector $\mathbf{C}_{comb}$ and $\mathbf{C}_{tk}$. The OR disjunction of the two rows $\mathbf{C}_{comb}$ and $\mathbf{C}_{tk}$ result in a new solution vector $\mathbf{C}_{comb}$.

Step 1, 2 and 3 will be repeated as long as the solution vector $\mathbf{C}_{comb}$ is fully occupied. Now the time vector $\mathbf{t}_{LC}$ has an entry for each required load flow case.

$$\mathbf{t}_{LC} = \left[t_1 \; \ldots \; t_k \right] \tag{10}$$

The described method to determine a set of load flow cases which are required to represent all critical situations in a grid and a given period is recapitulated in the pseudo code diagram shown in Fig. 1.

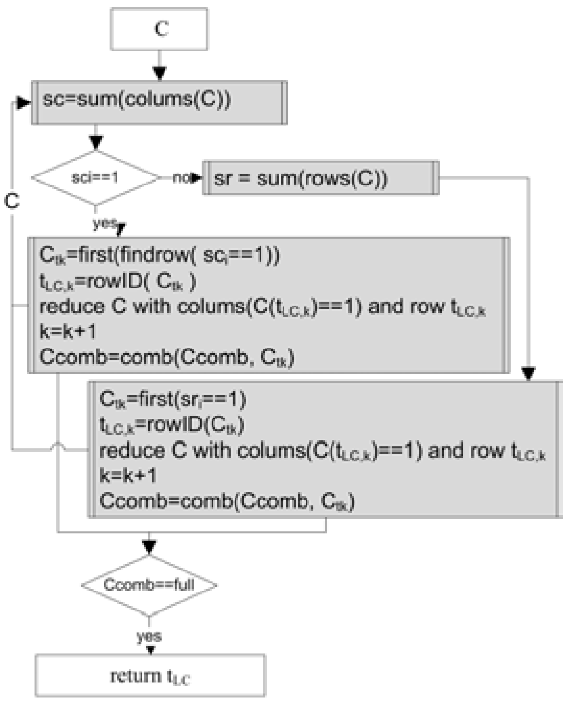

Fig. 1 Selection of load flow cases, search algorithm – pseudo code flow chart.

With this heuristic approach of selecting load flow cases, it is possible to find a minimum number of load flow cases required to express a maximum loading situation within the critical area for each line. Although this number of load flow cases is a local minimum but it is not guaranteed to be a global minimum. For small grids and few measured load flow cases, it is possible to check every combination of the rows of **C** and find the optimal solution, but with measurement of e.g. one year and ¼-hour resolution values it is not possible to solve this problem efficiently. A further problem is to select a load flow case in step 1 if there are more columns with single critical situations or in step 3 if the summation results in more then one maximum with the same value. In both cases this method selects the first detected entry.

F. Genetic algorithm

To improve the heuristic approach which mostly results in a local minimum number of load flow cases a genetic algorithm [1] is applied. In case of two or more load flow cases, represented by rows of the matrix C, with the same number of critical situations characterized by a non-zero entry at each respective positions we get a problem selecting a load flow case with the target to find a minimum number of required load flow cases. The heuristic approach, described above, will select always the first load flow case of the set of cases with the same maximal number of critical situations.

The implemented genetic algorithm includes the main variables Individual, Population, Parents, Children, Crossover rate, Permutation rate, Mutation rate, Number of iterations and Population size. These variables are used in the genetic algorithm consisting of the fitness function, the selection (tow round tournament) and the reproduction (crossover, permutation and mutation).

An individual as part of a population is the modified order of the time T. The start population and the following populations consist of single individuals with random respectively by the genetic algorithm modified orders of the time T. The improved selection of load flow cases starts with sorting the rows of C in the order of the individuals of the population vector in step 1.

The number of individuals of populations describes the population size (popsize).

The fitness is the number of required load flow cases, resulting from the load flow cases selection. This is commonly known as fitness function.

The selection of individuals for the next population is done with a two rounds tournament selection which is based on minimum numbers of load flow cases as result of the fitness function. These individuals make up the parents of the new population. The two rounds of tournament selection require a population size which is divisible by four.

The crossover swaps parts of the parents individuals of the new population. The random determined crossover points define the parts of the parents. The crossover probability (crossover rate) was set to 0.5.

The permutation changes a random number of elements in individuals of the new population consisting of parents and children. The permutation probability (permutation rate) is also 0.5.

The mutation is the randomized modification of single elements in the population with mutation probability defined by a mutation rate of 0.5.

The individuals are characterized by the order of the time steps. For crossover, permutation and mutation thus it is necessary to change the order in the individuals and not simply the elements and parts of the individuals.

Selection and reproduction (crossover, permutation and mutation) are repeated as long as a fixed number of iterations n_{iter} is reached.

G. Method overview

In Fig. 1 the implementation of the method to determine a set of load flow cases including the selection of load flow cases and the genetic algorithm is shown.

The result of the method, the time vector t_{LC} including the times of the selected load flow cases provides a set of load flow cases including load and generation values which are the bases for further load flow calculations in power system planning.

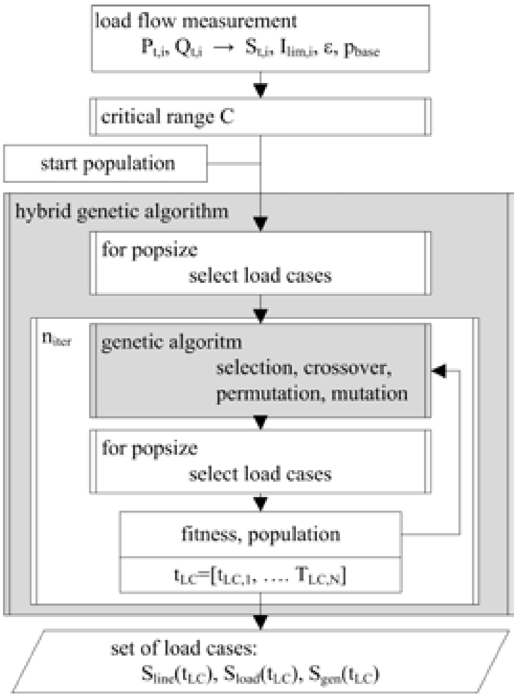

Fig. 1 Hybrid genetic algorithm[1] in combination with the selection of load flow cases – pseudo code flow chart.

III. Results

A. Test Grid I and Grid II

The application of the introduced method is demonstrated with two real high voltage girds (GRID I and GRID II). For both grids the measured active and reactive power is based on one year observation with a resolution of ¼ hours and includes

[1] Hybrid genetic algorithm uses the local search algorithm to calculate the fitness during the evolutionary cycles from a population to the next one.

values for each power line, each generation unit and each load.

The high voltage GRID I consists of 107 power lines. This grid is linked to the surrounding network at three points. GRID II consists of 124 power lines and is linked to the surrounding network at four points. Both power systems include also generation (thermal, run of and pump storage hydro) and load (urban, industrial and rural).

B. Population size and number of iterations

The coherence of required number of load flow cases depending on the number of iterations is investigated for different values of ε and s_{base}. After convergence of the fitness the genetic algorithm will be interrupted.

The population size was varied from minimum four to 200 (approximately the double number of lines) and shows also the same fast convergence. In most cases, when varying these parameters, the convergence value of the fitness function is already included in the initial population.

C. Set of load flow cases depending to the grid structure

To apply and evaluate the method of the determination of required load flow cases a set of different ε and s_{base} are used. The critical range was chosen as ε = 1 %, 5 %, 10 %, 15 % and 20 %. In order to determine whether a single line should be evaluated or not s_{base} = 20 %, 40 %, 50 % and 60 % are used.

Fig. 1 shows the determined, required 14 load flow cases for ε = 20 % and s_{base} = 60 % of Grid I. For these parameters the analyses of the full data set $\mathbf{S}_{meas}$ with L = 107 power lines results in 38 power lines under consideration with 44 critical situations. In the diagram display $s_{LC,k}$ the loading situations of the considered power lines. The 14 determined load flow cases have at least one situation within the critical range between $s_{i,upper/lower}$ and $s_{line,i,max/min}$ of each of the 38 power lines. The enveloping lines $s_{e.}$ are the minimal and maximal values of the set of required load flow cases. The lines for $s_{LC,k}$ and s_e in the diagram represent only the time interrelationship of a load flow case and not a linear function between 2 power lines.

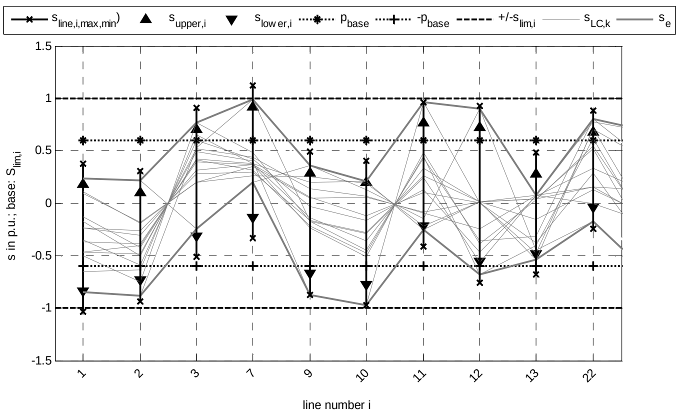

Fig. 1 Grid I, ε = 20 % and s_{base} = 60 % - 14 required load flow cases.

Table I gives a summary of the required load flow cases for different choices of ε and s_{base}. The number of the considered power lines L_k as function of limit loading s_{base} is also listed in Table I. The number of critical situations CS is the number of lines with loading situations within the critical range in both directions. For different grids the number of required load cases N_{LC} is also different. N_{LC} for different ε and s_{base} of Grid II (Table II) have the same tendency as Grid I and the absolute values are higher compared to Grid I. The biggest effect on this increase is the higher number of lines in Grid II.

TABLE I

GRID I / GRID II - REQUIRED NUMBER OF LOAD FLOW CASES $N_{LC,I}$ / $N_{LC,II}$ AND CONSIDERED NUMBER OF LINES (L_k) FOR ε AND S_{BASE}

s_{base}	20 %	40 %	50 %	60 %
$L_{k,I}$ / CS_I	81 / 135	74 / 100	60 / 70	38 / 44
$L_{k,II}$ / CS_{II}	115 / 170	88 / 104	71 / 83	50 / 54
ε = 1 %	71 / 88	53 / 57	42 / 44	30 / 28
ε = 5 %	47 / 56	38 / 37	34 / 33	26 / 22
ε = 10 %	37 / 38	30 / 28	27 / 26	22 / 19
ε = 15 %	30 / 32	26 / 25	24 / 21	18 / 17
ε = 20 %	24 / 26	20 / 21	18 / 20	14 / 16

D. Grid separation

Grid II can be separated according to geographical boundary conditions in two parts in the following part A and part B. Part A includes 50 and part B includes 74 of the 124 power lines in Grid II. For both parts the required load flow cases are determined for the same parameters ε and s_{base} as claimed above. The results (number of load flow cases) are summarized in table III.

TABLE III

GRID II - PART A / PART B - REQUIRED NUMBER OF LOAD FLOW CASES $N_{LC,A}$ / $N_{LC,B}$ AND CONSIDERED NUMBER OF LINES FOR ε AND S_{BASE}.

s_{base}	20 %	40 %	50 %	60 %
$L_{k,A}$ / CS_A	48 / 75	38 / 47	30 / 36	22 / 24
$L_{k,B}$ / CS_B	67 / 95	50 / 57	41 / 57	28 / 30
ε = 1 %	47 / 45	30 / 30	21 / 25	14 / 15
ε = 5 %	33 / 28	22 / 18	18 / 17	12 / 12
ε = 10 %	24 / 21	18 / 15	15 / 14	12 / 9
ε = 15 %	19 / 15	15 / 12	13 / 12	11 / 8
ε = 20 %	16 / 13	13 /11	13 / 11	10 / 8

The number of required load flow cases is also a function of the grid size and is related to the number of lines. The analysis of Grid II with ε = 1 % and s_{base} = 20 % and with L_k = 115 lines and CS = 170 critical situations results in N_{LC} = 88 required load flow cases. Part A analyzed with same ε and s_{base} and with $L_{k,A}$ = 48 and CS_A = 75 results in $N_{LC,A}$ = 47. Part B with $L_{k,B}$ = 67 and CS_B = 95 results in $N_{LC,B}$ = 45. For smaller grids fewer required load flow cases are necessary but mostly the parts of a split grid requires for them alone more characteristic load flow cases as for the parts combined. The numbers of required load flow cases for the parts of a splited grid also have a decreasing tendency as the complete grid depending to increasing p_{base} and ε.

Comparison peak load and peak generation case

On the hand of Grid I the power line loadings at peak load situation (Fig. 1) and at peak generation situation (Fig. 2) are compared with the critical range with ε = 20 % and s_{base} = 60 %).

The peak loading situation matches eight times the critical range. So it is possible to find 35 critical situations which are not matching with the required criteria. As well the peak generation load flow case matches only eight times (different lines) the critical range.

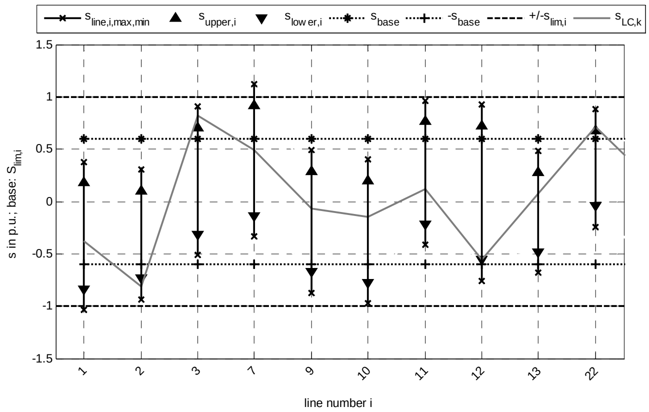

Fig. 1 Peak load situation $s_{LC,k}$ and load flow situation of 10 power lines in a meshed Grid I; critical range with ε = 20 % and s_{base} = 60 %.

The finding of these comparisons is that single load flow cases, e.g. peak load or peak generation, are not appropriate to give a worst case scenario of a power system. Most of the occurring critical situations are not taken into consideration with this single load flow cases.

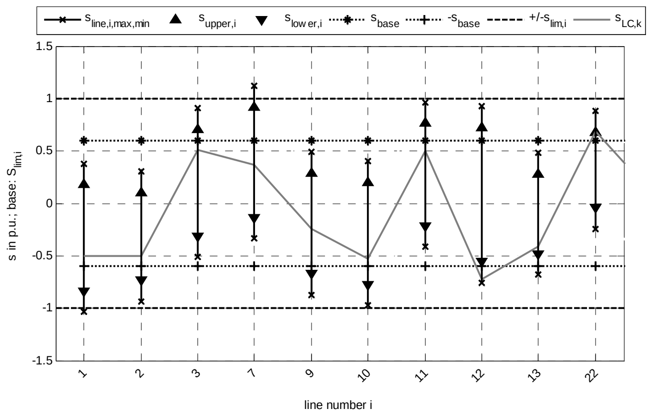

Fig. 2 Peak generation situation $s_{LC,k}$ and load flow situation of 10 power lines in Grid I; critical range with ε = 20 % and s_{base} = 60 %.

IV. Conclusion

With the presented method it is possible to find a minimum number of representative load flow cases which cover all critical power line loading situations in a given time frame.

The proposed method is primarily meant for use in meshed transmission grids and provides a systematical approach for power system planning.

For this method the bases data set, consisting of power flow measurement, is prepared with reducing double circuit lines, consideration of load flow direction, calculation line loadings, determination of a critical range of each power line, consideration only lines exceeding minimal loading and reducing of load flow cases with same characteristic.

The characteristic load flow cases are determined with a combination of a heuristic search algorithm and a genetic optimization algorithm to find a minimum number of load flow cases.

The number of required load flow case depends on the extension of the analyzed grid, on the claimed accuracy of the critical range ε and on the number of lines considered to be critical according to their maximal loading.

The classical planning scenarios e.g. peak load or peak generation usually only cover a limited number of critical loading situations. So they are not appropriate to give a worst case scenario of a power system.

V. References

[1] D. E. Goldberg, “Genetic algorithms for search, optimization, and machine learning,” 1 edition, Boston, Addison-Wesley Professional, 1989

[2] T.J. Hammons, ”Configuring the power system and regional power system developments in Southeast Europe,” in *Proc. 2004 Universities Power Engineering Conf.*, pp.1319 - 1326 vol. 2

[3] G. Latorre, R. D. Cruz, J. M. Areiza, and A. Villegas “Classification of publications and models on transmission expansion planning,” *IEEE Transaction on Power Systems*, vol. 18, no. 2, pp. 938–946, May 2003

[4] D. Oeding, “Elektrische Kraftwerke und Netze,” 6. edition, Berlin Heidelberg New York, Springer, 2004

[5] *EURELECTIC WG SYSTINT, European Interconnection State of the Art 2002 (SYSINT Annual Report)*, 2002-210-0002, 2002

[6] *UCTE Operational Handbook – Appendix 4: Coordinated Operational Planning, capture B. UCTE network calculations*, v2.5 E, 24.06.2004

[7] C. L. Krane, “Strukturbewertung elektrischer Übertragungsnetze,“ Dissertation, Dept. Electr. Power Syst. Univ. Aachen, 2007

[8] Wood, Allen J., and Wollenberg, BruceF., “Power Generation, Operation and Control”, 2nd edition, New York, Wiley-Interscience, 1984

VI. Biographies

Georg Rechberger was born in Linz, Austria, in 1978. He received his diploma degree in 2005 at Graz University of Technology, Graz, where he currently holds a position as scientific assistant at the Institute of Electrical Power System. His main work in research and teaching is in the field of electrical power system planning.

Herwig Renner was born in Graz, Austria, in 1965. He received his doctoral degree in 1995 at Graz University of Technology, Graz, where he currently holds a position as associate professor at the Institute of Electrical Power Systems. His main work in research and teaching is in the field of electrical power system analyses with special emphasis on quality, supply reliability and power system control and stability. In 2007 he was employed as visiting professor at Helsinki University of Technology.

Alexander Gaun was born in Kufstein, Austria, in 1979. He received his diploma degree in 2005 at Graz University of Technology, Graz, where he currently holds a position as scientific assistant at the Institute of Electrical Power Systems. His main work in research and teaching is in the field of electrical power system planning, reliability calculation, genetic algorithms and electromagnetic compatibility.

7.3 Annex C: Reference P3

Paper submitted to Electrical Engineering Journal and in reviewing status, pp. 1-27, Apr. 2010

Probabilistic Load Flow Computation Via Enhanced Cumulant based Gram-Charlier Expansion Considering Mutual Dependencies

Alexander Gaun [a], Georg Rechberger [a], Herwig Renner [a], Matti Lehtonen [b]

[a] *Institute of Electrical Power Systems, Graz, University of Technology,* Inffeldgasse 18/1, Graz, Austria, forename.surname@tugraz.at

[b] *Department of Electrical Power Systems, HUT University,* Otakaari 5 I, Espoo, Finnland, forename.surname@tkk.fi

Abstract:

The Probabilistic Load Flow method based on point estimation has become popular during the past years. In this paper it is shown how mutual dependencies between real measured sets of input random variables can be numerically implemented in the calculation by the means of joint cumulants. The modified method can be referred to as an “Enhanced Cumulant based Gram-Charlier Expansion”. First, it is proven by means of mutual information that the measured input data is non independent. Then, the accuracy of the proposed analytical method is demonstrated via comparing the results with discrete load flow simulations and a Point Estimate Method, by use of a synthetic example and measured load data of a real world 110 kV sub transmission power system. Lastly, a suitable framework to use the proposed method in fast probabilistic load flow studies is provided while maintaining a high degree of accuracy.

KEY WORDS

Gram-Charlier Typ A series expansion, Joint cumulants, Mutual information, Non-independent random variable, Probabilistic Load Flow;

Introduction

Electrical deterministic load flow (DLF) calculation is used to analyze and evaluate the planning, calculating and operating of power systems on a daily routine. DLF uses fixed values of bus injected power for generation and load demands of a selected network configuration to determine system states. The deterministic approach for power system expansion planning is usually performed with one or two worst case scenarios, winter peak load case and summer off peak load case [1], [2]. It can be shown [3], that this procedure does not contain all relevant loading conditions of power lines. The probabilistic load flow (PLF) approach can consider all relevant loading conditions and is may not only be faster as e.g. hourly step wise calculation of all loading conditions but also provides the consideration of uncertainties. It is independent whether an AC or DC based PLF approach is used.

The analytical PLF approach has a solid mathematical background and has been applied to power systems in different areas since 1974 [4] - [18]. The most common techniques are the convolution method [6], the cumulant based Gram-Charlier Type A series Expansion (GCE) method [9] and [10] and the point estimate (PE) method [12] - [15].

Many publications dealing with independent sets of random variables (RVs) based PLFs have been published [6] – [10], [15], [16], and [17]. As shown in the present paper, real world problems force to consider statistical dependence between RVs, as for example applied in [11] - [14] and [18]. The authors in [12] - [14] use a point estimate method combined with rotational transformation based on the eigenvector of the correlation matrix to transform the set of correlated RVs into an uncorrelated set of input RVs. In [18] a linear correlation model linked with a combination of convolution method with a Monte Carlo simulation (MCS) technique is applied to incorporate the mutual dependencies of input nodal powers. By using covariances of the input RVs the linear dependences between loads and generators are taken into account in [11], where no further remarks are justifying the proposed method.

In the present paper a simple formula for computation time inexpensive calculation of multivariate joint cumulants considering mutual dependencies of input nodal RVs is described and illustrated. The authors make no attempt to achieve total generality. Instead, the focus is mainly put on joint cumulants, likely to be of most use in applications, usually where the degree of cumulant order does not exceed four. The proposed method was first rudimentary introduced by [11] and is expanded in this paper to higher degrees of dependencies in terms of joint cumulants, with a different mathematical approach.

This paper is structured as follows: in the subsequent section a short discussion about mutual dependency of nodal powers is given, followed by the proof of dependencies between the used measured sets of input RVs. In the next section mathematical background of DC load flow, mathematical expectation, moments, cumulants and the Enhanced Cumulant based GCE are provided. In the following section the accuracy of the proposed method is demonstrated by the help of DLF's as a reference and by the 2m+1 PE scheme [13] and [15]. This calculation is based on the computational procedure described in the present paper via applying on a synthetically generated example and on real world measured data of a 110 kV sub transmission power system. Furthermore it is emphasized that it is sufficient to consider dependencies only up to a certain degree. Concluding remarks are given in the conclusions. Finally, in the Appendix, the mathematical deduction of the proposed equations is provided.

Dependence between nodal powers

There exist various reasons for dependencies between nodal powers. Mutual dependent loads are caused by common environmental factors such as temperature, sunset, rainfall, season, etc., and due to social factors such as sporting events, television programmes, meal times, working habits, wealth, educational background etc. Some of these factors affect all loads of a similar nature in a related manner and therefore a certain degree of linear dependence exists [18]. Dependencies involving generation are due to the utility's operating policy [18], market prices and environmental factors such as season, temperature, etc. Those dependencies might also be nonlinear, but usually can be linearized maintaining sufficient accuracy.

Independence among real measured signals indicates that there is no statistical dependence among the signals. Besides Gaussian RV, where independence means that the mutual correlations are zero, independence of higher order statistics is judged by joint cumulants, where dependence is indicated by nonzero mutual joint cumulants. Since the authors of this paper deal with non-Gaussian RV it is necessary to consider higher order statistics for independence. Therefore joint cumulants provide useful measures of the joint statistical dependence of real measured sets of RVs [19].

Mutual Information (MI) is also a measure of the dependence between RVs [20]. If two RVs are independent, the MI is zero; otherwise the value of the MI depends on the dependence between the two RVs. The difference to the linear correlation can be seen in the fact that MI can even evaluate dependence if any linear and nonlinear correlation between the two RVs exists, without requiring the specification of any kind of dependence model [21]. The MI algorithm "Calculation of I(X,Y) using an adaptive XY partition" from [22] can be used to estimate the mutual dependence. This algorithm returns $P_{null} \neq 0$ if the null hypothesis of statistical independence is not rejected [22], which indicates, if $P_{null} = 0$ that there is any statistical dependence. $P_{null} = Q\left(v/2, \chi^2/2\right)$ is the probability of the null hypothesis with v is the number of degrees of freedom, χ^2 is the value of the Chi-square test and Q is the incomplete gamma function with a upper limiting value of one.

The measured load data of the proposed grid in Figure 1 was analyzed with the conservative MI algorithm where each load and generation is compared with all 24 residual ones. The results of the calculation show clearly that there is statistical dependence between all nodes, regardless if generator and\or load bus. More than 50 % of all $\left|r_{(X,Y)}\right|$, a measures of linear dependence between two RVs, exceed 0.46, and there is at least a light correlation between all input RVs. Therefore it is necessary to include dependencies in the PLF calculations to achieve a higher accuracy of the obtained results compared with single DLF calculations.

Mathematical Formulation

Probabilistic DC Load Flow

The power flow equations of DLF analysis are nonlinear [9]. The input RVs of an PLF are density functions, which can be preferable applied with linear approximations. Therefore the solution of the problem is a linear combination of random input variables which means that the PLF in this paper is a DC-load flow based approach. Moreover if power transfer distribution factors (PTDF) are used to solve the load flow a very smart way can be proposed to perform the PLF.

The PTDF calculation, a sensitivity analysis of flows to injections, can be easily done by a matrix multiplication.

$$\mathbf{PTDF} = \mathbf{B^b} \cdot \mathbf{B}^{-1} \qquad (3.1)$$

where **B** is the reduced nodal susceptance matrix and $\mathbf{B^b}$ is a reduced matrix with the branch suceptances. In this case reduced means that the rows and columns corresponding to the slack bus (reference bus) are eliminated avoiding singularity. With eqn. (3.1) the DC load flow problem can be formulated with

$$\mathbf{P} = \mathbf{PTDF} \cdot \Delta\mathbf{P} \qquad (3.2)$$

where **P** is the vector of flows on the lines in the grid and **ΔP** is the vector of bus injected powers. Eq. (3.2) allows a fast DLF calculation for a set of input powers. Performing cumulant based Gram-Charlier Expansion (CGCE) PLF calculation, characteristic values (CV), such as moments and cumulants, of the independent input RVs are determined. Next, these characteristic values are multiplied with the PTDFs of the grid to obtain CV of the power flows on the lines. Subsequently the GCE factors are determined from the line flow CVs. Finally, applying the GCE method, the probability density function (PDF) or cumulative distribution function (CDF) of each branch active power flow is calculated.

Mathematical expectation

The mathematical expectation of a real-valued RV, expressed via a continuous function X, with probability density $f_X(x)$ is $\langle X \rangle = \int_{-\infty}^{\infty} x\, f_X(x)dx$ and for a discrete function $\langle X \rangle = \sum_i x_i p_i(X)$, where $\langle \bullet \rangle$ is the mathematical expectation operator and $p_i(X) = P(X = x_i)$ is the probability mass function. Furthermore considering the joint PDF of two or more continuous RVs $X_1,...,X_n$ the mathematical expectation is $\langle X_1...X_n \rangle = \int_{X_1}...\int_{X_n} x_1...x_n f(x_1,...,x_n)dx_1...dx_n$, where $f(x_1,...,x_n) = \mathrm{P}(a_1 \le X_1 \le b_1,...,a_n \le X_n \le b_n) = \int_{a_1}^{b_1}...\int_{a_n}^{b_n} f(x_1,...,x_n)dx_n...dx_1$. With the joint probability mass function $p_i(X_1,...,X_n) = P(X_1 = x_i,...,X_n = x_i)$ the expectation for multivariate discrete RVs $X_1,...,X_n$ can be calculated via $\langle X_1...X_n \rangle = \sum_i ... \sum x_1...x_n p_i(X_1,...,X_n)$.

Moments, Cumulants, Central Moments and Joint Cumulants

Via the definition of the r^th^ moment μ_r^X of one real-valued RV X, for integers r = 0,1, …

$$\mu_r^X = \langle X^r \rangle = \int_{-\infty}^{\infty} x^r f(x)dx \qquad (3.3)$$

and provided that it has a Taylor expansion about the origin, the moment generating function $M(\zeta)$ can be formulated $M(\xi)=\langle e^{\xi X}\rangle=\langle 1+\xi X+...+\xi^r X^r/r!+...\rangle=\sum_{r=0}^{\infty}\mu_r^X\xi^r/r!$ [23]. The cumulants κ_r^X of order r are the coefficients in the Taylor expansion of the cumulants generating function $K(\zeta)$ about the origin $K(\xi)=\log M(\xi)=\sum_{r=0}^{\infty}\kappa_r^X\xi^r/r!$. Extracting the coefficients from the above expansion the relationship between the first four cumulants and moments about an arbitrary point can be stated as $\kappa_1^X=\mu_1^X$, where μ_1^X is the mean value or mathematical expectation of the RV X, $\kappa_2^X=\mu_2^X-(\mu_1^X)^2$, $\kappa_3^X=\mu_3^X-3\mu_2^X\mu_1^X+2(\mu_1^X)^3$, $\kappa_4^X=\mu_4^X-4\mu_3^X\mu_1^X-3(\mu_2^X)^2+12\mu_2^X(\mu_1^X)^2-6(\mu_1^X)^4$ and in the reverse direction the moments are in terms of cumulants $\mu_2^X=\kappa_2^X+(\kappa_1^X)^2$, $\mu_3^X=\kappa_3^X+3\kappa_2^X\kappa_1^X+(\kappa_1^X)^3$ and $\mu_4^X=\kappa_4^X+4\kappa_3^X\kappa_1^X+3(\kappa_2^X)^2+6\kappa_2^X(\kappa_1^X)^2+(\kappa_1^X)^4$.
Expressions for higher order moments and cumulants (designated as κ) for a single RV can be better expressed with the recurrent formula

$$\kappa_{r+1}^X=\mu_{r+1}^X-\sum_{j=1}^{r}C_j^r\mu_j^X\kappa_{r-j+1}^X \qquad (3.4)$$

where the binomial coefficient C_a^b is defined by $a!/(b!(a-b)!)$. Since the mean of a single RV X can also be expressed in terms of $\langle X\rangle=\mu_1^X=\kappa_1^X$, the formulae for the cumulants in eqn. (3.4) are $\kappa_2^X=\sigma_X^2$, $\kappa_3^X=\langle XXX\rangle-\langle X\rangle^3-3\sigma_X^2\langle X\rangle$ and $\kappa_4^X=\langle XXXX\rangle+3\langle X\rangle^4-3\sigma_X^4-2\langle X\rangle(2\langle XXX\rangle-3\langle X\rangle\sigma_X^2)$, where σ_X^2 is the variance of RV X. The proof of the above equations can be found in the Appendix.
If a certain number of RV $X_1,...,X_n$ is independent, the cumulants of these RVs can be summed up [24]. Hence the formulae for a sum of a set of independent RVs $X_1,...,X_n$ for the first four cumulants is

$$\kappa_1^{(X_1+X_2+...+X_n)} = \sum_{i=1}^{n}\langle X_i\rangle = \sum_{i=1}^{n}\mu_1^{X_i} \qquad \kappa_2^{(X_1+X_2+...+X_n)} = \sum_{i=1}^{n}\sigma_{X_i}^2 = \sum_{i=1}^{n}\left(\mu_2^{X_i} - \left(\mu_1^{X_i}\right)^2\right)$$
$$\kappa_3^{(X_1+X_2+...+X_n)} = \sum_{i=1}^{n}\langle X_iX_iX_i\rangle - \sum_{i=1}^{n}\langle X_i\rangle^3 - 3\sum_{i=1}^{n}\sigma_{X_i}^2\langle X_i\rangle = \sum_{i=1}^{n}\left(\mu_3^{X_i} - 3\mu_2^{X_i}\mu_1^{X_i} + 2\left(\mu_1^{X_i}\right)^3\right)$$
$$\kappa_4^{(X_1+X_2+...+X_n)} = \sum_{i=1}^{n}\langle X_iX_iX_iX_i\rangle + 3\sum_{i=1}^{n}\left(\langle X_i\rangle^4 - \sigma_{X_i}^4\right) - 2\sum_{i=1}^{n}\langle X_i\rangle\left(2\langle X_iX_iX_i\rangle - 3\langle X_i\rangle\sigma_{X_i}^2\right) =$$
$$= \sum_{i=1}^{n}\left(\mu_4^{X_i} - 4\mu_3^{X_i}\mu_1^{X_i} - 3\left(\mu_2^{X_i}\right)^2 + 12\mu_2^{X_i}\left(\mu_1^{X_i}\right)^2 - 6\left(\mu_1^{X_i}\right)^4\right)$$
(3.5)

The n^th central moment m_n of RV X is defined as $m_n^X = \langle X - \langle X\rangle\rangle^n$. Furthermore the central moments m_n^X can be expressed as terms of the following binomial transform $m_n^X = \sum_{i=0}^{n} C_i^n (-1)^{n-i} \mu_i^X \left(m_1^X\right)^{n-i}$. In particular for the first four central moments expressed with moments are $m_1^X = 0$, $m_2^X = \mu_2^X - \left(m_1^X\right)^2$, $m_3^X = \mu_3^X - 3m_1^X\mu_2^X + 2\left(m_1^X\right)^3$ and $m_4^X = \mu_4^X - 4m_1^X\mu_3^X + 6\left(m_1^X\right)^2\mu_2^X - 3\left(m_1^X\right)^4$. In terms of cumulants by substituting the moments with eqn. (3.4) and dropping all terms with m_1^X and κ_1^X from these polynomials

$$m_2^X = \kappa_2^X + \left(\kappa_1^X\right)^2 - \left(m_1^X\right)^2 = \kappa_2^X$$
$$m_3^X = \kappa_3^X + 3\kappa_2^X\kappa_1^X + \left(\kappa_1^X\right)^3 - 3m_1^X\left(\kappa_2^X + \left(\kappa_1^X\right)^2\right) + 2\left(m_1^X\right)^3 = \kappa_3^X$$
$$m_4^X = \kappa_4^X + 4\kappa_3^X\kappa_1^X + 3\left(\kappa_2^X\right)^2 + 6\kappa_2^X\left(\kappa_1^X\right)^2 + \left(\kappa_1^X\right)^4 - 4m_1^X\left(\kappa_3^X + 3\kappa_2^X\kappa_1^X + \left(\kappa_1^X\right)^3\right) +$$
$$+6\left(m_1^X\right)^2\left(\kappa_2^X + \left(\kappa_1^X\right)^2\right) - 3\left(m_1^X\right)^4 = \kappa_4^X + 3\left(\kappa_2^X\right)^2$$
(3.6)

A joint cumulant *cum* considers various RV and is a measure of joint dependence of these RVs. Hence, the mutual dependence between the sets of RVs is included in the equation of the multivariate joint cumulants [19]. The joint cumulant up to the u^th number of elements is defined in terms of mathematical expectations by

$$cum^{(X_1,\ldots,X_u)} = \sum_{P}(-1)^{p-1}(p-1)!\langle v_1\rangle\ldots\langle v_p\rangle \qquad (3.7)$$

where the summation extends over all partitions of the indices, p is the number of blocks of the partition and $P = \{v_1,\ldots,v_p\}$ is a partition of a set of u indices into p non-empty blocks [25]. For example eqn. (A.17) gives one definition for a third order joint cumulant.

$$cum_n^{(Z)} = cum_n^{(X+Y)} = \sum_{i=0}^{n} C_i^n cum^{i(n-i)} \qquad (3.8)$$

where $cum^{i(n-i)} = cum^{(\underbrace{X,...,X}_{i},\underbrace{Y,...,Y}_{n-i})}$. Applying eqn. (3.7) and (3.8), one can imagine, that it is a straightforward but tedious job to calculate the joint cumulant of a high order considering a general number of RVs. Therefore in this paper this is done only up to the 4^{th} order, to consider the skewness and kurtosis of a sum of non independent distributions. If there is no mutual dependence between RVs, the joint cumulant of order n is equal to the cumulant of order n. This fact is used to propose new formulas for the 2^{nd}, 3^{rd} and 4^{th} order joint cumulant considering the mutual dependence between RVs by eqn. (3.5) and the to the case of "n" RVs extended eqn. (3.4) (see Eq. (A.12), (A.20) and (A.28)).

Enhanced Cumulant based GCE (ECGCE)

Gram-Charlier Expansion

The asymptotic GCE, where skewness and kurtosis directly appear as parameters of the power series used to obtain the PDF or CDF from CV of active power line flows, has become popular for instance in PLF studies [9], [10]. However, although GCE and cumulants have already been proposed at the end of the 19^{th} century [26], there are still questions of the convergence of such an expansion that arise in the case of distributions which are far from Gaussianity [27]. Furthermore, being polynomial approximations, GCE series are based on linearizations and have the drawback of resulting in negative values for certain parameters.

The method of GCE is built on moments. These moments up to the fourth order depend very largely upon contributions from the tails of a distribution [27], which leads to an error of the approximation. Besides these drawbacks the asymptotic GCE has the advantage of computationally inexpensiveness and that it can be easily generalized by the means of joint cumulants for dependent RVs as it is demonstrated in the present paper.

From [23], it follows that Tchebycheff-Hermite polynomials $H_r(R)$ are

$$H_r(R) = R^r - \frac{r^{[2]}}{2^1 1!} R^{r-2} + \frac{r^{[4]}}{2^2 2!} R^{r-4} - \frac{r^{[6]}}{2^3 3!} R^{r-6} + ... \qquad (3.9)$$

with $H_0 = 1$ and $r^{[z]} = r(r-1)(r-2)...(r-(z-1))$. In particular, by straightforward computations the first four polynomials are $H_1(R) = R$, $H_2(R) = R^2 - 1$, $H_3(R) = R^3 - 3R$ and $H_4(R) = R^4 - 6R^2 + 3$. Moreover, the well known GCE can be derived from

$$F(R) = \sum_{r=0}^{\infty} (c_r H_r(R)) \Phi(R)$$
$$f(R) = \sum_{r=0}^{\infty} (c_r H_r(R)) \varphi(R) \tag{3.10}$$

where $\Phi(R)$ represents the CDF of a normal distribution in the data range R, $\varphi(R)$ represents the PDF of a normal distribution and c_r can be found in virtue of the orthogonal relationship between Tchebycheff-Hermite polynomials by multiplying $H_r(R)$, integrating from $-\infty$ to ∞ and substituting in eqn. (3.9)

$c_r = \frac{1}{r!}\left(\mu_r^X - \frac{r^{[2]}}{2^1 1!}\mu_{r-2}^X + \frac{r^{[4]}}{2^2 2!}\mu_{r-4}^X - \frac{r^{[6]}}{2^3 3!}\mu_{r-4}^X + ...\right)$ with $c_0 = 1$. In particular, for the first four moments are $c_1 = 0$, $c_2 = (\mu_2^X - 1)/2$, $c_3 = \mu_3^X/2$ and $c_4 = (\mu_4^X - 6\mu_2^X + 3)/24$. Thus for the first four central moments

$$c_1 = c_2 = 0;\ c_3 = \frac{m_3^X}{3!};\ c_4 = \frac{1}{4!}(m_4^X - 3) \tag{3.11}$$

According to eqn. (3.10), [23] and [9], we find the formulae for the CDF $F(R)$ and the PDF $f(R)$, where the normal distribution is the leading term of the expansion and both $\Phi(R)$ and $\varphi(R)$ with $\mu_1^X = 0$ and $\sigma_X^2 = 1$. When we normalize the factors c_r of eqn. (3.11) by changing, for instance, m_3^X to m_3^X/σ_X^3, $F(R)$ and $f(R)$ for the ECGCE follows

$$F(R) = \Phi(R)\left(1 + \frac{1}{3!}\frac{m_3^X}{\sigma_X^3} H_3(R) + \frac{1}{4!}\left(\frac{m_4^X}{\sigma_X^4} - 3\right) H_4(R) + ...\right) \tag{3.12}$$

$$f(R) = \varphi(R)\left(1 + \frac{1}{3!}\frac{m_3^X}{\sigma_X^3} H_3(R) + \frac{1}{4!}\left(\frac{m_4^X}{\sigma_X^4} - 3\right) H_4(R) + ...\right) \tag{3.13}$$

where the superfixes of for example σ_X^3 are exponents, not indices.

Computational procedure ECGCE

The difference between the CGCE and ECGCE is that the proposed method considers dependencies between RVs. It is based on following steps.

1) Compute the PTDFs of the power system, using eqn. (3.1)

2) Determine the cumulants of the discrete bus injected powers applying eqn. (3.3), calculating the moments and eqn. (3.4) calculating the cumulants up to the specified order of moments.

3) According to $\kappa_{l,r}^{\Sigma} = \sum_{i=1}^{n} \left(h_{l,i}\right)^r \kappa_r^{X_i}$, via eqn. (3.2), calculate the cumulants of the power flows $\kappa_{l,r}^{\Sigma}$ on the power lines l with order r , where $h_{l,i}$ is the PTDF of a ΔP at bus i for line l and n is the maximum number of buses.

4) Compute the joint cumulants of the power flows $cum_{l,r}^{\Sigma}$ on the power lines l with order r = 2, 3 and 4 and with $x_l^r = X_i \left(h_{l,i}\right)^r$ based on

$$cum_{l,2}^{\Sigma} = \kappa_{l,2}^{\Sigma} + 2\sum_{i}^{n-1}\sum_{j>i}^{n} Cov(x_i^2, x_j^2) \qquad (3.14)$$

$$\begin{aligned} cum_{l,3}^{\Sigma} = \kappa_{l,3}^{\Sigma} &+ 3\sum_{i=1}^{n}\sum_{k\neq i}^{n}\left(\left\langle \underset{p=i,i,k}{\left\{x_p^3\right\}} \right\rangle - \prod_{p=i,i,k}\left\langle x_p^3\right\rangle - \left\langle x_i^3\right\rangle \sigma_{X_k}^2 \left(h_{l,k}\right)^3\right) + \\ &+6\sum_{i=1}^{n-2}\sum_{k>i}^{n-1}\sum_{j>k}^{n}\left(\left\langle \underset{p=i,k,j}{\left\{x_p^3\right\}} \right\rangle - \prod_{p=i,i,k}\left\langle x_p^3\right\rangle\right) - 6\sum_{i=1}^{n}\sum_{k=1}^{n-1}\sum_{j>k}^{n}\left\langle x_i^3\right\rangle Cov(x_k^3, x_j^3) \end{aligned} \qquad (3.15)$$

$$\begin{aligned} cum_{l,4}^{\Sigma} &= \kappa_{l,4}^{\Sigma} + 24\sum_{i=1}^{n-1}\sum_{k>i}^{n}\sum_{j=1}^{n-1}\sum_{u>j}^{n} Cov(x_i^4, x_k^4)\left\langle x_j^4\right\rangle\left\langle x_u^4\right\rangle - \\ &-24\sum_{i=1}^{n-2}\sum_{k>i}^{n}\sum_{j=i}^{n-1}\sum_{\substack{u>j;j\geq i+k-1\\ \{ik\}\neq\{ju\}}}^{n} Cov(x_i^4, x_k^4)Cov(x_j^4, x_u^4) + 12\sum_{i=1}^{n}\sum_{k=1}^{n-1}\sum_{j>k}^{n}\left(\left\langle x_i^4\right\rangle^2 - \sigma_{X_i}^2\left(h_{l,i}\right)^4\right)Cov(x_k^4, x_j^4) + \\ &+6\sum_{i=1}^{n-1}\sum_{k>i}^{n}\left(\left\langle \underset{p=i,i,k,k}{\left\{x_p^4\right\}}\right\rangle - \prod_{p=i,k}\left(\sigma_{X_p}^2\left(h_{l,p}\right)^4\right) - 2Cov(x_i^4, x_k^4)^2 + 3\left(\prod_{p=i,k}\left\langle x_p^4\right\rangle\right)^2\right) + \\ &+24\sum_{j>k}^{n-1}\sum_{u>j}^{n}\left(\sum_{i=1}^{n-3}\sum_{k>i}^{n-2}\left(\left\langle \underset{p=i,k,j,u}{\left\{x_p^4\right\}}\right\rangle + 3\prod_{p=i,k,j,u}\left\langle x_p^4\right\rangle\right) - \sum_{i=1}^{n}\sum_{k=1}^{n-2}\left\langle x_i^4\right\rangle\left\langle \underset{p=k,j,u}{\left\{x_p^4\right\}}\right\rangle\right) + \\ &+12\sum_{i=1}^{n}\sum_{k\neq i}^{n-1}\sum_{\substack{j>k\\ j\neq i}}^{n}\left(\left\langle \underset{p=i,i,k,j}{\left\{x_p^4\right\}}\right\rangle + \left(\sigma_{X_i}^2\left(h_{l,i}\right)^4 + 3\left\langle x_i^4\right\rangle^2\right)\prod_{p=k,j}\left\langle x_p^4\right\rangle\right) - 12\sum_{i=1}^{n}\sum_{k=1}^{n}\sum_{j\neq k}^{n}\left\langle x_i^4\right\rangle\left\langle \underset{p=k,k,j}{\left\{x_p^4\right\}}\right\rangle + \\ &+4\sum_{i=1}^{n}\sum_{k\neq i}^{n}\left(\left\langle x_i^4\right\rangle\left(3\prod_{p=i,i,k}\left\langle x_p^4\right\rangle + 3\left\langle x_k^4\right\rangle\sigma_{X_i}^2\left(h_{l,i}\right)^4 - \left\langle\left(x_k^4\right)^3\right\rangle + 1.5\left\langle x_i^4\right\rangle\sigma_{X_k}^2\left(h_{l,k}\right)^4\right) + \left\langle \underset{p=i,i,i,k}{\left\{x_p^4\right\}}\right\rangle\right) \end{aligned}$$

(3.16)

5) First compute the central moments of each line of the grid applying (3.17), adapted from (3.6)

$$m_{l,2}^X = cum_{l,2}^\Sigma \qquad m_{l,3}^X = cum_{l,3}^\Sigma \qquad m_{l,4}^X = cum_{l,4}^\Sigma + 3\left(cum_{l,2}^\Sigma\right)^2 \qquad (3.17)$$

and afterwards calculate the GCE coefficients c using (3.11).

6) Finally, the results from the ECGCE can be obtained by applying eqn. (3.12) and (3.13) to calculate the CDF and the PDF of the line flows.

Case Studies

Goodness of Expansion and reference PLF methods

In order to verify the approach taking into account joint cumulants, it is applied to synthetically generated and real world data sets. Six different PLF-methods, five of them considering different joint cumulants, are used to obtain the most appropriate one for PLF including mutual dependencies. The results of the six different PLF-methods[1] are compared with 8760 DLF calculations (Method R), representing each hour of a year. An overview of the different methods and their features is given in Table 1.

Method	max moment order	Computational effort	Mutual dependency included	Description
A	4	$(2m+1)\cdot DLF$	Yes	2m+1 PE scheme [13] + [15]
B	4	$1\cdot DLF$	No	–
C	4	$1\cdot DLF + n\cdot cum_2^\Sigma$	Yes	Up to cum_2
D	4	$1\cdot DLF + n\cdot \sum_{i=2,3,4} cum_i^\Sigma$	Yes	Up to cum_4
E	9	$1\cdot DLF$	No	[9]
F	9	$1\cdot DLF + n\cdot cum_2^\Sigma$	Yes	Up to cum_2
R	–	$8760\cdot DLF$	Yes	Sequential DLF

Table 1: Overview of the applied PLF-methods; m is the number of uncertain RVs and n is the number of power lines.

[1] All calculations in this paper are done in Matlab™

Two different measures are used to evaluate the accuracy of the PLF results. The first one is the Average Root Mean Square (ARMS) error $ARMS = \sqrt{\sum_{i=1}^{N}(GC_i - DC_i)^2} / N$ where DC_i is the i[th] point's value of the CDF obtained from the Method R, GC_i is the i[th] point's value of the CDF calculated with the proposed PLF method and N is the number of used points to represent the CDF. The second measure is the well known coefficient of determination R^2. It is the ratio of the sum of squares of the regression (SSR) to the total sum of squares (SST) $R^2 = SSR/SST$ with SSR being defined as $SSR = \sum_{i=1}^{N}\left(GC_i - \overline{DC}\right)^2$ and SST being defined as $SST = \sum_{i=1}^{N}\left(DC_i - \overline{DC}\right)^2$, where $\overline{DC}$ is the mean of Method R. R^2 can take any value between 0 and 1, with a value closer to 1 indicating a better fit while the ARMS tends to zero in that case.

Synthetic example

Let's consider a simple triangle network with three lines, two loads and one slack bus ($i = 1$) and three different load data sets $X_{1...3,i}$ to show the influence of different degree of mutual dependency on the proposed methods. The line reactance between buses 1 and 3 equals to 0.05 pu, between bus 1 and bus 2 equals to 0.025 pu and between bus 2 and bus 3 equals to 0.075 pu. The first load data set for bus $i = 2,3$ ($X_{1,i} = |\varepsilon_i|$) consists of the absolute value of a random normally distributed variable (ε_i) with zero mean and unit variance. The second data set is a linear dependent data set with $X_{2,i} = \varepsilon_1 \pm 0.2X_{1,i}$. The third data set describes a nonlinear correlation between the two loads via $X_{3,3} = \sqrt{3 + X_{2,3}}$ and $X_{3,2} = \log\left(3 + X_{2,2}\right)$.

The validation of the proposed method for all kind of dependencies can be found in Table 2 exemplarily for the line between slack bus and second sink bus. As expected, the first case with ideal independent load data results in a high goodness of fit ($R^2 \approx 1$) and an almost zero ARMS error, comparing PLF results with DLF results. Note that the 2m+1 PE scheme in Table 2 achieves the worst results. Since it is more realistic to have certain mutual dependencies, the authors propose to take joint cumulants into account. If linear dependent RVs are considered, the covariances between the RVs have the major impact on the adequacy increase.

Non-zero joint cumulants are a consequence of dependence, thus the marginal R^2 step from the 2nd joint cumulant to the 4th joint cumulant can be explained.
In the case of nonlinear dependence in Table 2, each additional joint cumulant, starting with B, increases R^2 and decreases the ARMS measure so that their trend converges against the ultimate values. The results in Table 2 point out, that Method A is superior to Method B. Considering the covariances between the RVs (Method A and Method C), the joint cumulant Method C achieves better result, which is similar to the linear dependent example. The asymptotic expansion of the ECGCE, including the 4th order joint cumulant (Method D), taking into account skewness and kurtosis, provides the best approximation to the DLF PDF of the four investigated methods with the smallest differences. Finally, disregarding the errors introduced by the (EC)GCE expansion already discussed, a remarkable accuracy increase of the obtained results with higher orders of joint cumulants can be detected in the linear dependent data set and in the non-linear dependent data set.

Method	Goodness measure	$X_{1,i}$ $I_{(X_{1,2},X_{1,3})}, P_{null} = 0.9178$	$X_{2,i}$ $I_{(X_{2,2},X_{2,3})}, P_{null} = 0$	$X_{3,i}$ $I_{(X_{3,2},X_{3,3})}, P_{null} = 0$
A	ARMS	0.0007	0.0030	0.0054
	R^2	0.9991	0.9842	0.9891
B	ARMS	0.0004	0.0031	0.0064
	R^2	0.9998	0.9828	0.9849
C	ARMS	0.0004	0.0022	0.0045
	R^2	0.9998	0.9906	0.9918
D	ARMS	0.0004	0.0022	0.0038
	R^2	0.9998	0.9908	0.9943

Table 2: Goodness measure for synthetic example, with $I(X,Y)$ being the MI measure.

Real World example

Table 3 summarizes the computational effort for selected methods, given in terms of real measured computation time. The main computational steps in joint cumulant calculation (e.g. see eqn.(3.16)) are the number of expectation summations calculation $\#\Sigma\langle\cdot\rangle$, hence providing platform independent comparisons for different joint cumulant orders in the caption of Table 3. It can be observed that the computational "burden" increases exponentially with the order of the considered joint cumulant. In order to put the approach on a competitive basis compared with the DLF calculations, on the one hand a minimization of the calculation time and on the other hand a suitable level of accuracy has to be achieved. Since nonlinear dependencies between real world input RVs of sub transmission grids can usually be linearized without loss of accuracy, the authors propose to perform the calculations including 2nd order joint cumulants because this is computationally inexpensive and contains the major enhancement, as shown with the linear dependent synthetic example (see Method B and C).

Method	A	B	C	E	F	R
t_{Comp} in sec	0.290	0.063	0.210	0.156	0.301	1.486

Table 3: Computation times for selected methods for the real world example (best of ten runs) performed on a Intel® Core™ 2 Duo CPU E 8600 @ 3.33 GHz with 3.50 GB RAM running on Windows XP Service Pack 3. Theoretical computational expensiveness for the proposed real world example for a single power line in terms of the number of expectation summations calculation $\#\sum\langle\cdot\rangle$: Method C: 253 and Method D: 234817

In Figure 1, the real world power network, consisting of 40 transmission lines, four generators, 23 buses with 21 loads is shown. The necessary system data and load data is listed in [28].

The proposed method in this paper is based on discrete RVs and on the calculation of multivariate probability densities. Considering joint cumulant discrete probability densities are calculated by partitioning the n-dimensional space into equal spaced intervals and counting the observed occupation number. Therefore the value of the mathematical expectation is depending on the partition of the n-dimensional space. In the proposed real world example 15 intervals are used, in order to achieve a compromise between calculation time and accuracy.

In Figure 2 the improvement of the fit obtained by the PLF calculations including the results obtained by the DLF calculations are shown. The improvement of the fit for the real world example is obvious (see Figure 2), but not as powerful as in the nonlinear dependent synthetic example and the 2m+1 PE scheme algorithm provides the worst results. Comparing the difference between the ECGCEs that are calculated via Methods B, C, and D in Figure 3 shows exclusively a marginal varied curve for the Method B. Hence, there is no real impact of the 3rd and 4th order joint cumulant in Figure 3. Thus, this is another indication that it is sufficient to consider only 2nd order joint cumulants in the proposed calculation.

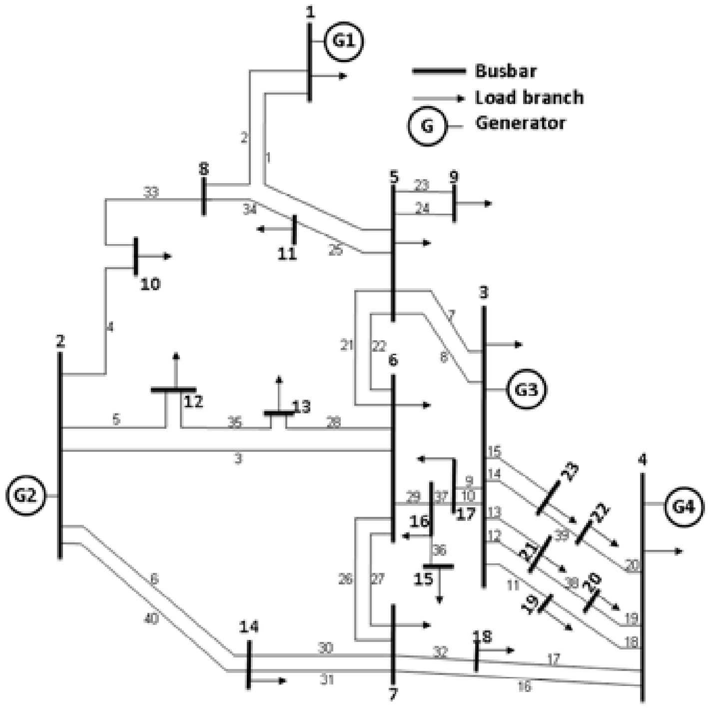

Figure 1: The proposed real world 110 kV sub transmission power system with line numbers

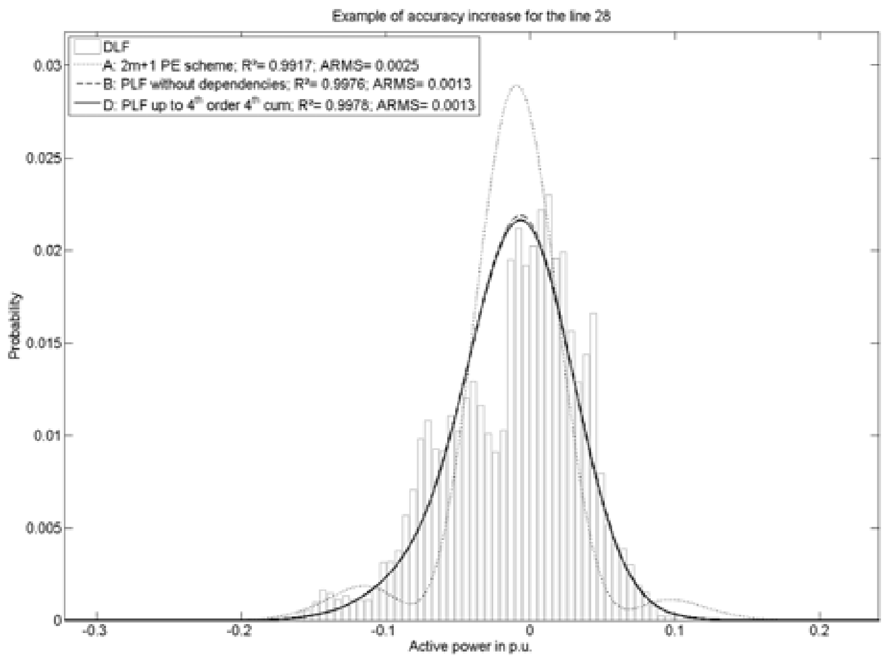

Figure 2: Improved PDF of the obtained results from the PLF based on ECGCE on line from bus 6 to bus 13; Results from Method D and C are congruent

The results of Figure 4 indicate a less accurate fit for this example if higher orders of moments are included in the expansion. Including the covariances of the RVs with moments up to the 9th order (F) achieves R² of 0.9976 and ARMS of 0.0020.

This is worse compared to the result of the expansion Method C where R^2 is 0.9985 and ARMS is 0.0015. In other words, the consideration of higher moments could partly compensate the corrective change of the covariances.

However this is not a general rule and it is only true for approximately the half of the lines in the proposed example (see Table 4).

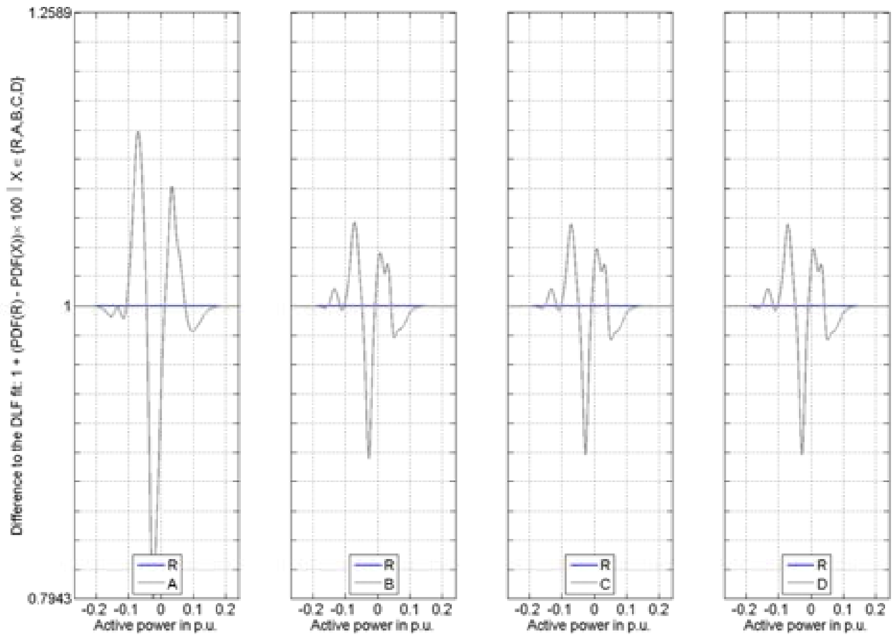

Figure 3: Influence of different PLF methods on line from bus 6 to bus 13

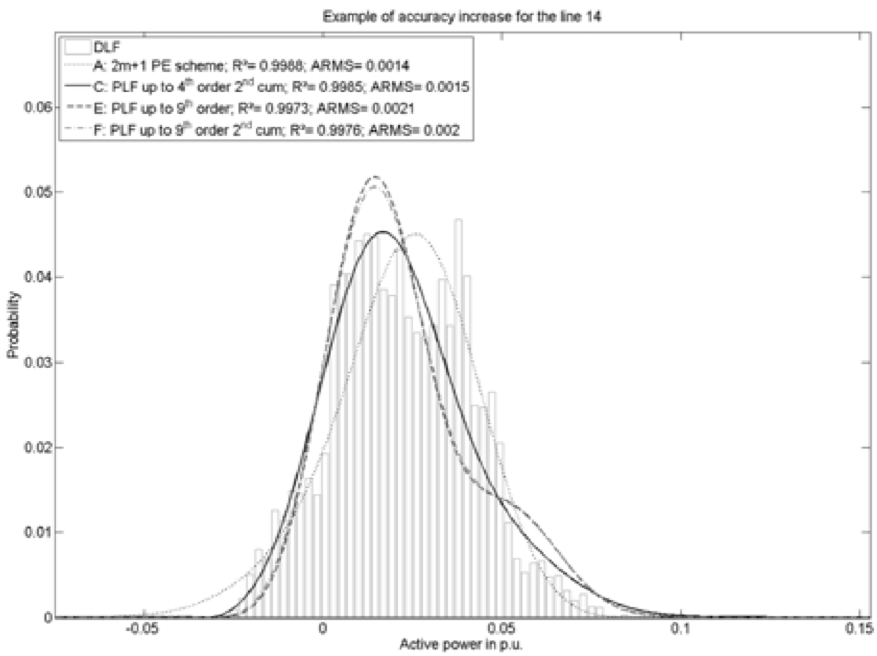

Figure 4: PDF of the obtained results for higher order moments on line from bus 3 to bus 23

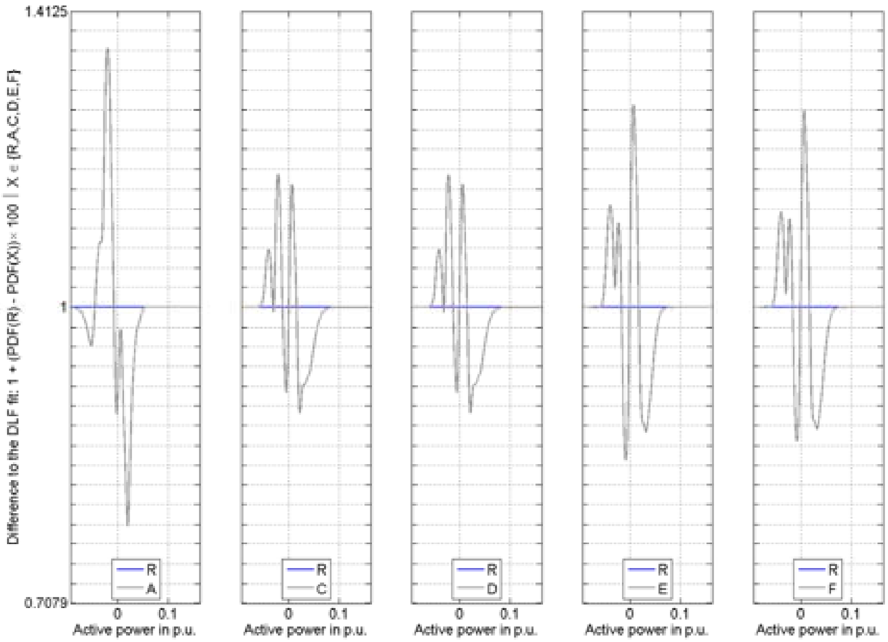

Figure 5: Influence of different PLF methods on line from bus 3 to bus 23

Line	A	B	C	F	Line	A	B	C	F
$\sum_1^{40}$	39.6181	39.6491	39.6605	39.6586	20	0.9948	0.9885	0.9886	0.9955
1	0.9972	0.9887	0.9907	0.9891	21	0.9934	0.9846	0.9850	0.9763
2	0.9961	0.9836	0.9878	0.9878	23	0.9927	0.9991	0.9991	0.9998
3	0.9934	0.9995	0.9995	0.9989	25	0.9964	0.9901	0.9910	0.9898
4	0.9965	0.9903	0.9906	0.9850	27	0.9891	0.9872	0.9872	0.9849
5	0.9967	0.9996	0.9996	0.9995	28	0.9917	0.9976	0.9978	0.9959
6	0.9920	0.9972	0.9972	0.9956	29	0.9738	0.9880	0.9886	0.9880
8	0.9900	0.9910	0.9911	0.9923	30	0.9890	0.9946	0.9946	0.9916
9	0.9595	0.9822	0.9826	0.9851	32	0.9854	0.9878	0.9878	0.9900
11	0.9987	0.9918	0.9918	0.9929	33	0.9944	0.9886	0.9888	0.9819
12	0.9988	0.9964	0.9964	0.9964	34	0.9957	0.9939	0.9944	0.9936
14	0.9988	0.9984	0.9985	0.9976	35	0.9939	0.9990	0.9990	0.9981
16	0.9799	0.9842	0.9841	0.9872	36	0.9988	0.9998	0.9998	0.9999
17	0.9789	0.9800	0.9801	0.9852	37	0.9652	0.9853	0.9860	0.9868
18	0.9923	0.9825	0.9825	0.9936	38	0.9969	0.9858	0.9858	0.9917
19	0.9961	0.9903	0.9903	0.9954	39	0.9984	0.9932	0.9932	0.9939

Table 4: R^2 measure of selected methods for the proposed real world power system (see Appendix for transmission line numbers)

This effect depends obviously strong on the shape of the line power flow distribution and is the more distinct the more non-Gaussian distribution are considered. Since the PDF of power flows on line 14 is almost Gaussian, the 2m+1 PE scheme ("Method A") achieves slightly better measures than all other methods. Figure 5 shows that E and F, taking into account the covariances up to the 9^{th} moment, achieve almost the same approximation of the power flow PDF on line bus 3 and 23. Moreover higher order joint cumulants greater two have again no impact on the accuracy of the result.

The selected summarized R^2 measures of Method A, Method B, Method C and Method F in Table 4 show that where no correlation affects the calculation results, for instance the power flow on stub lines 23, 24, and 36 Method F achieves the highest accuracy due to the best approximation of the stub line load. This is also the reason why the overall sum of all R^2 for G is higher as the one from B. Comparing Method A and C it can be noted that depending on the type of modeled power flow distributions, either for the half of the power lines Method A, in the case of a more Gaussian distribution is better or for the residual part the Method C in the case of a more Non-Gaussian distribution. In the proposed real world example Method C achieves the best overall approximation of all compared methods if all transmission lines except the stub lines are observed.

Distributions having left or right heavy tails, as for instance the PDF in Figure 2, force higher approximation errors, since the approximation of the corresponding non heavy tail is difficult by the means of a polynomial expansion with limited orders. Thus the error introduced by the GCE method itself is not insignificant but the benefits of the method, discussed in the preceding sections dominate, so the authors propose that the method is justified.

Conclusion

In this paper it is demonstrated by the means of a conservative mutual information algorithm that real measured load data of a 110 kV sub transmission grid are mutual dependent. Hence, using this set as input data for a probabilistic load flow calculation, these mutual dependencies have to be taken into account to improve accuracy of calculation results.

Based on the well known Gram-Charlier series expansion the authors have developed enhanced formulae to consider mutual dependencies up to the fourth order. Therefore the traditional formulae of cumulants have been extended to joint cumulants, where mutual dependencies are included. In addition the formulae of joint cumulants have been generalized to equations where “n” random variables of bus injected powers are involved to calculate the distribution of active power flows on power lines. This method is referred to as an “Enhanced Cumulant based Gram-Charlier Expansion (ECGCE)”. Furthermore, by considering the dependencies between RVs in formulae of joint cumulants the proposed approach can not only be applied with GCE but also with Cornish-Fisher expansion, Edgeworth Expansion or others.

The presented synthetic example clarifies the influence of different joint cumulants on the goodness of the distribution for linear and nonlinear dependent RVs. The real world calculation results are also verifying that, based on the proposed R^2 and ARMS measure, the accuracy of the obtained results enhances with increasing order of included joint cumulants. In the real world example and the synthetic example the major contributions of linear dependent RVs originate from the 2nd order joint cumulants.

The 2m+1 PE scheme, where correlation between RV is directly accounted for with a rotational transformation of the input RVs, achieves in an overall view an inferior accuracy than the proposed method and is computational more expensive, compared with the method where 2nd order joint cumulants are computed.

The computation time of higher order joint cumulants including all required input RVs is nearly growing exponentially. By taking into account an insignificant error the consideration of 2nd order joint cumulants is sufficient for the proposed example to achieve an equivalent level of accuracy while reducing the calculation time. In the authors view the proposed method with 2nd order joint cumulants is the best methodology with respect to mentioned performance measures.

References

1. UCTE Operational Handbook – Appendix 4: Coordinated Operational Planning, capture B. UCTE network calculations, v2.0 E, 03.05.2006
2. T.J., Hammons, ”Configuring the power system and regional power system developments in Southeast Europe,” in Proc. 2004 Universities Power Engineering Conf. (UPEC), Vol. 2, pp.1319 – 1326, 6-8. Sep. 2004

3 G., Rechberger, A., Gaun, and H., Renner, “Systematical Determination of Load Flow Cases for Power System Planning”, 2009 IEEE Bucharest PowerTech Conference Proceedings, 6 pages, 28. June 2009 – 2. July 2009

4 K., Kinsner, A., Serwin, and M. Sobierajski, “Practical Aspects of Stochastic Load Flow Calculations”, Archiv für Elektrotechnik 60, pp. 283-288, 1978

5 J., Biricz, and H., Müller, “Probabilistische Lastflußberechnung”, Archiv für Elektrotechnik 64, pp. 77-84, 1981 (in German)

6 J., Schwippe, O., Krause, and C., Rehtanz, “Probabilistic Load Flow Calculation Based on an Enhanced Convolution Technique” 2009 IEEE Bucharest PowerTech Conference Proceedings, 6 pages, 28. June 2009 – 2. July 2009

7 P., Chen, Z., Chen, and B., Bak-Jensen, “Probabilistic Load Flow: A Review“, Third International Conference on Electric Utility Deregulation and Restructuring and Power Technologies, 2008. DRPT 2008, pp. 1586-1591, 6-9 April 2008

8 B., Borkowska, "Probabilistic load flow"; IEEE Trans. Power Apparatus and Systems, Vol. PAS-93, No. 3, pp. 752-755, May-June, 1974

9 P., Zhang, and S. T., Lee, “Probabilistic Load Flow Computation Using the Method of Combined Cumulants and Gram-Charlier Expansion”, IEEE Trans. Power Systems, Vol.19, No 1, pp. 676-682, Feb. 2004

10 D.-J. Shin, J.O. Kimb, K.-H. Kimc, and C. Singh, “Probabilistic approach to available transfer capability calculation”, Electric Power Systems Research 77, pp. 813–820, 2007

11 L.A., Sanabria, and T.S., Dillon, “Power system reliability assessment suitable for a deregulated system via the method of cumulants”, Electrical Power & Energy Systems, Vol. 20, No. 3, pp. 203-211, 1998

12 M. E., Harr, “Probabilistic estimates for multivariate analysis”, Appl. Math. Modelling, Vol. 13, pp 313-318, May 1989

13 P. Caramia, G. Carpinelli, and P. Varilone, “Point estimate schemes for probabilistic three-phase load flow”, Electr. Power System. Res., Vol. 80, pp. 168-175, 2010

14 H.P. Hong “An efficient point estimate method for probabilistic analysis”, Reliability Engineering and System Safety, Vol. 59, pp 261-267, 1998

15 J.M. Morales, and J. Perez-Ruiz, “Point Estimate Schemes to Solve the Probabilistic Power Flow”, IEEE Trans. Power Systems, Vol.22, No 4, pp. 1594-1601, Nov. 2007

16 J., Usaola, "Probabilistic load flow with wind production uncertainty using cumulants and Cornish–Fisher expansion”, Int. J. of Electr. Power and Energ Syst., Vol. 31, pp 474-481, Feb. 2009

17 L.A., Sanabria, and T.S., Dillon, “Stochastic power flow using cumulants and Von Mises functions”, Int. J. of Electr. Power and Energy Syst., Vol. 8, No. 1, pp. 47-60, Jan. 1986

18 A.M., Leite da Silva, V.L., Arienti, and R.N. Allan, “Probabilistic Load Flow Considering Dependence Between Input Nodal Powers”, IEEE Transactions on Power Apparatus and Systems, Vol. PAS-103, No. 6, pp. 1524-1530, June 1984

19 H.W., Block, and Z., Fang, “A Multi Variate Extension of Hoeffding’s Lemma”, The Annals of Probability, Vol. 16, No. 4, pp.1803-1820, 1988

20 W., Li, “Mutual Information Functions Versus Correlation Functions”, Journal of Statistical Physics, Vol. 60, No. 5-6, pp 823–837, Sep. 1990

21 A., Dionisio, R., Menezes, and D. A., Mendes, “Mutual information: a measure of dependency for nonlinear time series”, Physica A, Vol. 344, pp. 326–329, 2004

22 C.J., Cellucci, A.M., Albano, and P.E., Rapp, “Statistical validation of mutual information calculations: Comparisons of alternative numerical algorithms”, Physical Review E 71, 066208, 14 pages, 2005

23 M. G., Kendall, “The advanced Theory of Statistics”, Charles Griffin & Company limited, Vol. 1, 2nd edition, 1945

24 D.R., Brillinger, “Time series: data analysis and theory”, San Francisco, Holden-Day, Inc., 1981

25 P., McCullagh, “Tensor Methods in Statistics”, Monographs on Statistics and Applied Probability, Chapman and Hall Ltd., University Press, Cambridge, 1987

26 A., Hald, “The Early History of the Cumulants and the Gram-Charlier Series” International Statistical Review, Vol. 68, No. 2, pp 137-153, 2000

27 E.S., Pearson, “Some problems arising in approximating to probability distributions, using moments”, Biometrika, Vol. 50, No. 1/2, pp. 95-112, June 1963

28 A., Gaun, G., Rechberger, H., Renner, “Enumeration Based Reliability Assessment Algorithm Considering Nodal Uncertainties”, IEEE PES GM 2010, 8 pages, 25-29 July 2010

29 T.P., Speed, “Cumulants and partition lattices”, Austral. J. Statist, Vol. 25, No. 2, pp. 378–388, 1983

Appendix

Definitions

First we start with some general definitions. The covariance of two RVs X and Y is per definition, using eqn. (3.7)

$$Cov\left(X,Y\right)=\left\langle XY\right\rangle-\mu_X\mu_Y=\left\langle XY\right\rangle-\left\langle X\right\rangle\left\langle Y\right\rangle=(-1)^1 1!\left\langle X\right\rangle\left\langle Y\right\rangle+(-1)^0 0!\left\langle XY\right\rangle=cum^{(X,Y)} \tag{A.1}$$

where e.g. $\langle Y\rangle=\mu_Y$ is the mean of RV Y. The variance of RV X is per definition, again applying eqn. (3.7),

$$\sigma_Y^2=Cov(Y,Y)=\left\langle YY\right\rangle-\mu_Y^2=\left\langle YY\right\rangle-\left\langle Y\right\rangle^2=(-1)^1 1!\left\langle Y\right\rangle\left\langle Y\right\rangle+(-1)^0 0!\left\langle YY\right\rangle=cum^{(Y,Y)} \tag{A.2}$$

One property of joint cumulants is multi-linearity [29], i.e.

$$cum^{(Y,X+Y,\ldots)}=cum^{(Y,X,\ldots)}+cum^{(Y,Y,\ldots)} \tag{A.3}$$

Another property is that joint cumulants are symmetrical [24], i.e.

$$cum^{(Y,X,\ldots)} = cum^{(X,Y,\ldots)} \qquad \text{(A.4)}$$

It follows with eqn. (A.4) that the covariance is also commutative, i.e.

$$Cov(X,Y) = Cov(Y,X) \qquad \text{(A.5)}$$

Furthermore it can be shown that

$$\langle XY\ldots\rangle = \langle YX\ldots\rangle \qquad \text{(A.6)}$$

Proof: Based on eqn. (A.5) and (A.1) it is obvious that $\langle XY\rangle = \langle YX\rangle$.

The mathematical expectation of $\langle XY\ldots\rangle$ can be reduced to $\langle XY\ldots\rangle = \langle X\rangle\langle Y\rangle\ldots$, by assuming that X is independent from Y. The 2nd order cumulant for a single RV X can be expressed with the mathematical expectation via eqn. (A.2)

$\kappa_2^X = \mu_2^X - \left(\mu_1^X\right)^2 = \langle XX\rangle - \langle X\rangle^2 = \sigma_X^2$, the 3rd order cumulant is

$\kappa_3^X = \mu_3^X - 3\mu_2^X\mu_1^X + 2\left(\mu_1^X\right)^3 = \langle XXX\rangle - 3\langle XX\rangle\langle X\rangle + 2\langle X\rangle^3 = \langle XXX\rangle - \langle X\rangle^3 - 3\sigma_X^2\langle X\rangle$

and the 4th order cumulant for a single RV X follows

$$\begin{aligned}\kappa_4^X &= \mu_4^X - 4\mu_3^X\mu_1^X - 3\mu_2^2 + 12\mu_2^X\left(\mu_1^X\right)^2 - 6\left(\mu_1^X\right)^4 = \\ &= \langle XXXX\rangle - 4\langle XXX\rangle\langle X\rangle - 3\langle XX\rangle^2 + 12\langle XX\rangle\langle X\rangle^2 - 6\langle X\rangle^4\end{aligned} \qquad \text{(A.7)}$$

Hence, with $\langle XX\rangle = \sigma_X^2 + \langle X\rangle^2$, simplifying eqn. (A.7) to

$$\begin{aligned}\kappa_4^X &= \langle XXXX\rangle - 4\langle XXX\rangle\langle X\rangle - 3\langle XX\rangle^2 + 12\langle XX\rangle\langle X\rangle^2 - 6\langle X\rangle^4 = \\ &= \langle XXXX\rangle - 4\langle XXX\rangle\langle X\rangle - 3\left(\sigma_X^2 + \langle X\rangle^2\right)^2 + 12\left(\sigma_X^2 + \langle X\rangle^2\right)\langle X\rangle^2 - 6\langle X\rangle^4 = \\ &= \langle XXXX\rangle - 4\langle XXX\rangle\langle X\rangle - 3\sigma_X^4 - 6\sigma_X^2\langle X\rangle^2 - 3\langle X\rangle^4 + 12\sigma_X^2\langle X\rangle^2 + 12\langle X\rangle^4 - 6\langle X\rangle^4 = \\ &= \langle XXXX\rangle + 3\cdot\langle X\rangle^4 - 3\cdot\sigma_X^4 - 2\langle X\rangle\left(2\langle XXX\rangle - 3\langle X\rangle\sigma_X^2\right)\end{aligned}$$

(A.8)

2nd order joint cumulants

Let us start with the proof of the 2nd joint cumulants, eqn. (3.14). Via eqn. (3.8) the 2nd joint cumulant for two RVs X and Y is

$$cum_2^{(X+Y)} = \binom{2}{0}cum^{(X,X)} + \binom{2}{1}cum^{(X,Y)} + \binom{2}{2}cum^{(Y,Y)} = cum^{(X,X)} + 2cum^{(X,Y)} + cum^{(Y,Y)}$$

(A.9)

Using eqn. (3.7) and by applying eqn. (A.2), eqn. (A.9) can be simplified to $cum_2^{(X+Y)} = cum^{(Y,Y)} + 2cum^{(X,Y)} + cum^{(X,X)} = \sigma_Y^2 + \sigma_X^2 + 2Cov(X,Y)$. Now eqn. (A.9) can be extended to the case of "n" non independent RVs. First consider eqn. (A.9) and RV Y is substituted with RVs A + B

$cum_2^{(X+Y)} = cum_2^{(X+(A+B))} = cum^{(A+B,A+B)} + 2cum^{(X,A+B)} + cum^{(X,X)}$. Applying eqn. (A.3) we get

$cum_2^{(X+A+B)} = cum^{(A,A)} + cum^{(A,B)} + cum^{(B,A)} + cum^{(B,B)} + 2cum^{(X,A)} + 2cum^{(X,B)} + cum^{(X,X)}$.

This eqn. can again be reduced with eqns. (A.1) and (A.2) to

$cum_2^{(X+A+B)} = \sigma_A^2 + Cov(A,B) + Cov(B,A) + \sigma_B^2 + 2Cov(X,A) + 2Cov(X,B) + \sigma_X^2$ and by using eqn. (A.5) to

$$cum_2^{(X+A+B)} = \sigma_A^2 + \sigma_B^2 + \sigma_X^2 + 2Cov(X,A) + 2Cov(X,B) + 2Cov(A,B) \qquad \text{(A.10)}$$

In the second step we extend eqn. (A.10) to the case of "n" dependent RV. It gives

$$cum_2^{(X_1+X_2+...+X_n)} = \sum_{i=1}^{n} \sigma_{X_i}^2 + 2\sum_{i}^{n-1}\sum_{j>i}^{n} Cov(X_i, X_j) \qquad \text{(A.11)}$$

If the variance term in eqn. (A.11) is substituted by the first line of eqn. (3.5) the following equation for the joint cumulant of order two for a set of $X_1,...,X_n$ RVs results

$$cum_2^{(X_1+X_2+...+X_n)} = \sum_{i=1}^{n}\left(\mu_2^{X_i} - \left(\mu_1^{X_i}\right)^2\right) + 2\sum_{i}^{n-1}\sum_{j>i}^{n} Cov(X_i, X_j) \qquad \text{(A.12)}$$

3rd order joint cumulants

The proof of eqn. (3.15) can be done in a similar way as already expressed. Applying eqn. (3.8) the 3rd joint cumulant for two RVs X and Y is:

$$\begin{aligned} cum_3^{(X+Y)} &= \binom{3}{0} cum^{(X,X,X)} + \binom{3}{1} cum^{(X,X,Y)} + \binom{3}{2} cum^{(X,Y,Y)} + \binom{3}{3} cum^{(Y,Y,Y)} = \\ &= cum^{(Y,Y,Y)} + 3cum^{(X,Y,Y)} + 3cum^{(X,X,Y)} + cum^{(X,X,X)} \end{aligned} \qquad \text{(A.13)}$$

Eqn. (A.13) can again be extended to the case of "n" non independent RVs by substituting RV X with RVs A + B and RV Y with RVs C + D and applying eqn. (A.3)

$$cum_3^{(A+B+C+D)} = cum^{(C+D,C+D,C+D)} + 3cum^{(A+B,C+D,C+D)} + 3cum^{(A+B,A+B,C+D)} + cum^{(A+B,A+B,A+B)} =$$
$$= cum^{(C,C,C)} + cum^{(C,C,D)} + cum^{(C,D,C)} + cum^{(C,D,D)} + cum^{(D,C,C)} + cum^{(D,C,D)} + cum^{(D,D,C)} +$$
$$+cum^{(D,D,D)} + 3cum^{(A,C,C)} + 3cum^{(A,C,D)} + 3cum^{(A,D,C)} + 3cum^{(A,D,D)} + 3cum^{(B,C,C)} +$$
$$+3cum^{(B,C,D)} + 3cum^{(B,D,C)} + 3cum^{(B,D,D)} + 3cum^{(A,A,C)} + 3cum^{(A,A,D)} + 3cum^{(A,B,C)} +$$
$$+3cum^{(A,B,D)} + 3cum^{(B,A,C)} + 3cum^{(B,A,D)} + 3cum^{(B,B,C)} + 3cum^{(B,B,D)} + cum^{(A,A,A)} +$$
$$+cum^{(A,A,B)} + cum^{(A,B,A)} + cum^{(A,B,B)} + cum^{(B,A,A)} + cum^{(B,A,B)} + cum^{(B,B,A)} + cum^{(B,B,B)}$$

(A.14)

Now calculating the joint cumulant $cum^{(Y,Y,Y)}$, $cum^{(X,Y,Y)}$ and $cum^{(X,Y,Z)}$ in terms of expectation values with eqn. (3.7) and using the Hasse diagrams for partition lattices for three items [25]:

$$cum^{(Y,Y,Y)} = (-1)^2\, 2!\langle Y\rangle\langle Y\rangle\langle Y\rangle + (-1)^1\, 1!\langle Y\rangle\langle YY\rangle + (-1)^1\, 1!\langle Y\rangle\langle YY\rangle + (-1)^1\, 1!\langle Y\rangle\langle YY\rangle +$$
$$+(-1)^0\, 0!\langle YYY\rangle = \langle YYY\rangle - 3\langle Y\rangle\langle YY\rangle + 2\langle Y\rangle^3$$

(A.15)

$$cum^{(X,Y,Y)} = (-1)^2\, 2!\langle X\rangle\langle Y\rangle\langle Y\rangle + (-1)^1\, 1!\langle X\rangle\langle YY\rangle + (-1)^1\, 1!\langle Y\rangle\langle XY\rangle + (-1)^1\, 1!\langle Y\rangle\langle YX\rangle +$$
$$+(-1)^0\, 0!\langle XYY\rangle = \langle XYY\rangle - 2\langle Y\rangle\langle XY\rangle + 2\langle X\rangle\langle Y\rangle^2 - \langle X\rangle\langle YY\rangle$$

(A.16)

$$cum^{(X,Y,Z)} = (-1)^2\, 2!\langle X\rangle\langle Y\rangle\langle Z\rangle + (-1)^1\, 1!\langle X\rangle\langle YZ\rangle + (-1)^1\, 1!\langle Y\rangle\langle XZ\rangle + (-1)^1\, 1!\langle Z\rangle\langle XY\rangle +$$
$$+(-1)^0\, 0!\langle XYZ\rangle = \langle XYZ\rangle + 2\langle X\rangle\langle Y\rangle\langle Z\rangle - \langle X\rangle\langle YZ\rangle - \langle Y\rangle\langle XZ\rangle - \langle Z\rangle\langle XY\rangle$$

(A.17)

Using eqn. (A.15), eqn. (A.16), eqn. (A.17) and eqn. (A.7), eqn. (A.14) results in

$$cum_3^{(A+B+C+D)} = \langle AAA\rangle + \langle BBB\rangle + \langle CCC\rangle + \langle DDD\rangle + 3\langle AAB\rangle + 3\langle AAC\rangle + 3\langle AAD\rangle + 3\langle AB$$
$$+6\langle ABC\rangle + 6\langle ABD\rangle + 3\langle ACC\rangle + 6\langle ACD\rangle + 3\langle ADD\rangle + 3\langle BBC\rangle + 3\langle BBD\rangle + 3\langle BCC\rangle +$$
$$+6\langle BCD\rangle + 3\langle BDD\rangle + 3\langle CCD\rangle + 3\langle CDD\rangle - \langle A\rangle^3 - \langle B\rangle^3 - \langle C\rangle^3 - \langle D\rangle^3 - 3\langle A\rangle^2\langle B\rangle - 3\langle A\rangle^2$$
$$-3\langle A\rangle^2\langle D\rangle - 3\langle B\rangle^2\langle A\rangle - 3\langle B\rangle^2\langle C\rangle - 3\langle B\rangle^2\langle D\rangle - 3\langle C\rangle^2\langle A\rangle - 3\langle C\rangle^2\langle B\rangle - 3\langle C\rangle^2\langle D\rangle -$$
$$-3\langle D\rangle^2\langle A\rangle - 3\langle D\rangle^2\langle B\rangle - 3\langle D\rangle^2\langle C\rangle - 6\langle A\rangle Cov(A,B) - 6\langle A\rangle Cov(A,C) - 6\langle A\rangle Cov(A,D)$$
$$-6\langle A\rangle Cov(B,C) - 6\langle A\rangle Cov(C,D) - 6\langle A\rangle Cov(B,D) - 6\langle B\rangle Cov(A,B) - 6\langle B\rangle Cov(A,C) -$$
$$-6\langle B\rangle Cov(A,D) - 6\langle B\rangle Cov(B,C) - 6\langle B\rangle Cov(B,D) - 6\langle B\rangle Cov(C,D) - 6\langle C\rangle Cov(A,B) -$$
$$-6\langle C\rangle Cov(A,C) - 6\langle C\rangle Cov(A,D) - 6\langle C\rangle Cov(B,C) - 6\langle C\rangle Cov(B,D) - 6\langle C\rangle Cov(C,D) -$$
$$-6\langle D\rangle Cov(A,B) - 6\langle D\rangle Cov(A,C) - 6\langle D\rangle Cov(A,D) - 6\langle D\rangle Cov(B,C) - 6\langle D\rangle Cov(B,D)$$
$$-6\langle D\rangle Cov(C,D) - 3\langle A\rangle\sigma_A^2 - 3\langle A\rangle\sigma_B^2 - 3\langle A\rangle\sigma_C^2 - 3\langle A\rangle\sigma_D^2 - 3\langle B\rangle\sigma_A^2 - 3\langle B\rangle\sigma_B^2 - 3\langle B\rangle\sigma_C^2$$
$$-3\langle B\rangle\sigma_D^2 - 3\langle C\rangle\sigma_A^2 - 3\langle C\rangle\sigma_B^2 - 3\langle C\rangle\sigma_C^2 - 3\langle C\rangle\sigma_D^2 - 3\langle D\rangle\sigma_A^2 - 3\langle D\rangle\sigma_B^2 - 3\langle D\rangle\sigma_C^2 -$$
$$-3\langle D\rangle\sigma_D^2 - 6\langle A\rangle\langle B\rangle\langle C\rangle - 6\langle A\rangle\langle B\rangle\langle D\rangle - 6\langle A\rangle\langle C\rangle\langle D\rangle - 6\langle B\rangle\langle C\rangle\langle D\rangle$$

(A.18)

$$cum_3^{(X_1+X_2+...+X_n)} = \sum_{i=1}^{n}\langle X_i X_i X_i\rangle - \sum_{i=1}^{n}\langle X_i\rangle^3 - 3\sum_{i=1}^{n}\langle X_i\rangle\sigma_{X_i}^2 + 3\sum_{i=1}^{n}\sum_{k\neq i}^{n}\langle X_i X_i X_k\rangle -$$
$$-3\sum_{i=1}^{n}\sum_{k\neq i}^{n}\langle X_i\rangle^2\langle X_k\rangle - 6\sum_{i=1}^{n}\sum_{k=1}^{n-1}\sum_{j>k}^{n}\langle X_i\rangle Cov(X_k,X_j) \quad (A.19)$$
$$-3\sum_{i=1}^{n}\sum_{k\neq i}^{n}\langle X_i\rangle\sigma_{X_k}^2 - 6\sum_{i=1}^{n-2}\sum_{k>i}^{n-1}\sum_{j>k}^{n}\langle X_i\rangle\langle X_k\rangle\langle X_j\rangle + 6\sum_{i=1}^{n-2}\sum_{k>i}^{n-1}\sum_{j>k}^{n}\langle X_i X_k X_j\rangle$$

Substituting the first three terms of eqn. (A.19) with the equivalent term in eqn. (3.5) the following formulae for the 3rd order joint cumulant with $\left\langle \{X_p\}_{p=i,k,j} \right\rangle = \langle X_i X_k X_j\rangle$ results

$$cum_3^{(X_1+X_2+...+X_n)} = \sum_{i=1}^{n}\left(\mu_3^{X_i} - 3\mu_2^{X_i}\mu_1^{X_i} + 2\left(\mu_1^{X_i}\right)^3\right) + 6\sum_{i=1}^{n-2}\sum_{k>i}^{n-1}\sum_{j>k}^{n}\left(\left\langle \{X_p\}_{p=i,k,j}\right\rangle - \prod_{p=i,k,j}\langle X_p\rangle\right) -$$
$$-6\sum_{i=1}^{n}\sum_{k=1}^{n-1}\sum_{j>k}^{n}\langle X_i\rangle Cov(X_k,X_j) + 3\sum_{i=1}^{n}\sum_{k\neq i}^{n}\left(\left\langle \{X_p\}_{p=i,i,k}\right\rangle - \prod_{p=i,i,k}\langle X_p\rangle - \langle X_i\rangle\sigma_{X_k}^2\right)$$
$$(A.20)$$

4th order joint cumulants

The following brief proof follows the same way as shown for the 3rd order joint cumulant, starting with RVs A, B, C and D. Applying eqn. (3.8), the 4th order joint cumulant, with $X = A + B$ and $Y = C + D$ is

$$cum_4^{(A+B+C+D)} = cum^{(C+D,C+D,C+D,C+D)} + 4cum^{(A+B,C+D,C+D,C+D)} +$$
$$+6cum^{(A+B,A+B,C+D,C+D)} + 4cum^{(A+B,A+B,A+B,C+D)} + cum^{(A+B,A+B,A+B,A+B)} \quad (A.21)$$

The joint cumulant of $cum^{(Y,Y,Y,Y)}$, $cum^{(X,Y,Y,Y)}$, $cum^{(X,X,Y,Y)}$, $cum^{(X,Y,Z,Z)}$ and $cum^{(W,X,Y,Z)}$ follows in terms of expectation values with eqn. (3.7), using the Hasse diagrams for four items in [25] and applying eqns (A.1), (A.2) and (A.6):

$$cum^{(Y,Y,Y,Y)} = ... = \langle YYYY\rangle - 4\langle Y\rangle\langle YYY\rangle + 3\langle Y\rangle^4 + 6\langle Y\rangle^2\sigma_Y^2 - 3\sigma_Y^4 \quad (A.22)$$

$$cum^{(X,Y,Y,Y)} = ... = \langle XYYY\rangle + 3\langle Y\rangle^2 Cov(X,Y) + 3\langle X\rangle\langle Y\rangle\sigma_Y^2 + 3\langle X\rangle\langle Y\rangle^3 -$$
$$-\langle X\rangle\langle YYY\rangle - 3\langle Y\rangle\langle XYY\rangle - 3Cov(X,Y)\sigma_Y^2 \quad (A.23)$$

$$cum^{(X,X,Y,Y)} = ... = \langle XXYY\rangle - 2\langle X\rangle\langle XYY\rangle - 2\langle Y\rangle\langle XXY\rangle + 3\langle X\rangle^2\langle Y\rangle^2 +$$
$$+4\langle X\rangle\langle Y\rangle Cov(X,Y) + \langle X\rangle^2\sigma_Y^2 + \langle Y\rangle^2\sigma_X^2 - 2\left(Cov(X,Y)\right)^2 - \sigma_Y^2\sigma_X^2 \quad (A.24)$$

$$cum^{(X,Y,Y,Z)} = \ldots =$$
$$= \langle XYYZ \rangle - \langle X \rangle \langle YYZ \rangle - \langle Z \rangle \langle XYY \rangle - 2\langle Y \rangle \langle XYZ \rangle + 3\langle X \rangle \langle Y \rangle^2 \langle Z \rangle + 2\langle X \rangle \langle Y \rangle Cov(Y,Z) +$$
$$+2\langle Y \rangle \langle Z \rangle Cov(X,Y) + \langle X \rangle \langle Z \rangle \sigma_Y^2 - Cov(X,Z)\sigma_Y^2 - 2Cov(X,Y)Cov(Y,Z) + Cov(X,Z)\langle Y \rangle$$

(A.25)

$$cum^{(W,X,Y,Z)} = \ldots = \langle WXYZ \rangle - \langle W \rangle \langle XYZ \rangle - \langle Y \rangle \langle WXZ \rangle - \langle Z \rangle \langle WXY \rangle - \langle X \rangle \langle WYZ \rangle +$$
$$+3\langle Y \rangle \langle Z \rangle \langle X \rangle \langle W \rangle + Cov(W,Z)\langle X \rangle \langle Y \rangle + \langle W \rangle \langle Z \rangle Cov(X,Y) + Cov(W,Y)\langle X \rangle \langle Z \rangle +$$
$$+Cov(W,X)\langle Z \rangle \langle Y \rangle + \langle W \rangle \langle X \rangle Cov(Z,Y) + \langle W \rangle \langle Y \rangle Cov(X,Z) - Cov(W,Z)Cov(X,Y) -$$
$$-Cov(W,Y)Cov(X,Z) - Cov(W,X)Cov(Z,Y)$$

(A.26)

Applying eqn. (A.3) and eqn. (A.4) on eqn. (A.21), replacing each term with the right term of eqn. (A.22) to (A.26), rearranging and summing up equal terms, the following extended equation to the case of "n" non independent RVs can be proposed if additionally eqn. (A.6) and (3.5) is considered:

$$cum_4^{(X_1+X_2+\ldots+X_n)} = \sum_{i=1}^{n}\left(\mu_4^{X_i} - 4\mu_3^{X_i}\mu_1^{X_i} - 3\left(\mu_2^{X_i}\right)^2 + 12\mu_2^{X_i}\left(\mu_1^{X_i}\right)^2 - 6\left(\mu_1^{X_i}\right)^4\right) +$$
$$+4\sum_{i=1}^{n}\sum_{k\neq i}^{n}\left(\langle X_i \rangle\left(3\prod_{p=i,i,k}\langle X_p \rangle + 3\langle X_k \rangle \sigma_{X_i}^2 - \langle X_k{}^3 \rangle + 1.5\langle X_i \rangle \sigma_{X_k}^2\right) + \left\langle \underset{p=i,i,i,k}{\{X_p\}} \right\rangle\right) -$$
$$-24\sum_{i=1}^{n-2}\sum_{k>i}^{n}\sum_{j=i}^{n-1}\sum_{\substack{u>j;j\geq i+k-1\\ \{ik\}\neq\{ju\}}}^{n} Cov(X_i,X_k)Cov(X_j,X_u) + 12\sum_{i=1}^{n}\sum_{k=1}^{n-1}\sum_{j>k}^{n}\left(\langle X_i \rangle^2 - \sigma_{X_i}^2\right)Cov(X_k,X_j) +$$
$$+6\sum_{i=1}^{n-1}\sum_{k>i}^{n}\left(\left\langle \underset{p=i,i,k,k}{\{X_p\}} \right\rangle - \prod_{p=i,k}\sigma_{X_p}^2 - 2Cov(X_i,X_k)^2 + 3\left(\prod_{p=i,k}\langle X_p \rangle\right)^2\right) - 12\sum_{i=1}^{n}\sum_{k=1}^{n}\sum_{j\neq k}^{n}\langle X_i \rangle \left\langle \underset{p=k,k,j}{\{X_p\}} \right\rangle$$
$$+24\sum_{j>k}^{n-1}\sum_{u>j}^{n}\left(\sum_{i=1}^{n-3}\sum_{k>i}^{n-2}\left(\left\langle \underset{p=i,k,j,u}{\{X_p\}} \right\rangle + 3\prod_{p=i,k,j,u}\langle X_p \rangle\right) - \sum_{i=1}^{n}\sum_{k=1}^{n-2}\langle X_i \rangle \left\langle \underset{p=k,j,u}{\{X_p\}} \right\rangle\right) +$$
$$+12\sum_{i=1}^{n}\sum_{k\neq i}^{n-1}\sum_{\substack{j>k\\ j\neq i}}^{n}\left(\left\langle \underset{p=i,i,k,j}{\{X_p\}} \right\rangle + \left(\sigma_{X_i}^2 + 3\langle X_i \rangle^2\right)\prod_{p=k,j}\langle X_p \rangle\right) + 24\sum_{i=1}^{n-1}\sum_{k>i}^{n}\sum_{j=1}^{n-1}\sum_{u>j}^{n} Cov(X_i,X_k)\prod_{p=j,u}\langle X_p \rangle$$

(A.28)

where $\{ik\} \neq \{ju\}$ means that the set of indices of variable X_i and X_k have to be different from the set of indices of variable X_j and X_u, irrespective of their sequence.

7.4 Annex D: Reference P4

Paper printed in conference proceedings and presented at EnInnov 2010, 11. Symposium Energieinnovation, February 10th – February 12th, 2010, Graz, Austria

Point Estimate Methods for Probabilistic Optimized Power Flow

A. Gaun, G. Rechberger, and H. Renner

Abstract— **A benchmark of four different Point Estimate (PE) methods is presented in this paper, in order to identify the most suitable method for a load curtailment and / or generation dispatch probabilistic optimized power flow (POPF) with correlated input random variables (RVs). The four different methods are benchmarked against a sequential Monte Carlo Simulation (MCS). The Benchmark is based on calculation time, calculation accuracy and problem applicability. By the use of a real world sub transmission power network with measured probability density distributions of nodal input active powers the calculation accuracy of the proposed PE methods is determined. Furthermore it is demonstrated that the rotational transformation, used to take into account the correlated nodal input RVs, is essential for the proposed methodologies in order to achieve sufficient calculation results compared with the MCS. Summarizing the results of the benchmark, Harr's algorithm is the most adequate method for the proposed short term congestion management POPF problem.**

Index Terms— **Correlated random variables, Expected energy not supplied, Optimized probabilistic power flow, Point Estimate schemes, Uncertainties;**

I. Introduction

Electric power system transmission expansion or restructuring planning is still mainly based on discrete off-peak load and peak load cases in different forecast scenarios. As demonstrated in [1] it is not sufficient to concentrate on these discrete load cases. Therefore, and in order to incorporate uncertainties, that for instance originate forced by high wind energy supply, uncertain load increase forecasts, new market and regulatory strategies and so on, the authors of this paper propose to use probabilistic approaches, e.g. the Probabilistic Load Flow (PLF) methodology [2]-[13] or the Probabilistic Optimized Power Flow (POPF) [14]-[18] approach. Hence, by the use of these probabilistic methodologies it is possible to optimize the load shedding and/or the economic generation dispatch of a power system based on input random variables (RVs). Furthermore, if the input RVs are derived from over one year measured load data, a calculation, including all relevant occurring power system loading situations, can be conducted.

There exist various reasons for dependencies between nodal input powers, for instance these dependencies are caused by common environmental factors such as temperature, sunset, rainfall, season, etc., [19]. To take these dependencies into account in the POPF calculation the authors in [18] use a heuristic approach, where the correlation is modeled in such a way that the demand of a given node is high if the demand of any other node is high as well. The authors of the present paper propose to use the rotational transformation based on the eigenvector of the correlation matrix to transform the set of correlated RVs into an uncorrelated set of input RVs, as it was already applied to the PLF calculation in [9]-[12].

The aim of this paper is to find an adequate methodology for a short term congestion management POPF with minimal load curtailment costs and / or economic generator dispatch costs in order to minimize the effect of outage costs obtained from a (n-1)-contingency analysis in a power system planning algorithm developed by the authors. In particular, four different Point Estimate (PE) methodologies are benchmarked by the help of problem applicability, computational expensiveness and calculation accuracy obtained from a real world example compared against a sequential Monte Carlo Simulation (MCS), taken as reference. The four PE methods are the algorithm of Harr [9], the 2m scheme, the 2m+1 scheme and the 4m+1 scheme (see [10] and [13]). The 3m scheme and the 3m+1 scheme are not investigated, since both usually provide useless complex solution values [13].

This paper is structured as follows. In the subsequent section a brief description of the used PE methodologies is given including relevant calculation formulae. In section III the basic formulas of the linear POPF and the applied solution strategy are presented. A real world sub transmission power system with real measured input RVs is used in section IV to verify the most applicable PE method for the POPF problem. Based on certain (n-1)-contingencies, the calculation accuracies of the proposed methodologies are compared by the use of this power system. Concluding remarks and the results of the PE method benchmark are given in section V. A list of used nomenclature can be found in section VII.

A. Gaun is with the Institute of Electrical Power Systems (IFEA), Graz University of Technology, 8010 Graz, Austria (e-mail: alexander.gaun@tugraz.at).

G. Rechberger is with IFEA, Graz, University of Technology, 8010 Graz, Austria (e-mail: georg.rechberger@tugraz.at).

H. Renner is with the IFEA, Graz University of Technology, 8010 Graz, Austria (e-mail: herwig.renner@tugraz.at).

II. Point Estimate Methods

A. Basic approach and consideration of uncertainties

The PE method is used to calculate moments, characteristics of RVs, of a random quantity that is a function F of RVs. The first PE method was developed by Rosenblueth [8] for Gaussian RVs and has later been extended to deal also with non-symmetric distributions. The PE methods concentrate the statistical information of input RVs on K points of each input $\varphi_{X,i}$. By using Hong's PE method [12], the calculation of F has to be evaluated only K times for each input RV, where the difference among the individual calculations is the deterministic value assigned to the i[th] RV $x_{i,k}$, while the remaining m-1 input RVs are fixed to their corresponding mean value. Hence, in the present paper the OPF for all schemes derived from Hong [12] can be calculated with following nodal input data $\mu_{X,\{1...m\}\forall m\neq i}$ and $x_{i,k}$, so that $Z_{i,k} = F\left(\mu_{X,1}, \mu_{X,2}, ..., x_{i,k}, ..., \mu_{X,m}\right)$. The deterministic value to be determined is

$$x_{i,k} = \mu_{X,i} + \xi_{X,i,k}\sigma_{X,i} \tag{1}$$

where the standard location $\xi_{X,i,k}$ and the weight $w_{i,k}$ are obtained by solving the nonlinear system of eqn. (2) [12]

$$\begin{aligned} &\sum_{k=1}^{K} w_{i,k} = \frac{1}{m} \\ &\sum_{k=1}^{K} w_{i,k}\left(\xi_{X,i,k}\right)^{j} = \eta_{X,i,j} \qquad j = 1, ..., 2K-1 \end{aligned} \tag{2}$$

with

$$\eta_{X,i,j} = \frac{M_{j,X_i}}{{\sigma_{X,i}}^{j}} \tag{3}$$

$$M_{j,X_i} = E\left(X_i - \mu_{X,i}\right)^j = \int_{-\infty}^{\infty}\left(X_i - \mu_{X,i}\right)^j \varphi_{X,i} dX_i \tag{4}$$

where $E(\cdot)$ is the mathematical expectation operator, $\eta_{X,i,1} = 0$, $\eta_{X,i,2} = 1$ and for instance the system of (2) can be solved by the procedure developed by the authors of [20].

Since correlations between input RVs are very likely in real world power systems, a rotational transformation can be applied to map the set of correlated RVs into a set of uncorrelated RVs. By the use of the eigenvalues of the correlation matrix $\mathbf{R_X}$ the RVs $\mathbf{U}$ in the standardized eigenspace can be obtained with the linear eqn. (5)

$$\mathbf{U} = \mathbf{T}\left(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}\right) \tag{5}$$

where $\mathbf{T} = \mathbf{V'L}^{-1/2}\mathbf{STD}^{-1/2}$.

The results of all deterministic OPFs for the schemes derived from Hong [12] are added to the moments $\overline{\mu}$ of the output RVs in eqn (6).

$$\overline{\mu}_j = E\left(Z^j\right) \cong \sum_{i=1}^{m}\sum_{k=1}^{K} w_{i,k}\left(Z_{i,k}\right)^j. \tag{6}$$

In Fig. 1 the flow diagram of the PE method, derived from Hong, is shown in order to clarify the methodology including eigenvalue based rotational transformation. The statistical output information, that is the obtained moments by eqn. (6), are expanded to PDF's by the use of a Gram-Charlier Expansion (GCE). The moments of eqn. (6) are converted to cumulants. By the help of the Gram-Charlier Expansion factors, obtained from the central moments based on cumulants, and calculated up to the j[th] order, the PDF of the POPF results can be determined. The details about the used formulae for the GCE can be found in [3] and [21].

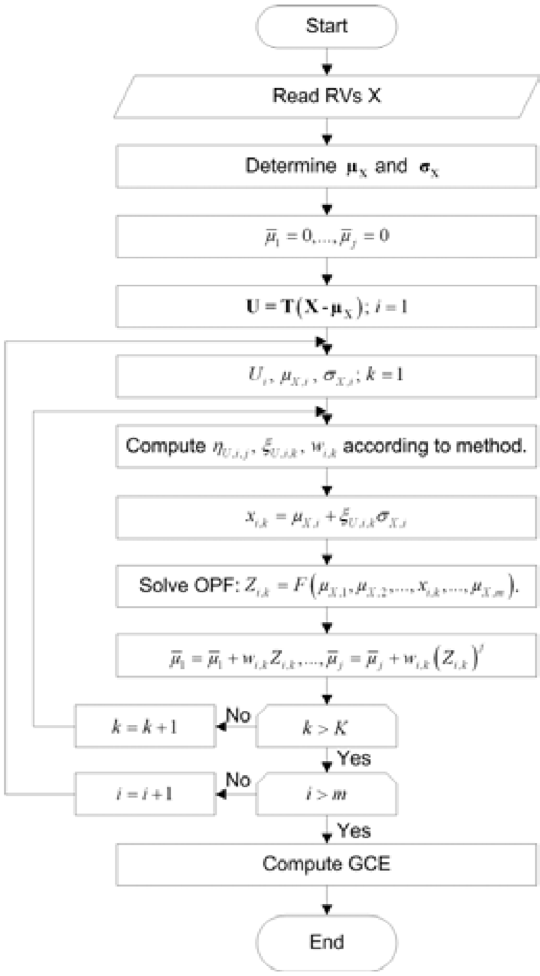

Fig. 1. Flow diagram of the $K \cdot m$ PE methodologies. For the $(K-1) \cdot m + 1$ PE methodologies a final calculation of $Z_{i,K} = F\left(\mu_{X,1}, \mu_{X,2}, ..., \mu_{X,i}, ..., \mu_{X,m}\right)$ before the GCE has to be performed. The weighted results are summed up with $\overline{\mu}_{1...j}$, where w_0 is the corresponding weighting factor. Harr's algorithm see [9].

B. *Harr's algorithm (K = 2, j = 2)*

Harr's PE method reduces the computational burden from 2^m to 2m [9] and thus circumvent the main disadvantage of Rosenblueth's algorithm [8]. Harr's algorithm, which is directly designed for correlated symmetric RVs, includes therefore the calculation of the first two moments $\overline{\mu}_1$ and $\overline{\mu}_2$ about the origin. The computation process is summarized as the following [11]:

The coordinates of the two intersection points are determined as

$$\mathbf{x}_{i\pm} = \boldsymbol{\mu}_\mathbf{X} \pm \sqrt{m}\mathbf{STD}^{-1/2}\mathbf{V}_i = \boldsymbol{\mu}_\mathbf{X} \pm \boldsymbol{\xi}_\mathbf{X} \tag{7}$$

where $i = 1, 2, ...m$. Based on these two points, the corresponding model output values $Z_{i\pm} = F\left(\mathbf{x}_{i\pm}\right)$ are computed. The j[th] (j = 1,2) order moment of the model output is then calculated as

$$\bar{\mu}_j = \frac{1}{m}\sum_{i=1}^{m}\lambda_i \overline{Z_i^j} \tag{8}$$

with j being an exponent and in which

$$\overline{Z_i^j} = \frac{Z_{i+}^j + Z_{i-}^j}{2}. \tag{9}$$

Harr's PE method has advantages due to its simplicity, its low computational burden and its ability to provide real valued solutions for the concentrations.

C. 2m scheme (K = 2, j = 3)

As it is obvious from the scheme's name, this scheme only uses two concentrations for each input RV. The main difference to Harr's algorithm is that the 2m schemes takes the skewness of the input RVs into account and can thus be adopted with non symmetrical input distributions. By analytically solving (2) for the standard locations and weights the subsequent formulae (10) and (11) can be obtained [10], [13]:

$$\xi_{i,k} = \frac{\eta_{U,i,3}}{2} + (-1)^{3-k}\sqrt{m + \left(\frac{\eta_{U,i,3}}{2}\right)^2} \tag{10}$$

with $k = 1,2$ and

$$w_{i,k} = (-1)^k \frac{1}{m}\frac{\eta_{U,i,3-k}}{\eta_{U,i,1} - \eta_{U,i,2}}. \tag{11}$$

As a main drawback of the 2m scheme, but also for all other $K \times m$ schemes and Harr's algorithm, could be seen that $\xi_{i,1}$ and $\xi_{i,2}$, respectively ξ_X depend on the number of uncertain RVs m. Hence, with increasing m, the locations $x_{i,1}$ and $x_{i,2}$ move away from the mean value to maybe unknown regions of the input RVs PDF, thus resulting in unknown values. On the other hand the 2m scheme has the same advantages as Harr's PE algorithm.

D. 2m+1 scheme (K = 3, j = 4)

This scheme is derived from the 3m scheme and this specific variant of Hong's PE method takes into account, as well as the 4m+1 scheme, one more evaluation of function F at the point of all input RVs means $Z_{i,3} = F(\mu_{X,1}, \mu_{X,2}, ..., \mu_{X,i}, ..., \mu_{X,m})$, hence, $(K-1)\cdot m+1$ calculations have to be performed. Solving (2) for K=3, with the standard location $\xi_{i,3}$ is set to zero, and by applying the method of [20] on the formula of the 3m scheme in [13], since there is no analytical solution for the problem, the other standard locations and weights are:

$$\xi_{i,k} = \frac{\eta_{U,i,3}}{2} + (-1)^{3-k}\sqrt{\eta_{U,i,4} - 3\left(\frac{\eta_{U,i,3}}{2}\right)^2} \tag{12}$$

$$w_{i,k} = \frac{(-1)^{3-k}}{\eta_{U,i,k}\left(\eta_{U,i,1} - \eta_{U,i,2}\right)} \tag{13}$$

with $k = 1,2$ and

$$w_{i,3} = \frac{1}{m} - \frac{1}{\eta_{U,i,4} - \left(\eta_{U,i,3}\right)^2}. \tag{14}$$

Setting $\xi_{i,3} = 0$ yields $x_{i,k} = \mu_{X,i}$ and therefore the third location is for all m points the same. Hence, computing this point once with the corresponding weight w_0

$$w_0 = \sum_{i=1}^{m} w_{i,3} = 1 - \sum_{i=1}^{m}\frac{1}{\eta_{U,i,4} - \left(\eta_{U,i,3}\right)^2} \tag{15}$$

is equal to m calculations with weights $w_{i,3}$. The advantages of this scheme are that the standard locations are independent of m and that it takes into account the kurtosis and skewness of all input RVs, while insignificantly increasing the computational expensiveness. Contrariwise for certain standard concentrations non-real valued locations could occur in (12) and hence providing useless unreal solutions. However, in the present real world example only real valued standard concentrations were calculated, as it is the case in most other power system problems, due to the types of input RV distributions [13].

E. 4m+1 scheme (K = 5, j = 6)

In this second considered $(K-1)\cdot m+1$ scheme, moments of input RVs higher than the fourth order are considered. This could increase the accuracy of the obtained calculation results but has also negative effects to the computational burden and a greater probability of obtaining non-real solutions arises. The order of considered moments j is bounded to six in this case, since higher orders lead to a high probability of negative density functions values when applying the GCE method. Note that this fact emerges especially with the use of the PE method due to the obtained standard locations and weights located away from the mean value of the distribution.

As in the 2m+1 scheme, this scheme has also a numerical non-analytical solution derived from the 5m scheme, with $\xi_{i,5} = 0$ and the other $\xi_{i,1...4}$ are the roots of the fourth order polynomial $p(\xi) = \xi^4 + C_3\xi^3 + C_2\xi^2 + C_1\xi + C_0$, with the coefficients being the solution of the linear equations system in (16)

$$\begin{pmatrix} \eta_{U,i,1} & \eta_{U,i,2} & \eta_{U,i,3} & \eta_{U,i,4} \\ \eta_{U,i,2} & \eta_{U,i,3} & \eta_{U,i,4} & \eta_{U,i,5} \\ \eta_{U,i,3} & \eta_{U,i,4} & \eta_{U,i,5} & \eta_{U,i,6} \\ \eta_{U,i,4} & \eta_{U,i,5} & \eta_{U,i,6} & \eta_{U,i,7} \end{pmatrix} \cdot \begin{pmatrix} C_0 \\ C_1 \\ C_2 \\ C_3 \end{pmatrix} = -\begin{pmatrix} \eta_{U,i,5} \\ \eta_{U,i,6} \\ \eta_{U,i,7} \\ \eta_{U,i,8} \end{pmatrix}. \tag{16}$$

Once, $\xi_{i,1...4}$ are obtained, the weights are calculated by solving the linear system in (17)

$$\begin{pmatrix} \xi_{i,1} & \xi_{i,2} & \xi_{i,3} & \xi_{i,4} \\ \xi_{i,1}{}^2 & \xi_{i,2}{}^2 & \xi_{i,3}{}^2 & \xi_{i,4}{}^2 \\ \xi_{i,1}{}^3 & \xi_{i,2}{}^3 & \xi_{i,3}{}^3 & \xi_{i,4}{}^3 \\ \xi_{i,1}{}^4 & \xi_{i,3}{}^4 & \xi_{i,3}{}^4 & \xi_{i,4}{}^4 \end{pmatrix} \cdot \begin{pmatrix} w_{i,1} \\ w_{i,2} \\ w_{i,3} \\ w_{i,4} \end{pmatrix} = \begin{pmatrix} \eta_{U,i,1} \\ \eta_{U,i,2} \\ \eta_{U,i,3} \\ \eta_{U,i,4} \end{pmatrix}. \tag{17}$$

Note that in this case it is also not necessary to perform 5m calculations. By setting $\xi_{i,5} = 0$ the computational burden can be reduced to 4m+1 with the four solutions of (17) and one calculation at $Z_{i,5} = F\left(\mu_{X,1}, \mu_{X,2}, ..., \mu_{X,i}, ..., \mu_{X,m}\right)$ with the corresponding weight

$$w_0 = \sum_{i=1}^{m} w_{i,5}. \tag{18}$$

III. Optimized Power Flow

A probabilistic optimized power flow (POPF) method, for short term congestion management, based on PE method and linear outage costs and economic generation dispatch is applied to minimize the effect of load curtailment and / or generation dispatch of relevant contingencies obtained from a (n-1)-contingency analysis in a reliability calculation algorithm for example designed for power transmission system planning. It is further assumed, that the utility of the power system has no load curtailment philosophy, as for instance the philosophy that has been adopted in [22]. Instead the optimization is based on CIC [23] and linear generation costs [24]. The authors of the present paper use a DC load flow based approach since the system planning tool has to be computationally efficient and the main goal of a system planning is to satisfy the customer's real power demand [3].

The load curtailment is modeled as a negative real power injection at load buses with associated costs adapted from the OPF from [24]. In the present real world example linear cost functions are assumed, although the OPF can also deal with piece wise linear cost functions. The OPF is implemented in that way, that optimal load curtailment and optimal generation dispatch is used to satisfy the given security constraints (21) - (24). Thus the linear programming minimization problem is formulated as follows:

$$\min_{\theta, P_{cur}} \sum_{i=1}^{n} C_{cur,i} \tag{19}$$

with $C_{cur,i} = P_{cur,i} c_1 + P_{cur,i} c_2 t$

subject to

$$\mathbf{B}_{\mathbf{bus}} \times \boldsymbol{\theta} = \mathbf{P}_{\mathbf{g}} \tag{20}$$

$$P_{g,j}^{\min} \le P_{g,j} \qquad \forall j = 1, 2, ..., n_v \tag{21}$$

$$P_{g,j} \le P_{g,j}^{\max} \qquad \forall j = 1, 2, ..., n_v \tag{22}$$

$$\mathbf{B}_{\mathbf{f}} \times \boldsymbol{\theta} \le P_{nr,k} \qquad \forall k = 1, 2, ..., n_t \tag{23}$$

$$-\mathbf{B}_{\mathbf{f}} \times \boldsymbol{\theta} \le P_{nr,k} \tag{24}$$

The linear Optimization Toolbox™ from MATLAB© with the implemented solver is used to solve the linear optimization problem in (19). Furthermore all other calculations in the present paper are also performed with MATLAB©.

IV. Case Study with Real World Example

A 110 kV transmission power system (Fig. 2) is used to verify the accuracy of the proposed methodologies, in order to use the results for a benchmark of the proposed PE methods. The obtained PE method calculation results are compared with a straightforward MCS. Therefore the R² measure is used, which is a statistical measure for how successful the fit of a curve is in explaining the variation of the data. R² is defined as the ratio of the sum of squares of the regression (SSR) to the total sum of squares (SST) $R^2 = SSR/SST$ with SSR is defined as $SSR = \sum_{i=1}^{N}\left(GC_i - \overline{DC}\right)^2$ and SST is defined as $SST = \sum_{i=1}^{N}\left(DC_i - \overline{DC}\right)^2$. R² can take any value between 0 and 1, with a value closer to one indicating that a greater proportion of variance is accounted for by the fit.

TABLE I

Power System Transmission Line Data of the Power Grid in Fig. 2

line	f-t	l	X	P_{nr}	line	f-t	l	X	P_{nr}
		km	mΩ/km	MW			km	mΩ/km	MW
1	1-5	8.60	266	410	21	5-6	4.66	252	225
2	1-8	3.93	262	410	22	5-6	4.70	252	225
3	2-6	7.72	287	281	23	5-9	4.10	141	82
4	2-10	8.24	324	281	24	5-9	4.04	141	82
5	2-12	8.37	291	281	25	5-11	3.02	260	410
6	2-14	8.37	291	281	26	6-7	6.16	390	110
7	3-5	5.90	282	281	27	6-7	6.20	390	110
8	3-5	5.81	282	281	28	6-13	3.64	287	281
9	3-17	1.95	118	120	29	6-16	4.50	114	104
10	3-17	2.24	128	134	30	7-14	3.71	279	281
11	3-19	3.60	123	134	31	7-14	3.72	279	281
12	3-21	3.90	119	104	32	7-18	3.61	285	281
13	3-21	4.54	128	134	33	8-10	7.55	294	281
14	3-23	2.08	133	104	34	8-11	1.76	256	410
15	3-23	2.59	126	134	35	12-13	3.64	287	281
16	4-7	6.99	296	281	36	15-16	0.45	114	134
17	4-18	3.37	297	281	37	16-17	1.64	120	120
18	4-19	3.10	116	134	38	20-21	0.96	116	134
19	4-20	2.26	136	104	39	22-23	1.38	142	104
20	4-22	3.15	129	104	40	2-14	8.37	291	281

f … from bus, t … to bus

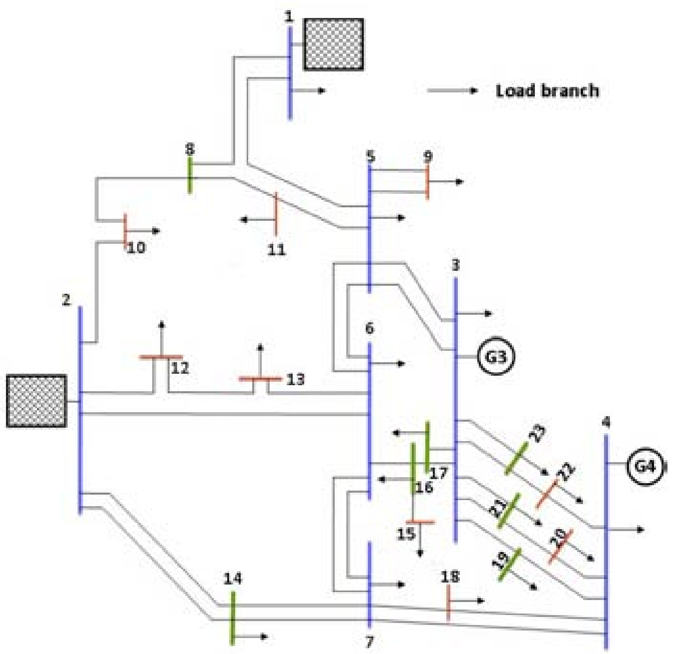

Fig. 2. Real world 110 kV sub transmission power system

Fig. 2 shows the 110 kV sub transmission grid with 23 buses, 23 over head transmission lines, 17 cables, two generators, two connections to other grids, 21 load buses and a maximum load in the base case of 791,11 MW, where bus one is the reference bus. In Table I the required line data is given and Table II contains the 1st, 2nd, 3rd and the 4th moment of the bus injected powers [25] for the base case. These input $\varphi_{X,i}$ are obtained from 8760 values of real measured hourly nodal powers and represent a whole year. Moreover the 25th, 50th and the 75th percentile of $\mathbf{R_X}$ for all input RVs are 0.23, 0.46 and 0.70, which verifies that there is at least a light correlation between all input nodal RVs and the application of the rotational transformation is necessary. With the use of the GCE and the according cumulants, the PDF of each nodal power can be reconstructed.

Table III shows the generation limits for the POPF and Table IV gives the load bus individual CIC for the load curtailment cost function and generation cost function with the two cost function factors c_1 and c_2, where the cost functions of buses 15 and 16 are modified to authors-defined values.

In order to show the applicability of the proposed PE based POPF the expected energy not supplied (EENS) at bus 15, obtained from the proposed methodologies, are compared with the sequential 8760 MC simulations, one for each hour of the year. For demonstration purposes it is assumed that power transmission line 36 is switched off (base case). The results of the calculations are depicted in Fig. 3 where the PDF of the obtained EENS results for an assumed outage of one hour are illustrated. Additionally the R² measure is quoted in the legend of Fig. 3. Basically, it is almost impossible with the moments limited GCE to model a distribution that has a heavy tail, as the one presented in Fig. 3, Fig. 4 and Fig. 5. The heavy tail is the result of having non-negative EENS (grey marked area in the depicted figures) with certain probability of zero EENS. For more information concerning this topic see [26]. A visual comparison of the obtained results shows that the 2m scheme yields in the worst fit of the MCS. All three remaining methods, the Harr's algorithm and the two $(K-1)\cdot m+1$ schemes, provide quite similar R² measures with the 4m+1 yielding in the best approximation of the MCS due to the consideration of moments with order higher than four. Since all input $\varphi_{X,i}$ could be modelled more or less

TABLE III
GENERATION DATA FOR THE POPF

ST	P_g^{max}	P_g^{min}
	MW	MW
G1	632	-210
G2	200	-90
G3	440	0
G4	280	-10

TABLE IV
CIC FACTORS [23] AND GENERATION COST FACTORS [24]

ST	c_1	c_2	ST	c_1	c_2
	€/kW	€/kWh		€/kW	€/kWh
1	2.90	23.46	11	1.38	13.38
G1	9.20	110.40	12	2.28	22.47
G2	9.20	110.40	13	2.28	22.47
3	1.88	18.43	14	2.28	22.47
G3	9.20	94.40	15	1.88	18.43
4	1.88	18.43	16	0.98	18.60
G4	9.20	94.40	17	1.88	18.43
5	0.92	9.34	18	1.88	18.43
6	0.66	7.32	19	1.88	18.43
7	2.90	23.46	20	1.88	18.43
9	1.38	13.38	21	2.28	22.47
10	1.38	13.38	22	2.28	22.47
			23	2.28	22.47

accurate by a normal distribution, Harr's algorithm yields in reasonable results. Due to this fact, the expectation E(X) of Harr's PDF is the best approximation of the MCS expectation with an absolute failure of 0.01 MWh. The 2m scheme, where the skewness of the input RVs is incorporated in the calculation, results in an overestimation at the zero EENS region and therefore in an underestimation of E(X).

Normally outages of power lines last longer than one hour. Hence it is important to calculate E(X) for such outages. In the following example, it is assumed that the outage of line 36 lasts 10 hours. The input RVs for the PE methods are transformed to 10-hours-average RVs values via a moving average calculation and the results of the deterministic OPF values are multiplied with 10 to compute the ENS. In the OPF calculation the duration of 10 hours is considered via a multiplication of the variable cost factor c_2 by ten (see eqn. (19)). The reference MCS is performed in such a way that for each measured time step of the original RVs the sum

TABLE II
MOMENTS [25] OF NODAL POWER DATA

Bus	μ_1^X	μ_2^X	μ_3^X	μ_4^X	Bus	μ_1^X	μ_2^X	μ_3^X	μ_4^X
	MW					MW			
1	-19.7	419.0	-9.6e3	2.3e5	11	-31.2	1.0e3	-3.7e4	1.4e6
G1	460.8	2.5e5	1.4e8	7.8e10	12	-31.4	1.1e3	-3.9e4	1.5e6
G2	19.8	1.1e4	1.1e6	4.7e8	13	-6.5	50.5	-451.9	4.5e3
3	-28.2	848.2	-2.7e4	9.1e5	14	-36.3	1.4e3	-6.0e4	2.7e6
G3	147.9	2.6e4	4.7e6	9.1e8	15	-4.0	21.9	-144.7	1.1e3
4	-24.8	655.9	-1.8e4	5.3e5	16	-30.1	953.9	-3.2e4	1.1e6
G4	114.1	1.6e4	2.3e6	3.3e8	17	-21.7	508.6	-1.3e4	3.4e5
5	-24.6	652.0	-1.8e4	5.4e5	18	-24.1	610.0	-1.6e4	4.5e5
6	-60.5	3.7e3	-2.4e5	1.5e7	19	-21.2	485.9	-1.2e4	3.1e5
7	-47.3	2.4e3	-1.2e5	6.8e6	20	-18.5	364.2	-7.6e3	1.7e5
9	-10.1	112.5	-1.4e3	1.8e4	21	-24.5	658.2	-1.9e4	6.0e5
10	-15.0	263.6	-5.3e3	1.2e5	22	-28.6	909.1	-3.2e4	1.2e6
					23	-23.2	579.6	-1.6e4	4.4e5

of the ENS of ten consecutive OPF calculations with the following 10 time steps are calculated. The obtained results are depicted in Fig. 4.

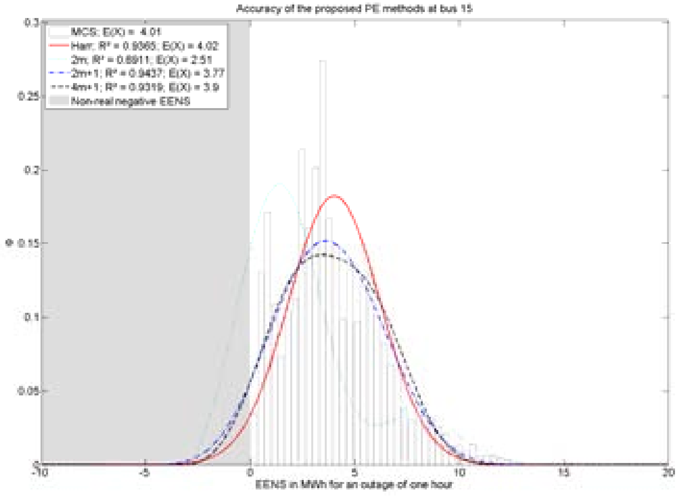

Fig. 3. Comparison of different PE methods with the load at bus 15, for an one hour outage of line 36.

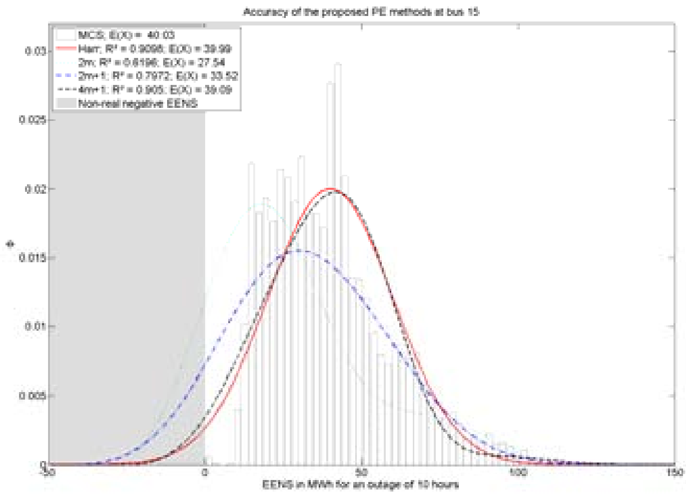

Fig. 4. Comparison of different PE methods for an assumed 10-hour-outage of line 36 from bus 15 to 16.

For the 10-hour-outage Harr's algorithm provides the best E(X) approximation and the highest R² measure. Again the 2m scheme is the worst method and the two $(K-1)\cdot m+1$ schemes result in practical fits.

The single outage of line 36 doesn't take into account any correlated input RV in the calculation. In order to consider correlation, the loadability of the real world power system is increased by a load factor of two for each load. Moreover, the "generation" at bus 2 is also increased by the factor 2 (high load case), as well as the limits of G2, which are set to + 400 MW and – 80 MW. With the increased loadability it is possible to incorporate load dependent overloading of power lines and thus to show the effect of correlated nodal input RVs. Therefore line 37 is switched off for one hour and the results of the different PE methods are shown in Fig. 5. In this example the two $(K-1)\cdot m+1$ schemes are not depicted, because these schemes don't result in any load curtailment at any load bus.

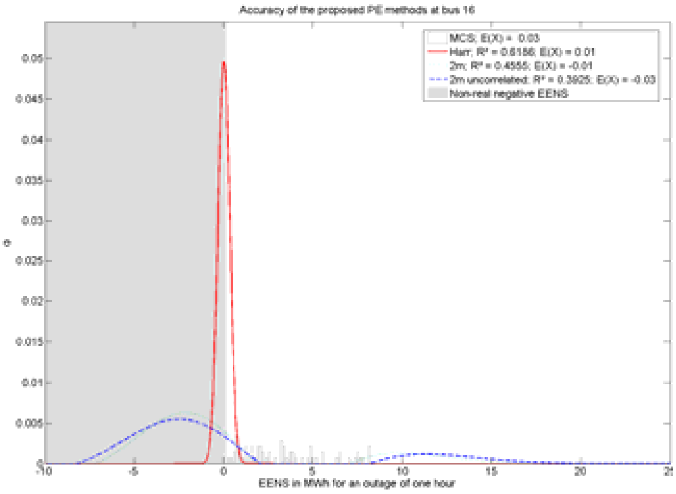

Fig. 5. Comparison of the different PE methods with higher loadability and switched off line 37. This results in a load curtailment at bus 16.

The worst fit of the MCS is from the 2m uncorrelated scheme, where no correlations between the input RVs were considered. Both 2m schemes result in a negative expectation of the ENS and therefore both schemes are inapplicable for the proposed POPF, since in most applications of the POPF similar EENS PDFs, with high probability of zero EENS, have to be expected. Harr's algorithm yields again in the best approximation of the MCS simulation with respect to the R² measure and to E(X).

Calculating the short term congestion management expectation of the power system loads, when line 37 is outaged for 8 hours, following computations have to be performed. With the sequential MCS OPF which is based on the average values of measured load data, the optimal load curtailment for different overloading durations is calculated. Due to the CIC it is optimal to curtail the load at buses 15 and 16, if a certain level of energy is reached. This occurs in the present example whenever the curtailment duration exceeds six hours. The E(X) for the curtailed load at bus 15 is 0.1639 MW and the E(X) at bus 16 is 0.2274 MW, based on the results of the sequential MCS.

In order to calculate the POPF results for longer durations, including dependencies between input RVs, several POPF calculations are necessary, one for each single hour of the present example, starting with t = 1 and ending with t = 8, using the previous described moving average approach. The obtained results from the POPF calculations are further used with a power transfer distribution factor based simulation with the measured input RVs, where the durations of the overloadings of power line 29 are calculated. Based on the frequency of these different long lasting overloading situations and on the expectation obtained from the POPF calculations, the E(X) of the load on bus 15 is 0.0042 MW and the E(X) at bus 16 is 0.0481 MW. This calculation could be enhanced with a MCS based computation of the outage durations and with consideration of the distribution of the

POPF results and not only the expectation, but this is not the main topic of this paper and is therefore not further explained. As one could see from the obtained results this calculation has the drawback of a high inaccurateness, especially with small probabilities of load curtailment or generation dispatch. Moreover the described POPF based calculation is computationally expensive for long durations of line outages. Therefore the authors propose to use the results of the one-hour-outage durations for the calculations in the by the authors developed genetic algorithm based power system planning algorithm. Although the results of the one-hour-outage calculation are only a rough estimation and don't calculate an overall optimal generation dispatch or load curtailment, the goal of the power system planning tool is to find power system configurations without violating any security constraints, and this can be realized with the developed presented short term congestion management algorithm via punishing configurations with load curtailment in the genetic algorithm.

V. Conclusion

From the point of view of computation time, the benchmarked methodologies can be ranked in the following way, starting with the lowest computational burden and ending with the computational most expensive scheme. Harr's PE algorithm and the 2m scheme have the same computational burden of twice the number of uncertain RVs. One calculation more is needed to perform the 2m+1 scheme POPF, followed by the 4m+1 scheme, which has a computational burden of almost twice the 2m+1 scheme.

A different ranking yields if the accuracy of the proposed methodologies is taken into consideration. Again, starting with the scheme with the highest R^2 measure for the high load case, Harr's algorithm is on the first place. Although Harr's PE method is only for symmetric input RVs it has major advantages when it is applied on the POPF algorithm with multi modal Non-Gaussian input RVs. It provides the best approximation (highest R^2 and best E(X)) for distributions with high probability of zero EENS and low probability of positive nonzero EENS, which is the most likely case for the proposed POPF application in real world power systems. Although the 2m+1 and 4m+1 scheme don't result in any load curtailment in the proposed example, other simulation results show that their R^2 measure is quite equal to Harr's R^2. Basically, the results of the POPF calculation depend on the input distributions and on the amount of the curtailed load or dispatched generation. The worst fit of the obtained PDFs origins when the 2m scheme is used.

Summarizing the problem applicability, Harr's algorithm and the 2m scheme has the drawback that its standard locations depend on the number of uncertain RVs and thus could result in undefined values of input RVs. The $(K-1)\cdot m+1$ schemes may compute non real solutions and in addition the 4m+1 scheme could have negative PDF values when applying the GCE. The last mentioned disadvantages don't affect Harr's PE algorithm.

Considering all listed benchmark inputs, following ranking of the proposed PE methods for the POPF problem could finally conclude this paper: Harr's algorithm is the only method that yields in reasonable results. All remaining three PE schemes are more or less inapplicable for the POPF problem because they have either a too high calculation error (2m scheme) or they don't find all curtailed loads ($(K-1)\cdot m+1$ schemes).

VI. References

[1] G. Rechberger, A. Gaun, and H. Renner, "Systematical Determination of Load Flow Cases for Power System Planning", *in. Proc. 2009 IEEE Bucharest PowerTech Conference*, 6 pages

[2] J. Usaola, "Probabilistic load flow with wind production uncertainty using cumulants and Cornish–Fisher expansion", *Int. J. Electr. Power Energ Syst.*, vol. 31, pp.474-481, Feb. 2009

[3] P. Zhang, and S. T. Lee, "Probabilistic Load Flow Computation Using the Method of Combined Cumulants and Gram-Charlier Expansion"; *IEEE Trans. Power Systems*, vol.19, no 1, pp. 676-682, Feb. 2004

[4] C.-L. Su, "Probabilistic Load-Flow Computation Using Point Estimate Method", *IEEE Transactions on Power Systems*, vol. 20, no. 4, pp.1843-1851, Nov. 2005

[5] ———, "A New Probabilistic Load Flow Method", in *Proc. IEEE Power Eng. Soc. General Meeting*, pp. 1795-1800, Jun. 2005

[6] A.M. Leite da Silva, V.L. Arienti, and R.N. Allan, "Probabilistic Load Flow Considering Dependence Between Input Nodal Powers", IEEE Transactions on Power Apparatus and Systems, vol. PAS-103, no. 6, pp. 1524-1530, June 1984

[7] L. Min, and P. Zhang, "A Probabilistic Load Flow with Consideration of Network Topology Uncertainties", in *Intelligent Systems Applications to Power Systems*, pp. 1-5, *ISAP 2007*

[8] E. Rosenblueth, "Point estimates for probability moments", in *Proc. Nat. Acad. Sci. U.S.A.*, vol. 72, no. 10, pp. 3812-3814, Oct. 1975

[9] M. E., Harr, "Probabilistic estimates for multivariate analysis", Appl. Math. Modelling, Vol. 13, pp 313-318, May 1989

[10] P. Caramia, G. Carpinelli, and P. Varilone "Point estimate schemes for probabilistic three-phase load flow", *Electr. Power System. Res.*, vol. 80, pp. 168-175, 2010

[11] C.-H. Chang, Y.-K. Tung, and J.-C. Yang, "Evaluation of probability point estimate methods", *Applied mathematical modeling*, vol. 19, no 2, pp. 95-105, Feb. 1995

[12] H. P. Hong, "An efficient point estimate method for probabilistic analysis", Reliability Engineering and System Safety, vol. 59, pp. 261-267, 1998

[13] J. M. Morales, and J. Perez-Ruiz, "Point Estimate Schemes to Solve the Probabilistic Power Flow", *IEEE Transactions on Power Systems*, vol. 22, no. 4, pp.1594-1601, Nov. 2007

[14] G. Verbic, and C. A. Canizares, "Probabilistic Optimal Power Flow in Electricity Markets Based on a Two-Point Estimate Method", *IEEE Transactions on Power Systems*, vol. 21, no. 4, pp.1883-1893, Nov. 2006

[15] A. Schellenberg, W. Rosehart, and J. Aguado, "Cumulant Based Probabilistic Optimal Power Flow (P-OPF)", *8th International Conference on Probabilistic Methods Applied to Power Systems*, Ames, Iowa, pp. 506-511, Sep. 2004

[16] ———, "Introduction to Cumulant-Based Probabilistic Optimal Power Flow (P-OPF)", *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp.1184-1186, May 2005

[17] M. Madrigal, K. Ponnambalam, and V. H. Quintana, "Probabilistic Optimal Power Flow" in Proc. 1998, *IEEE Canadian Conference on Electrical and Computer Engineering*, vol. 1, pp.385-388

[18] G. Verbic, A. Schellenberg, W. Rosehart, and C. A. Canizares, "Probabilistic Optimal Power Flow Applications to Electricity Markets", Probabilistic Methods Applied to Power Systems, PMAPS 2006, pp. 1-6, June 2006

[19] A.M., Leite da Silva, V.L., Arienti, and R.N. Allan, "Probabilistic Load Flow Considering Dependence Between Input Nodal Powers", *IEEE*

Transactions on Power Apparatus and Systems, vol. PAS-103, no. 6, pp. 1524-1530, June 1984
[20] A. C. Miller, and T. R. Rice, “Discrete approximations of probability distributions”, *Management Science*, vol. 29, no. 3, pp. 352-362, March 1983
[21] A., Hald, “The Early History of the Cumulants and the Gram-Charlier Series”, *International Statistical Review*, vol. 68, no. 2, pp 137-153, 2000
[22] N.A.A. Samaan, “Reliability assessment of electric power system using genetic Algorithms”, Dissertation, Texas A&M University, Aug. 2004
[23] M. Hyvärinen, "Electrical networks and economies of load density", TKK Dissertation 146, Department of Electrical Engineering, Helsinki University of Technology TKK, 2008
[24] MATPOWER, A Matlab System Simulation Package. [online] Available: http://www.pserc.cornell.edu/matpower/matpower.html.
[25] M. G. Kendall, *The advanced Theory of Statistics*, London: Charles Griffin & Company limited, 2nd ed., vol. 1, 1945
[26] E.S., Pearson, “Some problems arising in approximating to probability distributions, using moments”, *Biometrika*, vol. 50, no. 1/2, pp. 95-112, June 1963

VII. Nomenclature

The following nomenclature is used in the paper:

$\mathbf{B_{bus}}$	Nodal susceptance matrix.
$\mathbf{B_f}$	“from” bus susceptance matrix.
C_{cur}	Cost of the curtailed or dispatched active power.
$C_{0,\dots P}$	Coefficients of a (P+1)th order polynomial $p(\xi)$.
c_1	Constant factor either of the linear customer interruption cost (CIC) function or of the linear generation cost function.
c_2	Variable cost factor.
$\overline{DC}$	Mean value of the Monte Carlo Simulation (MCS).
DC_i	The ith point’s value of the cumulative distribution function (CDF), obtained from MCS.
E(X)	Mathematical expectation of RV X.
F	Function for calculating output values from input RVs.
GC_i	The ith point’s value of the CDF, obtained from Gram-Charlier Expansion (GCE).
j	Number of moments for the GCE.
K	Number of points for the PE methodologies.
L	Eigenvalue matrix, with $\mathbf{L} = diag(\lambda_1, \lambda_2, \dots, \lambda_m)$.
M_{j,X_i}	The jth order Central moment of ith RV X_i.
m	Number of uncertain random variables.
n	Total number of buses in the power system.
n_t	Total number of power transmission lines.
n_v	Total number of buses that have installed generation.
P_{cur}	Amount of curtailed or dispatched active power.
$\mathbf{P_g}$	Real power generation at buses n_v.
$\mathbf{P_g^{min}}$	Minimum available generation of the power system.
$\mathbf{P_g^{max}}$	Maximum available generation at buses n_v.
$\mathbf{P_{nr}}$	Nominal rating of all power system transmission lines.
$\mathbf{R_X}$	Correlation matrix of RVs X.
R^2	Goodness of fit measure.
STD	Diagonal matrix with σ_X^2 of X.
SSR	Sum of squares of the regression.
SST	Total sum of squares.
T	Transformation matrix for RVs from original space into standardized eigenspace $\mathbf{U}$.
t	Used duration of the line outage
V	Eigenvector matrix, with $V = [v_1, v_2, \dots, v_m]$.
w	Weight for the PE method.
X	Input random variables (RVs).
$\mathbf{x_i}$	Coordinate of the ith point location for PE method.
Z_i	Result of the ith deterministic Optimized Power Flow.
$\boldsymbol{\theta}$	Voltage angles of all buses in the power system.
$\varphi_{X,i}$	Probability density function (PDF) of ith RV X_i.
$\eta_{X,i}$	The ith standard central moment of RV X_i.
λ_i	Eigenvalue of the ith RV in the correlation matrix.
$\overline{\mu}_j$	The jth order moment of ith Z_i.
$\mu_{X,i}$	Mean value of ith RV X_i.
μ_i^X	The ith moment of RV X_i, with $\mu_1^X = E(X)$.
v_i	The eigenvector of RV X_i.
$\sigma_{X,i}$	Standard deviation of ith RV X_i.
ξ	Standard location of a point for the PE method.
$\boldsymbol{\xi_X}$	Vector of standard locations for Harr’s algorithm.

VIII. Biographies

Alexander Gaun was born in Kufstein, Austria, in 1979. He received his diploma degree in 2005 at Graz University of Technology, Graz. He is currently a scientific university assistant at the Institute of Electrical Power Systems. His research interests include electrical power system planning, reliability computation, genetic algorithms and electromagnetic compatibility.

Georg Rechberger was born in Linz, Austria, in 1978. He received his diploma degree in 2005 at Graz University of Technology, Graz, where he currently holds a position as scientific university assistant at the Institute of Electrical Power Systems. His main work in research and teaching is in the field of electrical power system planning.

Herwig Renner was born in Graz, Austria, in 1965. He received his doctoral degree in 1995 at Graz University of Technology, Graz, where he currently holds a position as associate professor at the Institute of Electrical Power Systems. His main work in research and teaching is in the field of electrical power system analyses with special emphasis on quality, supply reliability and power system control and stability. Since 2009 he is visiting docent at the Helsinki University of Technology (HUT), Department of Electrical Engineering.

7.5 Annex E: Reference P5

Paper printed in conference proceedings and presented at the 2010 IEEE PES General Meeting, July 26th - July 29th, 2010, Minneapolis, Minnesota, USA

Enumeration Based Reliability Assessment Algorithm Considering Nodal Uncertainties

A. Gaun, *Student Member, IEEE,* G. Rechberger, *Student Member, IEEE,* H. Renner, and M. Lehtonen

***Abstract*— A novel fast enumeration based reliability assessment algorithm based on a probabilistic load flow approach considering nodal uncertainties is proposed in this paper. The frequency and duration indices for each load point and the expected energy not supplied as well as the expected interruption costs are derived with the proposed methodology. By reducing the number of necessary system states a huge decrease of computation time is achieved while maintaining an adequate degree of accuracy. In order to verify the accuracy and the computational expensiveness of the proposed algorithm a commercial available program is used to compare the calculation results of a utility driven sub transmission power system. It is demonstrated that a fast calculation of the load indices, including station originated outages, can be performed, based on uncertain nodal powers, for instance nodal injections from wind farms or other stochastic power system loads, with a high degree of accuracy.**

***Index Terms*— Expected interruption cost, frequency and duration method, optimized probabilistic power flow, probabilistic load flow, reliability enumeration method, uncertainties**

I. Introduction

THE uncertainties that originate whenever power system expansions are planned or power systems are restructured, for instance forced by high wind energy supply, uncertain load increase forecasts, new market and regulatory strategies and so on force an adapted approach to solve these arising challenges. Reliability analysis models are useful in this case for comparative evaluation of expansion plans, justification of additional transmission lines, cost/benefit analysis of expansion plans and estimation of interruption costs [1]. Though in order to develop adequate methodologies it is not sufficient to concentrate on discrete off-peak load and peak load cases in different forecast scenarios as demonstrated in [2].

Methods for reliability assessment are classified into two categories: the Monte Carlo simulations (MCS) and enumerative methods. The MCS method [3] consists of randomly selecting system component states and load states, including uncertainties, and subsequently simulating contingency effects applied with the sequential approach or the random approach. Since the MCS method is computationally expensive there are different techniques to speed up the convergence of the simulation [1]. Enumeration methods calculate outage events that may have adverse effects on reliability indices of the power system and then analyze these events to quantify their impact. There are also techniques to obtain adequate annual reliability indices e.g. the load characteristic approach, where each event is analyzed with different load flow situations obtained by a (stepwise) reproduction of the load duration curve [4].

The aim of this paper is to present a novel fast methodology for power transmission system reliability assessment with uncertain nodal powers applying the enumeration method and probabilistic load flow (PLF) method. The frequency and duration indices (F&D) for each load point (lp) in the transmission system, as for instance demonstrated in [5] and [6], are obtained. A DC load flow approach for all calculations is applied, since the proposed system planning tool has to be computationally efficient and the main goal of system planning is to satisfy the customer's real power demand [7].

This paper is structured as follows: in the subsequent section the novel approaches are described, followed by a brief introduction to cumulant based PLF. The applied calculation methodologies are given in section IV with special focus on the implemented algorithms to perform the state enumeration, the effect analysis and to calculate probabilistic input data based reliability indices. Section V illustrates the method utilized on a 23 bus sub transmission power system and contains information about expected computation time for large power systems. The obtained results are verified and compared with a commercial reliability calculation program [8]. In section VI concluding remarks are given.

II. Reliability Assessment method

System outages, that may affect reliability indices, can occur due to conditions created by multiple component failures in interconnected power systems. The following outage types are typically included in reliability assessment algorithms: independent outages, dependent outages due to power transmission line capacity reasons, common mode outages and substation originated outages [9].

Probabilistic F&D assessment of transmission systems can

A. Gaun, G. Rechberger, and H. Renner are with the Institute of Electrical Power Systems, Graz University of Technology (TUG), 8010 Graz, Austria (e-mail: alexander.gaun@tugraz.at, georg.rechberger@tugraz.at, herwig.renner@tugraz.at).

M. Lehtonen is with the Department of Electrical Power Systems, Helsinki University of Technology (HUT), 02150 Espoo, Finland (e-mail: matti.lehtonen@tkk.fi)

give an indication of the frequency of load supply interruption and of its expected duration and thus provides results to distinguish system development scenarios. The basic F&D indices outage duration T and index f, which is defined as the transition frequency between failure and success states in 1/yr. T is calculated from the ratio $\Pr/f$ where $\Pr$ in h/yr is the transition frequency (average failure rate) λ in 1/yr of each failure state multiplied with the duration of each failure state (average repair duration) r in h. The F&D indices are felt to be more "natural" than other common indices from the customer point of view, and allow a better evaluation of the impact of load supply interruptions [6]. For these reasons, the assessment of F&D indices is implemented in this methodology. Furthermore the expected energy not supplied (EENS) and the expected interruption costs (EIC) are also calculated for each load point.

III. Probabilistic Load Flow

The following brief description provides a rough guidance how the PLF calculation is implemented in the proposed algorithm ([7], [10] and [11]).

Based on the probability density function (PDF) of independent input nodal random variables (RVs) a PLF can be performed. Real measured nodal powers have normally a certain degree of mutual correlation [12]. Methods to incorporate these correlations in the PLF calculation already exist [13], and application on PLF is analyzed by the authors, but this topic is out of the scope of this paper. Independent RVs are generated out of measured values of active power by applying the principal component approach [14]. The coordinates of the original time series in the new coordinate system defined by the principal components is summed with the mean of the original measured signal to maintain a realistic load point time series, taking into account the necessary balance between load and generation. Thus a set of annual load point data with no correlations is created, showing a daily load profile in an equivalent data range as the original measured data set in order to perform a PLF adapted from uncorrelated annual input data. Note that this input data is not valid for any justifiable statement about the analyzed transmission grid. However it can be used to demonstrate the proposed novel calculation methodology.

In a first step, the cumulants of the bus injected powers are determined from the moments of the independent input RVs. In the next step these cumulants are multiplied with the power transfer distribution factors (PTDF) of the grid to obtain the cumulants of the power flows on the lines. The PTDF calculation, a sensitivity analysis of flows to injections, can be easily done by a matrix multiplication.

$$\mathbf{PTDF} = \mathbf{B}^{-1} \times \mathbf{B}^{\mathbf{b}} \tag{1.1}$$

where $\mathbf{B}$ is the reduced nodal susceptance matrix and $\mathbf{B}^{\mathbf{b}}$ is a reduced matrix with the branch suceptances. In this case reduced means that the rows and columns corresponding to the slack bus (reference bus) are eliminated avoiding singularity. Subsequently the discrete line flow cumulants, obtained with the PTDFs, are used to calculate the mean, the variance and the central moments of the power flows on the branches. In the next step, the asymptotic Gram-Charlier Expansion (GCE) factors, up to the highest order of moments (normally between 6 and 9 [7]), are determined. Finally, applying the GCE method, the PDF or cumulative distribution function of each branch active power flow is calculated.

IV. Methodology

A. Fundamental approach

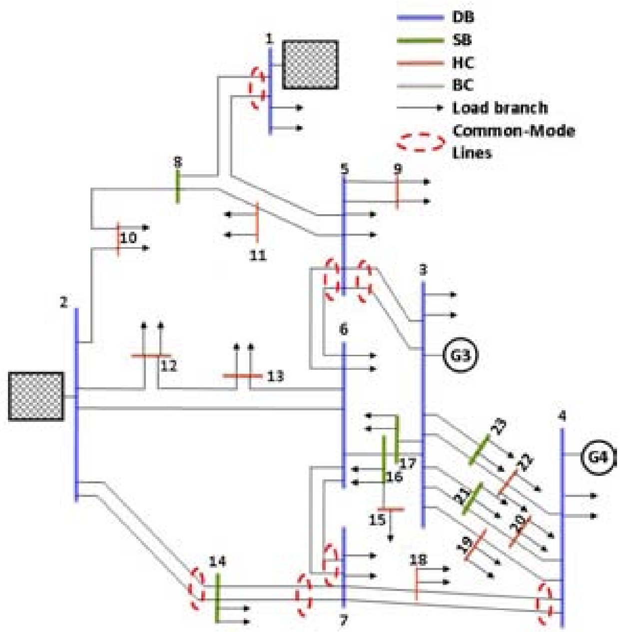

Fig. 1. 110 kV sub transmission power system

Traditionally, system state enumeration, which is for instance based on a depth first search combined with a minimum state level selection approach and with the minimum probability of occurrence approach [15], is used to identify possible system outages. These system states include normal states and failure states, which are all analyzed in the system state analysis.

In this paper a heuristic approach is applied for 1st order substation originated outages and in order to reduce the total set of system states, a minimal cutset (MCT) approach, considering independent 1st and 2nd order outages and common mode outages of power line components, is applied. A cutset is a set of branches for a grid such that, by removing these branches, there is no path in the power system graph from a source bus to a sink bus. A MCT is a cutset where no subset of it is a cut and therefore a MCT provides a list of outage events that cause interruptions to loads at substations. With the novel approach presented by the authors in [16], it is possible to calculate all MCTs up to the second order even for large planar interconnected networks in reasonable time. Applying this approach, only failure states have to be analyzed in the system state analysis, which reduces the total reliability assessment computation demand.

Let's consider the power system in Fig. 1, where a total number of 820 1st and 2nd order system states have to be analyzed. Applying the MCT approach reduces this number to 13 failure states. Since both, the number of system states and the MCT computation time have an exponential growth with increasing bus number, while the MCT has a linear growth with increasing line number when the bus number is fixed [16], the idea is that the decrease in the failure states computation has at least to compensate the computational effort of the MCT calculation.

The second new approach is the application of PLF techniques in the enumeration based reliability assessment in order to incorporate uncertainties [10]-[12], [17] and annual load flow data. By the use of probabilistic techniques annual loads can be modeled via one single load flow calculation. Hence, a lower computation time compared to multi level load model or multi interval load models can be expected [15]. This approach is used to obtain dependent outages, which are caused by (n-1)-contingencies.

B. Independent outage enumeration

The independent outages and common mode outages are calculated with a MCT approach. These MCTs are incorporated in the reliability indices calculation with a two-state Markov model, that is, the component is in success or failure state. In order to include power line common mode outages, the well known four state Markov model from [18] is used. For the sake of completeness the equations for an independent outage of series connected components (subscript j) and common mode (subscript cm) outages are listed:

$$f_j = \sum_{i=1}^{n} \lambda_i ; \qquad p_j = \sum_{i=1}^{n} \lambda_i \cdot r_i \tag{1.2}$$

$$f_{cm} = \lambda_1 \lambda_2 \left(r_1 + r_2 \right) + \lambda_{12} ; \quad p_{cm} = f_{cm} \cdot \frac{r_1 r_2 r_{12}}{r_1 r_2 + r_1 r_{12} + r_{12} r_2} \tag{1.3}$$

where subscript 12 specifies the common mode failure and p is the unavailability of a component in h/yr.

C. Substation originated outages

The consideration of substations (subscript SS) in reliability analysis is of high importance [19], therefore substations are included in this planning algorithm. The average failure rate λ_{SS} for independent substation components is small compared with conventional failure rates of electric power lines (see Table IV). Therefore it is sufficient to incorporate single component substation outages in the reliability indices calculation, since the probability of higher order events is rather marginal. The single component substation outages incorporated in this calculation algorithm are based on a heuristic approach taking into account corrective switching actions in different substation configurations. The analytically obtained equations, calculating the F&D indices of load feeders and source feeders in different substation types, have been derived with a fault tree analysis. These equations for each of the two loads are depending on the number of load feeders l ($l \geq 2$) at each substation and on the number of touching power line feeders (source feeder) s ($s \geq 2$).

TABLE I
MULTIPLICATION FACTORS FOR DIFFERENT SUBSTATION TYPES TO CALCULATE SUBSTATION F&D INDICES

ST	s, l	m_{ST}
BC		$\begin{bmatrix}0 & 0 & 1 & 0 & 2 & 0 & 1\end{bmatrix}$
HC		$\begin{bmatrix}0 & 0 & 2 & 0 & 2 & 0 & 1\end{bmatrix}$
SB	$\bar{s},\bar{l}$	$\begin{bmatrix}1 & 1 & s+l-1 & 1 & (s+l)/2 & (s+l)/2+2 & 1\end{bmatrix}$
	$\bar{s},\tilde{l}$	$\begin{bmatrix}1 & 1 & s+l-1 & 1 & (s+l\pm1)/2 & (s+l\mp1)/2+2 & 1\end{bmatrix}$
	$\tilde{s},\bar{l}$	$\begin{bmatrix}1 & 1 & s+l-1 & 1 & (s\pm1+l)/2 & (s\mp1+l)/2+2 & 1\end{bmatrix}$
	$\tilde{s},\tilde{l}$	$\begin{bmatrix}1 & 1 & s+l-1 & 1 & (s\pm1+l\pm1)/2 & (s\mp1+l\mp1)/2+2 & 1\end{bmatrix}$
DB	$\bar{s},\bar{l}$	$\begin{bmatrix}0 & 1 & (s+l)/2 & 1 & (s+l)/2 & 2 & 1\end{bmatrix}$
	$\bar{s},\tilde{l}$	$\begin{bmatrix}0 & 1 & (s+l\pm1)/2 & 1 & (s+l\pm1)/2 & 2 & 1\end{bmatrix}$
	$\tilde{s},\bar{l}$	$\begin{bmatrix}0 & 1 & (s\pm1)/2+1 & 1 & (s\pm1)/2+1 & l & 1\end{bmatrix}$
	$\tilde{s},\tilde{l}$	$\begin{bmatrix}0 & 1 & (s\mp1+l\pm1)/2 & 1 & (s\mp1+l\pm1)/2 & 2 & 1\end{bmatrix}$

where $\bar{s},\bar{l}$ are even, $\tilde{s},\tilde{l}$ are odd, $\pm$ and $\mp$ is for the two loads.

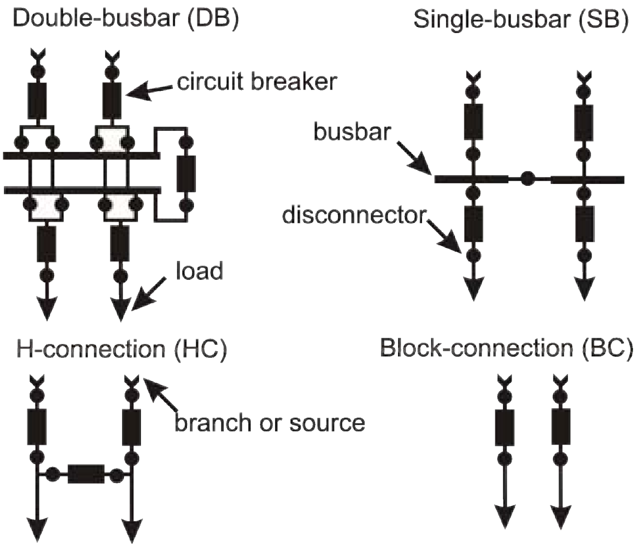

Fig. 2. Four different types of busbar connection for loads and sources in the proposed algorithm [20]

A source feeder s is for instance a branch to another interconnected substation or a connected generator. Note that, if any touching branch is a stub line without any sufficient generation at any substation on the stub line path, this branch has to be taken into consideration in the reliability indices calculation as a load feeder. Thus the multiplication factors of Table I permit together with equations (1.4) to deduce the F&D indices of each substation.

$$f_s = \boldsymbol{\lambda}_{\mathbf{SS}} \times \mathbf{m}_{\mathbf{ST}}^{\mathbf{T}} ; \quad p_s = \mathbf{U}_{\mathbf{SS}} \times \mathbf{m}_{\mathbf{ST}}^{\mathbf{T}} \tag{1.4}$$

with $\boldsymbol{\lambda}_{\mathbf{SS}} = \begin{bmatrix}\lambda_B & 0 & \lambda_C & 0 & \lambda_D & 0 & \lambda_T\end{bmatrix}$ and

$$\mathbf{U}_{\mathbf{SS}} = \begin{bmatrix}\lambda_B r_B & \lambda_B r_B^{CA} & \lambda_C r_C & \lambda_C r_C^{CA} & \lambda_D r_D & \lambda_D r_D^{CA} & \lambda_T r_T\end{bmatrix}$$

where U is the unavailability vector of components; m_{ST} represents the multiplication factor for either f or p; r^{CA} is the time for corrective actions; subscript B is the according data from the busbar; subscript C represent the according data from the circuit breaker; subscript D is the according data the disconnector and subscript T stands for the according data from the transformer.

Currently four substation types (ST) are included in the calculation algorithm (see Fig. 2). These are double busbar (DB), single busbar (SB), upper H-connection (HC) and the so called “Block”-connection (BC). Since the calculation is based on a heuristic approach any other substation type can be included in the calculation methodology in future.

D. Effect analysis

In order to achieve a sufficient accurate algorithm while keeping the computation time low, contingencies of overloaded lines are analyzed only for single line outages. Therefore the authors use the line outage distribution factors (LODF) to calculate all relevant outages. These factors give sensitivities of the redistributed prefault power flow over an outaged line to the remaining lines in the power system [11]. Computing LODFs, first a matrix $\mathbf{F} = \mathbf{PTDF} \times \mathbf{A'}$ of injection pairs to flows sensitivities has to be deduced, where $\mathbf{A}$ is the branch incidence matrix. The LODFs are then computed with

$$\mathbf{LODF} = \mathbf{F} \times \mathbf{S_F} \tag{1.5}$$

where $\mathbf{S_F}$ is a $n_t \times n_t$ matrix, and n_t is the total number of power system transmission lines, with $1/(1-diag(\mathbf{F}))$ at its diagonal elements and all other elements being zero. Modeling the impact of a single line outage as a linear function of power injections, equation (1.6) can be used. Each overloaded power line is analyzed as a relevant contingency if following inequality is true

$$\left|\mathbf{P}^{\mathbf{max}}_{\mathbf{diag}}\right| + \left|\mathbf{LODF} \times \left|\mathbf{P}^{\mathbf{max}}\right|\right| > \mathbf{P_{nr}} \tag{1.6}$$

where $\mathbf{P}^{\mathbf{max}}_{\mathbf{diag}}$ is a $n_t \times n_t$ matrix composed of the maximal loading of the power lines obtained from PLF, in any of the two possible power flow directions, with zeros at its diagonal, $\mathbf{P}^{\mathbf{max}}$ is a $n_t \times n_t$ matrix consisting of the maximal loading of the power lines, in any of the two possible power flow directions and $\mathbf{P_{nr}}$ is the vector of the power line nominal rating. A probabilistic optimized power flow (POPF) method based on point estimate method and linear outage and generation dispatch costs is applied to minimize the effect of load curtailment and generation dispatch of the relevant system failure states from (1.6). The basic point estimate method process, which enables the use of a deterministic optimized power flow (OPF), is explained in [21] and [22]. Contrary to [23], where a two point estimate method is used for a probabilistic load flow problem, a (2e + 1)-point estimate method [24], with e uncertain input variables, is applied to perform the POPF since this method can achieve a higher accuracy while not increasing the computational effort dramatically for the proposed real world example. The load curtailment is modeled as a negative real power injection with associated cost adapted from the OPF from [25]. The OPF is implemented in that way, that optimal load curtailment and optimal generation dispatch is used to satisfy the given security constraints (1.9) - (1.11). Thus the linear programming minimization problem is formulated as follows:

$$\min_{\theta, P_{cur}} \sum_{i=1}^{n} C_{cur,i} \tag{1.7}$$

with $C_{cur,i} = P_{cur,i} \cdot (c_1 + c_2)$

subject to

$$\mathbf{B_{bus}} \times \boldsymbol{\theta} = \mathbf{P_g} \tag{1.8}$$

$$P^{\min}_{g,j} \le P_{g,j} \le P^{\max}_{g,j} \qquad \forall j = 1, 2, ..., n_v \tag{1.9}$$

$$\mathbf{B_f} \times \boldsymbol{\theta} \le P_{nr,k} \qquad \forall k = 1, 2, ..., n_t \tag{1.10}$$

$$-\mathbf{B_f} \times \boldsymbol{\theta} \le P_{nr,k} \tag{1.11}$$

where n stands for the total number of buses in the power system; n_v represents the total number of buses that have installed generation; θ stands for the voltage angle at bus k; $P^{min}_{g,j}$ is the minimum available generation at bus j; $P^{max}_{g,j}$ stands for the maximum available generation at bus j; $P_{g,j}$ is the real power generation at bus j; B_{bus} represents the nodal susceptance matrix; B_f is the “from” bus susceptance matrix; $P_{nr,k}$ is the nominal rating of system transmission line k; P_{cur} stands for the amount of curtailed or dispatched active power; C_{cur} stands for the cost of the curtailed or dispatched active power; c_1 represents the constant factor either of the linear customer interruption cost (CIC) function or of the linear generation dispatch cost function and c_2 is the variable cost factor.

The results of the contingency evaluations are stored and subsequently used to calculate the load point reliability indices. The conditional probability $p_{c,lc}$ of a state, where load curtailment takes place, can be calculated with the probability of the single line outage $p_{j,lc}$ multiplied with the probability of the load curtailment p_{lc}. This p_{lc} is the ratio between the expected value of the total load at a lp and the expected value of the curtailed load at the same lp. To calculate an upper bound of EIC and EENS, the duration of each load curtailment failure state is the duration of the preceding single line outage. The transition frequency f_{lc} between a normal state and a failure state with load curtailment follows straightforward with $f_{lc} = p_{c,lc}/T_{lc}$.

E. Reliability Indices

The probability indices for a load point are the sum of all different identified outages that affect the load point. Each load indices, estimated with the gained states of the power transmission line effect analysis, is summed with the obtained indices of the upstream substation. Additionally, due to the heuristic substation approach used in this methodology, some special cases have to be considered. For instance, if a load point is connected with the power system generation buses over a stub line or if there is only one reference bus in the power system, it is necessary to add the relevant outage probability p_{sc} and the state cycle frequency f_{sc} of these special cases to the load point reliability indices. The three different classes of F&D reliability indices for load points, used in the proposed method, can be computed by the following equations:

$$\Pr_{lp} = \sum_{i=j,s,cm,sc} \sum_{i} p_{i,lp} \tag{1.12}$$

$$f_{lp} = \sum_{i=j,s,cm,sc} \sum_{i} f_{i,lp} \tag{1.13}$$

$$T_{lp} = \Pr_{lp} / f_{lp} \tag{1.14}$$

where j is the subscript for outages obtained by the MCT calculation and obtained in the effect analysis; s stands for the subscript for substation originated failure states; sc represents the subscript for additional special cases that affect a load point.

F. Monte Carlo Simulation

In order to estimate EENS and the reliability worth index EIC, including load shedding, for each load point and each failure state i a MCS is performed. Based on the cumulative distribution function, obtained from the PDF of the input nodal powers, used to calculate the PLF, and the obtained duration of the failure states, applying the inverse transform method [3], a set of active power random values $P_{RV,lp}$ can be computed. In the proposed real world example, hourly measured load values are used to generate the PDF of the input variables. Thus $P_{RV,lp}$ can be reshaped with the failure state duration T_i, to simulate a sufficient number of random time series and to calculate the ENS_{lp} distribution with the MCS approach. Parallel to this process the IC_{lp} distribution can be calculated based on the CIC. It follows with (1.15) and (1.16) the expectation of ENS and IC, adapted from the distribution of the nodal power and the individual failure state duration, where the expectation $\langle \cdot \rangle$ is used to estimate the EENS and the EIC for each load point. The expectation for a discrete distribution is $\langle X \rangle = \sum_i x_i p_i(X)$ with

TABLE II
POWER SYSTEM TRANSMISSION LINE DATA OF THE POWER GRID IN FIG. 1

line	f-t	l	X	P_{nr}	line	f-t	l	X	P_{nr}
		km	mΩ/km	MW			km	mΩ/km	MW
1	1-5	8.60	266	410	21	5-6	4.66	252	225
2	1-8	3.93	262	410	22	5-6	4.70	252	225
3	2-6	7.72	287	281	23	5-9	4.10	141	82
4	2-10	8.24	324	281	24	5-9	4.04	141	82
5	2-12	8.37	291	281	25	5-11	3.02	260	410
6	2-14	8.37	291	281	26	6-7	6.16	390	110
7	3-5	5.90	282	281	27	6-7	6.20	390	110
8	3-5	5.81	282	281	28	6-13	3.64	287	281
9	3-17	1.95	118	120	29	6-16	4.50	114	140
10	3-17	2.24	128	134	30	7-14	3.71	279	281
11	3-19	3.60	123	134	31	7-14	3.72	279	281
12	3-21	3.90	119	104	32	7-18	3.61	285	281
13	3-21	4.54	128	134	33	8-10	7.55	294	281
14	3-23	2.08	133	104	34	8-11	1.76	256	410
15	3-23	2.59	126	134	35	12-13	3.64	287	281
16	4-7	6.99	296	281	36	15-16	0.45	114	134
17	4-18	3.37	297	281	37	16-17	1.64	120	180
18	4-19	3.10	116	134	38	20-21	0.96	116	134
19	4-20	2.26	136	120	39	22-23	1.38	142	104
20	4-22	3.15	129	120	40	2-14	8.37	291	281

Grey highlighted lines are cables. f … from bus, t … to bus

TABLE III
CUMULANTS OF NODAL POWER DATA

ST	κ_1	κ_2	κ_3	κ_4	ST	κ_1	κ_2	κ_3	κ_4
		1e3	1e4	1e5					
1	110	9.39	-17	78.7	13	-6	4.42	-8.42	25.7
2	84	4.16	-21	-111	14	-36	3.22	4.66	51.9
3	162	3.17	9.86	30.2	15	-4	2.34	-5.78	50.6
4	109	1.96	-1.88	62.6	16	-30	2.16	-0.62	2.17
5	-27	0.29	-0.22	-0.23	17	-22	1.73	-0.34	23.3
6	-60	0.16	-0.08	0.12	18	-24	1.46	0.16	2.42
7	-47	0.04	0.02	0.01	19	-21	0.99	0.28	2.25
9	-10	0.01	-1e3	-7e-3	20	-18	0.93	1.21	8.68
10	-15	0.01	-4e-4	14e-3	21	-24	0.85	2.69	25.9
11	-31	8e-3	1e-3	7e-4	22	-19	0.56	1.13	24.8
12	-31	7e-3	1e-3	18e-3	23	-23	0.55	0.09	1.45

TABLE IV
COMPONENT FAILURE RATES AND REPAIR TIMES

Component	r	r^{CA}	λ
	h	h	1/yr km, 1/yr
Circuit breaker	100	0.167	0.00340
Busbar (including disconnector)	200	0.167	0.00680
Transformer	300		0.00300
Overhead line	48		0.00040
Cable	336		0.00100
Common mode overhead line	48		0.00004
Generator or transmission grid	0		ideal

TABLE V
GENERATION DATA FOR THE POPF

ST	P_g^{max}	P_g^{min}
	MW	MW
1	440	-210
2	200	-90
3	440	0
4	280	-10

TABLE VI
CIC FACTORS [26]

ST	c_1	c_2	ST	c_1	c_2
	€/kW	€/kWh		€/kW	€/kWh
1	2.90	23.46	13	2.28	22.47
2	2.28	22.47	14	2.28	22.47
3	1.88	18.43	15	2.28	22.47
4	1.88	18.43	16	1.88	18.43
5	0.92	9.34	17	1.88	18.43
6	0.66	7.32	18	1.88	18.43
7	2.90	23.46	19	1.88	18.43
9	1.38	13.38	20	1.88	18.43
10	1.38	13.38	21	2.28	22.47
11	1.38	13.38	22	2.28	22.47
12	2.28	22.47	23	2.28	22.47

$p_i(X) = P(X = x_i)$ being the probability mass function.

$$EENS_{lp} = \sum_{i=j,s,cm,sc} \left\langle \sum_{i} \mathbf{ENS}_{i,lp} \right\rangle \tag{1.15}$$

$$EIC_{lp} = \sum_{i=j,s,cm,sc} \left\langle \sum_{i} \mathbf{IC}_{i,lp} \right\rangle \tag{1.16}$$

V. CASE STUDY

A. Real world sub transmission power system

A 110 kV utility driven sub transmission power system is used to verify the accuracy of the proposed methodology.

Fig. 1 shows the 110 kV sub transmission grid with 23 nodes, 23 over head transmission lines, 17 cables, two generators and two power system connections, 21 load buses and maximum load of 791,11 MW, where bus one is the reference bus. In Table II the required line data is given and Table III contains the input nodal powers in terms of cumulants, where κ_n is the nth cumulant (see [10] and [27]). With the use of the GCE and the according cumulants, the PDF of each nodal power can be reconstructed. In Table IV the used failure rates and repair times of the system components are listed. Table V shows the generation limits for the POPF and Table VI gives the load bus individual CIC for load curtailment with the two cost function factors c_1 and c_2. Furthermore zero costs for dispatchable generation devices are used in the proposed study.

B. Results

The three F&D indices and the two energy based values for each load bus are calculated. The frequency distribution of hourly measured active power over one year for each load and generation bus, transformed with the principal component approach, is used as input RVs for the PLF calculation. The obtained results are compared with the calculation results of the commercial software NEPLAN [8] where the load characteristic approach (five intervals, each 20 % of a whole year) is used to calculate annual indices.

In Fig. 3 the obtained EIC results from the NEPLAN calculation and from the proposed algorithm are compared and the estimated differences between the obtained results, related to the NEPLAN results, are depicted. Table VII gives the exact values of the previous comparison for all remaining calculated indices. The differences range for f between 0 % (no load at the substation, bus 2 and 8) and 1.36 % (see Table VII). The difference is depending on the type of substation and on the connection degree to the gird. Since the proposed methodology does not calculate all 2nd order faults for substations, f is insignificantly higher at those substations where only first order faults are evaluated. In particular this circumstance has a small impact at substations of the type HC (bus 9, 10, 11, 12, 13, 15, 18, 20 and 22), where higher order faults, for example the independent outage of two circuit breakers at the same substation, changes the calculation result in the more accurate calculation of the NEPLAN algorithm. Since this kind of fault has a high number of possible combinations but a low probability of $4\cdot10^{-3}\cdot4\cdot10^{-3}=1.6\cdot10^{-5}$ outages per year, it can be neglected to speed up the calculation. The small differences in Table VII result mainly from those independent second order faults in substations.

The difference of the T indices mainly depends on the type of substation and therefore on the difference in the considered failure states. Although more 2nd order faults are unconsidered at the double busbar substations the difference of T is significant smaller at these load points. This is closely related with the higher number of possible switching actions at this type of substations.

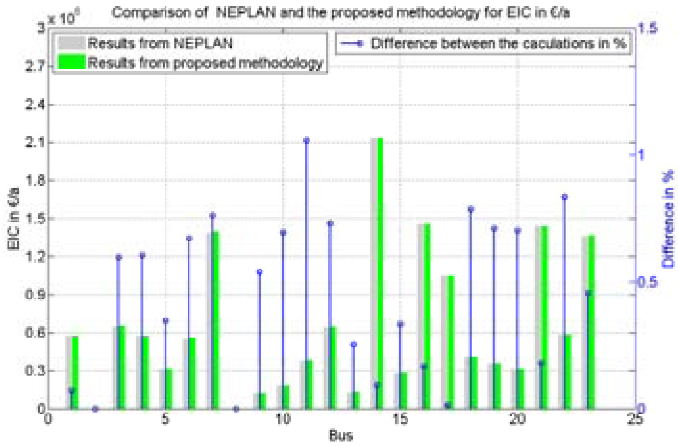

Fig. 3. Comparison of obtained EIC results and differences between the two algorithms

TABLE VII
DIFFERENCE BETWEEN NEPLAN RESULTS AND THE PROPOSED METHOD

ST	f	T	Pr	EENS	ST	f	T	Pr	EENS
	%					%			
1	0.03	0.05	0.01	0.20	13	1.15	0.38	0.77	0.92
2	0	0	0	0	14	0.48	0.18	0.30	0.11
3	0.50	0.10	0.59	0.59	15	0.38	0.15	0.23	0.14
4	0.59	0	0.59	0.60	16	0.41	0.14	0.27	0.23
5	0.57	0.03	0.59	0.30	17	0.50	0.21	0.29	0.09
6	0.59	0	0.59	0.80	18	1.35	0.52	0.82	0.83
7	0.63	0.03	0.59	0.69	19	1.20	0.41	0.78	0.75
9	0.96	0.2	0.75	1.00	20	1.20	0.43	0.77	0.58
10	1.36	0.57	0.79	0.52	21	0.51	0.21	0.29	0.16
11	1.33	0.56	0.77	1.02	22	1.18	0.41	0.76	0.87
12	1.17	0.40	0.77	0.67	23	0.51	0.21	0.30	0.48

Similar statements can be made for the outage probability Pr of the proposed example. Here also a high dependence on the type of substation can be recognized.

The calculation of the EENS in NEPLAN is based on the load case and on the load characteristics. This approximation of EENS is the more accurate the more detailed the load characteristics are, but it leads to a more extensive calculation time. The proposed methodology in the present paper calculates the EENS with an MCS approach. Together with the unconsidered 2nd order outages, for instance independent 2nd order faults in substations or certain combinations of independent outages and substation outages, the differences in Table VII can be explained. Hence, the first of the two occurring effects depend mainly on the type of substations. The second effect is the inexact modeling of the bus load characteristic in the NEPLAN algorithm and depends mainly on the characteristics of the load curve. This effect is demonstrated in Fig. 4, where for an assumed outage of one hour for each load point, calculation results of different methods are compared. In order to verify the EENS results accuracy of the NEPLAN calculation (3) with five intervals, the EENS of the random MCS approach (1), applied in the proposed methodology and the EENS obtained from a sequential MCS (2), adapted from the input measured load curve (exact EENS calculation), are contrasted in Fig. 4. The sum of the absolute difference introduced by the NEPLAN calculation is 2.31 % and the sum of the absolute difference introduced by the proposed MCS is 0.79 %, which justifies the selected approach in the proposed algorithm because it is more accurate than the load characteristic approximation.

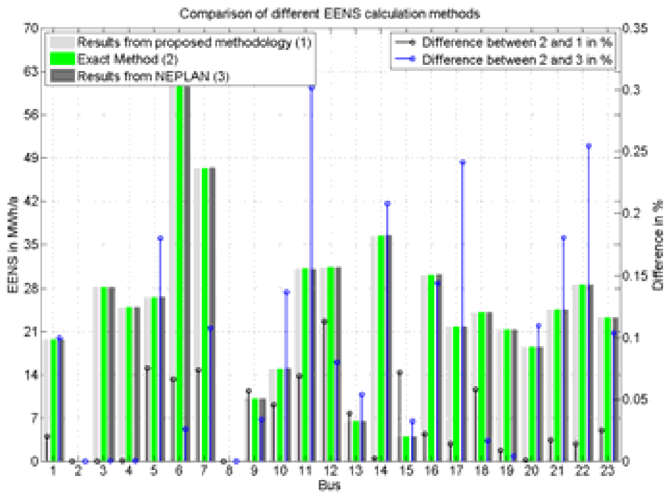

Fig. 4. Comparison of the different EENS calculation methods, where the difference is based on the exact method

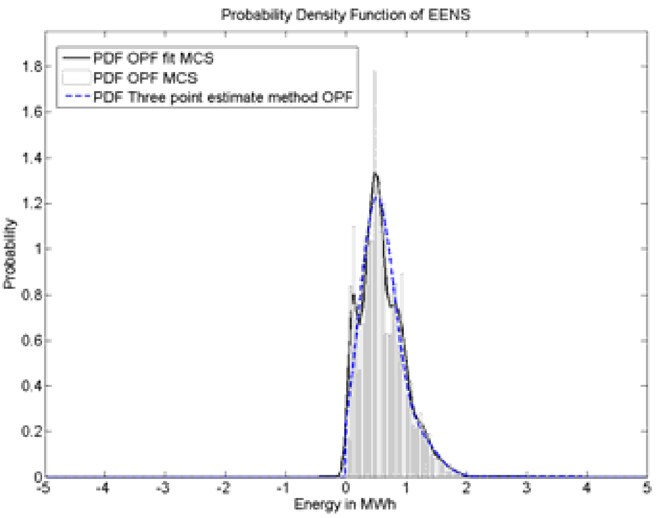

Fig. 5. PDF of the EENS at bus 15 from the POPF and the MCS

The EIC are calculated with a linear function based on the ENS. Therefore it is not surprising that the difference of EIC has a similar characteristic as the EENS difference.

In order to show the applicability of the implemented POPF, the EENS at bus 15, obtained from the proposed methodology, is compared with a MCS. In Fig. 5 the calculation results of the (2e+1)-point estimate method of the single line outage from bus 15 to bus 16 are depicted. The obtained results are an excellent fit of the exact results from the MCS based on the load curve.

C. Computation Performance

The calculation with NEPLAN (including all first order and second order faults) lasts 1792.82 sec for the proposed power transmission system on a Intel® Core™2 Duo CPU E 8600 @ 3.33 GHz with 3.50 GB RAM running on Windows XP Service Pack 3. The calculation on the same computer for the proposed methodology programmed in Matlab® lasts 24.33 sec for the second order point estimate method and 24.86 sec for the 2e + 1 point estimate method. In fact, this is a huge decrease in computation time. The proposed third order point estimate method methodology needs only 1.39 % of the calculation time in NEPLAN to perform the reliability assessment calculation.

For the proposed power system in Fig. 1 the overall MCT computation time is 8.1 sec and a single PLF of the power system for the state analysis needs 0.09 sec. Hence the applied MCT approach is much faster (~10 sec) as conventional system state enumeration and system state analysis (~74 sec).

The major amount (85 %) of computation time for the real word example is consumed by two procedures, the MCT analysis procedure (41 %), where approximately 80 % of this time are used for the MCT computation, and the POPF calculation (44 %) in the effect analysis. Depending on the number of induced cycles [16] in the power system, this computation time can diversify per each two terminal MCT (2TMCT), but for a rough estimation of overall computation time the 2TMCT computation time can be multiplied by number of source nodes times number of sink nodes. For example, the overall 2TMCT computation time for the IEEE 118 bus test system [25] is 1202 sec. The computation time of the effect analysis depends on the number of identified (n-1)-contingencies and approximately twice on the number of uncertain RVs in the power system due to the used POPF method. A single OPF calculation for the IEEE 118 bus test system lasts 0.12 sec, which results in a worst case computation time of 28.71 sec per investigated (n-1)-contingency. Using the MCT computation time of the real world example as reference, assumed a similar ratio between POPF computation time and 2TMCT computation time, and based on the suboptimal implementation in Matlab®, a total computation time for the IEEE 118 bus test system of 1 h can be estimated, which is very impressive for a very large system compared with NEPLAN computation time.

VI. Conclusion

A novel approach for assessing reliability indices, including substation originated outages, with uncertain nodal powers has been described in the paper. The methodology is based on probabilistic load flow and on probabilistic optimized power flow. The reliability assessment was performed with a utility driven 110 kV sub transmission power system with synthetically generated uncorrelated nodal power distributions based on realistic load characteristics. Comparing the results of the proposed methodology with the indices obtained with a commercial available program an adequate degree of accuracy can be observed. The maximum difference is 1.36 % of the obtained transition frequency. The authors hold that this difference is acceptable for a planning tool.

One major advantage of the proposed methodology can be seen in the fast computation time. Compared with the commercial program an impressive reduction in computation time follows. In order to emphasize this advantage it is now possible to use the proposed algorithm in a genetic optimization tool, to perform reliability optimized power system planning in reasonable time.

The implemented probabilistic optimized power flow used for optimal load curtailment calculation and generation dispatch in any contingency case is a computationally efficient way to consider uncertainties in the load curtailment process. Furthermore the obtained distributions show, compared to Monte Carlo simulations, a high goodness of fit.

As demonstrated in this paper, with the proposed methodology it is possible to include annual distributions of nodal powers in a reliability enumeration algorithm while maintaining an adequate accuracy for all used reliability indices.

VII. References

[1] Y. Fang, A.P. S. Meliopoulos, G. J. Cokkinides, and G. Stefopoulos, "A bulk power system reliability assessment methodology", *European Transactions on Electrical Power*, vol. 17, pp.713-425, April 2007.

[2] G. Rechberger, A. Gaun, and H. Renner, "Systematical Determination of Load Flow Cases for Power System Planning", *in. Proc. 2009 IEEE Bucharest PowerTech Conference*, 6 pages.

[3] R. Billinton, and W. Li, *Reliability assessment of electric power systems using Monte Carlo methods*, New York: Plenum Press, 1994.

[4] R. Billinton, and R. N. Allan, "*Reliability Evaluation of Power Systems*", 2nd ed., New York: Plenum Press, 1996.

[5] R. Billinton, and W. Zhang, "Algorithm for Failure Frequency and Duration Assessment of Composite Power Systems", IEE Proc. Generation, Transmission and Distribution, vol. 145, no. 2, pp. 117-122 March 1998.

[6] A.C.G. Melo, M.V.F. Pereira, and A.M. Leite da Silva, "Frequency and Duration Calculations in Composite Generation and Transmission Reliability Evaluation", *IEEE Transactions on Power Systems*, vol. 7, no. 2, pp. 469-475, May 1992.

[7] P. Zhang, and S. T. Lee, "Probabilistic Load Flow Computation Using the Method of Combined Cumulants and Gram-Charlier Expansion"; *IEEE Trans. Power Systems*, vol.19, no 1, pp. 676-682, Feb. 2004.

[8] BCP Busarello, Cott and Partner AG, NEPLAN Version 5.3.51, Bahnhofstr. 40, CH-8703 Erlenbach, Switzerland.

[9] R. Allan, and R. Billinton, "Probabilistic Assessment of Power Systems", *Proceedings of the IEEE*, vol. 88, no. 2, pp.140-162, Feb 2000.

[10] J. Usaola, "Probabilistic load flow with wind production uncertainty using cumulants and Cornish–Fisher expansion", *Int. J. Electr. Power Energ Syst.*, vol. 31, pp.474-481, Feb. 2009.

[11] L. Min, and P. Zhang, "A Probabilistic Load Flow with Consideration of Network Topology Uncertainties", in *Intelligent Systems Applications to Power Systems*, pp. 1-5, *ISAP 2007*.

[12] A.M. Leite da Silva, V.L. Arienti, and R.N. Allan, "Probabilistic Load Flow Considering Dependence Between Input Nodal Powers", IEEE Transactions on Power Apparatus and Systems, vol. PAS-103, no. 6, pp. 1524-1530, June 1984.

[13] P. McCullagh, *Tensor Methods in Statistics*, Monographs on Statistics and Applied Probability, Chapman and Hall Ltd., Cambridge: University Press, 1987.

[14] I.T. Jolliffe, *Principal component analysis*, New York: Springer, 1986.

[15] W. Zhang, "Reliability Evaluation of Bulk Power Systems using Analytical and Equivalent Approaches", Ph.D. dissertation, Dep. of Electrical Engineering, University of Saskatchewan, Saskatoon, 1998.

[16] A. Gaun, H. Renner, and G. Rechberger, "Fast Minimal Cutset Evaluation in Cyclic Undirected Graphs for Power Transmission Systems", *in. Proc. 2009 IEEE Bucharest PowerTech Conference*, 8 pages.

[17] C.-L. Su, "Probabilistic Load-Flow Computation Using Point Estimate Method", *IEEE Transactions on Power Systems*, vol. 20, no. 4, pp.1843-1851, Nov. 2005.

[18] R. Billinton, and R. N. Allan, *Reliability Evaluation of Engineering Systems, Concepts and Techniques*, 2nd ed., New York: Plenum Press, 1992, pp. 352-358.

[19] R. Nighot, and R. Billinton, "Reliability evaluation of the IEEE-RTS incorporating station related outages", *IEEE Power Engineering Society General Meeting*, vol. 1, pp. 118-123, June 2004.

[20] T. Paulun, "Strategische Ausbauplanung für elektrische Netze unter Unsicherheit", Ph.D. dissertation, ABEV, Band 115, ISBN 978-3-934318-77-9, Feb. 2007.

[21] E. Rosenblueth, "Point estimates for probability moments", in *Proc. Nat. Acad. Sci. U.S.A.*, vol. 72, no. 10, pp. 3812-3814, Oct. 1975.

[22] H. P. Hong, "An efficient point estimate method for probabilistic analysis", Reliability Engineering and System Safety, vol. 59, pp. 261-267, 1998.

[23] G. Verbic, and C. A. Canizares, "Probabilistic Optimal Power Flow in Electricity Markets Based on a Two-Point Estimate Method", *IEEE Transactions on Power Systems*, vol. 21, no. 4, pp.1883-1893, Nov. 2006.

[24] J. M. Morales, and J. Perez-Ruiz, "Point Estimate Schemes to Solve the Probabilistic Power Flow", *IEEE Transactions on Power Systems*, vol. 22, no. 4, pp.1594-1601, Nov. 2007.

[25] MATPOWER 3.2, A Matlab System Simulation Package. [online] Available: http://www.pserc.cornell.edu/matpower/matpower.html.

[26] M. Hyvärinen, "Electrical networks and economies of load density", TKK Dissertation 146, Department of Electrical Engineering, Helsinki University of Technology TKK, 2008.

[27] M. G. Kendall, *The advanced Theory of Statistics*, London: Charles Griffin & Company limited, 2nd ed., vol. 1, 1945.

VIII. Biographies

Alexander Gaun was born in Kufstein, Austria, in 1979. He received his diploma degree in 2005 at Graz University of Technology, Graz. He is currently a scientific university assistant at the Institute of Electrical Power Systems. His research interests include electrical power system planning, reliability computation, genetic algorithms and electromagnetic compatibility.

Georg Rechberger was born in Linz, Austria, in 1978. He received his diploma degree in 2005 at Graz University of Technology, Graz, where he currently holds a position as scientific university assistant at the Institute of Electrical Power Systems. His main work in research and teaching is in the field of electrical power system planning.

Herwig Renner was born in Graz, Austria, in 1965. He received his doctoral degree in 1995 at Graz University of Technology, Graz, where he currently holds a position as associate professor at the Institute of Electrical Power Systems. His main work in research and teaching is in the field of electrical power system analyses with special emphasis on quality, supply reliability and power system control and stability. Since 2009 he is a visiting professor at the Helsinki University of Technology.

Matti Lehtonen was born in 1959. Since 1999 he has been a professor at the Helsinki University of Technology, where he is now head of Power Systems and High Voltage Engineering. The main activities of Dr. Lehtonen include power system planning and asset management, power system protection and applications of information technology in distribution systems.

7.6 Annex F: Reference P6

Paper printed in conference proceedings and presented at 2010 Electric Power Quality and Supply Reliability Conference (PQ2010), June 16th – June 18th, Kuressaare, Estonia

Probabilistic Reliability Optimization using Hybrid Genetic Algorithms

A. Gaun, G. Rechberger, and H. Renner

Institute of Electrical Power Systems, Graz University of Technology, Austria

ABSTRACT: In this paper transmission power system structure optimization is performed via a minimal spanning tree based encoded fuzzy logic self-controlled hybrid genetic algorithm (GA). During the redundancy optimization of the power system network a binary encoded GA is used for a modified transmission network expansion problem, finding the optimal power line type with respect to the net present value (NPV) of minimal investment cost, operating costs and load flow constraints. Each individual is evaluated by a minimal state probability reliability estimation algorithm verifying a certain minimal reliability constraint. A developed improvement algorithm is used for individuals not satisfying a reliability constraint. A recently developed fast reliability calculation algorithm, computing energy not supplied, and the obtained NPV of the transmission network expansion problem are utilized as minimization function. The algorithm is applied to a real world sub transmission system in order to discuss strategies for future system expansions.

Keywords: Hybrid genetic algorithm, minimal spanning tree, redundancy optimization, reliability, transmission expansion planning, uncertainties

1. Introduction

REDUNDANCY optimization of transmission power systems is an NP-hard combinatorial problem, with computational effort growing exponentially with the number of substations and affecting reliability elements in the power system. Redundancy optimization, which is known as total investment-cost minimization with subject to reliability constraints in a multi state power system, is a special case of network design optimization. This optimization is of particular interest in various fields including operation research, engineering and computer sciences and has been studied via different solution strategies in many applications [1] - [6]. A deterministic reliability criterion with a branch and bound (B&B) algorithm is used in [2] to compute a optimal solution. Probabilistic reliability criteria for transmission system expansion planning, also using B&B approach, has been applied in [3] to solve the redundancy problem. In order to minimize the cost subject to a reliability level constraint, especially in the context of this nonlinear mixed-integer problem, genetic algorithms (GA) have been applied in many studies [1], [4] - [6]. A general overview about genetic algorithms in reliability engineering can be found in [7], and a more specialized and comprehensive overview is given in [8], where special focus is put on mathematical network modeling. In [1] a universal generating function technique is used to ensure that the system satisfies required reliability constraints whereas in [4] a system state sampling technique is used to obtain an expected energy not supplied (EENS) based reliability worth parameter for the GA fitness function. In [5] a GA is used for redundancy optimization in order to find the minimal cost system configuration that provides a desired reliability level and satisfy consumer demand modeled via piecewise cumulative load curve. A market based scenario approach, considering uncertainties is used in [6] where reliability costs (load curtailment and congestion costs) are taken into account in the optimization function.

However, by the knowledge of the authors none of the optimization algorithms for transmission systems considers both, uncertainties in the input data of the optimization problem and substation configurations in the reliability evaluation, which are important for significant reliability indices [9]. The authors of the present paper propose to incorporate reliability in the redundancy optimization by using expected interruption cost (EIC) of each load point, system expected interruption cost (SEIC), including substation originated outages. The objective of the proposed algorithm is to find the network structure with a minimum net present value (NPV), neglecting the (n-1)-criterion for grid structures. The fitness function consists of NPV of investment costs (ICs) and operating costs (OCs), including lifetime caused renewal of power system elements and transmission system SEIC based on uncertain load and generation data.

In this paper an algorithm for communication network design optimization with a k-terminal network reliability constraint [10] is modified and adjusted for evolutionary hybrid redundancy optimization of transmission networks. In particular, a simple binary coded transmission network expansion GA for optimal choice of power line types, the repair strategy, the reliability estimation function and the reliability evaluation algorithm for transmission networks, including uncertain load data [1], are implemented in the modified hybrid GA (HGA). Contrary to other GA for redundancy optimization [4] - [6], with the proposed enumeration algorithm it is now possible to consider following outage types in the non-linear optimization function: dependent outages forced by power line overloading and cleared with short time load congestion management [12], single and double line outages and their individual repair times, station originated outages and common mode (CM) outages.

This work is organized as follows. In the subsequent section the problem of the current graph structure optimization is described, followed by the modified transmission network expansion GA formulation including applied GA methodology. In section IV the basic modifications to the communication network design algorithm are given and the overall procedure of the developed HGA is provided. In section V the algorithm is applied to a real world 110-kV sub transmission system demonstrating the motivation of developing such an algorithm: providing useful information for strategic planning of future transmission expansion. Section VI concludes and the Appendix contain necessary data and informations for the real world example.

2. Problem statement

A transmission power system is modeled by a graph G = (V, E, L, C, {µ}, {λ}), where V is the set of given nodes, E represents the set of all possible branches, L stands for the set of branch lengths, C stands for the cost set of the entire graph G, with $\{\mu\}$ and $\{\lambda\}$, being the set of outage probability and outage duration respectively. Both reliability parameters can depend on the estimated power line type which could be a single or double circuit with different nominal ratings (NR) and resistance per unit length being either a cable (cl) or an overhead (oh) transmission line.

The optimization problem is to select a system topology with minimum cost, satisfying a certain level of reliability. The nonlinear combinatorial problem is defined via the problem assumptions, that planar graphs with locations of each substation are given, that the link costs, outage probability and outage duration are fixed and known. In this case, the fuzzy logic self-controlled (FLSC) redundancy optimization hybrid GA (FLSC HGA) problem is formulated as follows:

$$\min \left(FIT(x) + SEIC(x) \right) \quad (1)$$

subject to

$$R(x) > R_0 \quad (2)$$

$$x_{ij,k} \geq 0 \quad \forall (i,j) \in V, k \in K \quad (3)$$

where $x_{ij,k} \in x$ is the power line between substation i and j of power line type k. In the present paper problem (1) is solved via a hybrid GA, where the modified transmission network expansion problem, finding the optimal power line types for a given sub-graph (G' $\subseteq$ G) G' = (V ,E', L', C', {µ'}, {λ'}), is explained in the subsequent section and the exact reliability optimization problem is formulated by eqn. (6).

3. Transmission network expansion

A. Binary decoded GA

The classical nonlinear mixed integer formulation of modified transmission network expansion problem, finding the optimal branch type for a certain power system configuration, which is a major key task in finding an optimal solution for the redundancy problem, can be solved via linear programming, based on [13]-[17]. Since these approaches don't allow considering uncertainties in the optimization in this paper a GA is used to solve the mixed integer problem based on probabilistic load flow (PLF).

This transmission network expansion GA is implemented in the main reliability optimization GA as a hybrid optimization sub function minimizing the total investment NPV of the power system, including power lines and substation elements, such as disconnectors, transformers, circuit breakers and busbars, with subject to nominal ratings of power lines. This optimization is nonlinear. It is depending on $1\ldots|K|$ different power line types, where K is the set of different power line types and $|K|$ is the number of different power line types. Furthermore it is also depending on different substation types [1], which are depending on the node degree d_v and the kind of substation, in particular if it is a load and/or generation bus. Single busbars are used whenever more than two and less or equal four power lines are connected to the substation, an upper H-connection is used if two power lines touch the substation and Block-connection is for stub lines. Double busbars are used in all other cases and always if generation is located at the node.

The GA is a simple binary decoded algorithm, where the integers in the encoded individual represent one of the $1\ldots|K|$ different types of $1\ldots|E'|$ power lines for the as sub-graph G' represented power system, where G' is an individual of the main FLSC HGA. The proposed transmission network expansion GA uses a heuristic crossover and a binary mutation operator with mutation probability *pMBCGA* where randomly positions of chromosomes are changed from 1 to 0. The heuristic crossover operator with crossover probability *pCBCGA* creates a new child randomly, depending on the binary decoded parents and returns the best two of three (two parents and one child) possible individuals. A roulette wheel reproduction operator is used to selected *popSizeBCGA* individuals proportional to their fitness. As problem depending stop criterion is either used a maximal number of iterations *maxGenBCGA* or, when the obtained solution has not changed for certain *ncBCGA* generations. Start solutions, beside several randomly generated ones, are $|K|$ individuals, where each of the different $|K|$ individuals contains the identical power line type.

B. Mathematical Formulation

The fitness function (4) for the transmission network expansion is based on NPV calculation over a predefined period. ICs and OCs with cyclic re-investments are considered in the NPV [18]. Furthermore a penalty approach for those individuals who don't satisfy (5) is used in order to limit the search space to feasible solutions. These individuals are raised by an overloading dependent penalty factor times the chromosome IC for all power lines. The mathematical formulation of the minimization problem with DC load flow based PLF approach, calculating the cumulated distribution function of all power line flows and estimating $\mathbf{P}^{\mathbf{max}}$, is based on power transfer distribution factors (PTDF) method introduced in [19] and enhanced for dependent input variables by the authors, with bounds on circuit flows being shown below. $\mathbf{P}^{\mathbf{max}}$ stands for the maximal loading vector of all power lines $|E'|$, in any of the two possible power flow directions.

$$FIT = min \sum_{n=1,v=1}^{|V|,|E'|} NPV\left(C_{invest}\left(x\right)\right) \tag{4}$$

with $C_{invest}\left(x\right) = \left(C_{n,\,power\,line}\left(x\right) + C_{v,sub}\left(x\right) + C_{n,v,opera}\left(x\right)\right)$

subject to

$$\left|\mathbf{P}^{\mathbf{max}}\right| \le \mathbf{P}_{\mathbf{nr}} \tag{5}$$

where $\boldsymbol{C}_{\boldsymbol{n},\,\boldsymbol{power\,line}}$ is the IC of the n^th^ power line, $\boldsymbol{C}_{\boldsymbol{v},\boldsymbol{sub}}$ is the IC of the v^{th} substation, $\boldsymbol{C}_{\boldsymbol{n},\boldsymbol{v},\boldsymbol{opera}}$ are the operating costs of all substation elements and power lines. Since the dispatch of generation is not optimized, generation limits do not have to be set as boundary condition which simplifies the formulation.

4. FLSC reliability optimization

A. Problem formulation

The mathematical formulation of the reliability optimization problem, satisfying a certain reliability limit R_0 for the proposed algorithm is formulated as follows:

$$SEIC = min \sum_{1}^{|V|,|E'|} NPV\left(EIC\left(x\right)\right) \tag{6}$$

subject to

$$R\left(x\right) > R_0 \tag{7}$$

where $R\left(x\right)$, the reliability of each individual is estimated based on a state selection method. *SEIC* is calculated for certain individuals satisfying (7), due to computation burdens [1]. The IC of each individual is linearly depending on the length of each power line x_{ij}. Therefore the length l_{ij} of x_{ij} is used as weight for decoding the self-controlled GA, with an adjusted FLC, where its basics are given in [8], [10] and [20].

The applied crossover operator with crossover probability *pCFGA* is a slightly altered version of [10]. Randomly equal distributed either the first parent $|E'_1|$ edges, the second parent $|E'_2|$ edges, or, *pnC* edges are used for the size of the offspring in order to find a certain balance between exploration and exploitation of the search space. If $\min\left\{|E'_1|,|E'_2|\right\} > \hat{v}\cdot|V|+1$ then $pnC = \min\left\{|E'_1|,|E'_2|\right\}-1$, else $pnC = \min\left\{|E'_1|,|E'_2|\right\}$, where $\hat{v}$ is the minimum intermeshing degree for the start population (see section B).

The applied selection method is roulette wheel selection with included elitist selection which means that all individuals are selected proportional to their fitness except the best individual which is always a member of the new generation.

B. Start population and decoding

In order to generate feasible start population $Pop(t=0)$ a heuristic named *randG* is used. By removing randomly chosen edges from G = (V, E, L) as long as a certain bounded and randomly selected intermeshing degree v, with $v = \left[\hat{v}...1.5\right]$ [21], is not undercut and as long the obtained G' = (V, E', L') is a connected planar graph [22], [23] each of the start individuals is generated. G´ is decoded by the help of a minimal spanning tree (MST) [24], where the length of each edge l_{ij} is the edge weight. In this paper feasible means a planar power grid, where all substations are connected over at least one path, which is ensured with *randG* and the MST approach during the progress of the FLSC HGA. This edge-based decoding ensures that in every chromosome is at least one path from one node to all others which is a necessary requirement for redundancy optimization and permits the use of special mutation and crossover operators [10], generating again feasible individuals. Each individual is represented by the following chromosome, where the length of the chromosome is equal to four times $|E'|$ (see Fig. 1).

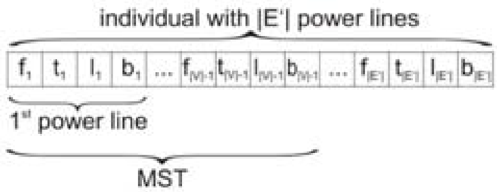

Fig. 1. Individual G', representing a feasible transmission network where f … from node, t … to node, l … length of power line, b ∈ 1… $|K|$.

The first $4\cdot(|V|-1)$ alleles of the individual represent the edges of the MST and the residual parts are the additional branches of the graph.

C. Reliability improvement algorithm

A reliability connection analysis of each encoded individual, based on the minimal probability state selection method considering single power line outages, double circuit power line outages and CM outages is used to estimate whether the individual satisfies a certain reliability bound $R(x) > R_0$ or not. In particular the maximal non-availability of each individual's load points is estimated based on an evaluation of all obtained system states within a 0.8-quantile, neglecting those various states with a certain minimal probability [25]. If this reliability is greater R_0, for instance the minimum non-availability of a single outaged power line, the individual is improved respectively repaired in order satisfy (7). Repair is made based on the set of nodes of the MST of the repairable individual with minimum node degree $\delta(V) < 2$, where $\delta(V) := \min\{d(s_i) | \forall s_i \in G'\}$, and the purpose is to realize maximum increase of reliability while reducing link ICs, which are proportional to l_{ij}.

The pseudo code of the modified repair algorithm is shown below where two strategies are used. The 1st one is to add the edge with the lowest cost if only one node has a $\delta(V) < 2$ and the 2nd strategy is to calculate the shortest path (SP) [23] between two randomly chosen nodes (i, j) in order to increase d_v.

procedure: Reliability improvement Algorithm
input: G' = (V,E',L´), G = (V,E, L)
output: repaired chromosome E'';

$E'_{MST} \leftarrow eval\ MST(V, E')$

$d_v \leftarrow deg(V,E'_{MST})\ \ \forall v \in V$; ok = true

while ok do

$\delta(V) \leftarrow min\{d_v | v=1,...,|V|\}$; $D_1 \leftarrow \{\delta(V) = d_v\}$;

$D_2 \leftarrow \{n_{i,j} | n_{i,j} \notin E'_{MST}\ \&\ n_{i,j} \in E\}$;

if $|D_1| = 1$ then $D_3 \leftarrow \{n_{i,j} | i \in D_1\ \&\ n_{i,j} \in D_2\}$;

else $D_3 \leftarrow eval\ SP\{(i, j) | (i, j) \in D_1\ \&\ n_{i,u_1;...;u_x,j} \in D_2\}$;

end

$n_{i,j,add} \leftarrow \min\{l_{i,j} | (i, j) \in D_3\}$; $E'_{MST} \leftarrow \{E'_{MST} \cup n_{i,j,add}\}$;

if $\delta(V) > 1$ then ok = false; $E'' \leftarrow E'_{MST}$ end

end

D. HDLC mutation routine

Mutation produces with mutation probability $pMFGA$ random changes in various $popSize \cdot pMFGA$ individuals, where $popSizeFGA$ is the size of each population. In this paper a heuristic highest degree, lowest cost (HDLC) mutation operator is applied. This operator, proposed in [10] and modified in this paper, uses a kind of randomized greedy local search. The authors of the present paper propose to use a random number pnM to decide with equal probability whether an edge with minimum l_{ij} is added to the altered individual, where an edge of the highest degree has not been removed, guaranteeing the exploitation of the search space. Furthermore the modified HDLC routine removes only those edges, linked to the node with the highest degree and nodes with degree greater two in order to ensure that the graph is connected.

E. Immigration routine

The immigration is used to update each population with new randomly generated individuals exploring the search space ignoring the exploitation of the promising regions of the search space [8]. The worst $pIFGA$ genomes in terms of FIT of the current population are replaced with new $randG$ randomly generated and BCGA optimized individuals, irrespective if their fitness value is better or not, with $pIFGA$ being the immigration probability.

F. Reliability calculation

Besides the minimal probability state selection method applied for reliability improving individuals a novel fast enumeration based reliability assessment algorithm based on a probabilistic load flow approach considering nodal uncertainties [1] applying the frequency and duration method is used to incorporate frequency and duration load point indices, like EENS or EIC, in the redundancy optimization. Generally, using certain R_0 as constraint for the optimization, the algorithm will optimize all configurations satisfying this limit, while no distinction is made between more reliable network structures and less reliable network structures. If the EIC of analyzed configurations, based on customer interruption costs (CIC) [26], including substations originated outages, is accessorily used in the optimization function, the algorithm will find the most reliable configuration.

Although the proposed algorithm, developed by the authors, includes several heuristics and modifications in order to reduce the computation time for individuals it is still cumbersome to compute reliability indices. Hence, this algorithm is applied only to the 0.4-quantile Q_{40} of FIT individuals satisfying R_0. The $SEIC$ is estimated and its NPV is added to the fitness function, obtaining a solution with minimal C_{invest} and minimal $SEIC$. Those individuals that are not investigated with the developed reliability algorithm are punished in terms of $SEIC$ via adding PEN with $PEN = \max(EIC) \cdot \max(FIT) / \min(FIT)$.

G. Overall procedure FLSC HGA

The following pseudo code of the FLSC redundancy optimization HGA, including transmission expansion planning BCGA, summarizes the overall procedure.

```
procedure:  FLSC HGA for redundancy optimization
input:      G = (V, E, L, C, {μ}, {λ}); GA parameters (popSizeFGA,
              maxGenFGA, pMFGA, pIFGA, pCFGA, u)
output:     network structure of the fittest individual
begin
  R_0 ← min{μ_ij · λ_ij · l_ij} ∀(i, j) = 1,2,...|E'|_pop(t);
  t ← 0;
  initialize Pop(t) by randG;
  FIT eval(Pop) by BCGA;
  reliability estimation of Pop eval(R);
  while t ≤ maxGenFGA (not termination condition) do
    crossover Pop(t) to yield Child(t);
    HDLC mutation Pop(t) to yield Child(t);
    immigration operation to yield Child(t);
    FIT eval(Child) by BCGA;
    reliability estimation of Child eval(R);
    if any( R < R_0 ) then
      [Pop;Child]' ← reliability improvement algorithm for all R < R_0;
      FIT eval([Pop;Child]') by BCGA;
      reliability estimation of all R > R_0 eval(R);
    end
    SEIC eval Q_40(R > R_0); SEIC ← PEN(R < R_0)
    eval(FIT + SEIC);
    if t ≥ u then
      SFLC to eval pMFGA, pIFGA, pCFGA;
    end
    Pop(t+1) from [Pop(t);Child(t)] by elitist roulette wheel selection;
    t ← t+1;
  end
end
```

5. Real world example

A 110-kV utility driven sub transmission power system is used to identify strategies for future power system expansions and power systems reconfigurations and to verify the effectiveness of the proposed GA. The present 110-kV sub transmission grid consists of 24 nodes, three generators and two power system connections and 21 load buses, with bus one being the slack bus. Based on 84 identified feasible power line connections (see Table A-I), which are clustered into two regions, one rural and industrial region with four different selectable over head transmission line types and one (sub-) urban region, with four selectable cable line types (see Table I), the sub transmission system is optimized. Table A-II contains the measured input nodal powers in terms of moments [1], [27] and gives the load bus individual CIC for load curtailment with the two cost function factors c_1 and c_2 and CIC for dispatchable generation devices, which are set to zero in this demonstrative study. Furthermore mutual correlations between the load buses and generation buses in the range between 0.23 and 0.70 are considered, based on measured load curves. In Table II additional data like cost and lifetime for substation devices and reliability data for the FLSC HGA is given. For the POPF [12] in the fast reliability enumeration algorithm [1] the generation limits are set to the minimum and maximum values of the generation active power distributions in Table A-II.

The settings of the FLSC HGA parameters for the real world sub transmission power system are:

population size, $popSizeFGA = 15$;

maximum generation, $maxGenFGA = 40$;

start crossover probability, $pCFGA = 0.45$;

start mutation probability; $pMFGA = 0.45$;

start immigration probability; $pIFGA = 0.1$;

FLSC parameter [10]:

u = 1.3; $\alpha = 0.5$;

$\mu_{mFGA} = (0.8 - 0.25) / (2^u - 2\alpha)$;

$\mu_{cFGA} = (0.8 \cdot (1 - pMFGA) - 0.15) / (2^u - 2\alpha)$;

$$pMFGA = \begin{cases} 0.25 \\ 0.25 + \mu_{mFGA} \cdot (\Delta f - \alpha) \\ 0.8 \end{cases};$$

$$pCFGA = \begin{cases} 0.8 \cdot (1 - pMFGA) \\ 0.8 \cdot (1 - pMFGA) - 0.15 - \mu_{cFGA} \cdot (\Delta f - \alpha) \\ 0.15 \end{cases}$$

minimum intermeshing degree, $\hat{\nu} = 1.15$;

and the settings of the transmission network expansion BCGA are:

population size, $popSizeBCGA = 40$;

maximum generation, $\max GenBCGA = 80$;

crossover probability, $pCBCGA = 0.9$;

mutation probability, $pMBCGA = 0.1$;

stop criterion, $ncBCGA = 10$;

The NPV is calculated over a period of 50 years with an interest rate of 5 %.

In order to benchmark the proposed FLSC HGA the existing power system is analyzed with the used data. In the following Fig. 2, the grid structure of the existing real world sub transmission power system is shown, including the optimized power line types via BCGA and slight modifications in terms of power line lengths and power line types making the sub transmission grid anonymous. The power system in Fig. 2 has an NPV FIT of 1.91e8 € including all transformers, substations, power lines, and renewals due to LT reasons and an NPV SEIC of 1.34e9 €. In total this results in an NPV of 1.52e9 €. In Fig. 3 a typical convergence plot of the proposed binary decoded BCGA is shown emphasizing BCGA's strength for solving the proposed problem. The obtained solution of the BCGA is the optimal solution for the proposed problem resulting in minimal costs while satisfying the load flow constraints.

The optimization of the modified transmission network expansion problem for the proposed realized real world sub transmission power system lasts on an Intel® Core™ 2 Duo CPU E 8600 @ 3.33 GHz with 3.50 GB RAM running on Windows XP Service Pack 3 on average 18 seconds and the SEIC calculation lasts on average 60 seconds, due to high number of dependent n-1 outages. These two functions represent the main impact on the computation time and a total time of approximately 80 seconds per individual can be estimated. In other words the computation for the whole generation will last approximately 14 h if we assume that all individuals satisfy $R(x) > R_0$. Therefore only the 0.4-quantile of FIT individuals satisfying (7) is analyzed in terms of reliability.

TABLE I
USED POWER LINE TYPES FOR TRANSMISSION EXPANSION GA

b-type	#	c-type	R	X'	NR	IC	OC	LT
			in Ω/km	in Ω/km	in MW	in €/km	in €/km	in a
oh	1	s	0.15	0.40	100	0.13e6	5e3	60
oh	2	s	0.05	0.30	260	0.20e6	5e3	60
oh	3	d	0.025	0.31	520	0.30e6	7e3	60
oh	4	d	0.013	0.21	1040	0.35e6	7e3	60
cl	1	s	0.075	0.15	100	0.65e6	5e3	40
cl	2	s	0.025	0.13	170	0.85e6	5e3	40
cl	3	d	0.025	0.14	230	1.35e6	7e3	40
cl	4	d	0.012	0.10	400	1.80e6	7e3	40

with b-type … branch type being either oh or cl; c-type … circuit type being either single circuit (s) or double circuit (d); LT … lifetime

TABLE II
COMPONENT FAILURE RATES, REPAIR TIMES, INVESTMENT COSTS, OPERATING COSTS, LIFETIME

Component	r	r^{CA}	λ	IC	OC	LT
	h	min	1/a km, 1/a	€	€/a	a
Circuit breaker	100	10	3.4e-3	-	-	-
Busbar (incl. disconnector)	200	10	6.8e-3	8e5[1]	3e4[1]	50[1]
Transformer	300	-	3.0e-3	45e4	500	45
Overhead line (oh)	48	-	0.4e-3	-	-	-
Cable (cl)	336	-	1.0e-3	-	-	-
CM overhead line	48	-	0.4e-4	-	-	-
Generator	0	-	ideal	-	-	-

[1] per feeder, including busbar, disconnector and circuit breaker;

In Fig. 4 the convergence plot of the grid structure optimization is shown. The minimal value of obtained NPV SEIC and NPV FIT is 1.50e9 €, which is reached after 30 generations, where NPV FIT is 1.79e8 €. The parameter tuning steps of the applied fuzzy logic are also shown in Fig. 4. The GA parameters, set by the fuzzy logic, are not fixed and therefore ensure a better exploration and exploitation of the search space. In Fig. 5 the obtained optimized transmission power system including optimized power line types, representing a better solution than the currently realized power system (Fig. 2) is depicted. Comparing realized power system and optimized structure both NPV, FIT and SEIC, of the optimized structure are better in terms of a decreased value, which results in less power lines with a lower value of SEIC. The outcome of the optimization process is an enhanced power system structure with a decreased overall NPV.

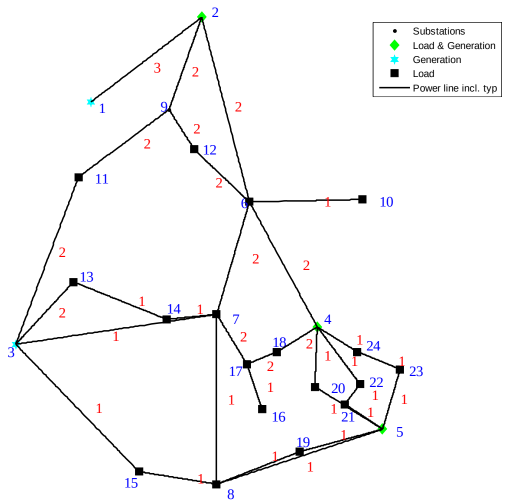

Fig. 2. Realized real world sub transmission power system optimized by BCGA with respect to power line types.

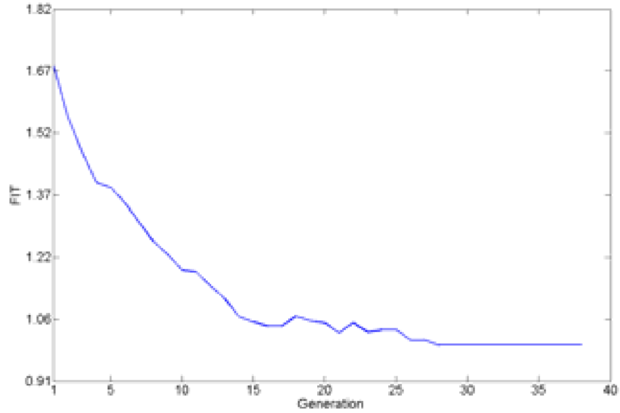

Fig. 3. Typical convergence plot of the BCGA for power line type optimization in transmission network expansion.

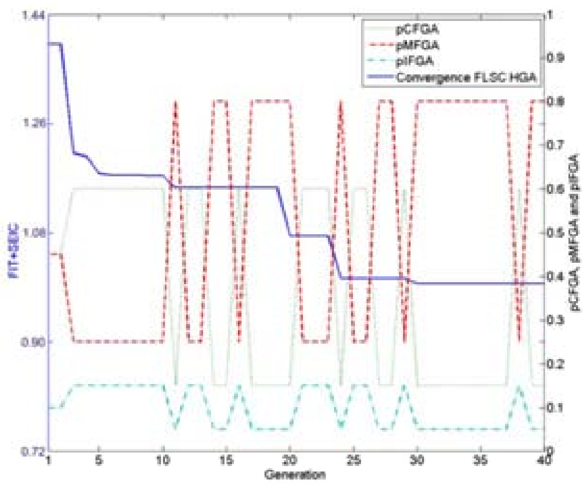

Fig. 4. Convergence plot of the hybrid FLSC HGA, applied on the real world grid with fixed positions of substations, including optimized power line type, satisfying (7). Additionally the parameter tuning of FLSC HGA is shown.

The average node degree $\bar{d}_v$ of touching power lines of the optimized power system structure is 2.54 and the average node degree of the realized power system is 2.67. Since substations have a remarkable influence on power system reliability, and there is an almost linear relationship between the number of power line feeders and outage probability a possible strategy for future power system expansions is to minimize the number of feeders which can be realized with $\bar{d}_v$ converging against two. In other words this means for the expansion strategy, obtained from the optimization results that power system planning engineers should concentrate on ring structures in transmission power system, since power line loops tend to be more reliable grid structures. Furthermore these ring structures utilize upper H-connection substations with particular high reliability and low IC.

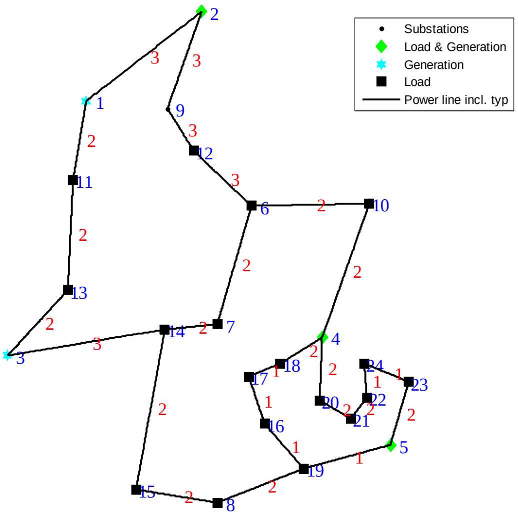

Fig. 5. Optimized transmission power system structure based on probabilistic load flow input data satisfying a certain reliability criterion with an NPV of 1.50e9 €. In red the optimal line type (see Table I) is given.

6. Conclusion

In this paper a fuzzy logic self-controlled hybrid genetic algorithm (FLSC HGA) is used for redundancy optimization of a transmission power system with uncertain load data. The transmission network expansion problem, finding the optimal power line type for a certain probabilistic load flow, is solved by a binary decoded genetic algorithm, including different substation types depending on the number of touching power lines. By the use of the present realized power system structure it is demonstrated, that the BCGA can be used for the nonlinear optimization task, resulting in an optimal or almost optimal grid structure satisfying all power flow constraints.

The redundancy optimization is done by a GA based on a reliability improvement algorithm that satisfies a minimum node degree of two while reducing the power system IC. By applying a shortest path algorithm and a heuristic each individual, having a non availability greater a certain limit, is repaired. This methodology, combined with special a minimal path based mutation and crossover routine provide an improvement of the overall investment, operating, renewal and expected outage costs in terms of NPV. As main expansion strategy for future power system expansions could be proposed to reduce the average node degree of the power system in order to reconfigure the grid structure which should consist of several loops, linking different substations. In [18] the cost optimal transmission power system, satisfying the (n-1)-criterion, estimated with a linear approach using a B&B algorithm, is also realized with different power line rings, connecting substations.

The proposed powerful algorithm can be used for green field studies, power system expansion planning's, benchmark studies and restructuring studies all including uncertainties and transmission network reliability.

Appendix

In Table A-I 84 identified potential power lines are listed and the measured input nodal powers in terms of moments [1], [27] and CIC are given for the real world sub transmission power system in Table A-II.

Acknowledgment

The authors of this paper would like to thank Prof. M. Lehtonen from the Department of Electrical Power Systems, Helsinki University of Technology (HUT), who provided the power system data used in this work.

References

[1] G. Levitin, "Redundancy Optimization for Multi-State System with Fixed Resource-Requirements and Unreliable Sources," *IEEE Trans. on Reliability*, vol. 50, pp. 52-59, March 2001.

[2] T. Tran, J.-S. Choi, D.-H. Joen, J.-B. Chu, R. Thomas and R. Billinton "A Study on Optimal Reliability Criterion Determination for Transmission System Expansion Planning," *KIEE International Trans. on Power Engineering*, vol. 5-A, no. 1, pp.62-69, 2005.

[3] J. Choi, T. Tran, A. A. El-Keib, R. Thomas, H. Oh, and R. Billinton, "A Method for Transmission System Expansion Planning Considering Probabilistic Reliability Criteria," *IEEE Trans. on Power Systems*, vol. 20, no. 3, pp. 1606-1615, August 2005.

[4] J.A. Gomez-Hernandez, J. Robles-Garcıa, and D. Romero-Romero, "Transmission project profitability including reliability optimization of bulk power systems," *Electrical Power and Energy Systems*, vol. 28, pp. 421-428, Feb. 2006.

[5] A. Lisnianski, G. Levitin, H. Ben-Haim, and D. Elmakis "Power system structure optimization subject to reliability constraints, " *Electric Power Systems Research*, vol. 39, pp. 145-152, July 1996.

[6] M.O. Buygi, G. Balzer, H.M. Shanechi, and M. Shahidehpour "Market-Based Transmission Expansion Planning," *IEEE Trans. on Power Systems*, vol. 19, no. 4, pp. 2060-2067, November 2004.

[7] G. Levitin, "Genetic algorithm in reliability engineering," *Reliability engineering & system safety*, vol. 91, pp. 975-976, 2006

[8] M. Gen, R. Cheng, and L. Lin, *Network Models and Optimization, Multiobjective Genetic Algorithm Approach*, London: Springer, 2008, p. 23.

[9] R. Nighot, and R. Billinton, "Reliability evaluation of the IEEE-RTS incorporating station related outages", *IEEE Power Engineering Society GM*, vol. 1, pp. 118-123, June 2004.

[10] L. Lin and M. Gen, "A Self-controlled Genetic Algorithm for Reliable Communication Network Design," in *Proc. 2006 IEEE Congress on Evolutionary Computation*, pp. 640-647.

TABLE A-I
IDENTIFIED POWER LINE CONNECTIONS FOR REAL WORLD SUB TRANSMISSION SYSTEM

f	t	l	b-type	f	t	l	b-type	f	t	l	b-type	f	t	l	b-type	f	t	l	b-type
		in km				in km				in km				in km				in km	
1	2	6.00	oh	4	7	4.80	oh	5	23	3.20	cl	8	14	7.10	oh	16	17	2.50	cl
1	3	10.50	oh	4	8	9.90	cl	5	24	4.10	cl	8	15	3.50	oh	16	18	3.00	cl
1	9	3.50	oh	4	10	6.50	oh	6	7	4.80	oh	8	16	4.60	cl	16	19	3.20	cl
1	11	3.20	oh	4	16	5.30	cl	6	9	5.40	oh	8	19	4.00	oh	17	18	1.80	cl
2	6	7.80	oh	4	17	4.50	cl	6	10	6.40	oh	9	11	4.90	oh	19	20	3.60	cl
2	9	4.00	oh	4	18	2.70	cl	6	11	7.80	oh	9	12	2.00	oh	19	21	3.60	cl
2	10	10.3	oh	4	19	6.60	cl	6	12	3.40	oh	9	13	9.30	oh	19	22	5.00	cl
2	12	5.70	oh	4	20	3.00	cl	7	8	7.00	oh	9	14	9.70	oh	20	21	1.90	cl
3	7	9.00	oh	4	21	4.30	cl	7	10	9.50	oh	10	12	8.50	oh	20	22	2.50	cl
3	8	10.60	oh	4	22	3.60	cl	7	12	6.90	oh	11	12	5.40	oh	20	24	2.90	cl
3	9	11.90	oh	4	23	5.00	cl	7	13	7.00	oh	11	13	4.70	oh	21	22	1.50	cl
3	11	7.40	oh	4	24	2.40	cl	7	14	2.20	oh	11	14	7.60	oh	21	23	3.60	cl
3	13	3.60	oh	5	8	7.80	oh	7	15	7.50	oh	12	13	8.20	oh	21	24	2.90	cl
3	14	7.00	oh	5	19	3.70	oh	7	16	5.60	cl	12	14	7.10	oh	22	23	2.30	cl
3	15	7.70	oh	5	20	4.20	cl	7	17	3.20	cl	13	14	4.50	oh	22	24	1.60	cl
4	5	6.30	cl	5	21	2.30	cl	7	18	4.00	cl	13	15	8.40	oh	23	24	2.50	cl
4	6	6.00	oh	5	22	2.60	cl	7	19	8.80	cl	14	15	6.50	oh				

TABLE A-II
MOMENTS OF NODAL POWER DATA [27] AND CIC FACTORS [26]

V	μ_1^X	μ_2^X	μ_3^X	μ_4^X	c_1	c_2	V	μ_1^X	μ_2^X	μ_3^X	μ_4^X	c_1	c_2
	MW	MW^2	MW^3	MW^4	€/kW	€/kWh		MW	MW^2	MW^3	MW^4	€/kW	€/kWh
G1	-211.3	6.1e4	-1.8e7	5.7e9	0	0	12	-31.2	1.0e3	-3.7e4	1.4e6	1.38	13.38
G2	460.8	2.5e5	1.4e8	7.8e10	0	0	13	-31.4	1.1e3	-3.9e4	1.5e6	2.28	22.47
G3	19.8	1.1e4	1.1e6	4.7e8	0	0	14	-6.5	50.5	-451.9	4.5e3	2.28	22.47
G4	147.9	2.6e4	4.7e6	9.1e8	0	0	15	-36.3	1.4e3	-6.0e4	2.7e6	2.28	22.47
G5	114.1	1.6e4	2.3e6	3.3e8	0	0	16	-4.0	21.9	-144.7	1.1e3	2.28	22.47
2	-19.7	419.0	-9.6e3	2.3e5	2.90	23.46	17	-30.1	953.9	-3.2e4	1.1e6	1.88	18.43
4	-28.2	848.2	-2.7e4	9.1e5	1.88	18.43	18	-21.7	508.6	-1.3e4	3.4e5	1.88	18.43
5	-24.9	655.9	-1.8e4	5.3e5	1.88	18.43	19	-24.1	610.0	-1.6e4	4.5e5	1.88	18.43
6	-24.6	652.0	-1.8e4	5.4e5	0.92	9.34	20	-21.2	485.9	-1.2e4	3.1e5	1.88	18.43
7	-60.5	3.7e3	-2.4e5	1.5e7	0.66	7.32	21	-18.5	364.2	-7.6e3	1.7e5	1.88	18.43
8	-47.3	2.4e3	-1.2e5	6.8e6	2.90	23.46	22	-24.5	658.2	-1.9e4	6.0e5	2.28	22.47
10	-10.1	112.5	-1.4e3	1.8e4	1.38	13.38	23	-28.6	909.1	-3.2e4	1.2e6	2.28	22.47
11	-15.0	263.6	-5.3e3	1.2e5	1.38	13.38	24	-23.2	579.6	-1.6e4	4.4e5	2.28	22.47

[11] A. Gaun, G. Rechberger, and H. Renner, "Enumeration Based Reliability Assessment Algorithm Considering Nodal Uncertainties", *accepted for presentation in. IEEE PES GM 2010*, Minnesota, 8 pages, July 25-29, 2010.

[12] A. Gaun, G. Rechberger, and H. Renner, "Point Estimate Methods for Probabilistic Optimized Power Flow," Proc. 11. Symposium Energieinnovation, pp.1-8, Feb. 2010

[13] L. Bahiense, G.C. Oliveira, M. Pereira, and S. Granville, "A Mixed integer Disjunctive Model for Transmission Network Expansion," *IEEE Trans. on Pow. Sys.*, vol. 16, no. 3, pp. 560-565, Aug. 2001.

[14] E.L. da Silva, H.A. Gil, and J.M. Areiza, "Transmission Network Expansion Planning Under an Improved Genetic Algorithm," *IEEE Trans. on Power Systems*, vol. 15, no. 3, pp. 1168-1175, August 2000.

[15] H.A. Gil, and E.L. da Silva, "A reliable approach for solving the transmission network expansion planning problem using genetic algorithms," *Electric Power Systems Research*, vol. 58, pp. 45-51, 2001.

[16] G. Latorre, R.D. Cruz, J.M. Areiza, and A. Villegas, "Classification of Publications and Models on Transmission Expansion Planning," *IEEE Trans. on Power Systems*, vol. 18, no. 2, pp. 938-946, May 2003.

[17] S.H.M. Hashimoto, R. Romero, and J.R.S. Montovani, "Efficient linear programming algorithm for the transmission network expansion planning problem," *IEE Proc.-Gener. Transm. Distrib.*, vol. 150, no. 5, pp. 536-542, Sept. 2003.

[18] H.-C.G. Maurer, "Integrierte Grundsatz- und Ausbauplanung für Hochspannungsnetze," Doctoral thesis, Fakultät für Elektrotechnik und Informationstechnik, RWTH Aachen, Okt. 2004.

[19] P. Zhang, and S. T. Lee, "Probabilistic Load Flow Computation Using the Method of Combined Cumulants and Gram-Charlier Expansion"; *IEEE Trans. Power Systems*, vol.19, no 1, pp. 676-682, Feb. 2004.

[20] Y.H. Song, G.S. Wang, P.Y. Wang, and A.T. Johns, "Environmental/economic dispatch using fuzzy logic controlled genetic algorithms, " *IEE Proc., Gener. Transm. Distrib.*, vol. 144, no. 4, p.377–382, July 1997

[21] D. Oeding, and B. R. Oswald, "*Elektrische Kraftwerke und Netze*," 6. Auflage, Springer-Verlag Berlin Heidelberg New York, p. 499, 2004.

[22] A. Gaun, H. Renner, and G. Rechberger, "Fast Minimal Cutset Evaluation in Cyclic Undirected Graphs for Power Transmission Systems", *in. PowerTech, 2009 IEEE Bucharest*, 8 pages, Okt. 2009.

[23] D. Gleich, "MatlabBGL A Matlab Graph Library", Institute for Computational and Mathematical Engineering, Stanford University. Available: http://www.stanford.edu/~dgleich/programs/matlab_bgl/, October 2008.

[24] G. R. Raidl and B. A. Julstrom, "Edge Sets: An Effective Evolutionary Coding of Spanning Trees, " *IEEE Trans. on Evol. Compu.*, vol.7, no.3, pp.225-239, June 2003.

[25] W. Zhang, "Reliability Evaluation of Bulk Power Systems using Analytical and Equivalent Approaches", Ph.D. dissertation, Dep. of Electrical Engineering, University of Saskatchewan, Saskatoon, 1998.

[26] M. Hyvärinen, "Electrical networks and economies of load density", TKK Dissertation 146, Department of Electrical Engineering, Helsinki University of Technology TKK, 2008.

[27] M.G. Kendall, "The advanced Theory of Statistics", Charles Griffin & Company limited, Vol. 1, 2nd edition, 1945

Authors mailing address: Institute of Electrical Power Systems, Graz University of Technology (TUG), 8010 Graz, Austria. E-mail: alexander.gaun@tugraz.at