NTNU

Norwegian University of
Science and Technology

Probabilistic Load Flow Studies: Analytical and Approximate Methods

Mari Holtet Eie

Master of Energy and Environmental Engineering
Submission date: June 2018
Supervisor: Vijay Vadlamudi, IEL

Norwegian University of Science and Technology
Department of Electric Power Engineering

Abstract

The power system is, in an increasing manner, subject to uncertainties, and the traditional method of analyzing power systems using the deterministic load flow (DLF) is increasingly proving to be insufficient to handle the challenges of the modern power system. The modern power system is characterized by integration of renewable energy sources whose generation is fluctuating and uncontrollable in nature; even on the demand side, an increased flexibility related to the development of smart grids and power systems is expected to increase the uncertainties. Probabilistic load flow (PLF) techniques provide engineers the opportunity to include these uncertainties in the analysis of power systems.

In this thesis, a review of the current state of the PLF research is presented, with focus on analytical and approximate methods available today. Two of the most commonly used and promising methods are the Cumulant method (CM) and the Point estimate method (PEM). These methods are studied in further detail and demonstrated on three different test systems. A major contribution of this study has been to provide a pedagogical presentation of the theoretical framework and methodological procedure, and thus one of the test cases includes thorough step-by-step illustrations of both methodologies. The aim of this master's project has also been to create in-house tools with which to conduct probabilistic load flow studies. This tool has been applied to yet another two test systems, in order to provide validation and to demonstrate the application on two test systems that are widely used in reliability studies.

The scope of this study has been the analytical and approximate methods of PLF, hence no simulation methodologies such as the Monte Carlo (MC) methods have been implemented. This has restricted the validation of the results from case studies in this thesis to comparison with results from previous research, whose availability is often limited.

The case studies of this thesis address uncertainties associated with load demand, as well as conventional and renewable generation. Correlation between different random variables in the system is an important part of these uncertainties, and is also studied in detail. Uncertainties related to outages and resulting changes in the network topology are aspects of the PLF that are not considered in this thesis, but rather left as a possibility of future work on the subject.

Sammendrag

Kraftsystemet er i stadig økende grad utsatt for usikkerheter og den tradisjonelle metoden for å analysere kraftsystemer ved å benytte deterministisk kraftsystemanalyse, viser seg i mange sammenhenger å være utilstrekkelig. Det moderne kraftsystemet kjennetegnes ved integrasjon av fornybare energikilder, hvis produksjon er fluktuerende og ukontrollerbar av natur. Også på lastsiden er økt fleksibilitet relatert til utviklingen av smarte nett og kraftsystemer forventet å bidra til å øke usikkerheten. Metoder for Probabilistisk Lastflytanalyse (PLF) gir ingeniører muligheten til å inkludere disse usikkerhetene i analyser av kraftsystemet.

Denne oppgaven vil gi en oversikt over forskning gjort innen probabilistisk lastflytanalyse, med fokus på analytiske og omtrentlige metoder. To av de mest brukte og lovende metodene er Kumulantmetoden (CM) og Pointestimatmetoden (PEM). Disse metodene vil bli studert nærmere og demonstrert på tre forskjellige testystemer. Et viktig bidrag fra denne studien har vært å gi en pedagogisk presentasjon av det teoretiske rammeverket og de metodologiske aspektene. Ett av case-studiene inkludert i oppgaven er derfor grundige trinnvise beskrivelser av metodene. Målet med denne masteroppgaven har også vært å implementere interne verktøy for å gjennomføre probabilistiske lastflytanalyser. Dette verktøyet har blitt benyttet på ytterligere to ulike testsystemer for validering og demonstrasjon av metodene på testsystemer som er mye brukt i pålitelighetsstudier.

Omfanget av denne studien har inkludert analytiske og omtrentlige metoder for PLF, derfor har ingen simuleringsbaserte metoder, slik som Monte Carlo (MC) simuleringer blitt implementert. Dette har begrenset mulighetene for validering av resultatene fra casestudiene i denne oppgaven, til sammenligning med resultater fra tidligere forskning, hvis tilgjengelighet ofte er begrenset.

Casestudiene i denne oppgaven adresserer usikkerhet knyttet til lastbehov, i tillegg til både konvensjonell kraftproduksjon og fornybare energikilder. Korrelasjon mellom ulike variable i et system er en viktig del av slike usikkerheter og blir også studert i detalj. Usikkerheter knyttet til utfall av linjer og andre nettverksendringer er forhold som ikke har blitt vurdert i denne oppgaven, men som blir foreslått som en mulighet for videre arbeid innen emnet.

Preface

This Master's thesis, written during the spring of 2018, concludes my education at the Norwegian University of Science and Technology (NTNU). The thesis gives an introduction to the field of probabilistic load flow, with an emphasis on analytical and approximate methods.

I would like to extend my sincerest gratitude to my supervisor at NTNU, Vijay Venu Vadlamudi. His enthusiasm for the field of power engineering has truly been an inspiration, and I have made great use of our many discussion throughout the year. I have come to value greatly his continuous focus on the importance of telling a good story to the reader, which I believe has greatly improved the readability of my thesis.

I would also like to thank Sigurd Strømsem for so generously helping me with the thesis, and finally Sondre Kvam, for all his help and continuous support.

Trondheim, June 2018

Mari Holtet Eie

Table of Contents

List of Tables

List of Figures

Abbreviations

PFA	=	Power Flow Analysis
NR	=	Newton-Raphson (methodology)
PEM	=	Point Estimate Method
CM	=	Cumulant Method
PDF	=	Probability Density Function
CDF	=	Cumulative Distribution Function
MGF	=	Moment Generating Function
CGF	=	Cumulant Generating Function
DLF	=	Deterministic Load Flow
PLF	=	Probabilistic Load Flow
MC	=	Monte Carlo (simulation technique)
GC	=	Gram-Charlier (expansion series)
CF	=	Corner-Fisher (expansion series)
FOSMM	=	First-order second-moment method
FOR	=	Forced Outage Rate
WT	=	Wind turbine

1 Introduction

1.1 Motivation

The modern power system is subject to a large amount of uncertainties, especially and increasingly so with regards to power injections due to load and generation at the system's buses. For a power system even approaching a realistic size, it is soon obvious that it is impractical to assess all possible values of these. If additionally the probabilities of the different states are not accounted for, the result will provide an unrealistic picture of the power system.

By adopting a probabilistic approach to the load flow analysis, engineers are able to provide a more realistic description of the operational state of the power system at a given time. Whichever method is used for PLF, simulation-based or analytical, the results are based on the probabilistic distribution of the random variables that represent the uncertainties in the system. States of the system that are most dominant from a probabilistic point of view, can be assigned the heaviest weight in planning and operational decision-making.

There are several factors that support predictions about increasing uncertainties in the future power system, and thus there is an increased need for probabilistic methods in power system assessment. The penetration of fluctuating, uncontrollable renewable sources such as wind power and solar energy is one of the main challenges. Introduction of smart grids is another, and the impact of this is perhaps even harder to predict. The increased flexibility that follows from smart grid solutions may be able to change the consumption pattern and thus create a far more complex demand pattern than what is seen today.

1.2 Scope of work

The objective of this thesis has been to investigate analytical and approximate methods of probabilistic load flow. Simulation methodologies (Monte Carlo techniques) represent an important field of research within probabilistic load flow, but are disregarded in the scope of this thesis. Leaving the simulation techniques out of the picture does however put some limitations on the case studies in this thesis, as simulation-based PLF is often used as a source of validation of the results obtained from approximate and analytical methods. Using an MC technique and running enough simulations for the method to converge and thus provide the

exact solution to the problem, serves as a validation of the non-simulation technique as well as a measurement of the accuracy of the results. Leaving out the MC tecniques in this thesis thus limits the validation of the case studies in this thesis to making use of previously published and validated work.

There are several uncertainties associated with the modern power system, and this thesis will look into a number of them. Uncertainties associated with load demand and conventional generation are discussed and modelled in detail. Correlation between these variables imposes an additional factor of uncertainty and is also investigated in case studies. Generation by renewable energy sources represent an increasingly important uncertainty factor in the power system, and probabilistic modelling of wind power is discussed and implemented to some extent. However, through all case studies in this thesis, a constant network topology is assumed, i.e. uncertainties associated with line outages or other incidents causing the physical network to change, are not taken into account in the case studies here.

1.3 Contributions

- This thesis aims to provide a framework for analytical and approximate methods for PLF by investigating the current state of the research field and providing closer insight into a select-few representative methodologies from the literature. Emphasis has been laid on providing pedagogical clarity of the fundamental and methodological aspects.

- Two well-established, yet fundamentally different methods have been chosen for close investigation in this thesis. These are the Point Estimate method (PEM) and the Cumulant method (CM). Together they represent both main groups of non-simulation methodologies - the approximate and analytical methods, respectively. The following aspects of uncertainties associated with the power system have been studied in this thesis:

  - *Load demand uncertainties.* Both continuous and discrete distributions describing the uncertainties of load demands have been included in the case studies of chapter 4.

  - *Generation uncertainties.* Modelling of the uncertainty related to forced outage of conventional generators has been included in case studies. The fluctuating production and large uncertainty associated with wind turbines have also been modelled with appropriate random distributions and included in case studies.

  - Ways of handling correlation between variables, both loads and generation units

have been studied in detail in chapter 3 for both methodologies, and are also included in several case studies carried out in chapter 4.

- In-house programming codes have been developed for all case studies in this thesis and are used for further use and research at the Department of Electric Power Engineering.

1.4 Report structure

Chapter 1 - *Introduction*, provides an introduction to the motivational background and a presentation of the scope, limitations and contributions of this project.

Chapter 2 - *Literature review*, presents a review of existing research and literature available on the subject of probabilistic load flow methods. It also includes an introduction to basic concepts of power flow analysis and statistics that are considered necessary theoretical background for the reader.

Chapter 3 - *General methodology*, presents the theoretical background and derivation of the two PLF methods that are studied thoroughly in this thesis – the Point estimate method (PEM) and the Cumulant method (CM).

Chapter 4 - *Case studies* applies the PEM and CM to three different test systems in order to demonstrate different aspects of uncertainties and methodologies. Discussions of the results from the different cases are also included in this chapter.

Chapter 5 - *Conclusions*, includes a discussion of the conclusions drawn with regards to the two methods studied in this thesis, from case studies and existing literature.

The thesis builds on a specialization project undertaken during Autumn 2017. I order to make the thesis self-contained, for narrative clarity, portions of the specialization report have been extensively made use of in the presentation of chapter 2 and portions of chapter 3.

2 | Literature Review

The first proposal of PLF was done in 1974 [5][29], and was based on the motivation of obtaining a better foundation for decision making for planning problems, by computing a probabilistic distribution of branch power flows corresponding to probabilistic nodal power injections. All the while keeping a fixed network configuration, such that only variation in load demand and generation were considered uncertainties in the system. Since this first proposal, the PLF has been developed further to be applied also in normal operation, short- and long-term planning, and in several other areas [8]. Various methodologies exist now, being able to also handle uncertainties such as branch outages and other network changes, and include aspects like correlation between the stochastic variables. In recent years, the focus has also increasingly been towards improving efficiency of algorithms, application in power system planning and inclusion of voltage control devices [8].

The basic problem that all PLF methods try to solve is the same; namely to most efficient, yet sufficiently accurate obtain the probabilistic distribution of the state vectors and line flows for each node and branch in the system [29]. The various PLF methods can now be divided into three main groups according to whether they are aiming to solve the problem numerically (i.e. using a Monte Carlo method), analytically (i.e. performing a conventional convolution technique) or approximately (i.e. using a combination of the two preceding techniques) [17].

MC simulation methods involves selecting a number of values for the input variables from their respective probability distribution, performing DLF for each of these, and finally obtaining the probabilistic distribution of the state variables from the results of the repeated simulations. Several thousand simulations are usually necessary to obtain meaningful results. Although accurate and relatively simple, the computational burden of this method is a major drawback for practical purposes. Real power systems are usually of such a size that MC methods are impractical and costly, and most researchers only use it for comparison purposes [29]. This thesis will not go further into MC methods, but rather focus on analytical and approximate methods

2.1 Analytical methods

The analytical methods are far more computationally effective than the numerical methods, but obviously require some mathematical assumptions to be made, in order to simplify the

problem. In the wake of the first proposal of PLF in 1974, several different methods were proposed to solve it. The use of conventional convolution techniques were dominant in the first methods that were developed. Other techniques such as Fast Fourier Transform (FFT) and the cumulant method were later also proposed [20].

The conventional convolution techniques require that the output variables (i.e. state vector and line power flows) are represented as linear combinations of the input variables. Assuming also independence between variables, a convolution technique can be applied to obtain the PDFs of the output variables. The major drawback of the convolution technique is the great amount of storage and time necessary to handle the large amount of impulses resulting from functions represented by many impulses being convoluted with each other [29]. The impact of this problem increases with the size of the power system, so for real-size power systems, the convolution methods can therefore be unattractive. Reference [2] addresses this problem and propose to use a convolution technique by applying FFT, and with that achieve greater accuracy and reduction in computational speed. However, the convolution technique still forms the basis of this method, and the problem can thus not entirely be overcome.

The cumulant method is another analytical method that quite remarkably avoids the convolution technique by utilizing some important properties of cumulants. In this way, the CM is able to obtain statistical parameters of state variables and line flows by algebraic operations rather than convolution operations. In recent literature, the cumulant method has been recommended for large transmission systems due to the low computation burden and relatively high accuracy [3][24]. Th CM is therefore chosen as the analytical method to be investigated further in this thesis.

In its most fundamental approach, the CM is dependent upon both the linearization and independence between variables that is also seen in the convolution methods. The linearization of the power flow equations will obviously cause the method to perform best for random variables that have distributions concentrated around the mean. In fact, [10] shows that the CM performs rather poorly when the PDFs of the input variables have dominant points away from the mean. For load demand distributions, this will rarely pose as a problem, as the load variation, whether it being a short-time daily variation or a longer yearly variation, can often be approximated by distributions close to the normal distribution. However, on the generation side of the transmission network, the ever increasing integration of renewable sources like wind turbines and photovoltaic systems represent a challenge to this assumption. Due to the uncontrollable nature of the primal sources (wind speed and solar irradiation, respectively), the PDFs of these power injections tend to indeed have dominant points far from the mean. As the presence of renewable sources in transmission system are generally expected to be only increasing, several proposals have been made to overcome this issue.

Reference [20] proposes an enhanced cumulant method where variables whose distributions are active at points far away from the mean are treated separately and decomposed into component parts. The results from this study prove very promising, at least if one is willing to make a compromise to some point between speed and accuracy when choosing the cumulant method over MC or approximate methods with much higher computational burden.

Several enhancements have also been proposed to attempt to work around the independence requirement of the CM. Considerable correlation exist in both the demand and generation part of the transmission system, and is sometimes of such importance that it can not be overlooked. On the load demand side, the seasonal and daily variations of temperature, social factors such as the daily cycle, holidays and so on, will cause residential loads in a neighbouring area to increase and decrease in a similar manner. On the generation side, the correlation is, like for the linearization issues discussed above, mainly an issue with the uncontrollable renewable sources. Especially in neighbouring wind farms the generation correlation is very strong [6], but this obviously also accounts for solar irradiation. In an extensive review of ways to handle variable correlation in PLF, [6] summarizes two methods possible for the cumulant method. The most common one being to model the correlated input variables as a function of independent random variables by some methods further described in references and in section 3.2 where one such methodology is introduced. The second possibility outlined in [6] is to make use of, in addition to self cumulants, also joint cumulants of the input random varibles. A drawback of the latter is that the procedure is quite computationally complex and may not serve as the best option, as the CM is often chosen as the preferred method because of its high computational efficiency. A method for handling correlation based on cumulants and DC load flow is proposed in [25] and also show promising results, in this case leaving the computation time unaffected when correlation is taken into account.

The raw output of the CM is, as previously mentioned, cumulants of the output variables, that is the state vector and the line power flows. The different order cumulants in themselves does usually not represent the statistical information that is of interest for the output. However, a lot of statistical data lies in the cumulants, and the more cumulants (i.e. of higher order) that are computed for each variable, the more statistical information can be extracted for that variable. In some cases, basic properties like the expected value and standard deviation will give a sufficient representation of the output variables. These properties are easily determined from the first few cumulants, as described in section 2.4.5. More often however, a more complex representation of the distribution is desirable and thus the CM is often used in conjunction with some extension series to form a PDF or CDF of the output random variables. Depending on the type of distribution, different expansion

schemes are recommended. Gram-Charlier and Corner-Fisher expansion are two possibilities, and [24] concludes with the first mentioned performing best for unimodal distributions that relate well to the normal distribution. Thus, the GC would be a good choice for expanding results from a PLF with power injections modelled with the normal distribution. For a PLF including input variables with non-Gaussian distributions (for instance for a system having wind power generation), the CF series is recommended in [24].

2.2 Approximate methods

First-order second-moment methods (FOSMM) and point estimate methods (PEM) stand out in this group of PLF methods [17].

The FOSMM starts off in a similar manner as the analytical methods, but it avoids the mathematical complexity of the convolution technique and can thus only approximate the output variables. Compared to MC methods, it drastically reduces the computation time, as only one DLF is required. The FOSMM outlined in [27], linearizes the power flow equations by using the first two terms of the Taylor expansion series, and is thus able to approximate the mean and standard deviation of the output variables directly, without the use of any convolution operations. [27] show promising results for this method, and despite its necessitity for linearization, the method is capable of handling correlated variables directly. Drawbacks of the method is the linearization requirement, and also its inability to produce more statistical data than only the mean and standard deviation of the output variables.

PEMs pose as well-suited alternatives to the FOSMM, though they have a quite different approach to the approximation of output variables. While the FOSMM can be seen closely related to the analytical methods, the PEM shares its main concepts with MC methods. As an analytical method has already been chosen to be more closely studied, the PEM will be the approximate method to be investigated further in this thesis. In the PEM, a limited number of simulations are done in strategic points for each of the input random variables to provide an approximate description for the statistical parameters of the output. The aim of the PEM is to drastically reduce the computation burden compared to a MC method, but to still run as many simulations necessary to obtain a sufficient accuracy for the results. The major difference from MC methods thus lies in the selection of points in which to run the DLFs, as they are not chosen randomly.

Because DLF routines are used to solve the PLF, the PEM does not have the challenges related to linearization of the power flow equations that the analytical methods and the FOSMM has. Also it is advantageous in the sense that perfect knowledge of the distribution of the input variables (i.e. a PDF) is not necessary; only the mean, variance, skewness and

kurtosis of the input variables are utilized in order to approximate the PDFs of the output variables.

Several different methods have been proposed in the family of PEMs. They mainly differ in number of DLFs to be performed (i.e. computation burden) and in types of variables they are able to handle (correlation and asymmetric variables) [17]. The general idea common for all these methods, is that each input random variable will produce a number of locations in which DLFs are being conducted. Each location will be assigned a weight based on the current variable's statistical properties, and the weighted sum of the results from the DLFs in all the different locations represent the final results for the output variables. The first method was developed by Rosenbleuth in 1975, and the final Rosenbleuth method proposed in 1981 was able to handle both correlated and unsymmetric random variables. However, the number of DLFs required with this method increase exponentially with the number of probabilistic inputs, so for real-size power systems the computational burden (which can in some cases even exceed the MC methods), make this method unattractive.

Since Rosenbleuth's method, several different PEMs have been developed. In this thesis, possible application on large power systems is emphasized, so only those methods that show efficiency in large scale systems are considered further. In that department, [17] points out two methods after a comparison between the main available techniques, namely Harr's and Hong's method, first proposed in 1989 and 1998, respectively. They both show high efficiency for large systems, but differ in what types of variables they are able to handle. Harr's method handles correlated variables, but not unsymmetric ones, Hong's method works the other way around. In this thesis, as in [17] and the dominant part of more recent literature [12][13][9], Hong's method is chosen for further investigation.

The main challenges with Hong's method is now its efficiency (as for all PEMs), its accuracy, and its inability to handle correlated variables. The computation burden (i.e. efficiency) and accuracy are obviously closely related and depend on the type of simulation scheme that is used (i.e. how many locations are produced and thus how many simulations necessary). Several schemes are investigated in [17] and are further discussed in section 3.1 of this thesis.

As it was outlined in the review of analytical methods, the ability to handle correlated variables is increasingly important when handling the modern power system. As for the analytical methods, several proposals have been made as to how to work around the independence requirement in the PEM. In [18] an improvement to the method later used in this thesis, is made by the use of orthogonal transformation and Cholesky decomposition. This method is adopted for the handling of correlation in this thesis, and is outlined in detail in section 3.1.2. Studies on the use of this method show promising results, especially for systems

where the uncertainty level is rather small [18]. An alternative to the orthogonal transformation is to adopt a discrete PEM as proposed in [1]. This method makes the PEM able to directly handle correlated variables without the need of transformation to the orthogonal space, but is in return rather computationally complex.

As for the CM, the raw output of the PEM will not necessarily be of the form that is of interest. The PEM produces various order moments of the output variables, so also here expansion series like the GC or CF series are often utilized. As is shown in section 2.4, the cumulants of a random variable are directly related to the moments and vice versa, so the discussion of extension series in section 2.1 equally applies to the PEM.

Based on this literature review, the following two sections will present some basics of power flow and statistics that are considered to be necessary theoretical background for the reader.

2.3 Basic power flow

Even in the probabilistic environment, most methods of power flow analysis still require running a number of deterministic iterations at different stochastic locations. This section will therefore present some relevant basics of deterministic power flow analysis, much inspired by the presentation of the subject in [22].

In PFA, four quantites, namely the voltage magnitude $|V|$, voltage angle δ, and real and reactive net power injections P and Q respectively, are associated with each bus in the system. The buses are then classified into one of the following three categories:

1. The **Slack bus** is the bus chosen as reference for the voltage angle δ, and has specified value for voltage magnitude $|V|$. The slack bus makes up for the differences in scheduled loads and generated power in the network.

2. A **PV bus** (also known as a *generator bus*) has specified values for active power injection P and voltage magnitude $|V|$. The injected reactive power Q and bus voltage angle δ are unknown variables to be determined in the analysis.

3. A **PQ bus** (also known as a *load bus*) has specified values for active and reactive power injections P and Q, respectively. The bus voltage magnitude $|V|$ and angle δ are unknown variables to be determined.

Now let $\mathbf{Y}$ denote the admittance matrix describing the network typology and relating the currents and voltages of all buses such that

$$\mathbf{I} = \mathbf{Y} \times \mathbf{V} \tag{2.1}$$

where $\mathbf{I}$ and $\mathbf{V}$ are vectors of injected bus currents and bus voltages, respectively. Each element Y_{ij} of $\mathbf{Y}$ can be written in terms of its magnitude and angle as $|Y_{ij}|\angle\theta_{ij}$. Equivalently, each entry V_j of $\mathbf{V}$ is expressed as $|V_j|\angle\delta_j$. From (2.1), the current injection at bus i may then be written as a sum of these polar representations of Y_{ij} and V_j;

$$I_i = \sum_{j=1}^{n} |Y_{ij}||V_j|\angle(\theta_{ij} + \delta_j) \tag{2.2}$$

Defining the complex power injection at bus i as $S_i = P_i - jQ_i$ and making use of the relationship $S = V_i^* I_i$, we then achieve:

$$P_i - jQ_i = |V_i|\angle(-\delta_i) \times \sum_{j=1}^{n} |Y_{ij}||V_j|\angle(\theta_{ij} + \delta_j) \tag{2.3}$$

Separating into real and imaginary part produces the two equations much referred to as the *power flow equations*, making the base for power flow analysis:

$$P_i = \sum_{j=1}^{n} |V_i||V_j||Y_{ij}| \cos(\theta_{ij} - \delta_i + \delta_j) \tag{2.4}$$

$$Q_i = -\sum_{j=1}^{n} |V_i||V_j||Y_{ij}| \sin(\theta_{ij} - \delta_i + \delta_j) \tag{2.5}$$

Now consider a simple, three-bus system as the one in figure 2.1 to support the further explanation of PFA. The specified values for each bus are indicated, making bus 1 the slack bus, and buses 2 and 3 load and generator buses respectively.

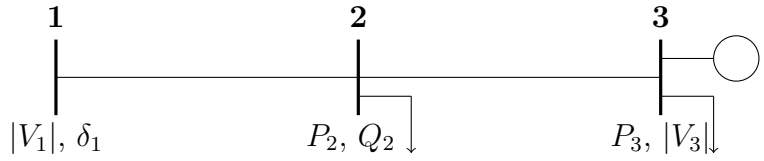


Figure 2.1: A simple, three-bus system with specified parameters indicated for each bus

The power flow equations are a set of non-linear equations and several different solution strategies exist for solving such equation systems simultaneously. The Newton-Raphson method is the most widely used method for this purpose, and it has also proved to be the most efficient and practical method for large power systems [22]. Based on this, the NR is the only methodology that will be introduced and used for deterministic power flow in this thesis.

The NR is an iterative method which takes an initial estimate of the unknown variables (hereafter *state variables*), expanding the power flow equations in Taylor's series above this estimate and obtaining a set of linear equations (refer to appendix A.4 for the linearization procedure). This linear set is then solved, and a new estimate for the state variables is obtained. This process is iterated until an acceptable value of mismatch between specified and calculated values of powers is reached. The linear equation set may be formulated as

$$\begin{bmatrix} \mathbf{\Delta P} \\ \mathbf{\Delta Q} \end{bmatrix} = \begin{bmatrix} \mathbf{J} \end{bmatrix} \begin{bmatrix} \mathbf{\Delta}\delta \\ \mathbf{\Delta}|\mathbf{V}| \end{bmatrix} \quad \Leftrightarrow \quad \mathbf{\Delta W} = \mathbf{J\Delta X} \tag{2.6}$$

where $\mathbf{J}$ denotes the Jacobian matrix of the state variables. The left-hand side matrix is the power mismatch vector $\mathbf{\Delta W}$ and contains the difference between scheduled and calculated powers, for the specified injected powers. $\mathbf{\Delta X}$ represents the incremental change to be made to the state variables before moving to the next iteration. In the case of the system in figure 2.1, we get $\mathbf{\Delta W} = [\Delta P_2 \quad \Delta P_3 \quad \Delta Q_2]^T$. The right-hand side vector is the state vector $\mathbf{\Delta X}$, consisting of the unknown voltage magnitudes and powers. For figure 2.1, $\mathbf{\Delta V} = [\Delta\delta_2 \quad \Delta\delta_3 \quad \Delta|V_2|]^T$.

The solution of the load flow problem is obtained after running a sufficient number of iterations such that the mismatch $\mathbf{\Delta W}$ is at an acceptable value after the last iteration. The iterative procedure can be formulated as follows:

$$\mathbf{\Delta X}^{(k)} = \mathbf{J}^{-1}\mathbf{\Delta W} \tag{2.7}$$

$$\mathbf{X}^{(k+1)} = \mathbf{X}^{(k)} + \mathbf{\Delta X}^{(k)} \tag{2.8}$$

where k is the iteration step. Starting at iteration step 0, the incremental increase of the state variables are found from eq. (2.7), then the state variables are updated with these values according to eq. (2.8). When each entry of the power mismatch vector $\mathbf{\Delta W}$ has reached a value less than the maximum accepted error, the load flow calculation is concluded with the state variables set to the value of the current iteration step.

2.4 Basic statistics

This section will present some basics of statistics that is considered necessary background for the reader. The emphasis is laid on giving a general introduction of the most important concepts relevant for this thesis. Some additional expressions for selected probability distributions relevant for the types of random variables used in the case studies of chapter 4 can be found in appendix A.1.

2.4.1 Random variables

In the field of statistics, results of repeated experiments on selected individuals, are used to make inferences of the behaviour of a large group of individuals. The smaller group of randomly selected variables are often referred to as *samples*, and constitute the information that is used together with fundamental laws of probability and statistical interference, to draw conclusions about the larger group of individuals. This group is the collection of all individuals from which the samples were taken, often referred to as a *population* or a *sample space*. [26, 4]

It is often advantageous to provide a numerical description of the outcome of the repeated experiments, and hence the concept of *random variables* is introduced. A random variable, let it be denoted X, is actually rather a function, that assigns a real number $X(s)$ to each sample s in the sample space S. Dependent on the underlying sample space, the random variables can further be either discrete or continuous [4]. For a discrete random variable, the sample space consists of finite number of discrete values of which the random variable can take. The sample space of a continuous random variable however, contains an infinite number of values, represented by the number of points on a line segment. [26]

2.4.2 Distribution functions

Having defined a random variable X, the rule for describing the probability measures associated with X is called a *probability distribution*. There are different ways in which to describe the probability distribution, one is by using the cumulative distribution function (CDF) [4]

$$F_X(x) = P(X \leq x) \tag{2.9}$$

Where the uppercase X denotes the random variable and the corresponding lowercase x denotes particular values in the range of X. $P(\cdot)$ represents the probability of an event. The probability is by definition a value restricted between 0 and 1, and consequently so will the CDF $F_X(x)$. The limits of 0 and 1 are approached by the CDF when the values of x approach the extremas $-\infty$ and ∞, respectively. [4]

The CDF can be defined for both discrete and continuous random variables. For the discrete variable, the CDF will change its value in jumps corresponding to the discrete values x that the random variable X can take. The CDF of a continuous variable is a continuous function, which must be obtained from integration of some function able to describe $P(X = x)$ for all x.

This function is the *probability density function* (PDF) for the continuous random variable,

and the *probability mass function* (PMF) for the discrete random variable. These functions aim to describe the probability of X taking the value of x, i.e. $P(X = x)$, and to gather all these probabilities in a function. For the discrete variable, only able to take a finite number of different values, the PMF will be a sequence of probabilities $P(X = x_i)$. For convenience, this sequence will be denoted $p(x_i)$. The PMF of any discrete random variable will then satisfy the following two simple properties [4]:

$$\begin{aligned} &0 \leq p(x_i) \leq 1, \quad i = 1, 2, ..., n \\ &\sum_i p(x_i) = 1 \end{aligned} \tag{2.10}$$

where n is the finite number of discrete values x_i that the discrete random variable X is able to take.

If X is a continuous random variable, the probabilities are associated with *intervals* on the real number line, rather than points as for the discrete variable. Consequently, the probability of X taking the discrete value x is equal to zero [4]. So for the continuous variable, it makes more sense to describe the probability of X taking a value within some interval (a, b). From this reasoning, the probability density function $f_X(x)$ of the continuous random variable X is defined as the function that, when being integrated from a to b, yields the probability of X taking a value in the interval (a, b):

$$P(a < X \leq b) = \int_a^b f_X(x)dx \tag{2.11}$$

From the definition of the PDF $f_X(x)$, it has to satisfy the following conditions: [4]

$$\begin{aligned} &f_X \geq 0, \quad \text{for all } x \\ &\int_{-\infty}^{\infty} f_X(x)dx = 1 \end{aligned} \tag{2.12}$$

Some expressions for the CDF, PMF and PDF of distributions relevant for the case studies in this thesis are provided in appendix A.1.

2.4.3 Characteristic measures of random variables

The distribution functions introduced in the previous section provide a complete description of the probability distribution of the random variable X. Some times, however, it is neither necessary nor desired to describe a random variable by its density function. In practical engineering problems it is often more convenient to describe the random variable by some

measures that give a better insight into the properties of the variable's distribution that are actually of interest.

The *mean* or *expected value* is one such measure, and as the name implies, the mean is the value that the random variable is expected to take in the long run, i.e. when a sufficient number of experiments are performed on a sample such that the difference from using different samples is negligible. In a more mathematical sense, the mean is the weighted average of the values of the random variable, where the "weights" are the probabilities. Note that by using an average value, the mean does not necessarily need to be in the range of possible values of X; this will often occur when the random variable is discrete. It will however give an indication as to the location of the probability distribution or density, as the mean value will be the value of X that has the highest probability of occurring [4].

Although the mean value gives some indication about the probability distribution of X, it is alone not necessarily very descriptive, as it provides no information about the dispersion of the distribution around the mean. Such information can be obtained from the *variance* of the distribution. Denoting the mean of X by μ_X, the variance is defined as a weighted average of $(x - \mu_X)^2$. The more likely an occurrence of x is, the larger weight is assigned to it. Note that squaring this measure places the variance in a different dimension as the random variable itself, and hence the variance is not always an easily intuitive measure of the dispersion. The square root of the variance is defined as the *standard deviation* of X, often denoted σ_X. Being a weighted sum of the quantity $(x - \mu_X)$, it is often more convenient to use, at it has the same dimension at the random variable [4].

The mean and standard deviation will in many cases provide a sufficient amount of information to describe the randomness of a variable for a particular purpose. In this thesis, quantitative results for output random variables in the case studies will be provided by means of these two measures only. Much more information can be provided for a random variable by for instance calculating its moments and cumulants, which will be introduced in the subsequent section. However, these measures does not always have an intuitive meaning, and to create such a meaning some manipulation has to be performed to the moments and cumulants calculated directly from the variable's probability distribution. Two such measures, easily obtained from the moments of the distribution, are the *skewness* and *kurtosis*. The skewness is a measure of the distribution's asymmetry about the mean, where a nonzero value indicates that there exists asymmetry. The *kurtosis* makes some indication as to the shape of the tails of the distribution, i.e. when the values of x approach $-\infty$ and ∞. [26]

In this thesis, the mean and standard deviation will be used to present quantitative results, and the skewness and kurtosis will be used to obtain more information about the shape of the distribution, in order to approximate the CDFs of the random variables obtained by PLF.

2.4.4 Moments

The analytical and approximate methods presented in this thesis greatly makes use of the concepts of statistical moments, thus this will be the main focus of the following introduction. The moments of a probability distribution are important properties of the distribution, and in many cases a certain number of moments are enough to give a sufficient description of the distribution for that specific purpose. Let X be a continuous or discrete probabilistic variable with its probability distribution described by $f(x)$. Let further μ, σ and ν denote the expected value, standard deviation and variation of X, respectively. The r-th order raw moment of X is defined as:

$$\mu_r = E[X^r] = \begin{cases} \int_{-\infty}^{\infty} x^r f(x)dx, & X \textit{ continuous} \\ \sum_X x^r p(x), & X \textit{ discrete} \end{cases} \tag{2.13}$$

where the first-order raw moment equals the definition of the expected value or mean, $m_1 = \mu$. We further define the r-th order central moment as the raw moment about the mean:

$$m_r = E[(x-\mu)^r] = \begin{cases} \int_{-\infty}^{\infty} (x-\mu)^r f(x)dx, & X \textit{ continuous} \\ \sum_X (x-\mu)^r p(x), & X \textit{ discrete} \end{cases} \tag{2.14}$$

The second-order central moment is the coefficient of variance of X, $m_2 = \nu$. The standard deviation is then found from the second central moment as $\sigma = \sqrt{\nu}$. An alternative, often preferred way of determining the variance (and consequently also the standard deviation) uses only the first two orders of the raw moments: [26]

$$\nu = \sigma^2 = E[X^2] - \mu^2 \tag{2.15}$$

Other important properties of a distribution can be achieved from the n-th order standard central moment, defined as the ratio of central moment to the standard deviation,

$$\lambda_n = \frac{m_n}{\sigma^n} \tag{2.16}$$

By definition, the third an fourth-order standard central moment correspond to the distribution's skewness and kurtosis, respectively.

2.4.5 Cumulants

Cumulants are another sort of property of a probabilistic distribution, posing as an alternative to the moments. The moments and cumulants determine each other in the sense that two distributions with the same set of moments will also have identical sets of cumulants, and vice versa. The cumulants can be found from the moments, or directly from the cumulant generated function C_X. The latter is defined as the logarithm of the moment generating function M_X, and the cumulant of order n is found by evaluating the derivative of the CGF at $t = 0$:

$$M_X(t) = E[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} f(x) dx \tag{2.17}$$

$$C_X(t) = ln(M_X) \tag{2.18}$$

$$\kappa_n = \left. \frac{d^{(n)} K_X(t)}{dt^n} \right|_{t=0} \tag{2.19}$$

Alternatively, the cumulants can be found from the distribution's moments from the following relationship:

$$\begin{aligned} \kappa_1 &= m_1 \\ \kappa_n &= m_n - \sum_{i=1}^{n-1} \binom{n-1}{i} \cdot m_i \cdot \kappa_{n-i} \end{aligned} \tag{2.20}$$

In this thesis, the first six cumulants of random distributions will be used, and they are found from equation (2.20) an expressed in terms of the first six raw moments below [28]:

$$\begin{aligned} \kappa_1 &= \mu_1 \\ \kappa_2 &= \mu_1 - \mu_2 \\ \kappa_3 &= \mu_3 - 3\mu_2\mu_3 - 2\mu_1^3 \\ \kappa_4 &= \mu_4 - 4\mu_3\mu_1 - 3\mu_2^2 + 12\mu_2\mu_1^2 - 6\mu_1^4 \\ \kappa_5 &= \mu_5 - 5\mu_4\mu_1 - 10\mu_3\mu_2 + 20\mu_3\mu_1^2 + 3 + \mu_2^2\mu_1 - 60\mu_2\mu_1^3 + 24\mu_1^5 \\ \kappa_6 &= \mu_6 - 6\mu_5\mu_1 - 15\mu_4\mu_2 + 30\mu_4\mu_1^2 - 10\mu_3^2 + 120\mu_3\mu_2\mu_1 - 120\mu_3\mu_1^3 \\ &\quad - 270\mu_2^2\mu_1^2 + 360\mu_2\mu_1^4 - 120\mu_1^6 \end{aligned} \tag{2.21}$$

The much used measures mean and standard deviation can easily be found directly from

the two first cumulants by the relationships [28]

$$\begin{aligned} \mu &= \kappa_1 \\ \sigma &= \sqrt{\kappa_2} \end{aligned} \tag{2.22}$$

3 | General methodology

This chapter will present the theoretical background and derivation of the two PLF methods that are studied thoroughly in this thesis: the Point estimate method (PEM) and the Cumulant method (CM). Each method is first presented in its simplest form with uncorrelated variables, then modified to include the possibility of having correlated input random variables. The output of each method is raw moments or cumulants for the random variables, and some adjustment to this raw output is usually of interest. Common statistical properties like mean and standard deviation are easily obtained from the moments and cumulants, as shown in the following sections. If an approximation to the CDF of the random variable is desired, this can be achieved by using an expansion series, the procedure for which is outlined in section 3.3.

3.1 Point estimate method

The point estimate method will first be introduced in section 3.1.1 with the assumption of uncorrelated input random variables. Then a modified version able to handle correlated input variables is outlined in section 3.1.2.

3.1.1 Independent input random variables

The input to the PEM is a set of random variables, let these be denoted $x_i (i = 1, 2, ..., n)$, where n is the total number of input random variables. Assume for now these to be independent. In the scope of this thesis, x_i can be a generation unit or a load.

As in the deterministic load flow described in section 2.3, the output variables are the unknown state variables and the power flow in each line of the system. Unlike in the DLF, these will be of random nature when the input variables are random. The PEM is an approximate method, so based on knowledge of the random nature of the input variables, it approximates the random behaviour of the output variables. This approximation is achieved on the basis of a weighted sum of the results achieved by a number of DLFs run for strategic states with regards to the input variables.

These strategic states are determined based on information about the random distribution of the input variables. Based on certain criteria, each input variable x_i will produce K points

for which to run a DLF. Each of these K points is further made up of two parts; one is the value, or the new *location* for the input variable in question, the other one is the corresponding *weight* that the results from the DLF run at this location is to be given when all DLF results are summed up to give the final solution.

Now let a probabilistic output Y be defined by $Y = F(x_1, x_2, ..., x_n)$ such that F is the function expression relating the input and output variables. Following the above conceptual introduction to the PEM, the random output Y is approximated by evaluating F (which then represents a regular DLF) in K points for each of the n input variables $(x_1, ..., x_n)$.

Several different schemes have been proposed for the PEM, depending on how many points K are generated for each input variable. As the function F has to be evaluated $K \times n$ times, the value of K is obviously strongly influencing the efficiency of the method. Reference [17] investigates several different such schemes, finding that a $2 \times n$ scheme is not well suited for systems of a scale of realistic power systems, while a $2 \times n + 1$ scheme overcomes many of these issues by only requiring one additional evaluation of F. Thus, the latter strikes a good balance between accuracy and efficiency when a large number of input variables is used, and will be the scheme used further in this thesis. The following derivation and solution method of the PEM is referred to as Hong's method, and the derivation is much inspired by [14] and [17].

Each point k of the total K points generated from input variable x_i can be viewed as a pair consisting of a location $x_{i,k}$ and a weight $w_{i,k}$. Let the k-th location be defined by

$$x_{i,k} = \mu_i + \xi_{i,k}\sigma_i \tag{3.1}$$

where μ_i and σ_i are the mean and standard deviation of x_i, and $\xi_{i,k}$ is the standard location, yet to be determined.

Now let r denote the order of the moment of x_i, such that the r-th standard central moment is written as

$$\lambda_{i,r} = \frac{m_{i,r}}{\sigma_i^r} \tag{3.2}$$

where, in accordance with section 2.4.4, $m_{i,r}$ is the r-th order central moment of x_i and σ_i^r is the standard deviation of x_i. The weight of the location (i, k) is then found by combining

the expressions of (3.3)

$$\begin{aligned} \sum_{k=1}^{K} w_{i,k} &= \frac{1}{n} \\ \sum_{k=1}^{K} w_{i,k}\xi_{i,k}^{r} &= \lambda_{i,r} \quad (r = 1, 2, ..., 2K-1) \end{aligned} \tag{3.3}$$

Once the location and weight for all input variables are found, the r-th order moment of the output Y is found from

$$m_{Y,r} = E[Y^r] = \sum_{i=1}^{n}\sum_{k=1}^{K} w_{i,k} Y_{i,k}^{r} \tag{3.4}$$

where $Y_{i,k}$ is the function expression evaluated at the mean of all variables except x_i, whose mean is replaced by location $x_{i,k}$, i.e. $Y_{i,k} = F(\mu_1, \mu_2, ..., x_{i,k}, ..., \mu_n)$.

As previously mentioned, the $2n+1$ scheme is chosen for the PEM in this thesis. In this scheme, $K = 3$ because three locations are considered for each variable x_i. However, the third location $x_{i,3}$ is set to be equal to the mean value of x_i. To obtain this situation, it is obvious from eq. (3.1) the third standard location, $\xi_{i,3}$ has to be zero. Using $K = 3$ and $\xi_{i,3} = 0$ in eq. (3.3), yields, for the standard locations and weights of x_i:

$$\begin{aligned} \xi_{i,k} &= \frac{\lambda_{i,3}}{2} + (-1)^{3-k}\sqrt{\lambda_{i,4} - \frac{3}{4}\lambda_{i,3}^2}, \quad (k = 1, 2) \\ \xi_{i,3} &= 0 \end{aligned} \tag{3.5}$$

$$\begin{aligned} w_{i,k} &= \frac{(-1)^{3-k}}{\xi_{i,k}(\xi_{i,1} - \xi_{i,2})}, \quad (k = 1, 2) \\ w_{i,3} &= \frac{1}{n} - \frac{1}{\lambda_{i,4} - \lambda_{i,3}^2} \end{aligned} \tag{3.6}$$

The location of the k-th point produced by x_i is found from (3.1), but note that setting $\xi_{i,3} = 0$ results in $x_{i,3} = \mu_i$ for all values of i; in other words, for every x_i, the third location is fixed to its mean value. According to (3.4), F will then be evaluated at the same point once for each input variable. Obviously, this is unnecessary, so instead F will be evaluated once with *all* variables at their expected values $\mu_i (i = 1, ..., n)$, with the corresponding weight w_0 equal to the sum of the third-location weights of all variables:

$$w_0 = \sum_{i} w_{i,3} = 1 - \sum_{i=1}^{n} \frac{1}{\lambda_{i,4} - \lambda_{i,3}^2} \tag{3.7}$$

The above result reduces the required number of evaluations of F to $2n + 1$, even though

three locations are used for each of the n input variables.

When all locations and weights are found from equations (3.1) and (3.5) - (3.6), respectively, the random output variables can be formulated as the weighted sum of the DLF results. The solution of the DLF at location k for variable x_i is defined as

$$Y_{i,k} = F(\mu_1, \mu_2, ..., x_{i,k}, ..., \mu_n) \tag{3.8}$$

where $Y_{i,k}$ is the output variable of interest, which in the context of PLF can be a voltage magnitude, a voltage angle or the active or reactive power flow in a line. The current output $Y_{i,k}$ is now used to estimate the raw moments of the final output random variable Y:

$$E[Y^j] = \sum_{i,k} w_{i,k}(Y_{i,k})^j \tag{3.9}$$

where E represents the expected value, such that $E[Y^j]$ is the j-th order raw moment of Y.

The statistical characteristics of Y can be obtained from its raw moments of various order according to equations (2.13) - (2.16). In many cases the expected value (equal to the first raw moment) and the standard deviation (equal to the second central moment) will provide sufficient amount of information about the probability distribution for that specific purpose. Calculating even higher-order moments will provide more information and gives a more correct and detailed image of the distribution.

Implementation procedure

A qualitative description of the computational procedure is presented here. A flow chart of the procedure is also included in fig. 3.1.

1. Quantify the statistical properties of the input random variables that are necessary to calculate the locations and weights. Let the expected values be represented by the vector

   $$\boldsymbol{\mu} = [\mu_1, ..., \mu_n]^T \tag{3.10}$$

   and equivalently the standard deviation, skewness and kurtosis by the vectors $\boldsymbol{\sigma}$, $\boldsymbol{\lambda_3}$ and $\boldsymbol{\lambda_4}$, respectively.

2. By use of the available information in the above four vectors, calculate for all variables x_i

   (a) The standard locations $\xi_{i,k}$ according to (3.5)

   (b) The weights $w_{i,k}$ and w_0 according to (3.6)

(c) The locations $x_{i,k}$ according to (3.1)

Save the locations and weights to appropriate vectors.

3. For each value of i and k, replace the value of x_i with location $x_{i,k}$ and run a DLF. For each DLF, update the raw moments of the random output variables Y of interest, according to equation (3.9).

4. Run one DLF with all variables at their expected values. Update the raw moments of each Y according to equation (3.9), replacing $w_{i,j}$ with w_0.

5. If an approximation of the PDF and/or CDF of the output random variables is desirable, make use of an extension series to accomplish this. Extension series are discussed in section 3.3.

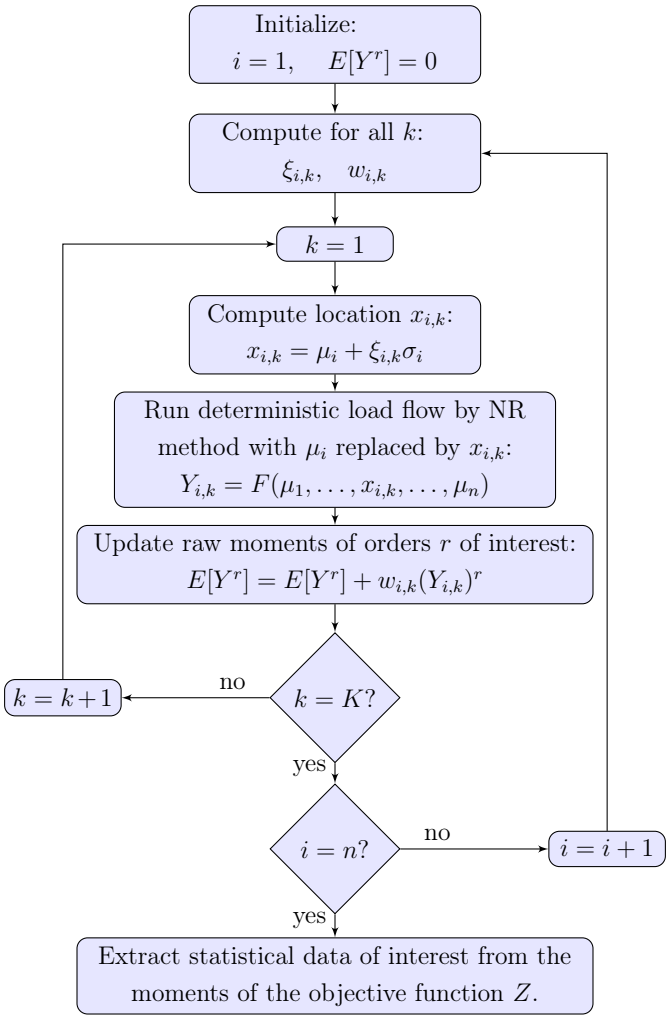


Figure 3.1: Flow chart of the point estimate method for independent input random variables

3.1.2 Correlated input random variables

Before starting off the derivation, some assumptions regarding the correlation schemes in the system are made in order to keep the derivation comprehensive and relevant for the case studies in this thesis. In principle, correlation can exist between power injections at the same bus, and/or between power injections at different buses. Additionally, correlation can exist between generation units and loads. In this thesis, only correlation of power injections between different buses is studied, so this assumption is made throughout the whole report. Also, generation units and loads are always assumed independent of each other, i.e. correlation can only exist between a number of loads in the system and/or between a number of generation units. Thus, if there exists correlation between both loads and generation units, there are two sets of correlated variables, each set described by a specific correlation matrix. The following derivations of modified PEM and eventually CM, will address one such set, i.e. the correlated random variables are assumed to be either loads or generation units. The other set is treated equivalently in a separate process.

Hong's method as outlined in section 3.1.1 requires the input random variables to be independent of each other. When correlation exists, some modification needs to be done to the methodology for it to give an accurate approximation of the randomness of the output variables. In this thesis, an analytical approach based on orthogonal transformation of the correlated variables [18] is chosen for handling correlated variables in the PEM.

By orthogonal transformation, the set of correlated input random variables are converted into a corresponding set of independent variables. The points in which to run DLFs are calculated from equations (3.1) and (3.5) - (3.6), but using the statistical properties of the independent variables rather than the original correlated ones. Once all these points are properly defined, they are transformed back into the original space before the DLFs are run in each point.

Let $\mathbf{x}$ be a vector of correlated random input variables such that

$$\mathbf{x} = [x_1, ..., x_m]^T, \quad m \leq n \tag{3.11}$$

where T denotes the transpose of a vector, m is the number of correlated random variables and n is the total number of input random variables. As in the regular PEM, the expected values, standard deviations, skewness and kurtosis of the now correlated input random variables may be represented by the vectors $\boldsymbol{\mu}_\mathbf{x}$, $\boldsymbol{\sigma}_\mathbf{x}$, $\boldsymbol{\lambda}_{\mathbf{3},\mathbf{x}}$ and $\boldsymbol{\lambda}_{\mathbf{4},\mathbf{x}}$, respectively.

The correlation between the input variables is described by a covariance-matrix $\mathbf{C}_\mathbf{x}$ (refer to appendix A.2). By definition, this matrix is symmetric, and thus there always exists a matrix $\mathbf{B}$ through which the set $\mathbf{x}$ of correlated variables can be orthogonally transformed

into a set $\mathbf{z}$ of independent variables:

$$\mathbf{z} = \mathbf{B} \cdot \mathbf{x} \tag{3.12}$$

Decomposing $\mathbf{C_x}$ into a lower and an upper triangular matrix by Cholesky decomposition, yields

$$\mathbf{C_x} = \mathbf{L}\mathbf{L}^T \tag{3.13}$$

Now the inverse of the lower triangular matrix $\mathbf{L}$ is exactly the matrix $\mathbf{B}$ that is needed for the orthogonal transformation, i.e.

$$\mathbf{B} = \mathbf{L}^{-1} \tag{3.14}$$

The covariance matrix $\mathbf{C_z}$ of the new set of independent variables is equal to the identity matrix $\mathbf{I}$. Please refer to appendix A.3 for an introduction to the orthogonal transformation and derivation of eqs. (3.12)–(3.14).

Now, to find the locations and weights of the independent variables in accordance with equations (3.1) and (3.5) - (3.6), $\boldsymbol{\mu_z}$, $\boldsymbol{\sigma_z}$, $\boldsymbol{\lambda_{3,z}}$ and $\boldsymbol{\lambda_{4,z}}$, i.e. the expected values, standard deviations, skewness and kurtosis of the independent variables, $\mathbf{z}$ needs to be determined. Under the assumption that the crossed-order moments of an order higher than two are equal to zero, the coefficients of skewness and kurtosis of the *correlated* variables can be represented in matrices such that the values of $\boldsymbol{\lambda_{3,x}}$ and $\boldsymbol{\lambda_{4,x}}$ are located on the diagonal of an elsewhere zero matrix:

$$\begin{aligned} \boldsymbol{\lambda_{3,x}} &= \mathrm{diag}(\lambda_{3,x_1}, ..., \lambda_{3,x_m}) \\ \boldsymbol{\lambda_{4,x}} &= \mathrm{diag}(\lambda_{4,x_1}, ..., \lambda_{4,x_m}) \end{aligned} \tag{3.15}$$

Then, according to [18], all the necessary statistical properties of the transformed, independent variables can be found from the following formulas:

$$\boldsymbol{\mu_z} = \mathbf{B}\boldsymbol{\mu_x} \tag{3.16}$$

$$\lambda_{3,z_i} = \sum_{r=1}^{m} B_{ir}^3 \lambda_{3,x_r} \sigma_{x_r}^3 \tag{3.17}$$

$$\lambda_{4,z_i} = \sum_{r=1}^{m} B_{ir}^4 \lambda_{4,x_r} \sigma_{x_r}^4 \tag{3.18}$$

where $\boldsymbol{\mu_z}$ is a vector of the expected values of the transformed variables, λ_{3,z_i} and λ_{4,z_i} are the skewness and kurtosis of variable z_i, respectively. Recall that the covariance matrix $\mathbf{C_z}$

is the identity matrix and according to appendix A.2, the standard deviation of z_i is then

$$\sigma_{z_i} = \sqrt{\mathbf{C_z}(i,i)} = 1 \tag{3.19}$$

When locations and weights have been generated from the transformed variables, $2 \times m + 1$ points are created in the independent space, each point having one variable z_i replaced by location $z_{i,k}$ and the remaining variables of $\mathbf{z}$ at their expected values given by $\boldsymbol{\mu_z}$. Then these points are transformed back into the original space by the inverse of the orthogonal transformation from equation 3.12 before DLFs are run in the resulting transformed points.

Implementation procedure - correlated variables

1. Given a correlation coefficient matrix $\mathbf{C}_\rho$, create the covariance matrix $\mathbf{C_x}$ according to the expressions provided in appendix A.2. Calculate matrices $\mathbf{L}$ and $\mathbf{B}$ by equations (3.13) and (3.14), respectively.

2. Obtain for the transformed variables $\boldsymbol{\mu_z}$, $\boldsymbol{\sigma_z}$, $\boldsymbol{\lambda_{3,z}}$ and $\boldsymbol{\lambda_{4,z}}$ from eqs. (3.16)–(3.19). If $m < n$, calculate these properties also for the uncorrelated input random variables $(x_{m+1}, ..., x_n)$.

3. For all variables, including the transformed variables $z_i (i = 1, ..., m)$ and the uncorrelated variables $x_i (i \neq 1, ..., m)$, calculate standard locations $\xi_{i,k}$, weights $w_{i,k}$ and w_0 and locations $z_{i,k}$ and $x_{i,k}$, respectively, according to eqs. (3.1), (3.5) and (3.6).

4. For the transformed variables, create $2 \times m + 1$ points on the form $(\mu_{z_1}, ..., z_{i,k}, ..., \mu_{z_m})$ and store the points in appropriate vectors.

5. Transform each point in the independent space defined above, back to the original space by applying the inverse orthogonal transformation, $\mathbf{x} = \mathbf{B}^{-1}\mathbf{z}$

6. For $i = 1, ..., m$, run $2 \times m$ DLFs in each of the points defined in point 5 above, all the while updating random output variable Y with $w_{i,j}$ according to eq. (3.9). For $i > m$, run DLFs according to the procedure outlined in section 3.1.1, continuing to update Y.

7. Run one DLF for all variables at their expected values, updating Y with w_0. Note that the expected values for the correlated variables were obtained in point 5.

8. If an approximation of the PDF and/or CDF of the output random variables is desirable, make use of an extension series to accomplish this. Extension series are discussed in section 3.3.

A flow chart of the point estimate method as it was outlined in this section, is included in figure 3.2.

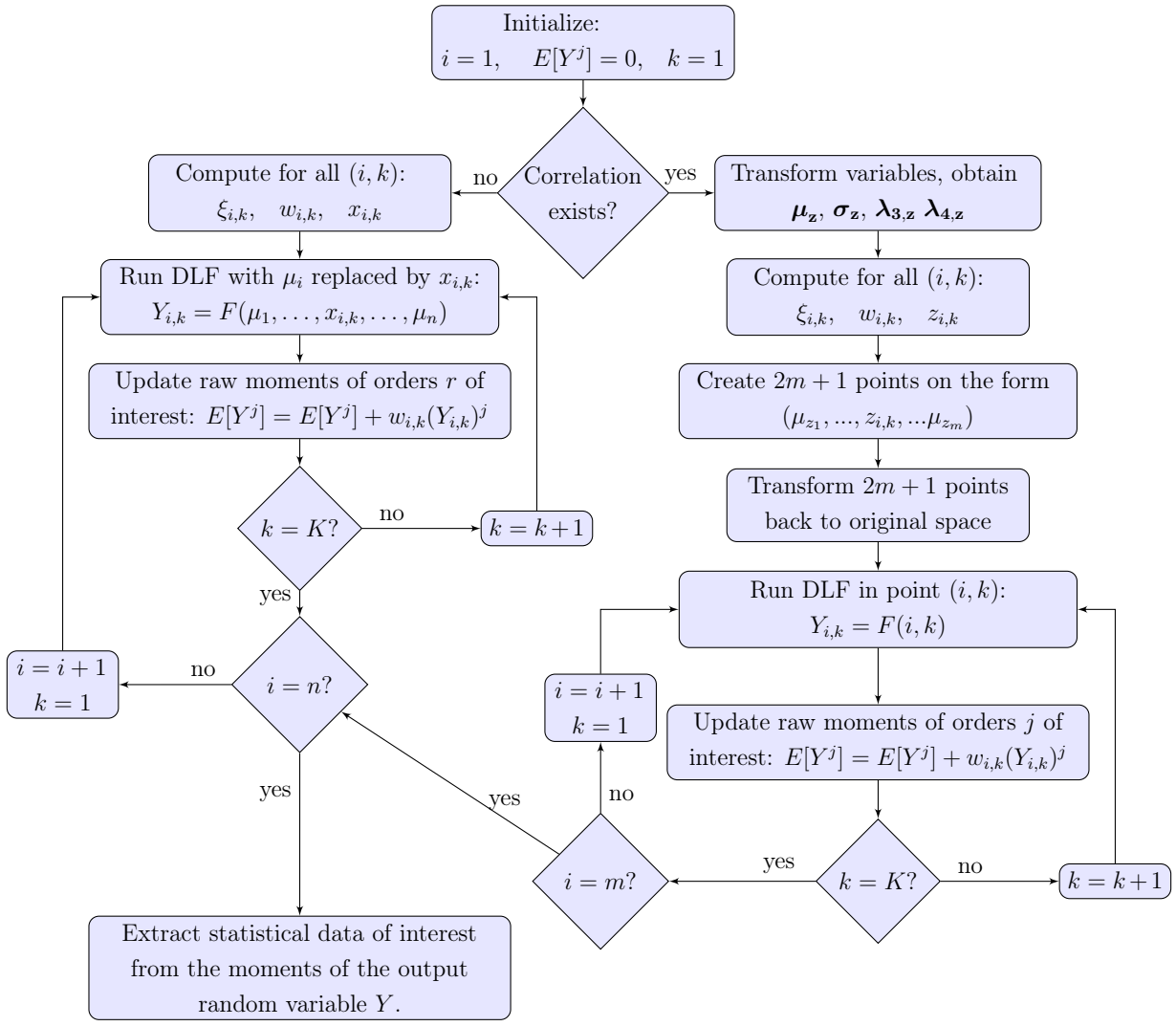


Figure 3.2: Flow chart of the point estimate method, including the option of correlated variables.

3.2 Cumulant method

Equivalently as for the PEM, the CM is introduced in section 3.2.1 assuming independent input random variables, then in section 3.2.2 the method is modified to accept correlated input variables.

3.2.1 Independent input random variables

The CM is based on linearization of the power flow equations. Recall that the linearized power flow equations are also what make up the basis for each iterative step of the NR methodology of solving the deterministic load flow problem. With reference to section 2.3, the linearized power flow equations can therefore be expressed as

$$\mathbf{\Delta X} = \mathbf{S}\mathbf{\Delta W} \tag{3.20}$$

where $\mathbf{S} = \mathbf{J^{-1}}$. In the deterministic load flow described in section 2.3, $\mathbf{\Delta X}$ represented the incremental change in the state variables necessary to reduce the power mismatches of $\mathbf{\Delta W}$ before the next iteration step. In this context, $\mathbf{\Delta X}$ will be interpreted in a slightly different way. The following derivation of the CM is mainly inspired by [28].

Let $\mathbf{X}$ still represent the state (output) variables and $\mathbf{W}$ the input variables (known power injections). However, the power injections at n buses are now random variables, let them be denoted $u_i (i = 1, 2, .., n)$ and initially assumed independent of each other. Thus, some (or all) entries of $\mathbf{W}$ are now random variables and consequently the output variables $\mathbf{X}$ are random as well.

Now let $\mathbf{W_0}$ denote the input variable vector in the case that all random input variables are located at their respective mean values. Assume that there exists a vector $\mathbf{X_0}$ which is the solution to the DLF problem with $\mathbf{W} = \mathbf{W_0}$, thus representing the mean of the random output variables. Further, the randomness of the input variables $\mathbf{W}$ can be modelled as disturbances $\mathbf{\Delta W}$ around their respective means $\mathbf{W_0}$. According to eq. (3.20), $\mathbf{\Delta X}$ represents the change in the output variables due to this disturbance $\mathbf{\Delta W}$ in the input variables. In other words, $\mathbf{\Delta X}$ describes the randomness of the output variables due to the randomness of the input variables:

$$\mathbf{\Delta X} = \mathbf{S_0}\mathbf{\Delta W} \tag{3.21}$$

where $\mathbf{S_0}$ is the sensitivity matrix relating $\mathbf{X_0}$ and $\mathbf{W_0}$, i.e. the inverse of the Jacobian matrix corresponding to the solution of the DLF problem with $\mathbf{W} = \mathbf{W_0}$.

The above paragraph summarizes the general idea on which the CM is built, and the

linearization requirement that was briefly discussed in section 2.1 becomes evident at this point. The randomness of the output variables is calculated using a sensitivity matrix $\mathbf{S_0}$ that is actually a linear approximation to the original load flow equations at the mean of each random input variable. Consequently, the relationship in eq. (3.21) is only valid for disturbances, i.e. random variations, relatively close to the mean. This fact can explain the generally poor performance of the CM in its most basic form, if random variables with large variation and/or dominant points are far from the mean. Thus, caution should be taken regarding the random distribution of the input variables when the CM is used.

When performing load flow analysis, the active and reactive power flows in the branches of the system are usually of interest. Hence an equivalent relationship as for the voltage magnitudes and angles in eq. (3.21) is of interest for the line power flows as well. Equivalently as for the bus power injections, the line power flows can be described by the following set of nonlinear equations:

$$P_{ij} = V_i V_j (G_{ij} cos\theta_{ij} + B_{ij} sin\theta_{ij}) - t_{ij} G_{ij} B_{ij} V_i^2 \tag{3.22}$$

$$Q_{ij} = V_i V_j (G_{ij} sin\theta_{ij} - B_{ij} cos\theta_{ij}) + (t_{ij} B_{ij} - b_{ij0}) V_i^2 \tag{3.23}$$

where t and b_0 represent the off-nominal transformer tap value and ground susceptance, respectively. Notice that these equations can be expressed as functions of the state variables $\mathbf{X}$, so let the active and reactive power flows in each line of the system be represented by the vector $\mathbf{Z} = \mathbf{g}(\mathbf{X})$. Linearizing the line power flow equations around the mean of the state variables $\mathbf{X_0}$ yields (refer to appendix A.4 for the derivation of eq. (3.24)):

$$\mathbf{\Delta Z} = \mathbf{D_0 \Delta X} \tag{3.24}$$

where $\mathbf{D_0}$ is the sensitivity matrix relating a random variation in the state variables $\mathbf{X}$ to the corresponding change in the power flows $\mathbf{Z}$. Because $\mathbf{\Delta X}$ in this sense are deviations away from the mean $\mathbf{X_0}$, $\mathbf{D_0}$ is calculated with respect to these mean values:

$$\mathbf{D_0} = \left. \frac{\partial \mathbf{Z}}{\partial \mathbf{X}} \right|_{\mathbf{X}=\mathbf{X_0}} \tag{3.25}$$

In order to get a direct relation between the output variables that are now the line power flows $\mathbf{Z}$, and the input random variables $\mathbf{W}$, eq. (3.20) is substituted into eq. (3.24), yielding

$$\mathbf{\Delta Z} = \mathbf{T_0 \Delta W} \tag{3.26}$$

with $\mathbf{T_0} = \mathbf{D_0 S_0}$. Now finally, the following set of linear equations describes the random

variation of state variables and line power flows respectively, due to random variation in the bus power injections:

$$\mathbf{\Delta X} = \mathbf{S_0 \Delta W} \tag{3.27}$$

$$\mathbf{\Delta Z} = \mathbf{T_0 \Delta W} \tag{3.28}$$

The above set of equations forms the foundation of the CM, but the way in which to describe the randomness $\mathbf{\Delta W}$ of the input variabels is yet to be introduced. The distribution of a random variable can be quantified in different ways, depending on the amount of information available and the necessary accuracy of the description. The $2n + 1$ scheme of the PEM, as it was outlined in section 3.1, represents each random variable by four quantities obtained from its first four raw moments. In the CM, cumulants of the distribution of each random variable are used to quantify the randomness of that variable. The reason for this choice is related to some important properties of cumulants that will be explained shortly. Cumulants were introduced in section 2.4.5 and are directly related to the moments of a distribution. Properties like mean, standard deviation, skewness and kurtosis can thus easily be obtained from this relation.

Now to the properties of the cumulants that make the CM able to perform the calculation of the randomness of the output variables in a single operation. The mathematical relationships are summarized below: [28, 14]

1. Let u_t be a random variable that is a sum of two independent random variables u_1 and u_2. Let the r-th order cumulants of the latter two be given by $K_{u_1}^{(r)}$ and $K_{u_2}^{(r)}$ $(r = 1, 2, ..., k)$. Then the r-th order cumulant of u_t is

$$K_{u_t}^{(r)} = K_{u_1}^{(r)} + K_{u_2}^{(r)} \tag{3.29}$$

   The above relationship can be generalized to account for a situation where u_t is a sum of n independent random variables, such that its r-th order cumulant is given by

$$K_{u_t}^{(r)} = \sum_{i=1}^{n} K_{u_i}^{(r)} \tag{3.30}$$

2. Let $h = au + b \quad (a \neq 0)$ where u is a random variable. Then the r-th order cumulant of h is

$$K_h^{(r)} = \begin{cases} aK_u^{(r)} + b & (r = 1) \\ a^r K_u^{(r)} & (r \geq 2) \end{cases} \tag{3.31}$$

Recall at this point that the random variables $u_i (i = 1, 2, ..., n)$ that constitute the input to the CM are net bus power injections. The net active injection to bus i is the difference between generated power P_{G_i} and load demand power P_{D_i} at that bus; similarly for the reactive injection. If both the generation and load are random variables, the active power injection P_i is the sum of the two random variables P_{G_i} and P_{D_i}. When cumulants are used to quantify the randomness of these variables, property 1 above simplifies this process to a simple summation of the various-order cumulants of P_{G_i} and P_{D_i}. Generalizing this case to include active and reactive power injections at all buses, gathered in the input vector $\mathbf{\Delta W}$ yields:

$$\mathbf{\Delta W}^{(r)} = \mathbf{\Delta W_g}^{(r)} + \mathbf{\Delta W_d}^{(r)} \tag{3.32}$$

where $\mathbf{\Delta W}^{(r)}$ contains the r-th order cumulant of all entries of the input (net power injection) vector $\mathbf{W}$. Equivalently, the r-th order cumulants of the corresponding generation and load demands are represented in $\mathbf{\Delta W_g}^{(r)}$ and $\mathbf{\Delta W_d}^{(r)}$, respectively.

Now that the cumulants of the input variables have been found, property 2 above can be used to obtain an expression for the cumulants of the output variables. Recall that the relations between the input variables and the two sets of output variables are described by eqs. (3.27) and (3.28), respectively. Looking closely at these equations, each output variable is in fact a linear sum of all random input variables. Thus, combining properties 1 and 2 above, the following relationship between the r-th order cumulants of input and output variables is valid: [28]

$$\begin{aligned} \mathbf{\Delta X}^{(r)} &= \mathbf{S_0}^r \mathbf{\Delta W}^{(r)} \\ \mathbf{\Delta Z}^{(r)} &= \mathbf{T_0}^r \mathbf{\Delta W}^{(r)} \end{aligned} \tag{3.33}$$

where $\mathbf{S_0}^r$ and $\mathbf{T_0}^r$ are obtained by element-wise operations, i.e. any element (i, j) of $\mathbf{S_0}^r$ is equal to

$$S_0^r(i, j) = (S_0(i, j))^r \tag{3.34}$$

and correspondingly for $\mathbf{T_0}^r$.

The CM is concluded with eq. (3.33) as the final output, which is the cumulants of state variables and line flow, are obtained. Statistical data like standard deviations of these output variables can easily be obtained by taking the square root of the second cumulant (according to section 2.4.5). The expected values are already found in $\mathbf{X_0}$ and $\mathbf{Z_0}$, respectively. Knowledge of higher-order cumulants is necessary in order to obtain an even better image of the probability distribution. In those cases extension series like the Gram-Charlier series may be used to obtain an approximation to the distribution.

Implementation procedure

A qualitative description of the computational procedure is presented below. Refer to section 4.1.3 for a more detailed step-by-step demonstration.

1. Obtain the expected values $\mathbf{X_0}$ and $\mathbf{Z_0}$ of the output variables by running one DLF with all random variables at their expected values. Save the last version of the Jacobian matrix $\mathbf{J_0}$ and obtain $\mathbf{S_0}$ by inverting $\mathbf{J_0}$.

2. Compute the cumulants of the random input variables $u_i (i = 1, 2, ..., n)$, u_i being the active or reactive part of the power generation or load demand at bus i. The cumulants can be found directly from the random variable's raw moments, as described in section 2.4.5.

3. Compute the cumulants of the net power injections according to eq. (3.32). Save the cumulants of all input variables to a matrix with increasing cumulant order along the columns.

4. Obtain sensitivity matrix $\mathbf{D_0}$ according to eq. (3.25). In this thesis, ready-made software in MATPOWER was used to do this calculation.

5. Calculate the cumulants of the output variables according to eq. (3.33), starting with the second cumulant and repeating the matrix operation for as many orders of cumulants that was calculated for the input variables in step 2. The first-order cumulants of the output variables are equal to their respective mean values, which were obtained in step 1.

6. If an approximation of the PDF and/or CDF of the output variables is desirable, make use of an extension series to obtain this.

A flow chart of the cumulant method as it was outlined in this section, is included in figure 3.3 (inspired by [28]).

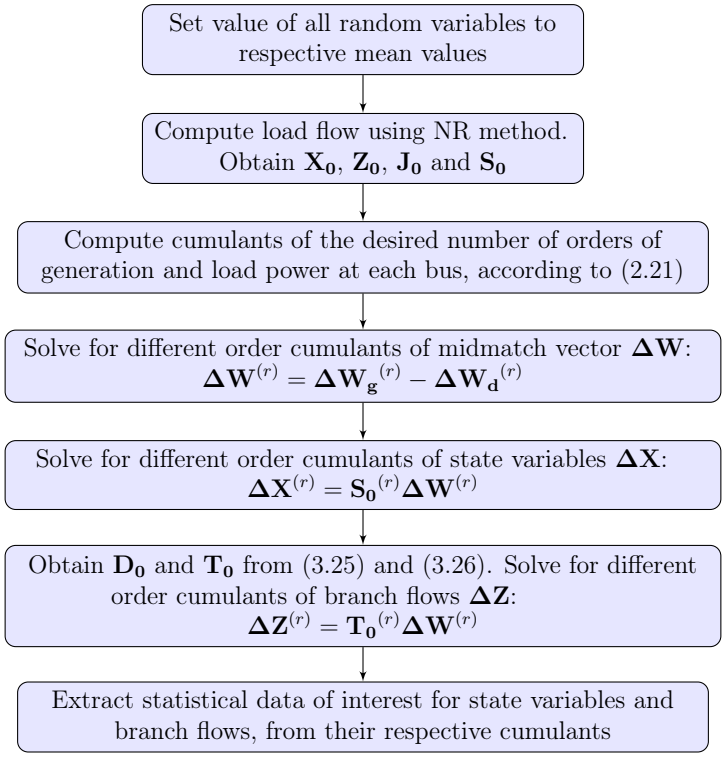


Figure 3.3: Flow chart of the CM for independent input random variables

3.2.2 Correlated input random variables

The methodology outlined in section 3.2 requires independent input variables, and in the following section some enhancements will be introduced to make the CM able to handle correlated variables. This enhanced CM was first introduced in [7] and the following derivation is much inspired by this. Note that this derivation is for the most general case, i.e. it includes the possibility of having some input variables that are not random, and also correlation between only some of the random variables.

Let $\mathbf{W} = [w_1, ..., w_l]$ be the power injections at each bus in the system, l being the total number of buses. Each power injection consists of a generation and load part, i.e. $w_i = w_{G_i} - w_{L_i} (i = 1, ..., l)$.

Now let $u_i (i = 1, ..., n)$ denote the random input variables, n being the total number of random variables and u_i representing either a generation unit or a load. Let further m of these random variables be correlated, such that

$$\mathbf{u} = [u_1, ..., u_m], \quad m \leq n \tag{3.35}$$

denote the the correlated random variables.

The idea of the modified methodology is to model the correlated variables as a function of several independent ones. The independent variables are obtained through orthogonal transformation that will be explained shortly. In the calculation of the cumulants of the output variables, the independent variables are used together with modified sensitivity matrices $\mathbf{S_0}$ and $\mathbf{T_0}$.

The process of orthogonal transformation in this modified CM is similar to the methodology used to transform the correlated variables in the modified PM described in section 3.1.2, as it also done on the basis of decomposing a correlation matrix by Cholesky's decomposition. However, in this section, the *correlation coefficient* matrix $\mathbf{C_u}$ (as opposed to the *covariance* matrix utilized in the modified PEM) is decomposed by Cholesky decomposition, yielding

$$\mathbf{C_u} = \mathbf{L}\mathbf{L}^T \tag{3.36}$$

where $\mathbf{L}$ is an inferior (lower) triangular matrix and T denotes the transpose of a matrix. $\mathbf{C_u}$ is a symmetric matrix, so by definition there always exist a matrix $\mathbf{B}$ that will transform the correlated variables of $\mathbf{u}$ to a vector of independent variables $\mathbf{v}$:

$$\mathbf{v} = \mathbf{B}\mathbf{u} \tag{3.37}$$

The transformation matrix $\mathbf{B}$ is, according to appendix A.3, equal the the inverse of the inferior matrix of the decomposed correlation matrix:

$$\mathbf{B} = \mathbf{L^{-1}} \tag{3.38}$$

Now the orthogonal transformation can also be interpreted in the way that the vector of correlated random variables $\mathbf{u}$ is modelled as a function of several independent random variables

$$\mathbf{u} = \mathbf{B^{-1}}\mathbf{v} = \mathbf{L}\mathbf{v} \tag{3.39}$$

Again, refer to appendix A.3 for a more detailed derivation of the above relations. The orthogonal transformation is the basis of the methodology of the modified CM, and up until this step, the modifications to the CM are equivalent to the modifications done to the PEM in section 3.1.2. In section 3.1.2, the necessary statistical properties of the transformed input variables were calculated analytically. In the case of the CM however, the statistical properties, that is the cumulants, of the transformed input variables will be calculated from a large selection of random numbers generated from the probability distribution of each variable. This methodology avoids complex analytical cumulant calculation and can in principle handle the cumulant calculation of input variables with complex distribution functions.

The general idea is to first obtain random numbers of each of the correlated variables. Then these numbers are transformed by the orthogonal transformation outlined above, to obtain random numbers of the corresponding independent variables. Raw moments and finally cumulants of the independent variables are then easily obtained from the random numbers representing the distribution of each random independent variable. The transformed, independent variables are used directly as input variables when the cumulants of the output variables are calculated, together with modified sensitivity matrices. The final, modified CM equations will be presented in their entirety by the end of this section. First, the methodology for generating random numbers will be presented.

The methodology for generation of random numbers used in this modified CM was first proposed by [18], then adopted to the CM in [7]. In [18] the obtained random numbers are used as input to MC simulation, so the methodology is referred to as an MC technique in [7]. The random numbers of the correlated variables are generated by first making use of a transformation methodology known as the Nataf transformation [15]. The Nataf transformation transforms a set of correlated random variables to a corresponding set of correlated standard normal variables [1]. Random numbers are then generated from each of these standard normal distributions before being transformed back to the original space. The purpose of performing the transformation to the standard normal space is to be able to handle any type of probability distribution in the input variables; the random numbers are generated from the simple standard normal distribution in any case. This makes this modified CM a very versatile methodology, as it is able to handle correlation between variables described by a large variety of different probability distributions. A possible drawback of the methodology is that perfect knowledge of the CDFs of the input variables is required. This can sometimes pose a challenge when the data on which the probabilistic modelling is based, is limited. From this qualitative introduction follows a more detailed step-by-step description of the MC technique. The derivation is mainly inspired by [7, 18].

1. For the correlated random variables $\mathbf{u}$ with correlation matrix $\mathbf{C_u}$, let $\mathbf{F}$ denote a vector containing the marginal CDFs of the variables in $\mathbf{u}$. Further, let $\mathbf{z} = [z_1, ..., z_m]$ denote the set of correlated standard normal variables obtained by marginal transformation of $\mathbf{u}$ [18]:

$$\mathbf{z} = \mathbf{\Phi}^{-1}(\mathbf{F}(\mathbf{u})) \tag{3.40}$$

   where $\mathbf{\Phi}$ contains the marginal CDFs of $\mathbf{z}$, i.e. for each variable the CDF of the standard normal distribution. As will turn out obvious from the successive steps, it is not necessary to actually carry out the calculation of eq. (3.40), it is merely included

[1] The standard normal distribution is a normal distribution with $\mu = 0$ and $\sigma = 1$

here to illustrate the relation between $\mathbf{u}$ and $\mathbf{z}$.

2. Start with determining the correlation matrix $\mathbf{C_z}$. It has the same physical interpretation as any other correlation matrix, so the elements along the diagonal are all equal to 1. The off-diagonal elements turn out to be related to the corresponding elements of the original correlation matrix, $\mathbf{C_u}$. Let, in accordance with appendix A.2, the elements of $\mathbf{C_u}$ and $\mathbf{C_z}$ be denoted $\rho_{u_{ij}}$ and $\rho_{z_{ij}}$, respectively. Then the following relationship can be established [15]:

$$\rho_{z_{ij}} = H\rho_{u_{ij}} \tag{3.41}$$

where H is a function that is dependent on $\rho_{u_{ij}}$ and the marginal distributions of u_i and u_j. Reference [15] provides expressions for calculating H for a number of different probability distributions associated with u_i and u_j. In the special case that the two variables in question are normal distributed, which is the only case relevant for the case studies on the CM in this thesis, the value of H is equal to unity.

3. Once the elements of $\mathbf{C_z}$ have been obtained, the process of orthogonal transformation as outlined in the beginning of this section is again made use of. In this case, it will be used to transform a set of samples $\mathbf{E_s}$ generated from a vector $\mathbf{e}$ of *independent* standard normal variables, into the corresponding set of samples $\mathbf{Z_s}$ from the vector $\mathbf{z}$ of *correlated* standard normal variables.[2] The relationship between correlated and independent variables in the orthogonal transformation was described by eq. (3.39), and can be adopted to also account for this case, yielding:

$$\mathbf{Z_s} = \mathbf{L_z}\mathbf{E_s} \tag{3.42}$$

where $\mathbf{L_z}$ is an inferior, triangular matrix obtained by Cholesky decomposition of the correlation matrix $\mathbf{C_z}$ of the correlated standard normal variables $\mathbf{z}$. At this point, make note of the fact that the actual generation of random numbers is done for a set of *independent standard normal variables*. Considering a set of m variables, m random samples, each of the desired sample size N, are generated independently from the standard normal distribution to obtain the matrix $\mathbf{E_s}$. This type of sampling is a common function that is included in most available software for statistical analysis [18], including MATLAB that was used for implementation in this thesis. By using expressions for H provided in [15], the correlation matrix $\mathbf{C_z}$ and hence the transformation matrix

[2]The subscript s will in the context of this section indicate that the matrix in question contains random numbers (samples) generated from each of the random variables represented in the matrix. As an example, if N samples were to be generated for m random variables, these would be represented in an $m \times N$ matrix, denoted with subscript s.

$\mathbf{L_z}$ can be obtained for input variables characterized by a large selection of probability distributions.

4. Once the random numbers $\mathbf{Z_s}$ of correlated standard normal variables are found, random numbers $\mathbf{U_s}$ of the original correlated variables $\mathbf{u}$, can be obtained by marginal transformation of $\mathbf{Z_s}$. Recall that the CDFs of $\mathbf{u}$ and $\mathbf{z}$ were given by $\mathbf{F}$ and $\mathbf{\Phi}$, and that the relation between the two sets of variables were given by eq. (3.40). Then, using the inverse of the marginal transformation of eq. (3.40), with the random numbers of the standard normal variables as input to $\mathbf{\Phi}$, yields:

$$\mathbf{U_s} = \mathbf{F}^{-1}(\mathbf{\Phi}(\mathbf{Z_s})) \tag{3.43}$$

$\mathbf{U_s}$ is now a $m \times N$ matrix containing N random numbers for each of the total of m correlated random variables. If a MC method was used to perform the PLF, these numbers would constitute the points in which to run the simulations [18]. In the modified CM however, these samples will be used to finally obtain cumulants of an independent set of random variables.

Orthogonal transformation and cumulant calculation

Given a matrix $\mathbf{U_s}$ of random numbers of the correlated input variables $\mathbf{u}$, the transformation from correlated to independent variables outlined in the beginning of this section can finally be performed. The goal of this transformation is to obtain a corresponding matrix $\mathbf{V_s}$ of random numbers representing the probability distribution of the independent variables $\mathbf{v} = [v_1, ..., v_m]$. The orthogonal transformation from $\mathbf{u}$ to $\mathbf{v}$ is described by eq. (3.37) and repeated below:

$$\mathbf{v} = \mathbf{B}\mathbf{u} \tag{3.44}$$

In $\mathbf{V_s}$ and $\mathbf{U_s}$, each column contains one random number of each of the corresponding random variables $\mathbf{v}$ and $\mathbf{u}$, respectively. Adopting eq. (3.44) to this situation, each column of $\mathbf{V_s}$ is obtained by performing the orthogonal transformation on the corresponding column of $\mathbf{U_s}$. Thus, N simple matrix calculations are performed on each of the columns in $\mathbf{U_s}$, to obtain the final matrix $\mathbf{V_s}$ of random numbers of the independent variables.

Each row i of $\mathbf{V_s}$ can now be interpreted as a discrete distribution describing the randomness of independent random variable v_i. Thus, the cumulants of v_i can easily be obtained by the expressions provided in section 2.4. Let the r-th order cumulant of the elements of $\mathbf{v}$ be denoted $\mathbf{v}^{(r)}$.

At this point, recall that the input variables to the CM are various-order cumulants of the

net power injections at each bus, that is the difference between generation and demand. In accordance with the assumptions made in the introduction of section 3.1.2, $\mathbf{u}$ and $\mathbf{v}$ represent a set of variables that are all *either* generation units or loads, and they are all located at different buses. Assume for now that $\mathbf{u}$ (and consequently also $\mathbf{v}$) are load variables and denote their elements $u_{L_i}(i = 1, ..., m)$ and $v_{L_i}(i = 1, .., m)$, respectively. Assume that no correlation exist between the generating units in the system, but that the generating units are considered random variables, denoted u_{G_i}. Then, the r-th order cumulant of the net power injection at bus i in the system is given by:

$$\Delta w_i^{(r)} = \begin{cases} u_{G_i}^{(r)} + v_{L_i}^{(r)}, & i = 1, ..., m \\ u_{G_i}^{(r)} + u_{L_i}^{(r)}, & i = m, ..., n \end{cases} \tag{3.45}$$

where the first case refers to buses at which a correlated load is located, and the second case to the buses with random variables that are not correlated to any other variables. Note that in both cases, the random variables involved in the subtraction are independent of each other, thus the r-th order cumulants can be added or subtracted (refer to section 3.2.1). When $\Delta w_i^{(r)}$ is calculated for all values of i, the vector of r-th order cumulants of all bus power injections is represented by $\mathbf{\Delta W}^{(r)}$, as before. Note that the power injections at all buses are included in $\mathbf{W}$, not only the ones corresponding to the state variables, as is the case for the deterministic NR load flow calculation and also the original CM of section 3.2.1. This is done in order to simplify the calculation of the modified sensitivity matrices, and will be explained further in the next subsection.

If there now was correlation also between c generating units at different buses, the procedure that has been outlined in this section for handling correlation would have to be repeated separately for these variables to finally obtain a set of independent variables. Then, at the buses where these units are located, the correlated generation unit u_{G_i} in eq. (3.45) would be replaced by its corresponding independent variable v_{G_i}.

Modified CM equations

As for the case of uncorrelated variables outlined in section 3.2.1, the first step of this modified CM is also to run deterministic DLF for all variables located at their expected values. From this step, $\mathbf{X_0}$, $\mathbf{Z_0}$, $\mathbf{J_0}$, $\mathbf{S_0}$ and $\mathbf{G_0}$ are obtained, exactly as before.

However, in order to be able to use the transformed, independent variables $\mathbf{v}$ directly in the input variables, some modifications need be made to the sensitivity matrices $\mathbf{S_0}$ and $\mathbf{T_0}$ that are used in the calculation of the cumulants of the output variables. Let the modified matrices be denoted $\mathbf{S_1}$ and $\mathbf{T_1}$. At this point an important note should be made: in the

modified CM as proposed in [7], the voltage magnitudes and angles at *all* buses in the system are included in the output vector $\mathbf{X}$, just as it was earlier noted that all net power injections are included in the input vector $\mathbf{W}$. This is simply an alternative way of doing the calculations (instead of including only the state variables in $\mathbf{X}$), but it requires the sensitivity matrices to be modified accordingly. The sensitivity matrix $\mathbf{S_0}$ describes the sensitivities of the cumulants of voltage magnitudes and angles to changes in the cumulants of power injections. Including all voltage magnitudes and angles in $\mathbf{X}$ instead of just the state variables, corresponds to inserting zeros into the rows and columns corresponding to the non-state variables. Similar steps are carried out for $\mathbf{G_0}$, to be able to produce the correct matrix $\mathbf{T_0}$.[3]

Once the full sensitivity matrices have been obtained, the following expressions can be used to obtain the elements of the modified sensitivity matrices $\mathbf{S_0}$ and $\mathbf{T_0}$:[7]

$$\begin{aligned} s_{ij1} = \sum_{k=j}^{m} s_{ik0} l_{km}, \quad t_{ij1} = \sum_{k=j}^{m} t_{ik0} l_{km} \quad j = 1, ..., m \\ s_{ij1} = s_{ij0}, \quad t_{ij1} = t_{ij0}, \qquad j \neq 1, ..., m \end{aligned} \tag{3.46}$$

where s_{ij0}, s_{ij1}, t_{ij0} and t_{ij1} are elements of $\mathbf{S_0}$, $\mathbf{S_1}$, $\mathbf{T_0}$ and $\mathbf{T_1}$, respectively. l_{km} is element (k, m) of the triangular matrix $\mathbf{L}$ obtained by Cholesky decomposition of the correlation matrix.

Eventually, the final equations of the modified CM can be formulated. Having obtained the elements of the r-th order input variable vector $\mathbf{\Delta W}$ by eq. (3.45) and the full sensitivity matrices by eq. (3.46), the r-th order cumulants of the output variables are finally obtained by [7]:

$$\begin{aligned} \mathbf{\Delta X}^{(r)} &= \mathbf{S_1}^r \mathbf{\Delta W}^{(r)} \\ \mathbf{\Delta Z}^{(r)} &= \mathbf{T_1}^r \mathbf{\Delta W}^{(r)} \end{aligned} \tag{3.47}$$

where $\mathbf{S_1}^r$ and $\mathbf{T_1}^r$ are obtained by element-wise operations, i.e. any element (i, j) of $\mathbf{S_1}^r$ is equal to

$$S_1^r(i, j) = (S_1(i, j))^r \tag{3.48}$$

[3]The use of the full sensitivity matrices (and including all voltage magnitude, angles and input power injections in the equations) is not explicitly stated in [7], but after some investigation of the equations this conclusion is drawn for the purpose of this thesis. This conclusion is also supported by [28], where the basic CM (without correlated variables) is described. Here, the sizes of the sensitivity matrices are explicitly given as $2n \times 2n$ and $2b \times 2n$ for $\mathbf{S_0}$ and $\mathbf{T_0}$ respectively, b being the number of branches in the system. Using the full sensitivity matrices as described above and including all variables in the CM equations for the uncorrelated case as described in section 3.2.1, will produce the exact same results as using the reduced sensitivity matrices. Both these cases have been implemented for the uncorrelated case, and can be investigated in the available Matlab codes.

and correspondingly for $\mathbf{T_1}^r$.

Implementation procedure - modified CM

A qualitative description of the computational procedure is presented below. Refer to section 4.1.3 for a more detailed step-by-step demonstration.

1. Obtain the expected values $\mathbf{X_0}$, $\mathbf{Z_0}$, $\mathbf{J_0}$ and $\mathbf{S_0}$ from DLF by the NR method, with all random variables located at their respective mean values. Obtain $\mathbf{D_0}$ by eq. (3.25). Expand $\mathbf{S_0}$ and $\mathbf{D_0}$ to full sensitivity matrices.

2. Identify the correlation coefficient matrix $\mathbf{C_u}$ of the correlated variables. Calculate the correlation matrix $\mathbf{C_z}$ of the correlated standard normal variables by eq. (3.41). Obtain $\mathbf{L_z}$ by Cholesky decomposition of $\mathbf{C_z}$.

3. Generate random numbers $\mathbf{E_s}$ of m independent standard normal variables. Obtain random numbers $\mathbf{Z_s}$ of correlated standard normal variables by the orthogonal transformation in eq. (3.42).

4. Obtain random numbers $\mathbf{U_s}$ of the original, correlated random variables $\mathbf{u}$ by the marginal transformation in eq. (3.43).

5. Obtain random numbers $\mathbf{V_s}$ of independent variables $\mathbf{v}$ by the orthogonal transformation in eq. (3.44). Compute the cumulants of $\mathbf{v}$ from the random numbers representing each v_i in $\mathbf{V_s}$.

6. Compute the r-th order cumulant $\Delta w_i^{(r)}$ of the net power injection w_i at bus i by eq. (3.45). Represent $\Delta w_i^{(r)}$ for all buses $i = 1, ..., n$ in $\mathbf{\Delta W}^{(r)}$.

7. Obtain the modified sensitivity matrices $\mathbf{S_1}$ and $\mathbf{T_1}$ by eq. (3.46).

8. Calculate the higher-order cumulants of the output variables according to eq. (3.47). The first-order cumulants of the output variables are equal to their respective mean values obtained in step 1.

9. If an approximation of the PDF and/or CDF of the output variables is desirable, make use of an extension series to obtain this.

3.3 Gram-Charlier expansion

The Gram-Charlier expansion is chosen in this thesis to perform the approximation of the CDFs of the output random variables.

Let u be a random variable, which in the context of this thesis will be a voltage magnitude, voltage angle, active or reactive power flow. Denote the mean and standard deviation of u by μ and σ, respectively. As described in appendix A.1.1, u can be expressed as a standardized variable z, by denoting

$$z = \frac{u - \mu}{\sigma} \tag{3.49}$$

Let the CDF and PDF of the standardized variable be denoted $F(x)$ and $f(x)$, respectively. Further let the CDF and PDF of the standard normal distribution (refer to appendix A.1.1) be denoted by Φ and ϕ, respectively. By Gram-Charlier expansion (GCE), the distribution functions of the standardized random variable are [29]:

$$\begin{aligned} F(z) &= \Phi(z) + \frac{c_1}{1!}\Phi'(z) + \frac{c_2}{2!}\Phi''(z) + \frac{c_3}{3!}\Phi^{(3)}(z) + ... \\ f(z) &= \phi(z) + \frac{c_1}{1!}\phi'(z) + \frac{c_2}{2!}\phi''(z) + \frac{c_3}{3!}\phi^{(3)}(z) + ... \end{aligned} \tag{3.50}$$

The constants c_i are found from the central moments m_r (where r denotes the moment order) of the random variable u. The first six such constants are given below as [29]:

$$\begin{aligned} c_0 &= 1 \\ c_1 &= c_2 = 0 \\ c_3 &= -\frac{m_3}{\sigma^3} \\ c_4 &= -\frac{m_4}{\sigma^4} - 3 \\ c_5 &= -\frac{m_5}{\sigma^5} + 10\frac{m_3}{\sigma^3} \\ c_6 &= \frac{m_6}{\sigma^6} - 15\frac{m_4}{\sigma^4} + 30 \end{aligned} \tag{3.51}$$

Now the constants c_i can be found directly from eq. (3.51), but as will soon be shown, they can also be expressed in terms of the cumulants of u. In the context of this thesis, using the cumulants rather than the central moments is very advantageous. When using the PEM to conduct the PLF, the output random variables will be expressed in terms of their raw moments of various orders. The cumulants of that variable is then easily obtained from eq. (2.21), which relates the cumulants and the raw moments of a distribution. When conducting PLF by the CM, the output random variables are represented by their various-

order cumulants, which can then be used directly as input to the GCE. The relationship between central moments m_r and cumulants κ_r are [21]:

$$\begin{aligned} m_1 &= 0 \\ m_2 &= \kappa_2 = \sigma^2 \\ m_3 &= \kappa_3 \\ m_4 &= \kappa_4 + 3\kappa^2 \\ m_5 &= \kappa_5 + 10\kappa_2\kappa_3 \\ m_6 &= \kappa_6 + 15\kappa_2\kappa_4 + 10\kappa_3^2 + 15\kappa_2^3 ... \end{aligned} \tag{3.52}$$

Now inserting eq. (3.52) into eq. (3.51), eq. (3.50) can be expressed as [21]:

$$\begin{aligned} F(z) &= \Phi(z) - \frac{C_1}{6}\Phi^{(3)} + \frac{C_2}{24}\Phi^{(4)}(z) - \frac{C_3}{120}\Phi^{(5)} + \frac{C_4 + 10C_1}{720}\Phi^{(6)} ... \\ f(z) &= \phi(z) - \frac{C_1}{6}\Phi^{(3)} + \frac{C_2}{24}\Phi^{(4)}(z) - \frac{C_3}{120}\Phi^{(5)} + \frac{C_4 + 10C_1}{720}\Phi^{(6)} ... \end{aligned} \tag{3.53}$$

where the the parameters G_1 and G_2 turn out to be the skewness and kurtosis of the standardized varibale. These are determined in terms of the original variable u as [21]:

$$\begin{aligned} C_1 = \kappa_{3,z} &= \frac{\kappa_{3,u}}{(\kappa_{2,u})^{3/2}} \\ C_2 = \kappa_{4,z} &= \frac{\kappa_{3,u}}{(\kappa_{2,u})^2} \end{aligned} \tag{3.54}$$

where κ_r, u denotes the r-th order cumulant of the random variable u.

Reference [21] further uses Hermite polynomials to improve the approximation of the unknown distribution, and is then able to get an expression for the PDF and CDF of u directly in terms of the standardized variable z:

$$\begin{aligned} f(z) &= \phi(z)\left[1 + \frac{C_1}{6}(z^3 - 3z) + \frac{C_2}{24}(z^4 - 6z^2 + 3) + \frac{C_1^2}{72}(z^6 - 15z^4 + 45z^2 - 15)\right] \\ F(z) &= \Phi + \phi(z)\left[1 + \frac{C_1}{6}(z^2 - 1) + \frac{C_2}{24}(z^3 - 3z) + \frac{C_1^2}{72}(z^5 - 10z^3 + 15z)\right] \end{aligned} \tag{3.55}$$

4 | Case studies

This chapter presents the case studies done in this thesis. Three different test systems have been used; a simple 3 bus system, the IEEE 14 bus system, and the IEEE 24 bus Reliability test system (RTS). All case studies were implemented in Matlab [16] and MATPOWER 6.0 [30] was utilized to do deterministic load flow computations and some other operations.

4.1 Step-by-step demonstration

This chapter will utilize a small 3-bus system simple enough to demonstrate the PEM and CM in a step-by-step manner. For both methodologies, it is the modified versions able to handle correlation between variables (described in sections 3.1.2 and 3.2.2, respectively) that are demonstrated here. However, both methods are also implemented for the case of independent variables, according to the procedures outlined in sections 3.1.1 and 3.2.1. The results for the exact same system and probabilistic modelling, but with no correlation between variables, are presented in appendix D.1. A quantitative validation of the methodology with independent variables is provided in section 4.2.

4.1.1 3-bus test system modelling

The base case case system (without probabilistic modelling) is the same as the one in [19] and its one-line diagram is presented in figure 4.1. It includes one load bus and two generator buses where bus 1 is defined as the slack bus. Base case data and line data for the system are given in tables 4.1 and 4.2, respectively.

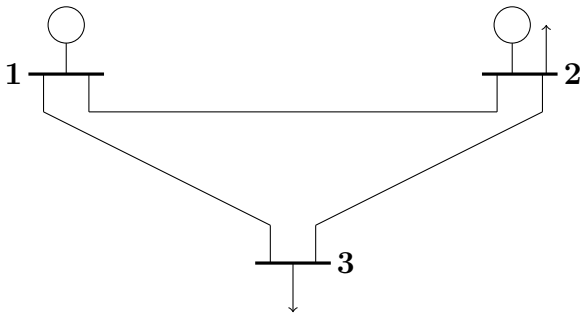


Figure 4.1: 3-bus test system for step-by-step demonstration

Table 4.1: Base case data for 3-bus test system

Bus i	Type	$\lvert V_i \rvert$	δ_i	P_{Li}	Q_{Li}	P_{Gi}	Q_{Gi}
1	Slack	1.05	0	0	0	-	-
2	PV	1.03	-	50	20	20	-
3	PQ	-	-	60	25	0	0

Table 4.2: Line data for 3-bus test system

From bus	To bus	R (p.u.)	X (p.u.)	b_{ij0} (p.u.)
1	2	0.08	0.24	0
1	3	0.02	0.06	0
2	3	0.06	0.018	0

The probabilistic modelling is based on the following:

- *Loads:* the loads at buses 2 and 3 are considered to be random variables. They are normal distributed with mean values μ equal to the base case values from table 4.1 and standard deviations σ equal to 10% of such mean values.
- *Generating units:* The generator at bus 2 is modelled with a capacity of 22 MW and a forced outage rate (FOR) of 0.09.
- *Correlation:* The loads at buses 2 and 3 are correlated with correlation factor $\rho = 0.4$. Loads and generating units are assumed independent.

4.1.2 Point Estimate Method: step-by-step solution

Having one generation unit and two complex loads as input random variables yields a total of $n = 5$ probabilistic inputs, of which $m = 4$ of these are correlated. According to the $2n+1$ scheme, DLF needs to be performed in 11 different locations.

Step 1: Quantify necessary statistical information

The expected value, standard deviation, skewness and kurtosis are the statistical properties of the input random variables that are necessary to calculate in order to apply the PEM. These are summarized for loads and generation in table table 4.3 below, based on the statistical modelling presented in section 4.1.1 and the properties of the various distributions, which can be found in appendix A.1. For convenience, the input random variables have been assigned numbers in the leftmost column of table 4.3.

Step 2: Calculate transformation matrices

Table 4.3: Statistical data for application of PEM

i	x_i	μ_{x_i}	σ_{x_i}	λ_{3,x_i}	λ_{4,x_i}
1	P_{L_2}	50	5	0	3
2	P_{L_3}	60	6	0	3
3	Q_{L_2}	20	2	0	3
4	Q_{L_3}	25	2.5	0	3
5	P_{G_2}	20.02	6.3	-0.1337	1.3662

The correlation matrix between buses 2 and 3 is given as:

$$\mathbf{C}_\rho = \begin{bmatrix} 1 & 0.4 \\ 0.4 & 1 \end{bmatrix} \tag{4.1}$$

For simplicity, the active and reactive parts of the complex loads at each bus are treated separately. Recall that as long as the same standard locations and weights are used for P_{L_i} and Q_{L_i}, they can be treated as if they were independent of each other [14]. In this case, both loads are normal distributed and therefore the active and reactive parts will always produce the same standard locations and weights.

The covariance matrix between the buses is calculated according to appendix A.2. For the active and reactive parts of the loads respectively, we get

$$\mathbf{C_{xP}} = \begin{bmatrix} \sigma^2_{P_{L2}} & \sigma_{P_{L2}P_{L3}} \\ \sigma_{P_{L2}P_{L3}} & \sigma^2_{P_{L3}} \end{bmatrix} = \begin{bmatrix} 25 & 12 \\ 12 & 36 \end{bmatrix} \tag{4.2}$$

$$\mathbf{C_{xQ}} = \begin{bmatrix} \sigma^2_{Q_{L2}} & \sigma_{Q_{L2}Q_{L3}} \\ \sigma_{Q_{L2}Q_{L3}} & \sigma^2_{Q_{L3}} \end{bmatrix} = \begin{bmatrix} 4 & 2 \\ 2 & 6.25 \end{bmatrix} \tag{4.3}$$

where the values of σ are in MW.

Further, the matrix $\mathbf{L}$, whose inverse is the transformation matrix $\mathbf{B}$, is found through Cholesky's decomposition [1]. For the active parts of the load:

$$\mathbf{L_P} = \begin{bmatrix} 5 & 0 \\ 2.4 & 5.5 \end{bmatrix}, \quad \mathbf{B_P} = \mathbf{L_P}^{-1} = \begin{bmatrix} 0.2 & 0 \\ -0.0837 & 0.1818 \end{bmatrix} \tag{4.4}$$

Equivalently, the matrices $\mathbf{L_Q}$ and $\mathbf{B_Q}$ for the reactive parts of the loads are easily found from decomposition of $\mathbf{C_{xQ}}$.

Step 3: Obtain statistical data of transformed input random variables

[1] In this thesis, MATLAB is used to do the calculations, and the built-in MATLAB function `chol` is used for the Cholesky's decomposition in this example

$\boldsymbol{\mu_z}$, $\boldsymbol{\sigma_z}$, $\boldsymbol{\lambda_{3,z}}$ and $\boldsymbol{\lambda_{4,z}}$, i.e. the statistical data for the transformed variables are found from eqs. (3.16)–(3.19). The calculations are shown for selected variables below, and summarized for all variables in table table 4.4.

$\boldsymbol{\mu_z}$ is easily obtained directly for all the transformed variables through a simple matrix operation. For the active parts:

$$\boldsymbol{\mu_z} = \mathbf{B}\boldsymbol{\mu_x} = \mathbf{B_P}\begin{bmatrix}\mu_{x_1}\\ \mu_{x_2}\end{bmatrix} = \begin{bmatrix}10\\ 6.55\end{bmatrix} \tag{4.5}$$

Further, σ_{z_i} is equal to 1 for all i, due to the covariance matrix of the transformed variables being equal to the identity matrix $\mathbf{I}$.

The skewness and kurtosis of the transformed variables are calculated according to eqs. (3.17) and (3.18), respectively. The calculation is shown for $z_1 = P_{L_2}$ below, and is done equivalently for all other variables:

$$\lambda_{3,z_1} = \sum_{r=1}^{2} B_{1r}^3 \lambda_{3,x_1}\sigma_{x_1}^3 = (0.2)^3(0)(5)^3 + (0)^3(0)(5)^3 = 0 \tag{4.6}$$

$$\lambda_{4,z_1} = \sum_{r=1}^{2} B_{1r}^4 \lambda_{4,x_1}\sigma_{x_1}^4 = (0.2)^4(3)(5)^4 + (0)^4(3)(5)^4 = 3 \tag{4.7}$$

Table 4.4: Statistical data for transformed variables

i	z_i	μ_{z_i}	σ_{z_i}	λ_{3,z_i}	λ_{4,z_i}
1	P_{L_2}	10	1	0	3
2	P_{L_3}	6.55	1	0	4.3605
3	Q_{L_2}	10	1	0	3
4	Q_{L_3}	6.55	1	0	4.3605

Step 4: Calculate locations and weights for all variables

Due to the fact that the active and reactive parts of the load have equal correlation matrices and equal percentage-wise random variation, the orthogonal transformation produces equal random variables for the active and reactive parts in the independent space, as is obvious from table 4.4. Thus, as the same standard locations and weights are used for the active and reactive parts of normal distributed loads, the final locations $z_{i,k}$ will also turn out equal in the independent space, according to eq. (3.1).

Calculation of standard locations, weights and locations are shown for the transformed variable z_1, representing the active part of the load at bus 2, is shown below. Calculation

is done equivalently for the other variables. For the non-correlated variable that is the generation at bus 2, the locations and weights are calculated based on its statistical data of table 4.3. The locations and weights are all summarized in table table 4.5.

$$
\begin{aligned}
\xi_{z_1,1} &= \frac{\lambda_{3,z_1}}{2} + (-1)^2\sqrt{\lambda_{4,z_1} - \frac{3}{4}\lambda_{3,z_1}^2} = \frac{0}{2} + \sqrt{3 - \frac{3}{4}\cdot 0^2} = 1.7321 \\
\xi_{z_1,2} &= \frac{\lambda_{3,z_1}}{1} + (-1)^2\sqrt{\lambda_{4,z_1} - \frac{3}{4}\lambda_{3,z_1}^2} = \frac{0}{2} - \sqrt{3 - \frac{3}{4}\cdot 0^2} = -1.7321 \\
w_{1,1} &= \frac{(-1)^2}{\xi_{z_1,1}(\xi_{z_1,1} - \xi_{z_1,2})} = \frac{1}{1.7321(1.7321 - (-1.7321))} = 0.1667 \\
w_{1,2} &= \frac{-1}{\xi_{z_1,2}(\xi_{z_1,1} - \xi_{z_1,2})} = \frac{-1}{-1.7321(1.7321 - (-1.7321))} = 0.1667 \\
z_{1,1} &= \mu_{z_1} + \xi_{z_1,1}\sigma_{z_1} = 10 + 1.7321 \cdot 1 = 11.7321 \\
z_{1,2} &= \mu_{z_1} + \xi_{z_1,2}\sigma_{z_1} = 10 - 1.7321 \cdot 1 = 8.2679
\end{aligned}
\tag{4.8}
$$

Table 4.5: Locations and weights for all variables

i	z_i	$z_{i,1}$	$w_{i,1}$	$z_{i,2}$	$w_{i,2}$
1	P_{L_2}	11.73	0.1667	8.27	0.1667
2	P_{L_3}	8.63	0.1147	4.45	0.1147
3	Q_{L_2}	11.73	0.1667	8.27	0.1667
4	Q_{L_3}	8.63	0.1147	4.45	0.1147
5	P_{G_2}	26.92	0.3921	12.28	0.3495

Step 5: Create $2m + 1$ points

Based on the locations and weights from table 4.5, $2m + 1$ points are created on the form $(\mu_{z_1}, ..., z_{i,k}, ..., \mu_{z_m})$ and are all summarized in table 4.6.

Step 6: Transform points back to the original space

Recall now that the active and reactive parts of the loads have different transformation matrices $\mathbf{B_P}$ and $\mathbf{B_Q}$, respectively. Thus, the columns representing active parts, in this case variables z_1 and z_2, are transformed using $\mathbf{B_P^{-1}}$, while variables z_3 and z_4 are transformed with $\mathbf{B_Q^{-1}}$.

Table 4.6: Transformed points

Point	P_{L2} (z_1)	P_{L3} (z_2)	Q_{L2} (z_3)	Q_{L3} (z_4)	Weight
1	$z_{1,1}$ 11.73	μ_{z_2} 6.55	μ_{z_3} 10	μ_{z_4} 6.55	$w_{1,1}$ 0.1667
2	$z_{1,2}$ 8.27	μ_{z_2} 6.55	μ_{z_3} 10	μ_{z_4} 6.55	$w_{1,2}$ 0.1667
3	μ_{z_1} 10	$z_{2,1}$ 8.63	μ_{z_3} 10	μ_{z_4} 6.55	$w_{2,1}$ 0.1147
4	μ_{z_1} 10	$z_{2,2}$ 4.45	μ_{z_3} 10	μ_{z_4} 6.55	$w_{2,2}$ 0.1147
5	μ_{z_1} 10	μ_{z_2} 6.55	$z_{3,1}$ 11.73	μ_{z_4} 6.55	$w_{3,2}$ 0.1667
6	μ_{z_1} 10	μ_{z_2} 6.55	$z_{3,2}$ 8.27	μ_{z_4} 6.55	$w_{3,2}$ 0.1667
7	μ_{z_1} 10	μ_{z_2} 6.55	μ_{z_3} 10	$z_{4,1}$ 8.63	$w_{4,1}$ 0.1147
8	μ_{z_1} 10	μ_{z_2} 6.55	μ_{z_3} 10	$z_{4,2}$ 4.45	$w_{4,2}$ 0.1147
9	μ_{z_1} 10	μ_{z_2} 6.55	μ_{z_3} 10	μ_{z_4} 6.55	w_0 -0.8670

As an example, the transformation of point 1 from table 4.6 is shown below.

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{B_P^{-1}} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} 5 & 0 \\ 2.4 & 5.5 \end{bmatrix} \begin{bmatrix} 11.73 \\ 6.55 \end{bmatrix} = \begin{bmatrix} 58.66 \\ 64.16 \end{bmatrix} \tag{4.9}$$

$$\begin{bmatrix} x_3 \\ x_4 \end{bmatrix} = \mathbf{B_Q^{-1}} \begin{bmatrix} z_3 \\ z_4 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 1 & 2.29 \end{bmatrix} \begin{bmatrix} 10 \\ 6.55 \end{bmatrix} = \begin{bmatrix} 20.00 \\ 25.00 \end{bmatrix} \tag{4.10}$$

Performing equivalent calculations for the remaining eight points of table 4.6, yields the actual points in which to run the DLFs, transformed back to the original space. These are summarized in table 4.7. Recall that there is yet one probabilistic input in the system that is not correlated to the others, namely the generation at bus 2, P_{G_2}, of which locations and weights were included in table 4.5. In table 4.7 P_{G_2} is included at its expected value for all the back-transformed points, while points 9-10 are added with P_{G_2} at the two different locations, keeping the final point 11 to all variables at their expected values, weighted with w_0. The locations and weights of P_G2 are calculated from eqs. (3.5) and (3.6) just as for the load variables. The mean, standard deviation, skewness and kurtosis constituting the input to eqs. (3.5) and (3.6) are found for generators by the expressions provided in appendix A.1.2.

Table 4.7: Points for which to run DLFs. All values of active and reactive power in MW and MVAr, respectively.

Point	P_{L2} (x_1)	P_{L3} (x_2)	Q_{L2} (x_3)	Q_{L3} (x_4)	P_{G_2} (x_5)	Weight ($w_{i,j}$)
1	58.66	64.16	20	25	20.02	0.1667
2	41.34	55.84	20	25	20.02	0.1667
3	50	71.48	20	25	20.02	0.1147
4	50	48.52	20	25	20.02	0.1147
5	50	60	23.46	26.73	20.02	0.1667
6	50	60	16.54	23.27	20.02	0.1667
7	50	60	20	29.79	20.02	0.1147
8	50	60	20	20.22	20.02	0.1147
9	50	60	20	29.79	26.92	0.3921
10	50	60	20	29.79	12.28	0.3495
11	50	60	20	29.79	20.02	-0.8670

Step 7: Run DLF at all points and update random output variable $Y_{i,j}$

In this thesis, MATPOWER has been utilized to run the DLFs. For each DLF, the results are used to update the first five raw moments of the random output variable Y, as these are used further in the Gram-Charlier expansion. A selection of load flow results is shown in table 4.7.

Table 4.8: Selected DLF results at selected points

Point	(i,k)	$V_{3_{i,k}}$ (p.u.)	$\delta_{3_{i,k}}$ (degrees)	$P_{13_{i,k}}$ (MW)	$Q_{13_{i,k}}$ (MVAr)	$w_{i,k}$
1	(1,1)	1.0334	-2.0929	78.4308	4.5914	0.1667
2	(1,2)	1.0301	-1.7490	60.2177	15.5907	0.1667
3	(2,1)	1.0303	-2.4004	78.7933	9.8383	0.1147
6	(3,2)	1.0320	-2.1027	69.4343	9.5008	0.1667
9	(5,1)	1.0300	-1.9068	64.7730	14.3909	0.3921

After each DLF, the random output variable Y is updated with the weighted result of the current DLF added to its previous value according to eq. (3.9). Note that Y can take the form of the state variables V_3, δ_2 and δ_3, and the active and reactive power flow in each line. Further, let the raw moment order j of Y be denoted $Y^{(j)}$. To illustrate the calculation of output variables Y, consider the first and second raw moments of $V_3^{(j)}$, which after running

DLF in the first three points have the values:

$$\begin{aligned} V_3^{(1)} &= w_{1,1}V_{3_{1,1}} + w_{1,2}V_{3_{1,2}} + w_{2,1}V_{3_{2,1}} \\ &= 0.1667 \cdot 1.0334 + 0.1667 \cdot 1.0301 + 0.1147 \cdot 1.0303 = 0.4622 \\ V_3^{(2)} &= w_{1,1}V_{3_{1,1}}^2 + w_{1,2}V_{3_{1,2}}^2 + w_{2,1}V_{3_{2,1}}^2 \\ &= 0.1667 \cdot 1.0334^2 + 0.1667 \cdot 1.0301^2 + 0.1147 \cdot 1.0303^2 = 0.4767 \end{aligned} \tag{4.11}$$

Continuing to update V_3 and the other output variables of interest in an equivalent manner for the remaining 7 DLF runs, yields the final raw moments of the various output random variables. For V_3, the final first and second order raw moments are

$$\begin{aligned} V_3^{(1)} &= \sum_{i=1}^{5}\sum_{k=1}^{2} w_{i,k}V_{3_{i,k}} = 1.0317 \\ V_3^{(2)} &= \sum_{i=1}^{5}\sum_{k=1}^{2} w_{i,k}V_{3_{i,k}}^2 = 1.0645 \end{aligned} \tag{4.12}$$

Step 8: Extract statistical data of output random variables

The previous step concludes the PEM in the sense that the output from the method, which is the raw moments of the output random variables, is obtained. These raw moments can then be utilized to present the results in the desired format, in this thesis by the expected values and standard deviation of each output variable. Then PDFs and CDFs have been approximated by Gram-Charlier expansion series, described in the next step.

Calculation of expected value and standard deviation is shown for output variable V_3 below, and summarized for all output variables in table 4.9. Please refer to eq. (4.12) for the values of first and second order raw moments of V_3 and to section 2.4 for the expressions for mean and standard deviation:

$$\mu_{V_3} = V_3^{(1)} = 1.0317 \text{ p.u.} \tag{4.13}$$

$$\sigma_{V_3} = \sqrt{V_3^{(2)} - (V_3^{(1)})^2} = 0.00201 \text{ p.u.} \tag{4.14}$$

Table 4.9: PLF results by PEM on 3-bus test system

		V_i		δ_i	
	Bus i	μ	σ	μ	σ
	1	1.05	0.0	0.0	0.0
	2	1.03	0.0	-2.7441	0.5073
	3	1.03173	0.00201	-2.0928	0.300
From	To	P_{ij}		Q_{ij}	
bus i	bus j	μ	σ	μ	σ
1	2	22.2010	3.6444	1.8840	1.0230
1	3	69.3084	8.0871	10.0949	5.1189
2	3	-8.1548	4.8421	17.6988	5.8059

4.1.3 Cumulant Method: step-by-step solution

Step 1: Run a NR load flow to obtain $\mathbf{X_0}$, $\mathbf{Z_0}$, $\mathbf{J_0}$ and $\mathbf{S_0}$. Obtain $\mathbf{D_0}$ and expand $\mathbf{S_0}$ and $\mathbf{D_0}$ to full sensitivity matrices.

A regular NR load flow analysis is run for the base case, i.e. the case of table 4.1. From this load flow, obtain the mean values of the state variables $\mathbf{X_0}$ [2] and line power flows $\mathbf{Z_0}$, and the Jacobian and sensitivity matrices $\mathbf{J_0}$ and $\mathbf{S_0}$, respectively. The variables included in the vectors $\mathbf{X}$ and $\mathbf{Z}$ in the NR load flow, as well as values of $\mathbf{X_0}$ and $\mathbf{Z_0}$ are given below:

$$\mathbf{X} = \begin{bmatrix} \delta_2 \\ \delta_3 \\ V_3 \end{bmatrix}, \quad \mathbf{Z} = \begin{bmatrix} P_{12} \\ P_{13} \\ P_{23} \\ Q_{12} \\ Q_{13} \\ Q_{23} \end{bmatrix}, \qquad \mathbf{X_0} = \begin{bmatrix} -2.7430 \\ -2.0923 \\ 1.03174 \end{bmatrix}, \quad \mathbf{Z_0} = \begin{bmatrix} 22.1885 \\ 69.2745 \\ -8.1713 \\ 1.8701 \\ 10.0677 \\ 17.6634 \end{bmatrix} \tag{4.15}$$

where the values of V and δ are given in p.u. and degrees, respectively. Equivalently, the values of P and Q are in MW and MVAr.

$\mathbf{S_0}$ is obtained by inverting $\mathbf{J_0}$, as shown below.

$$\mathbf{J_0} = \begin{bmatrix} 8.6762 & -4.6899 & -15.8019 \\ -5.0590 & 21.1003 & 20.3530 \\ 16.1927 & -22.1990 & 19.9665 \end{bmatrix}, \quad \mathbf{S_0} = \mathbf{J_0}^{-1} = \begin{bmatrix} 0.0951 & 0.0484 & 0.0259 \\ 0.0469 & 0.0467 & -0.0105 \\ -0.0250 & 0.0127 & 0.0174 \end{bmatrix} \tag{4.16}$$

The matrix $\mathbf{D_0}$ is found by evaluating $\partial\mathbf{Z}/\partial\mathbf{X}$ at $\mathbf{X} = \mathbf{X_0}$. In this thesis, ready-made

[2] Please note the meaning of *state variables* ($\mathbf{X_0}$) in this context as being consistent with the definition stated in section 2.3 (i.e. V and δ for PQ-buses and δ for PV-buses).

software [30] is used to do this calculation.[3] Note that only partial derivatives with respect to the state variables are included in the calculation of this matrix:

$$\mathbf{D_0} = \begin{bmatrix} \frac{\partial P_{12}}{\partial \delta_2} & \frac{\partial P_{12}}{\partial \delta_3} & \frac{\partial P_{12}}{\partial V_3} \\ \frac{\partial P_{13}}{\partial \delta_2} & \frac{\partial P_{13}}{\partial \delta_3} & \frac{\partial P_{13}}{\partial V_3} \\ \frac{\partial P_{23}}{\partial \delta_2} & \frac{\partial P_{23}}{\partial \delta_3} & \frac{\partial P_{23}}{\partial V_3} \\ \frac{\partial Q_{12}}{\partial \delta_2} & \frac{\partial Q_{12}}{\partial \delta_3} & \frac{\partial Q_{12}}{\partial V_3} \\ \frac{\partial Q_{13}}{\partial \delta_2} & \frac{\partial Q_{13}}{\partial \delta_3} & \frac{\partial Q_{13}}{\partial V_3} \\ \frac{\partial Q_{23}}{\partial \delta_2} & \frac{\partial Q_{23}}{\partial \delta_3} & \frac{\partial Q_{23}}{\partial V_3} \end{bmatrix} = \begin{bmatrix} -4.1157 & 0 & 0 \\ 0 & -16.4360 & -4.6715 \\ 4.6899 & -4.6899 & -15.8019 \\ 1.1562 & 0 & 0 \\ 0 & 4.8198 & -15.9312 \\ -16.3034 & 16.3034 & -4.5456 \end{bmatrix} \tag{4.17}$$

Now, having established the matrices in the reduced form that is used in the NR load flow, the sensitivity matrices are expanded to include sensitivities with respect to *all* variables in the system. This is done mainly in order to simplify the calculation of the modified sensitivity matrices $\mathbf{S_1}$ and $\mathbf{T_1}$. The non-state variables are constant throughout the whole LF calculation, so including those in the output variables is equivalent to inserting zeros into the rows and columns representing the non-state variables in a full sensitivity matrix. Investigating the construction of $\mathbf{D_0}$ in eq. (4.17), it is easily seen that one column of zeros needs to be inserted before the first column (corresponding to δ_1), and two columns of zeros between the second and third column (corresponding to V_1 and V_2) of $\mathbf{D_0}$ as defined in eq. (4.17). Doing an equivalent analysis of the construction of $\mathbf{J_0}$ yields the following final versions of $\mathbf{S_0}$ and $\mathbf{G_0}$:

$$\mathbf{S_0} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.0951 & 0.0484 & 0 & 0 & 0.0259 \\ 0 & 0.0469 & 0.0467 & 0 & 0 & -0.0105 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -0.025 & 0.0127 & 0 & 0 & 0.0174 \end{bmatrix}, \quad \mathbf{D_0} = \begin{bmatrix} 0 & -4.1157 & 0 & 0 & 0 & 0 \\ 0 & 0 & -16.4368 & 0 & 0 & -4.6715 \\ 0 & 4.6899 & -4.6899 & 0 & 0 & -15.8019 \\ 0 & 1.1562 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4.8198 & 0 & 0 & -15.9312 \\ 0 & -16.3034 & 16.3034 & 0 & 0 & -4.5456 \end{bmatrix} \tag{4.18}$$

The matrix $\mathbf{T_0}$, equal to the product of $\mathbf{S_0}$ and $\mathbf{D_0}$ is then

$$\mathbf{T_0} = \mathbf{S_0}\mathbf{D_0} = \begin{bmatrix} 0 & -0.3914 & -0.1992 & 0 & 0 & -0.1067 \\ 0 & -0.6542 & -0.8277 & 0 & 0 & 0.092 \\ 0 & 0.6209 & -0.193 & 0 & 0 & -0.1033 \\ 0 & 0.11 & 0.056 & 0 & 0 & 0.03 \\ 0 & 0.6241 & 0.0228 & 0 & 0 & -0.3273 \\ 0 & -0.6723 & -0.085 & 0 & 0 & -0.6731 \\ 0 & -0.025 & 0.0127 & 0 & 0 & 0.0174 \end{bmatrix} \tag{4.19}$$

[3] The MATPOWER function dSbr_dV computes partial derivatives of power flows w.r.t. voltages (angles and magnitudes). Please see Matlab code enclosed to this thesis for the full implementation.

Step 2: Initialize $\mathbf{u}$ and correlation matrix $\mathbf{C_u}$. Obtain $\mathbf{C_z}$ and decompose it by Cholesky decomposition

The correlated variables $\mathbf{u}$ are the loads at buses 2 and 3, each load consisting of an active and reactive part. Both the active and reactive parts of each load are correlated by the same correlation coefficient matrix to the other load. Let the active and reactive parts of the correlated variables be denoted $\mathbf{u_P}$ and $\mathbf{u_Q}$, respectively:

$$\mathbf{u_P} = \begin{bmatrix} u_{P_1} \\ u_{P_2} \end{bmatrix} = \begin{bmatrix} P_{L2} \\ P_{L3} \end{bmatrix}, \quad \mathbf{u_Q} = \begin{bmatrix} u_{Q_1} \\ u_{Q_2} \end{bmatrix} = \begin{bmatrix} Q_{L2} \\ Q_{L3} \end{bmatrix} \tag{4.20}$$

Equivalently, the correlated standard normal variables $\mathbf{z}$ will have an active part $\mathbf{z_P}$ and a reactive part $\mathbf{z_Q}$.

With a correlation coefficient of 0.4 between the two buses, the correlation coefficient matrix $\mathbf{C_u}$ of the original correlated variables is: (refer to appendix A.2)

$$\mathbf{C_u} = \begin{bmatrix} 1 & 0.4 \\ 0.4 & 1 \end{bmatrix} \tag{4.21}$$

For the correlation matrix $\mathbf{C_z}$ of the correlated standard normal variables $\mathbf{z}$, recall from section 3.2.2 that the diagonal elements of $\mathbf{C_z}$ are equal to 1, while the off-diagonal elements are related to the elements of the original correlation matrix $\mathbf{C_u}$ by a function H. As also stated in section 3.2.2, H has a value of 1 when the variables are normal distributed, so eq. (3.41) yields:

$$\rho_{z_{ij}} = \rho_{u_{ij}} \quad \Rightarrow \quad \mathbf{C_z} = \mathbf{C_u} = \begin{bmatrix} 1 & 0.4 \\ 0.4 & 1 \end{bmatrix} \tag{4.22}$$

$\mathbf{C_z}$ is now decomposed by Cholesky decomposition according to eq. (3.36), in order to obtain the inferior triangular matrix $\mathbf{L}$. The elements of $\mathbf{L}$ can be calculated from the expressions provided in appendix A.3 or by using some ready-made function in available software: [4]

$$\mathbf{L_z} = \begin{bmatrix} 1 & 0 \\ 0.4 & 0.9165 \end{bmatrix} \tag{4.23}$$

Step 3: Generate random numbers $\mathbf{E_s}$, and get random numbers $\mathbf{Z_s}$ by orthogonal transformation.

[4] In this thesis, the built-in Matlab function `chol` was used to decompose $\mathbf{C_z}$

Recall from section 3.2.2 that $\mathbf{E_s}$ are random numbers of independent, standard normal variables. These numbers will in their entirety be transformed between different spaces to finally obtain random numbers $\mathbf{V_s}$ of independent variables to be used directly as input to the CM equations. What first needs to be established, is the sample size (i.e. the number of random numbers) N. In this thesis a value of $N = 10\ 000$ is chosen for all the case studies on CM involving correlation. A sample size of about 10 000 is a widely-used choice for the MC simulation used for comparison in many papers investigating analytical and approximate methods. In [7], which is the main source for the derivation of the modified CM, 10 000 samples are used for both the comparative MC simulations and for the modified CM. In [18] and [14], two other important sources for the derivation of methodologies, 15 000 and 10 000 samples, respectively, are used for the MC simulations.

The generation of the samples making out $\mathbf{E_s}$ is in this thesis done using a built-in function in Matlab[5]. Due to space limitations, the matrices containing the random numbers will not be written out explicitly, only the matrix operations will be shown.

$\mathbf{E_s}$ is in this case a 2×10000 matrix because there are two loads that are correlated. Each row represents the distribution of a standard normal variable, and the two variables represented in $\mathbf{E_s}$ are independent. After $\mathbf{E_s}$ is obtained, the next step is to take into consideration the correlation of the two loads. This is done by orthogonal transformation of the samples in $\mathbf{E_s}$, resulting in the matrix of random numbers of the *correlated* standard normal variables $\mathbf{Z_s}$. Recall that the transformation matrix of the orthogonal transformation is the inferior triangular matrix obtained by Cholesky decomposition of the transformed variable's correlation matrix. The latter is now the correlation matrix $\mathbf{C_z}$, which was decomposed to obtain $\mathbf{L_z}$ in the previous step. Then, according to eq. (3.42), the random numbers of correlated standard normal variables are obtained by

$$\mathbf{Z_s} = \mathbf{L_z}\mathbf{E_s} \tag{4.24}$$

where $\mathbf{Z_s}$ is a 2×10000 matrix containing the random numbers of each correlated standard normal variable z_i along the rows.

Step 4: Obtain random numbers $\mathbf{U_s}$ by marginal transformation and get random numbers $\mathbf{V_s}$ by orthogonal transformation

The random numbers $\mathbf{U_s}$ of the original correlated variables $\mathbf{u}$ are now obtained through a marginal transformation utilizing the CDFs of the variables in $\mathbf{u}$ and $\mathbf{v}$. The CDFs of the variables in $\mathbf{v}$ are the CDFs $\mathbf{\Phi}$ of the standard normal distributions, which can be found

[5]The built-in Matlab function `normrnd` generates random numbers for the normal distribution, taking expected value μ and standard deviation σ as input variables. These are specified to the values of 0 and 1 respectively, to obtain the standard normal distribution.

in appendix A.1.1. The variables in $\mathbf{u}$ are in this case normal distributed, and CDFs of the normal distribution, also provided in appendix A.1.1, are gathered in $\mathbf{F}$. Note that in $\mathbf{F}$, unlike in $\mathbf{\Phi}$, the elements will be different functions, because the CDF of the normal distribution depends on the actual values of μ and σ for that particular variable. Due to different values of μ and σ, the active and reactive parts of each load variable are treated in separate matrices. Denote the vector of CDFs of the active part by $\mathbf{F_P}$ and the CDFs of the reactive parts $\mathbf{F_Q}$. Then, according to eq. (3.43), the random numbers of the original correlated variables are obtained by:

$$\mathbf{U_{sP}} = \mathbf{F_P}^{-1}(\mathbf{\Phi}(\mathbf{Z_s})) \tag{4.25}$$

$$\mathbf{U_{sQ}} = \mathbf{F_Q}^{-1}(\mathbf{\Phi}(\mathbf{Z_s})) \tag{4.26}$$

where $\mathbf{U_{sP}}$ and $\mathbf{U_{sQ}}$ are the random numbers of $\mathbf{u_P}$ and $\mathbf{u_Q}$, respectively. Again, explicit expressions are not provided here due to space limitations.

The random numbers of $\mathbf{U_{sP}}$ and $\mathbf{U_{sQ}}$ still represent correlated variables, so the numbers need to be transformed from the correlated space to the independent space before being used as input to the cumulant method. The transformation between correlated and independent space is once again performed by orthogonal transformation. This time, the correlation matrix of whose decomposition yields the transformation matrix, is $\mathbf{C_u}$. Decomposing $\mathbf{C_u}$ yields:

$$\mathbf{L} = \begin{bmatrix} 1 & 0 \\ 0.4 & 0.9165 \end{bmatrix} \tag{4.27}$$

Then, according to eq. (3.44), the random numbers $\mathbf{V_{sP}}$ and $\mathbf{V_{sQ}}$ of the independent variables are obtained by

$$\mathbf{V_{sP}} = \mathbf{L}^{-1}\mathbf{U_{sP}} \tag{4.28}$$

$$\mathbf{V_{sQ}} = \mathbf{L}^{-1}\mathbf{U_{sQ}} \tag{4.29}$$

$$\tag{4.30}$$

where $\mathbf{V_{sP}}$ and $\mathbf{V_{sQ}}$ are now 2×10000 matrices.

Step 5: Compute the r-th order cumulants of the independent variables $\mathbf{v}$ represented by $\mathbf{V_s}$

The random numbers in each row of $\mathbf{V_{sP}}$ and $\mathbf{V_{sQ}}$ can now be considered to be discrete distributions representing the corresponding independent variables v_{Pi} and v_{Qi}. The easiest way to obtain the cumulants of these variables is then to first find the raw moments by eq. (2.13), and then find the cumulants from those raw moments according to eq. (2.21). In

this case study, the final presentation of the results will be in the form of mean and standard deviation of each variable. Notice that the mean values of the output variables have already been obtained in step 1, so actually, in accordance with eq. (2.22), only the second order cumulant needs to be calculated for each random variable.

For the active power at bus 2, denoted v_{P1}, the first- and second-order raw moments are calculated from eq. (2.13), yielding

$$\begin{aligned} E[v_{P_1}] &= -0.5001 \\ E[v_{P_1}^2] &= 0.2525 \end{aligned} \tag{4.31}$$

where the values are given in p.u. Then, according to eq. (2.21), the second-order cumulant of v_{P_1} is

$$\Delta v_{P_1}^{(2)} = E[v_{P_1}^2] - E[v_{P_1}] = 0.0025 \tag{4.32}$$

The second-order cumulants of the active parts of both loads is then given by the $\mathbf{v_P}^{(2)}$, similarly for the reactive parts:

$$\begin{aligned} \mathbf{\Delta v_P}^{(2)} &= \begin{bmatrix} -0.0025 \\ -0.0035 \end{bmatrix} \\ \mathbf{\Delta v_Q}^{(2)} &= \begin{bmatrix} -0.0004 \\ -0.0006 \end{bmatrix} \end{aligned} \tag{4.33}$$

where all values are given in p.u.

Step 6: Compute the cumulants $\Delta w_i^{(r)}$ of the net power injections at each bus and represent $\Delta w_i^{(r)}$ for all buses in $\mathbf{\Delta W}^{(r)}$

The cumulants of the net power injections are calculated by eq. (3.45). At this point, recall that there is yet another random input variable in the system that is not correlated to any other variables, that is the generation at bus 2, u_{G_2}. So to obtain the cumulants of the net power injection at bus 2, the cumulants of u_{G_2} need be obtained. Referring to appendix A.1.2, the first- and second order raw moments of u_{G_2} are calculated from eq. (A.10):

$$\begin{aligned} E[u_{G_1}] &= 0.22 \cdot (1 - \text{FOR}) = 0.22 \cdot (1 - 0.09) = 0.2002 \\ E[u_{G_1}^2] &= 0.22^2 \cdot (1 - \text{FOR}) = 0.22^2 \cdot (1 - 0.09) = 0.0440 \end{aligned} \tag{4.34}$$

Then, using eq. (2.21), the second-order cumulant of u_{G_1} is

$$\Delta u_{G_1}^{(2)} = E[u_{G_1}^2] - E[u_{G_1}] = 0.0040 \tag{4.35}$$

Then finally the cumulants $\Delta w_i^{(r)}$ of the net power injections can be calculated by adding the cumulants of generation and load. For the active power at bus 2:

$$\Delta w_2^{(2)} = \Delta u_{G_1}^{(2)} + \Delta v_{P_L 2} = 0.0040 + 0025 = 0.0065 \tag{4.36}$$

For this particular system, at all buses except bus 2, $\Delta w_i^{(r)} = \Delta v_i^{(r)}$. Gathering the second-order cumulants for the active and reactive net power injections at all buses in $\mathbf{\Delta W}^{(2)} = [\mathbf{\Delta W_P}^{(2)} \mathbf{\Delta W_Q}^{(2)}]^T$, yields:

$$\mathbf{\Delta W}^{(2)} = \begin{bmatrix} 0 \\ 0.0064 \\ 0.0035 \\ 0 \\ 0.0004 \\ 0.0006 \end{bmatrix} \tag{4.37}$$

where rows 1 and 3 correspond to the slack bus.

Step 7: Obtain the modified sensitivity matrices $\mathbf{S_1}$ and $\mathbf{T_1}$

The elements of the modified sensitivity matrices are obtained from eq. (3.46), resulting in:

$$\mathbf{S_1} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.1145 & 0.0444 & 0 & 0.0104 & 0.0238 \\ 0 & 0.0656 & 0.0428 & 0 & -0.0042 & -0.0096 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -0.0199 & 0.0116 & 0 & 0.0069 & 0.0159 \end{bmatrix},$$

$$\mathbf{T_1} = \begin{bmatrix} 0 & -0.4711 & -0.1826 & 0 & -0.0427 & -0.0978 \\ 0 & -0.9853 & -0.7586 & 0 & 0.0368 & 0.0843 \\ 0 & 0.5437 & -0.1769 & 0 & -0.0413 & -0.0947 \\ 0 & 0.1324 & 0.0513 & 0 & 0.012 & 0.0275 \\ 0 & 0.6332 & 0.0209 & 0 & -0.1309 & -0.2999 \\ 0 & -0.7063 & -0.0779 & 0 & -0.2693 & -0.6169 \end{bmatrix} \tag{4.38}$$

Step 8: Calculate the cumulants of the output variables

The cumulants of the output variables are now calculated according to eq. (3.47). Recall that the first-order cumulants, equal to the mean values, were obtained in step 1. The second-order cumulants are then calculated by

$$\begin{aligned} \mathbf{\Delta X}^{(2)} &= \mathbf{S_1}^2 \mathbf{\Delta W}^{(2)} \\ \mathbf{\Delta Z}^{(2)} &= \mathbf{T_1}^2 \mathbf{\Delta W}^{(2)} \end{aligned} \tag{4.39}$$

Resulting in, for the voltages and line power flows:

$$\mathbf{\Delta X}^{(2)} = \begin{bmatrix} \Delta\delta_1^{(2)} \\ \Delta\delta_2^{(2)} \\ \Delta\delta_3^{(2)} \\ \Delta V_1^{(2)} \\ \Delta V_2^{(2)} \\ \Delta V_3^{(2)} \end{bmatrix} = \begin{bmatrix} 0 \\ 0.0052 \\ 0.0020 \\ 0 \\ 0 \\ 0.0002 \end{bmatrix}, \quad \mathbf{\Delta Z}^{(2)} = \begin{bmatrix} \Delta P_{12}^{(2)} \\ \Delta P_{13}^{(2)} \\ \Delta P_{23}^{(2)} \\ \Delta Q_{12}^{(2)} \\ \Delta Q_{13}^{(2)} \\ \Delta Q_{23}^{(2)} \end{bmatrix} = \begin{bmatrix} 0.0016 \\ 0.0083 \\ 0.0020 \\ 0.0001 \\ 0.0026 \\ 0.0035 \end{bmatrix} \tag{4.40}$$

where the units for V, P and Q are p.u. and the unit for δ is radians. The standard deviations for all variables are now easily found as the square root of the second-order cumulant. Converting to conventional units of MW, MVAr and degrees for P, Q and δ after this operation, yield the final results presented in table 4.10.

Table 4.10: PLF results by CM on 3-bus test system

		V_i		δ_i	
	Bus i	μ	σ	μ	σ
	1	1.05	0.0	0.0	0.0
	2	1.03	0.0	-2.743	0.5481
	3	1.03174	0.00179	-2.0923	0.3351
From	To	P_{ij}		Q_{ij}	
bus i	bus j	μ	σ	μ	σ
1	2	22.1885	3.9373	1.8701	1.1061
1	3	69.2745	9.0997	10.0677	5.1368
2	3	-8.1713	4.4892	17.6634	5.9064

4.2 IEEE 14 bus validation

The IEEE 14 bus test system has been used in order to validate the implementation of the two methods for the base case, which is a system with no correlation between variables and only conventional generation. Validation is based on the results obtained in [28], where PLF by the Cumulant method, as it was outlined in section 3.2 of this thesis, was performed on an identical system. The complete raw and probabilistic data of the system is therefore identical to [28] and provided in appendix C.1. A summary of the test system modelling is given below:

- *Loads:* The system includes 11 independent loads. 10 of these are normal distributed with mean and standard deviation specified in table C.2 for the active and reactive parts of each load separately. The load at bus 9 is given by a discrete distribution for its active and reactive parts in table C.3.
- *Generation:* The system includes 12 generators distributed over buses 1 and 2. Each generator is modelled by a Bernoulli distribution based on its FOR (refer to appendix A.1.2). Probabilistic data for generators may be found in table C.4. Reactive power limits on generators are not enforced.
- *Correlation:* No correlation exists between loads, between generators, or between generators and loads anywhere in the system.

PEM and CM are run for the same test system, and some results for voltages and line flows are presented in tables 4.11 and 4.12, respectively. All values of active and reactive powers are given in MW and MVAr, respectively. ϵ represents the percentage-wise error relative to the results from [28]. The complete results from [28] can be found in appendix D.2.

Table 4.11: Voltage results for IEEE 14 bus system

Var.	PEM		ϵ_{PEM}		CM		ϵ_{CM}	
	μ	σ	ϵ_μ	ϵ_σ	μ	σ	ϵ_μ	ϵ_σ
V_1	1.06	0	0.00	0.00	1.06	0	0.00	0.00
V_2	1.045	0	0.00	0.00	1.045	0	0.00	0.00
V_3	1.01	0	0.00	0.00	1.01	0	0.00	0.00
V_4	1.01708	0.002	0.01	0.99	1.01714	0.00202	0.00	0.00
V_5	1.01868	0.00162	0.00	1.22	1.01873	0.00164	0.00	0.00
V_6	1.07	0	0.00	0.00	1.07	0	0.00	0.00
V_7	1.06121	0.00283	0.01	1.05	1.06128	0.00286	0.00	0.00

Table 4.11 – continued

Var.	PEM		ϵ_{PEM}		CM		ϵ_{CM}	
	μ	σ	ϵ_μ	ϵ_σ	μ	σ	ϵ_μ	ϵ_σ
V_8	1.09	0	0.00	0.00	1.09	0	0.00	0.00
V_9	1.05561	0.00515	0.01	0.98	1.0557	0.00519	0.00	1.76
V_{10}	1.05072	0.00438	0.01	0.68	1.05079	0.00441	0.00	0.00
V_{11}	1.05676	0.00229	0.00	0.87	1.05681	0.00231	0.00	0.00
V_{12}	1.05518	0.00069	0.00	0.00	1.05519	0.00069	0.00	0.00
V_{13}	1.05034	0.00119	0.00	0.83	1.05035	0.00120	0.00	0.00
V_{14}	1.03532	0.00366	0.01	0.54	1.03538	0.00368	0.00	0.00
δ_1	0	0	0.00	0.00	0	0	0.00	0.00
δ_2	-4.9855	0.443	0.02	0.00	-4.9855	0.4432	0.02	0.05
δ_3	-12.7343	0.9976	0.03	0.00	-12.7317	0.9978	0.01	0.02
δ_4	-10.3113	0.6898	0.03	0.00	-10.3097	0.6903	0.01	0.07
δ_5	-8.7672	0.5791	0.03	0.05	-8.7658	0.5796	0.01	0.13
δ_6	-14.2241	0.8483	0.04	0.14	-14.2201	0.85	0.01	0.06
δ_7	-13.3603	0.974	0.03	0.13	-13.3572	0.9756	0.01	0.03
δ_8	-13.3603	0.974	0.03	0.13	-13.3572	0.9756	0.01	0.03
δ_9	-14.94	1.1476	0.03	0.17	-14.936	1.1499	0.01	0.03
δ_{10}	-15.099	1.0957	0.03	0.16	-15.095	1.0978	0.01	0.03
δ_{11}	-14.7928	0.9697	0.03	0.15	-14.7889	0.9715	0.01	0.04
δ_{12}	-15.0794	0.8818	0.04	0.14	-15.0753	0.8835	0.01	0.05
δ_{13}	-15.1592	0.9071	0.03	0.15	-15.1552	0.9089	0.01	0.05
δ_{14}	-16.0362	1.0596	0.03	0.15	-16.032	1.0616	0.01	0.03

Table 4.12: Line power flow results for IEEE 14 bus system

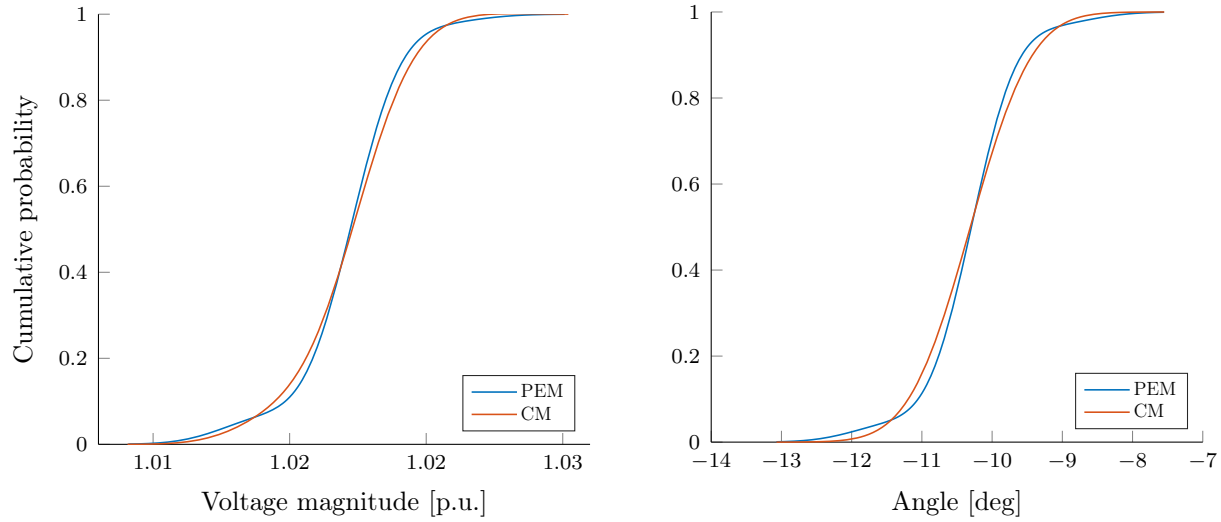
Var.	PEM		ϵ_{PEM}		CM		ϵ_{CM}	
	μ	σ	ϵ_μ	ϵ_σ	μ	σ	ϵ_μ	ϵ_σ
P_{1-2}	156.9366	13.3959	0.00	0.01	156.9714	13.4	0.02	0.04
P_{1-5}	75.4682	4.7904	0.00	0.07	75.4748	4.7931	0.01	0.13
P_{2-3}	73.2721	5.7569	0.00	0.00	73.2709	5.7575	0.00	0.01
P_{2-4}	56.1419	3.3316	0.00	0.01	56.1393	3.3343	0.00	0.08
P_{2-5}	41.522	2.4122	0.00	0.12	41.5187	2.4155	0.01	0.25
P_{3-4}	-23.2535	4.4614	0.00	0.01	-23.2545	4.4623	0.00	0.01
P_{4-5}	-61.0946	4.4956	0.00	0.13	-61.0969	4.4931	0.00	0.07
P_{4-7}	28.0606	3.572	0.00	0.01	28.0599	3.5716	0.00	0.00
P_{4-9}	16.0705	2.0382	0.00	0.07	16.07	2.0367	0.00	0.00

Table 4.12 – continued

Var.	PEM		ϵ_{PEM}		CM		ϵ_{CM}	
	μ	σ	ϵ_μ	ϵ_σ	μ	σ	ϵ_μ	ϵ_σ
P_{5-6}	44.111	2.6566	0.00	0.26	44.112	2.6636	0.00	0.00
P_{6-11}	7.3663	1.4715	0.00	0.29	7.3668	1.4758	0.01	0.00
P_{6-12}	7.789	0.4152	0.00	0.07	7.7909	0.4156	0.02	0.02
P_{6-13}	17.7556	1.2323	0.00	0.14	17.7543	1.2338	0.01	0.02
P_{7-8}	0	0	0.00	0.00	0	0	0.00	0.00
P_{7-9}	28.0607	3.572	0.00	0.01	28.0599	3.5716	0.00	0.00
P_{9-10}	5.215	1.5646	0.00	0.05	5.2145	1.5654	0.01	0.00
P_{9-14}	9.4161	1.2497	0.00	0.05	9.4153	1.2503	0.01	0.00
P_{10-11}	-3.7978	1.4453	0.00	0.10	-3.7982	1.4467	0.01	0.00
P_{12-13}	1.6172	0.3681	0.00	0.11	1.6192	0.3686	0.12	0.03
P_{13-14}	5.654	1.0977	0.00	0.22	5.6548	1.1001	0.01	0.00
Q_{1-2}	-18.8902	3.1279	0.23	0.05	-18.9418	3.1305	0.04	0.03
Q_{1-5}	5.5466	0.4912	0.81	2.81	5.502	0.5058	0.00	0.08
Q_{2-3}	4.7826	0.5663	0.63	0.21	4.7527	0.5652	0.00	0.02
Q_{2-4}	-0.3723	0.6449	9.04	1.78	-0.4075	0.6566	0.44	0.00
Q_{2-5}	2.6202	0.4909	1.11	1.90	2.5931	0.5005	0.07	0.02
Q_{3-4}	4.6012	2.0624	1.12	0.15	4.5517	2.0654	0.04	0.00
Q_{4-5}	16.0677	1.4392	0.07	0.01	16.0785	1.4392	0.00	0.01
Q_{4-7}	-9.8102	0.8538	0.19	0.54	-9.827	0.8586	0.02	0.02
Q_{4-9}	-0.4712	0.7833	3.66	0.79	-0.4871	0.7895	0.41	0.00
Q_{5-6}	12.0911	0.5577	0.10	0.16	12.1022	0.5603	0.00	0.30
Q_{6-11}	3.6289	1.0282	0.65	0.23	3.608	1.0306	0.07	0.00
Q_{6-12}	2.5182	0.1954	0.37	0.05	2.515	0.1954	0.24	0.05
Q_{6-13}	7.2496	0.6792	0.14	0.19	7.2366	0.6804	0.04	0.01
Q_{7-8}	-17.3371	1.6614	0.20	0.75	-17.3058	1.674	0.02	0.00
Q_{7-9}	5.7912	2.335	0.47	0.35	5.7699	2.3431	0.10	0.00
Q_{9-10}	4.1659	1.0617	0.23	0.05	4.1726	1.0612	0.06	0.00
Q_{9-14}	3.5768	0.6855	0.14	0.06	3.5792	0.6851	0.07	0.00
Q_{10-11}	-1.6704	1.009	0.71	0.02	-1.6612	1.0088	0.15	0.00
Q_{12-13}	0.7679	0.1823	1.12	0.05	0.7652	0.1822	0.76	0.00
Q_{13-14}	1.7904	0.6718	0.86	0.10	1.7778	0.6725	0.15	0.00

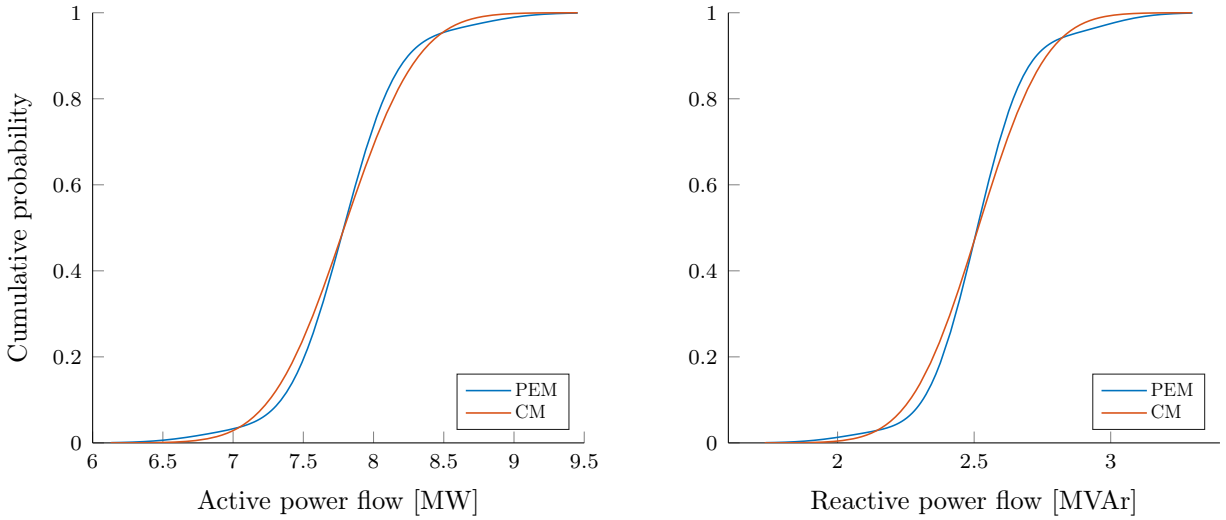
The results of the CDF approximations by Gram-Charlier expansion (GCE) for a select-few variables are included below. The GCE is not carried out in [28], but the figures are

included here for illustrative purposes.



(a) Voltage magnitude at bus 4 (V_4) **(b)** Voltage angle at bus 4 (δ_4)

Figure 4.2: CDF approximations by GCE for the voltage at bus 4



(a) Active power flow in line $6-12$ (P_{6-12}) **(b)** Reactive power flow in line $6-12$ (Q_{6-12})

Figure 4.3: CDF approximations by GCE for power flows in line $6-12$

4.2.1 Discussion

As a first note it should be emphasized that [28] runs PLF by the cumulant method, such that only the results from the CM in this case study are directly comparable by means of

both test system modelling and methodology. The results from the PEM case study are therefore compared to [28] on the basis of using a fundamentally different methodology to solve the same problem for an identical system model.

As can be observed from tables 4.11 and 4.12, the errors with respect to [28] are generally small for both methodologies. For the PEM, the relative errors are below 10%, and for the CM all errors are below 2%. This tendency is expected, as results from conducting the PEM are not available in [28], but how much deviation that should actually be expected from comparison of the two methods, is not necessarily easy to establish. It should also be mentioned here that a validation of the results obtained by the CM in [28] is not available, that is the results are not compared to those obtained by a simulation method providing the exact solution. Being an analytical method limited by the linearization of the power flow equations, the results obtained from the CM will always to some degree deviate from the definite solution. Thus, the results obtained from the PEM in this thesis are compared to results in [28] that are already to some extent deviating from the exact solution, so the percentage-wise errors in tables 4.11 and 4.12 do not necessarily provide a correct indication to the accuracy compared to the exact solution. As no discussion of the accuracy of the results is included in [28], it is not possible to determine in which direction the deviation occurs.

A second source of validation to this particular case study is reference [14], and although the extent of quantitative results available in this paper are very limited, it is useful to include in this discussion. In [14] case studies are performed for the PEM and CM, including also a MC method for validation of the results. One of the systems on which these studies are performed is the exact same system as the one in [28]. The results from these studies show that both the PEM and CM provide good results compared to the MCS. By investigation of the limited selection of numerical results provided in [14] it is apparent that the two methods often deviate from the exact solution in different directions. A maximum relative error of about 10% between the PEM and CM is present in the particular selection of results provided in [14], which can indicate that the results obtained by the PEM in this thesis are deviating from the CM by an acceptable amount. Nonetheless, as the complete numerical results for all variables in the system are not included in [14] this is merely an indication, and no definite conclusions can be drawn with regards to an acceptable level of relative error between the two methodologies from this source.

If a MC simulation method had been implemented for this particular case study, analyses of sensitivities with regards to the load variation level (i.e. the ratio of standard deviation to mean value for the loads) could have posed as an additional source of validation. In reference [14] the impacts from varying the load variation level are investigated in terms of accuracy of

the results obtained from the two non-simulation methodologies. A source of validation of the in-house implementation in this thesis, could be to investigate whether the same tendencies could be observed when the load variation level was adjusted equivalently. Reference [14] shows that the average error increases for both methodologies when the load variation level is increased, but that the increase is more obvious for the CM. This phenomenon is closely related to the fundamentals of the methodologies, as the CM is based on linearization of the load flow equations, and thus will be expected to struggle when the random variation about the mean increases. Having available in this case study a tool like the MC simulation to obtain the exact solution, would make a qualitative validation of these tendencies possible.

In conclusion for this case study, the deviation between the two methodologies are generally below a few percent, which should in most cases be considered acceptable. Comparison to source providing the exact results would serve as an even better validation, making possible an evaluation of which method actually performs best in terms of accuracy with respect to the exact solution.

4.3 IEEE RTS base case

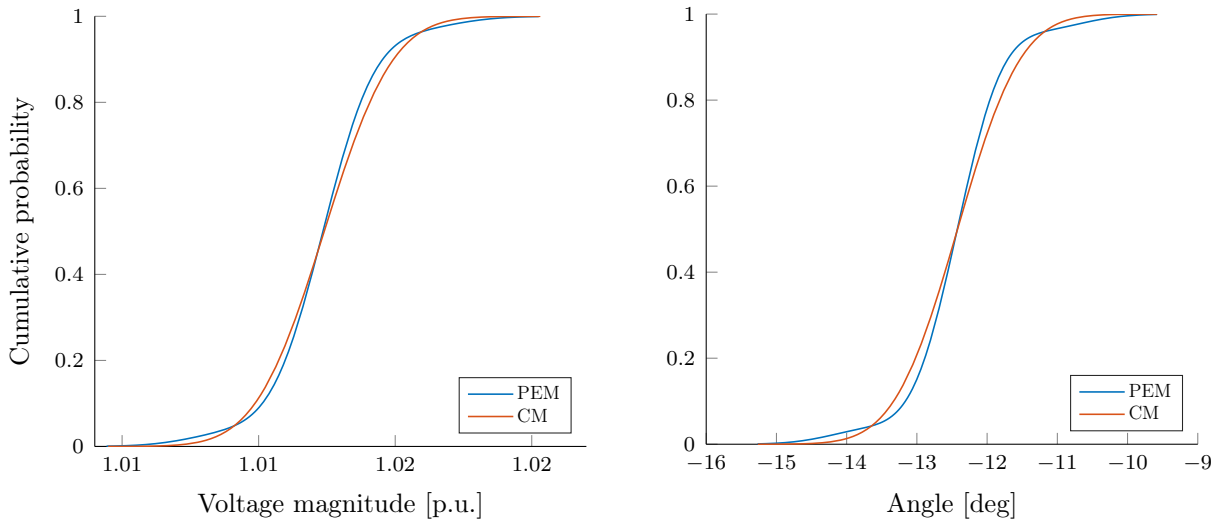
The IEEE Reliability test system (RTS) [11] is utilized in the remaining case studies, in order to demonstrate the methodologies on a slightly larger system and to have available a tool to conduct PLF on a widely-used test system for reliability studies. In this section, results for the *base case*, i.e. the case with no correlation between variables or unconventional generation is presented. The implementation of the methodologies is equal as in the previous case study of section 4.2, and is described in sections 3.1.1 and 3.2.1.

The line data and raw load data of the IEEE RTS are provided in appendix C.2. In this section, only load demands will be modelled as probabilistic inputs in order to keep the base case model simple. In the case studies of the two subsequent sections, this model will be kept as a constant basis when correlation between loads and correlated unconventional generation units are introduced in sections 4.4 and 4.5, respectively.

The probabilistic modelling can be summed up as follows:

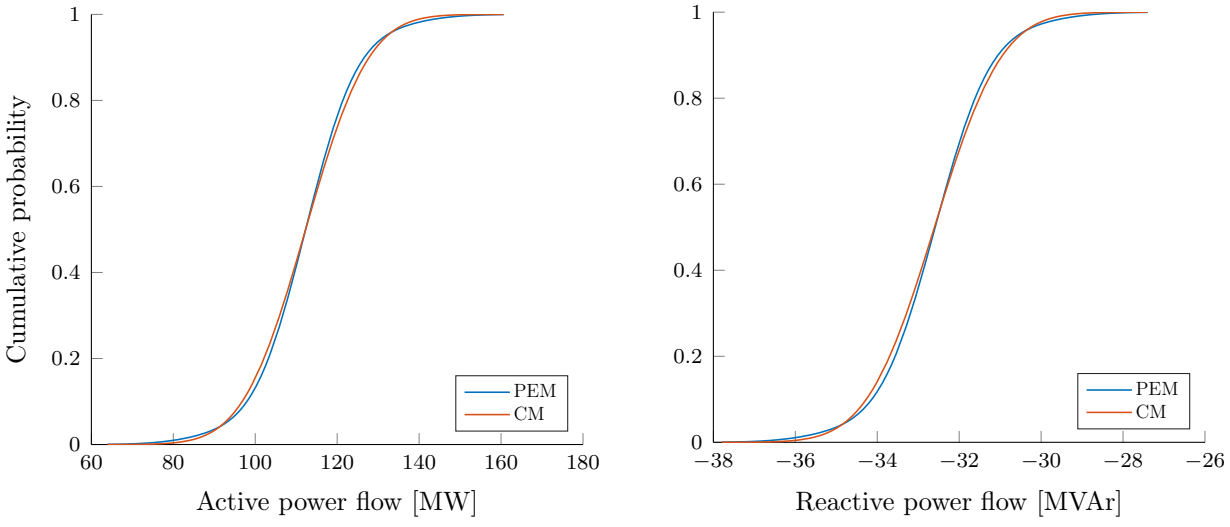
- *Loads:* All loads in the system are considered random variables. They are normal distributed with means equal to the values of the deterministic base case system, which can be found in table C.5 of appendix C.2. The standard deviation of each load is equal to 5% of the corresponding mean value. This load model is inspired by [18].
- *Generation:* No probabilistic modelling of generation in this base case model.
- *Correlation:* No correlation between any random variables in this base case model.

Complete results represented by mean values and standard deviations of voltages and line power flows are included in appendix D.3. The results of the CDF approximation by GCE is provided for selected variables below.



(a) Voltage magnitude at bus 6 (V_6) **(b)** Voltage angle at bus 6 (δ_6)

Figure 4.4: CDF approximations by GCE for the voltage at bus 6



(a) Active power flow in line $15-16$ (P_{15-16}) **(b)** Reactive power flow in line $15-16$ (Q_{15-16})

Figure 4.5: CDF approximations by GCE for power flows in line $15-16$

4.3.1 Discussion

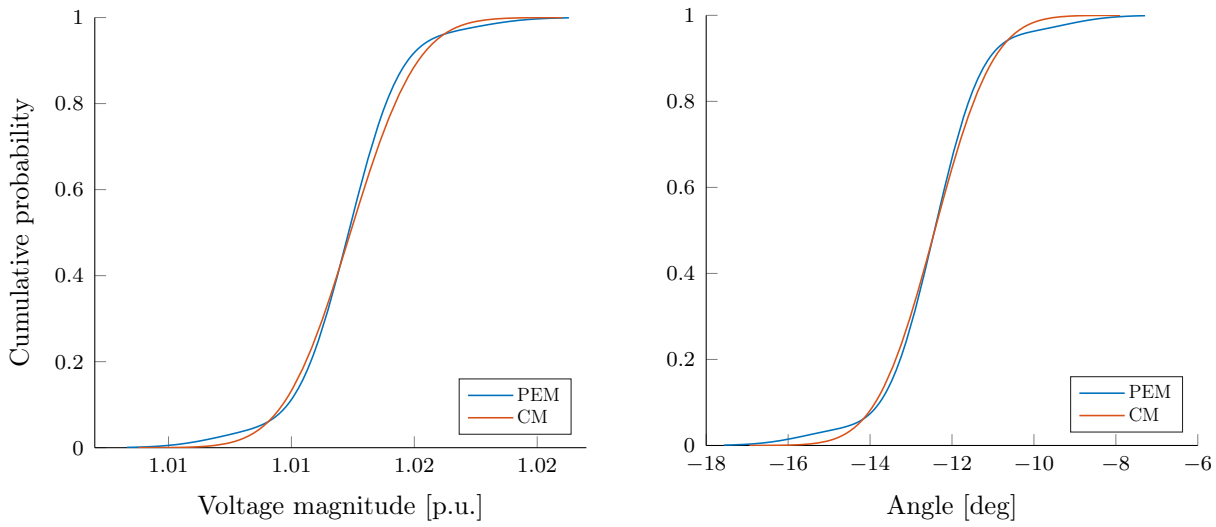
The IEEE RTS is a slightly larger system than the IEEE 14 bus system of the previous section. Nonetheless, the exact same procedure has been used in the implementation of the two methods, and in that sense, the results from this case study can pose as an additional validation of the implementation of the base case methodology.

Investigating the numerical results included in appendix D.3, the deviation between the two methods is negligibly small; never exceeding 0.1% in either mean values or standard deviations. Despite not having available the exact results for this PLF problem, the fact that two methodologies that are so fundamentally different produce results with differences that are almost negligible, is a source of validation in itself. In addition, the selected results for CDF approximations in figs. 4.4 and 4.5 gives some indication to the deviation between the higher-order characteristics (moments and cumulants) obtained by the two methods. The GCE implemented in this thesis makes use of the mean, standard deviation, skewness and kurtosis in order to make an approximation of the CDF. The relatively small deviations in the approximated functions indicate that also these quantities, obtained from the third- and fourth-order moments or cumulants, differ within what can be considered acceptable limits.

4.4 IEEE RTS Correlation

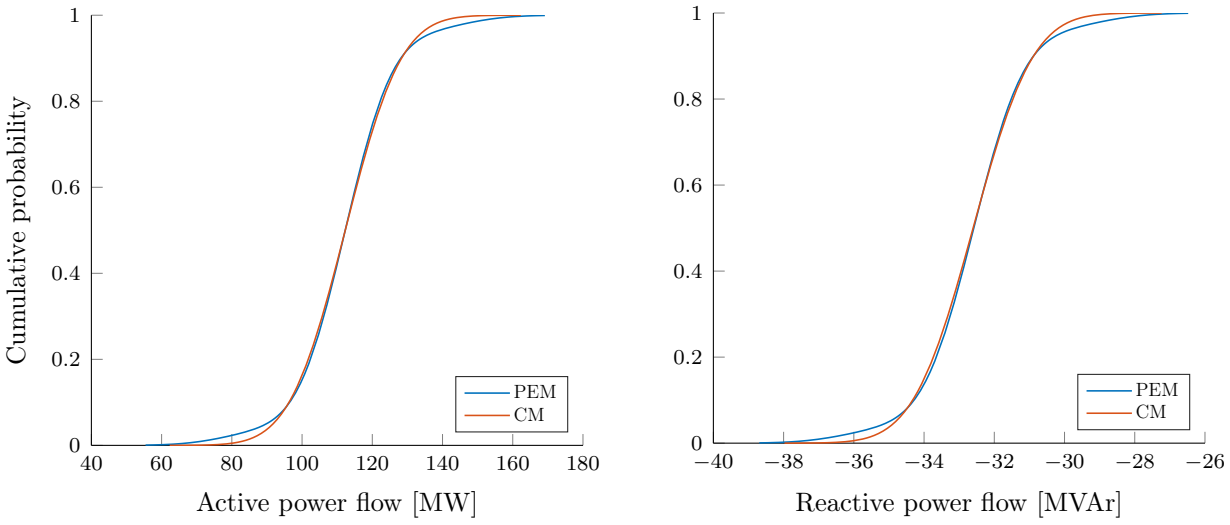
In this section, case studies on the IEEE RTS with correlation between variables are carried out. The implementation is done in accordance with the descriptions of the modified versions of the PEM and CM, outlined in sections 3.1.2 and 3.2.2, respectively. The step-by-step demonstrations in section 4.1 also describes the exact same procedure that is performed for the case of this study, excluding the probabilistic modelling of generation. The base case system from section 4.3 is retained in terms of the probabilistic modelling. Correlation is initially added between the loads and extended also to generation when wind turbines are added to the system in section 4.5.

Now, while retaining the probabilistic modelling of the system in section 4.3 correlation is added to the system by defining a value ρ_{ij} by which the loads at buses i and j are correlated (refer to appendix A.2). For the case studies in this section, all loads are correlated by the same correlation coefficient ρ in order to retain a simple system that can also be used to demonstrate some simple sensitivity analyses. Consider first all loads to be correlated by a coefficient of $\rho = 0.2$. The CDF approximations to the same variables as were presented in section 4.3 are given in figures (REF) below.



(a) Voltage magnitude at bus 6 (V_6) **(b)** Voltage angle at bus 6 (δ_6)

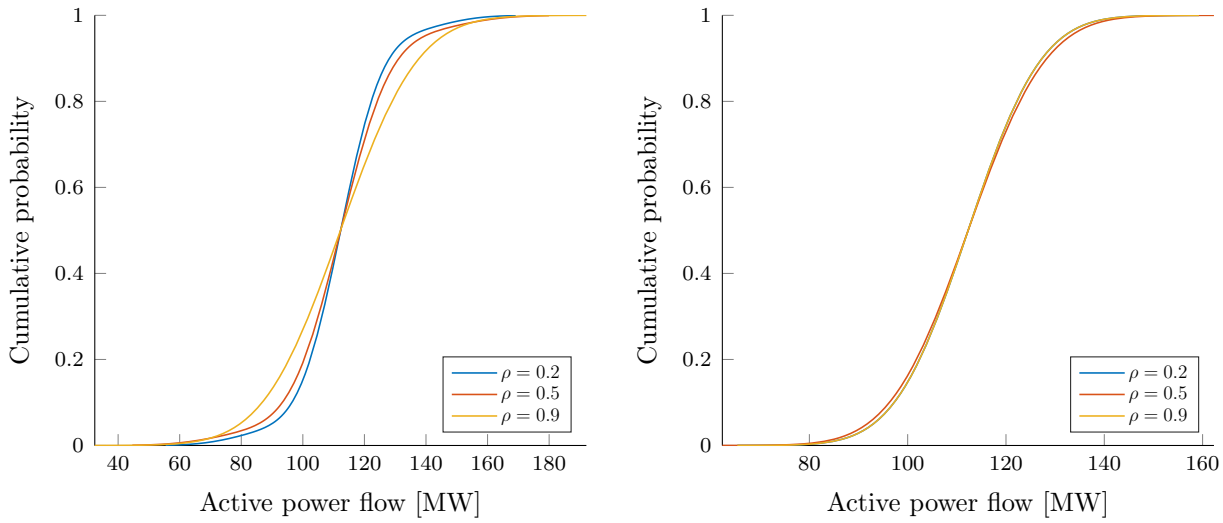
Figure 4.6: CDF approximations by GCE for the voltage at bus 6



(a) Active power flow in line $15-16$ (P_{15-16}) **(b)** Reactive power flow in line $15-16$ (Q_{15-16})

Figure 4.7: CDF approximations by GCE for power flows in line $15-16$

In order to study the impact of increasing correlation between the variables, a simple, qualitative sensitivity analysis with regards to the correlation coefficient ρ will be performed. The results using the original value of $\rho = 0.2$ will be compared to the results when $\rho = 0.5$ and $\rho = 0.9$. The results for the active power flow in line $15-16$ are presented for the PEM and CM in figs. 4.8a and 4.8b, respectively.



(a) PLF by the Point estimate method

(b) PLF by the Cumulant method

Figure 4.8: Sensitivities of ρ on the active power flow in line $15 - 16$

4.4.1 Discussion

The correlation scheme selected for these case studies is a simple one; all loads are correlated by the same correlation coefficient ρ. It is of course possible to relate different buses by different coefficients to create a more complex correlation scheme, for instance by dividing the RTS into two regions as is done in the case studies of [18]. Creating a more complex scheme is a fairly simple task, as the changes are only made to the correlation coefficient matrix, whose construction is intuitive and straight-forward. In order to keep the system simple, and to be able to perform a sensitivity analysis that can be interpreted in an qualitative and intuitive way, this simplest possible correlation scheme was chosen for these case studies.

The deviation between the results obtained from the PEM and CM is definitely increasing in this case compared to the previous one. This observation can be made directly from investigating the CDF approximations in figs. 4.6 and 4.7. For all variables, the CDFs have been plotted in the range of -4σ to $+4\sigma$ with respect to the mean value. By observing that in all the plots of figs. 4.6 and 4.7, the plotted line corresponding to the CM results is shorter than the corresponding PEM line, it is obvious that the standard deviation and thus in general the random variation of the results obtained from the CM is smaller than for the PEM. Investigation of the tails of the distribution, i.e. the shape of the curves when the values approach infinity, indicates that also the higher-order characteristics of the distribution are deviating more than in the previous case study of the base case.

In figs. 4.8a and 4.8b sensitivities with regards to the correlation coefficient ρ are inves-

tigated. For the PEM case of fig. 4.8a it can easily be observed that the variation, by means of the standard deviation, is increasing in a linear manner, as the length of the plotted lines increases with the correlation coefficient. This behaviour is expected, as a higher correlation coefficient must lead to a higher level of uncertainty in the random variables.

However, by inspection of fig. 4.8b, the same tendency is not apparent for the CM, which can indicate that the modified CM, as it was implemented in these case studies, is not behaving as it should when the coefficient of correlation is varied. When varying ρ and comparing the results from using the modified PEM and the modified CM, the differences between the obtained results tend to increase with ρ. In itself, this is not necessarily an unexpected phenomena, as the two methodologies are fundamentally different and are expected to behave accordingly. However, by investigation of fig. 4.8, it is obvious that this increased difference is mainly caused by the PEM results changing according to the intuitive understanding of the correlation coefficient sensitivity, while the CM is reacting significantly slower to the variation in correlation coefficient. In addition, there is the non-intuitive development of the coefficient of variation subject to the changes in the correlation coefficient that was noted above. These observations added together, gives a fairly strong indication that the implementation of the modified CM, outlined in section 3.2.2 and demonstrated on the small test system in section 4.1.3, is not entirely correct.

From literature, this suspicion is supported by a similar analysis on the correlation coefficient performed in [7], the main source of the modified CM utilized in these case studies. By using the same general procedure for modified CM, the standard deviation is found to be increasing in a manner almost proportional to the increasing correlation coefficient. This tendency is apparent for all output variables, including the active power flows.

The modified CM outlined in section 3.2.2 is a complex methodology and the possible errors in the implementation of the method will not be discussed in any further detail here. The extensive description of the general methodology in section 3.2.2 together with the detailed step-by-step demonstration should provide sufficient insight to the interpretation of the methodology and the computational steps to make possible a more extensive investigation of the possible sources of error in the implementation.

4.5 IEEE RTS Wind power

In this section, the system model developed in the two preceding sections is retained in terms of the probabilistic modelling of the loads (section 4.3) and the correlation between those loads (section 4.4). The inclusion of wind farms into the IEEE RTS is done with inspiration from [24], where two wind farms are placed at buses 17 and 22, replacing the conventional

generation at those buses. This is the only adjustment that is made to the RTS system as it is described in [11]. Inspired by [18], the two wind farms are correlated with a correlation coefficient of $\rho = 0.9$.

The wind farms are modelled as probabilistic inputs in accordance with the procedure outlined in appendix B. Inspired by [24], the rated power P_r of the two wind farms are 600 MW (bus 17) and 900 MW (bus 22) respectively. The parameters used to model the wind speed (i.e. the Weibull distribution shape and scale parameters) and the parameters describing the speed-power curve (i.e. the cut-in, cut-out and rated speeds) are not explicitly given in [24]. Inspiration as to the determination of these values is rather obtained from [23], where all the aforementioned parameters are provided, along with a presentation of the procedure for handling the uncertainty of WTs which was the main inspiration for appendix B.

In [23], the Weibull parameters determining the distribution describing the wind speed at the two sites are $c = 8.78$ and $k = 1.75$ (refer to appendix A.1.3). The characteristics of the wind turbines utilized in [23] are

$$\begin{aligned} v_{ci} &= 3 \text{ m/s} \\ v_r &= 13 \text{ m/s} \\ v_{co} &= 25 \text{ m/s} \end{aligned} \tag{4.41}$$

where v_{ci}, v_r and v_{co} are, in accordance with appendix B, the cut-in, cut-out and rated speeds respectively, of the given wind turbines.

Now, to model the uncertainty, a random sample of size N is generated from the Weibull distribution describing the wind speed at the site, and then passed through the speed-power curve of eq. (B.1). In [23] a sample size of N=10 000 is used, and is also adopted for the case of this study. A histogram of the wind speed at the site, represented by those 10 000 samples from the Weibull distribution with $c = 8.78$ and $k = 1.75$ is included in figure fig. 4.9. After the 10 000 values making out the distribution of fig. 4.9 are passed through the speed-power curve of eq. (B.1), an equivalent histogram can be used to describe the distribution of the power output of the wind farm. The power output of the 600 MW wind farm at bus 17 is presented in fig. 4.10 together with the speed-power curve resulting from using the values in eq. (4.41) with the expression for the idealized speed-power curve in eq. (B.1).

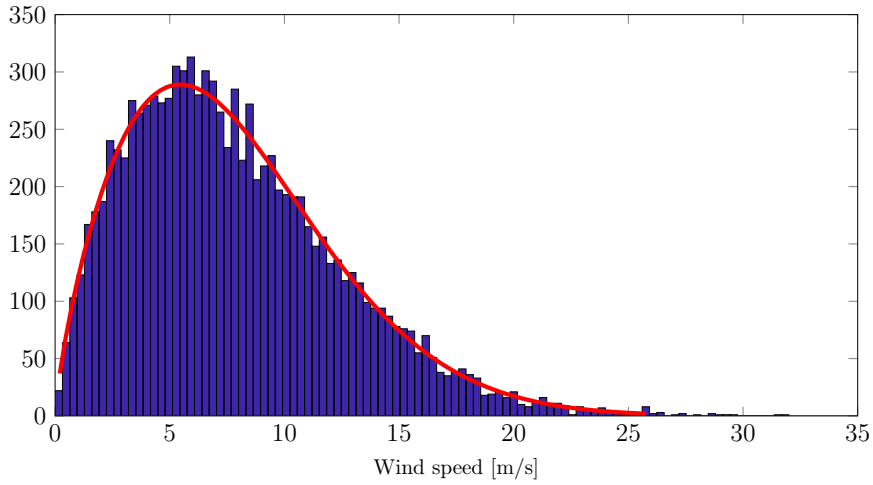
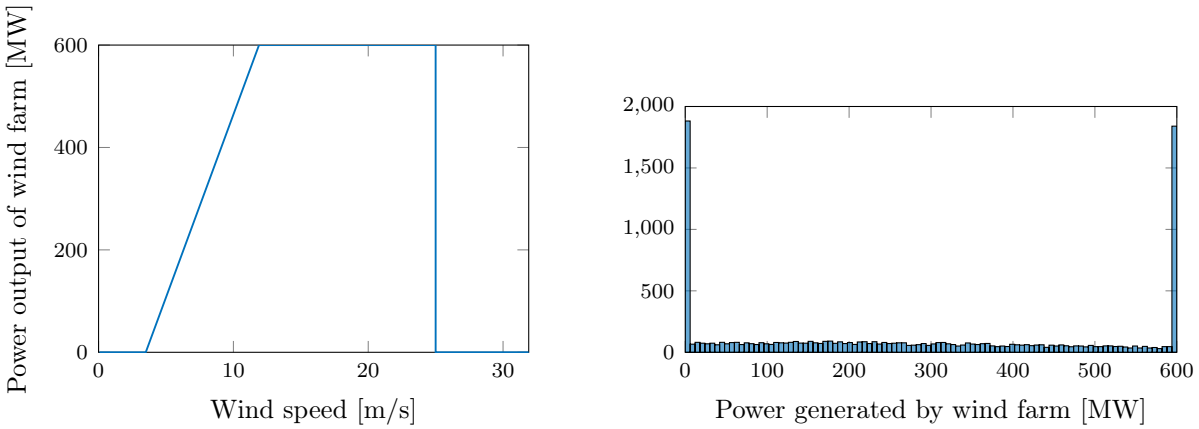


Figure 4.9: Histogram of the Weibull distributed wind speed, obtained from N samples

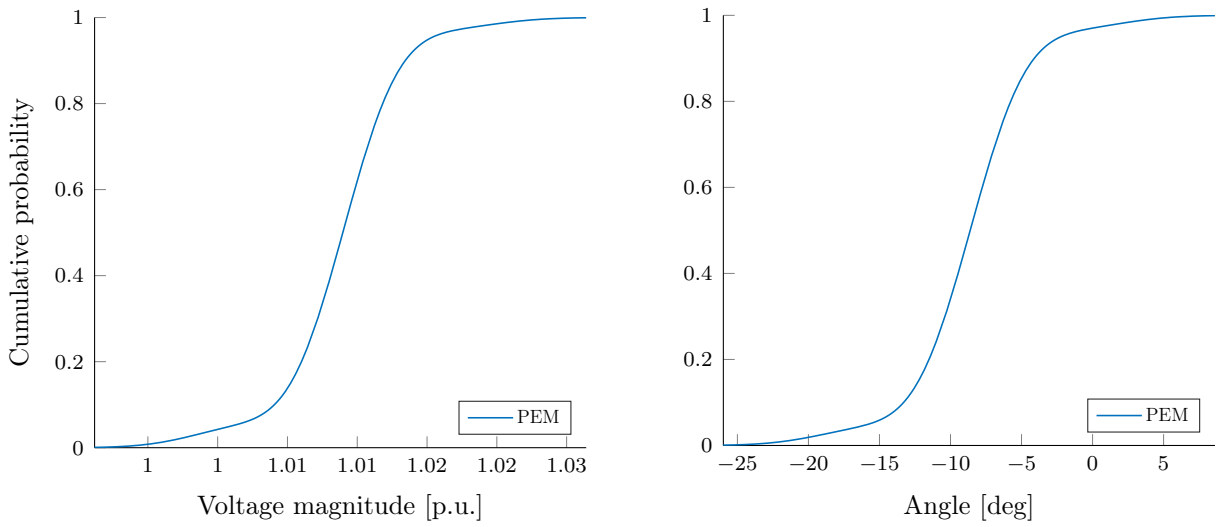


(a) Speed-power curve

(b) Histogram of the power output from wind farm at bus 17

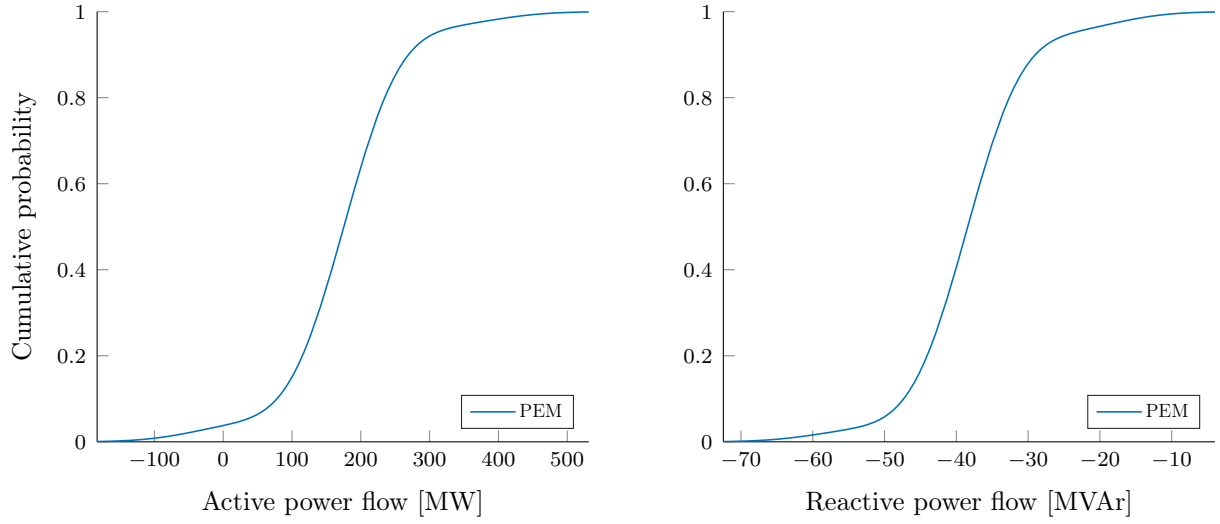
Figure 4.10: Calculation of generated power from wind farm

Now that the moments, skewness and kurtosis are easily found from the distribution in fig. 4.10 (as described in appendix B), the wind farm is included into the PEM just as any other random variable with given values of mean, standard deviation, skewness and kurtosis. The wind power case has only been implemented for the PEM in this thesis. The results for the selected variables also presented in the previous sections are included in (REF)



(a) Voltage magnitude at bus 6 (V_6) **(b)** Voltage angle at bus 6 (δ_6)

Figure 4.11: CDF approximations by GCE for the voltage at bus 6



(a) Active power flow in line $15-16$ (P_{15-16}) **(b)** Reactive power flow in line $15-16$ (Q_{15-16})

Figure 4.12: CDF approximations by GCE for power flows in line $15-16$

4.5.1 Discussion

The case for the system including wind power has only been implemented for the PEM in this thesis, and the implementation for the CM is left as an option for future work. Due to the trouble with the behaviour of the CM in the presence of correlated variables, as discussed in section 4.4, meaningful results were not obtained for the CM when the same procedure

as outlined above was performed. The trouble with the CM for this kind of case model can of course also be partly due to the linearization requirement of the method. As discussed in chapter 2, the CM is expected to perform poorly when handling input variables with a large coefficient of variation, as the equations for nodal power injections and line power flows are linearized around the mean value. The Weibull distribution used to model the wind speed in this section is a distribution with a large coefficient of variation it is expected that the results from the CM are prone to be less accurate than those from the PEM where original load flow equations are retained. However, with the correct implementation, it should be possible to produce meaningful results, as case studies of wind farm correlation is exactly what is studied in reference [7], which was the main source for the derivation of the modified version of the CM in section 3.2.2. In these studies, the Weibull distribution was also used to model the wind speeds, but due to lacking information about the parameters used in the speed-power curve and insufficient data of the test system, it was not considered possible to replicate these results.

Instead the test system used in [24] was chosen for these case studies, as it is a modified version of the IEEE RTS that was already used for the cases in the two preceeding sections. However, in [24] the probabilistic model of the wind speed is not explicitly given, only the rated an mean values of the final electrical output of the wind farm were provided. As this is not sufficient information to input to the model of the wind farm, the data for the wind speed modelling was collected elsewhere, as described in the results section above.

In all, there are no meaningful grounds of comparison for validation of the results obtained from these case studies, from [7, 24]. However, the implementation of the PLF methodology is unchanged, and thus validation of the methodology in itself is based on the discussions in the previous sections of this chapter. Only the generator inputs on those buses containing wind farms are changed compared to the case study model in section 4.4. Further, those inputs are treated in the PEM exactly equally as any other random inputs: by means of the mean, standard deviation, skewness and kurtosis. Hence no changes have been done with regards to the methodological steps.

The probabilistic modelling of each wind farm can be validated on a qualitative level using reference [23]. Recall that both the parameters of the Weibull distribution used to model the wind speed, and the cut-in rate, rated speed and cut-out speed that constitute the speed-power curve are identical in this case study and [23]. Only the rated electrical power output P_r is different as a single wind turbine is modelled in [23], but the output of a full wind farm is modelled in this case study. However, investigating eq. (B.1), the rated power is only a scaling factor and will not affect the shape of the distribution. The distribution of the wind speed, the speed power curve and the resulting electrical power output of figs. 4.9, 4.10a

and 4.10b, respectively, are included also in [23]. Recall that figs. 4.9 and 4.10b are obtained from using a random sample from the Weibull distribution, so the results from two different samples will always deviate slightly; but from a qualitative point of view, differences in the the shapes of the curves and the number of occurences in the histograms are negligible.

The general conclusion that can be drawn from the results of this case study is that the uncertainty in the output variables has increased drastically. The plots in figs. 4.11 and 4.12 are still plotted in the range of 4σ away from the mean in both directions. By comparison of figs. 4.11 and 4.12 to the corresponding results in figs. 4.6 and 4.7, it is obvious that the standard deviation and hence also the coefficient of variation has indeed increased drastically, as the range of the x-axis is much larger. These results are expected, as the Weibull distribution and more intuitively, the wind speed, is associated with high levels of uncertainty. This will consequently affect all variables in the system and result in a higher risk of limitation of transmission line capacities and other security assessment measures in the power system.

4.6 Overall discussion

In the preceding sections of this chapter, different case studies have been performed on various test systems in order to demonstrate the performance of the in-house implementation of the PEM and CM. In the case studies, emphasis has been laid on obtaining validation of the results, at least to the extent possible with available data from existing literature. This form of validation has proved challenging, due to the difficulties of encountering numerical results and sufficient information about test systems and input data in existing research. A quantitative validation of the base case methodologies outlined in sections 3.1.1 and 3.2.1 has been possible due to available results, data and sufficiently detailed methodology description in reference [28]. This validative case study was performed in section 4.2 and it has been concluded that both methodologies provide good results when compared to the results in [28]. An additional source of validation for the implementation of the most basic methodologies, was obtained in section 4.3, where a very simple probabilistic model of the IEEE Reliability Test System was utilized, and the average deviation between the results obtained from the two different methods was close to negligible.

The final two case studies of sections 4.4 and 4.5 investigated the possibility of including correlation between random variables in the methodologies. These studies show promising results for the PEM, but indicate some challenges with the implemented modified version of the CM. Possible sources for this were not discussed in any detail, but further investigation of the modified version of the CM that was chosen for this thesis, will rather be left as a

possibility for future work. In the final case study of section 4.5, two correlated wind farms were included in the system model from section 4.4 in order to investigate the impact of introducing a variable associated with very high levels of uncertainties into the system. The consequences were, as expected, that the uncertainty associated with the output voltages were increased drastically.

5 | Conclusions

The aim of this thesis has been to create in-house tools for conducting probabilistic load flow studies by analytical methods. Two methodologies that together represent the main groups of non-simulation techniques - the Point estimate method and the Cumulant method, have been investigated in their methodological concepts and through application in case studies using various test systems. Besides creating the in-house programming tools, emphasis has been laid on providing pedagogical clarity of the methodological theory, and to demonstrate the calculation procedures in sufficient detail to make possible reproduction of the results obtained in the case studies of chapter 4.

The possibility of validation of the results from the case studies in chapter 4 has been somewhat limited due to the fact that no MC simulation methodology has been implemented to serve as a source of validation. Quantitative validation of the results has only been found possible for the base case system, i.e. where there is no correlation between random variables and no unconventional generation sources. As the modified version of the PEM and CM that are able to handle correlation between variables have not been validated on a quantitative level, a case study including detailed step-by-step demonstrations of these modified methodologies have been performed, in order to at least provide transparency of the computational procedure.

Due to the difficulties related to validation of results, very few conclusions can be drawn from the case studies in this thesis, as to which method is best suited for a given purpose. To be able to do such analyses, a MC technique should be implemented to provide the possibility of irrefutable validation. Then extensive analyses could be performed on the sensitivities of load variation levels, correlation between variables and different types of distributed random variables, in order to investigate which method performs best compared to the exact solution of the problem.

However, several other researchers have addressed these issues, and some general conclusions as to the applications of the two methods can be obtained from existing literature. Reference [14] investigates the two methodologies studied in this thesis, including comparison with the results of a MC method providing what can be considered exact results of the problem. An important conclusion drawn by [14] is that when the level of uncertainty increases, so does the average error in the results obtained from both methodologies. However, the phenomena is more obvious for the CM, as it is limited by linearization of the load flow

equations. Consequently, if high levels of uncertainty are associated with the input variables, it is generally recommended that the PEM be used. A possible drawback of the PEM however, can be with regards to the computation time. In the choice between the PEM and CM, one can assume that the computational efficiency is of importancee, or else the more accurate MC simulation methods should be chosen to perform the PLF. The computational effort of the PEM is directly related to the system size, while the CM solves the problem in a single operation. So dependent on the size of the system, in terms of the number of random variables considered, different conclusions could be made to whether the PEM or CM is recommended.

In conclusion, for non-simulation methods like the PEM and CM, some compromise between speed and accuracy always has to be made. The best suited method for a specific purpose is dependent on the system size, the weightage of accuracy vs efficiency, and also on the level of uncertainty associated with the random variables in the system.

5.1 Future work

The most obvious suggestion for future work following from the discussions of the case studies, would be to make available a MC simulation method for the different cases studied in chapter 4. This would make possible a much wider range of case studies and an efficient way to validate results. More importantly, the work done in this thesis on the cumulant method, for the case with correlated variables, could be fine tuned for *reliable* results. Tendencies in the results obtained from PEM and CM by different sensitivity analyses can then be analyzed, providing better grounds for actually determining which method is best suited for different purposes.

In chapter 3, some assumptions were made regarding the possibilities of correlation between variables. To provide a relatively simple introduction to the complex methodologies, correlation was assumed to exist only between power injections at different buses, and only between loads or between generators (i.e. not between loads and generators). The applications of the methodologies would be extended even further if possibilites to include correlation between power injections at the same bus, or between loads and generators.

Finally, an important source of uncertainty in the power system that has not been addressed in this thesis are uncertainties related to network outages and other types of structural changes in the physical network. Studies including these types of uncertainties are important in reliability assessments and studies of security of supply.

Bibliography

[1] X. Ai and J. Wen. "A Discrete Point Estimate Method for Probabilistic Load Flow Based on the Measured Data of Wind Power". *IEEE Transactions on Industry Applications*, 49(5):2244–2252, 2013.

[2] R. N. Allan, A. M Leite da Silva, and R. C. Burchett. "Evaluation methods and accuracy in probabilistic load flow solutions". *IEEE Transactions on Power Apparatus and Systems*, PAS-100(5):2539–2546, 1981.

[3] P. Amid and C. Crawford. "Cumulant-based probabilistic load flow analysis of wind power and electric vehicles". *2016 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)*, (1):1–6, 2016.

[4] G. J. Anders. *Probability concepts in electric power systems*. John Wiley & Sons, 1990.

[5] B. Borkowska. "Probabilistic Load Flow". *IEEE Transactions on Power Apparatus and Systems*, PAS-93(3):1–6, 1974.

[6] D. Cai, X. Li, K. Zhou, J. Xin, and K. Cao. "Probabilistic load flow algorithms considering correlation between input random variables: A review". *Proceedings of the 2015 10th IEEE Conference on Industrial Electronics and Applications, ICIEA 2015*, pages 1139–1144, 2015.

[7] Defu Cai, Jinfu Chen, Dongyuan Shi, Xianzhong Duan, Huijie Li, and Meiqi Yao. Enhancements to the cumulant method for probabilistic load flow studies. *IEEE Power and Energy Society Gneral Meting*, pages 1–8, 2012.

[8] P. Chen, Z. Chen, and B. Bak-Jensen. "Probabilistic load flow: A review". *3rd International Conference on Deregulation and Restructuring and Power Technologies, DRPT 2008*, (April):1586–1591, 2008.

[9] C. Delgado and J. A. Domínguez-Navarro. "Point estimate method for probabilistic load flow of an unbalanced power distribution system with correlated wind and solar sources". *International Journal of Electrical Power and Energy Systems*, 61:267–278, 2014.

[10] J. F. Dopazo, O. A. Klitin, and A. M. Sasson. Stochastic load flows. *IEEE Transactions on Power Apparatus and Systems*, 94(2):299–309, 1975.

[11] C. Grigg and P. Wong. The IEEE reliability test system -1996 a report prepared by the reliability test system task force of the application of probability methods subcommittee. *IEEE Transactions on Power Systems*, 14(3):1010–1020, 1999.

[12] A.B. Krishna, N. Gupta, K.R. Niazi, and A. Swarnkar. "Probabilistic power flow in radial distribution systems using point estimate methods". *2017 4th International Conference on Advanced Computing and Communication Systems, ICACCS 2017*, 2017.

[13] A. Kumar. "Comparative analysis of Linearization and Point Estimate Methods for Solving Probabilistic Load Flow". *2016 IEEE 7th Power India International Conference (PIICON)*, 2016.

[14] G. Li and X. Zhang. "Comparison between Two Probabilistic Load Flow Methods for Reliability Assessment". *IEEE Power Energy Society General Meeting*, pages 1–7, 2009.

[15] Pei Ling Liu and Armen Der Kiureghian. Multivariate distribution models with prescribed marginals and covariances. *Probabilistic Engineering Mechanics*, 1(2):105–112, 1986.

[16] MATLAB. *version 9.2.0 (R2017a)*. The MathWorks Inc., Natick, Massachusetts, 2017.

[17] J. M. Morales and J. Pérez-Ruiz. "Point estimate schemes to solve the probabilistic power flow". *IEEE Transactions on Power Systems*, 22(4):1594–1601, 2007.

[18] J.M. Morales, L. Baringo, A.J. Conejo, and R. Minguez. "Probabilistic power flow with correlated wind sources". *IET Generation, Transmission & Distribution*, 4(5):641, 2010.

[19] P.S.R. Murty. *Power System Analysis*. Elsevier Ltd., 2 edition, 2017.

[20] O. A. Oke and D. W.P. Thomas. "Enhanced cumulant method for probabilistic power flow in systems with wind generation". *2012 11th International Conference on Environment and Electrical Engineering, EEEIC 2012 - Conference Proceedings*, pages 849–853, 2012.

[21] Purobi Patowary and Neeraj K Goyal. Security Assessment of Bus Voltages in Electric Power Systems. 56(Icpes):35–40, 2012.

[22] H. Saadat. *Power System Analysis*. PSA Publishing, 3 edition, 2010.

[23] Alireza Soroudi, Morteza Aien, and Mehdi Ehsan. A probabilistic modeling of photo voltaic modules and wind power generation impact on distribution networks. *IEEE Systems Journal*, 6(2):254–259, 2012.

[24] J. Usaola. "Probabilistic load flow in systems with wind generation". *IET Generation, Transmission & Distribution*, 3(12):1031, 2009.

[25] D. Villanueva, A. E. Feijóo, and J. L. Pazos. "An analytical method to solve the probabilistic load flow considering load demand correlation using the DC load flow". *Electric Power Systems Research*, 110:1–8, 2014.

[26] R. E. Walpole, R. H. Myers, S. L. Myers, and K. Ye. *Probability and statistics for engineers and scientists*. Pearson Education Limited, 9 edition, 2016.

[27] C. Wan, Z. Xu, Z. Y. Dong, and K. P. Wong. "Probabilistic load flow computation using first-order second-moment method". *IEEE Power and Energy Society General Meeting*, pages 1–6, 2012.

[28] X. Wang and J.R. McDonald. *Modern Power System Planning*. McGraw Hill Book Company Europe, 1994.

[29] P. Zhang and S. T. Lee. "Probabilistic load flow computation using the method of combined cumulants and Gram-Charlier expansion". *Power Systems, IEEE Transactions on*, 19(1):676–682, 2004.

[30] R. D. Zimmerman, C. E. Murillo-Sanchez, and R. J. Thomas. "MATPOWER: Steady-State Operations, Planning, and Analysis Tools for Power Systems Research and Education". *IEEE Transactions on Power Systems*, 26(1):12–19, 2011.

A | Some background theory

A.1 Some probability distributions

A.1.1 Normal distribution

The *normal* or *gaussian* distribution is the most important continuous distribution in the whole field of statistics [26]. In this thesis, it is used to model the majority of the nodal load demands present in the test systems utilized in the case studies. The following introduction to the normal distribution is inspired by [4, 26].

Let a random variable X be normal distributed, with expected value μ and standard deviation σ (refer to section 2.4 for an introduction of these quantities). The PDF of X is then [4]

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-\mu)^2}{2\sigma}\right), \quad -\infty \leq x \leq \infty, \quad \sigma > 0 \tag{A.1}$$

Note from eq. (A.1) that the PDF of the normal distribution is dependent on μ and σ. A short term for the distribution is N(μ,σ). This makes it hard to obtain the CDF from integration of eq. (A.1), because it cannot be integrated in closed form for every pair of limits a and b, respectively [4]. Instead, probabilities of the normal distribution are usually rather obtain from the CDF of the *standard normal distribution*. The standard normal distribution is a normal distribution with $\mu = 0$ and $\sigma = 1$, i.e. N(0,1). To demonstrate this relation, make a change of variables by denoting

$$z = \frac{x-\mu}{\sigma} \tag{A.2}$$

Inserting this into eq. (A.1) and integrating up to a value of x to obtain the CDF, yields:

$$F(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\frac{x-\mu}{\sigma}} \exp\left(-\frac{z^2}{2}\right) dz \tag{A.3}$$

From eq. (A.3), we can express $F(x)$ when $\mu = 0$ and $\sigma = 1$, i.e. the CFD Φ of the standard normal distribution as:

$$\Phi(z) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z} \exp\left(-\frac{t^2}{2}\right) dt \tag{A.4}$$

Now, by combining eqs. (A.3) and (A.4), the CDF of the normal distribution N(μ,σ) is

$$F_X(x) = \Phi\left(\frac{x-\mu}{\sigma}\right) \tag{A.5}$$

A.1.2 Bernoulli distribution

The Bernoulli distribution is a special case of the discrete *binomial* distribution. In the binomial distribution n experiments or *trials* are performed, with the outcome of each trial being either *success* or *failure*. The trials often referred to as *Bernoulli trials* must satisfy the following assumptions: [4]

1. There are only two possible outcomes for each trial; success and failure.
2. The probability of success of each event is the same for each trial.
3. Different trials of a given type are statistically independent.

Now assume that the above assumptions are satisfied and that from n trials, the probability of success is p and conversely the probability of failure is $q = 1 - p$. The probability mass function (PMF) expressing the probability k of successes in those n trials, is given by [4]

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \tag{A.6}$$

where $\binom{n}{k} = n!/[k!(n-k)!]$ is the binomial coefficient. The mean and variance of the binomial distribution is further

$$\begin{aligned} \mu &= np \\ \sigma^2 &= npq \end{aligned} \tag{A.7}$$

Now consider the case where the number of trials n, is equal to 1. Then the probability of success in this single trial is

$$P(X = 1) = \binom{1}{1} p^1 (1-p)^{1-1} = p \tag{A.8}$$

The only other possible outcome of this single trial is failure, i.e. the number of successes k is equal to zero. eq. (A.6) yields for this case:

$$P(X = 1) = \binom{1}{0} p^0 (1-p)^{1-0} = 1 - p \tag{A.9}$$

The binomial distribution with only $n = 1$ trial is what is known as the Bernoulli disribution [26]. From the expressions provided in section 2.4, raw moments, cumulants and other statistical measures can easily be found for the Bernoulli distribution. Let Y be a random variable that is a function of the Bernoulli distributed random variable X such that $Y = bX$. Referring to eq. (2.13), the r-th order raw moment of the Bernoulli distributed variable Y is:

$$E[y^r] = \sum_Y y^r p(y) = \sum_X (bx)^r p(x) = b^r \sum_X x^r p(x) = b^r(1 \cdot p + 0 \cdot (1-p)) = b^r p \tag{A.10}$$

From the raw moments, the r-th order cumulants of Y can easily be obtained from eq. (2.21). From eq. (A.7), the mean and standard deviation of Y are

$$\begin{aligned} \mu_Y &= bp \\ \sigma_Y &= b\sqrt{pq} \end{aligned} \tag{A.11}$$

Using eq. (2.16), the skewness λ_3 and kurtosis λ_4 are easily obtained as

$$\begin{aligned} \lambda_{3_Y} &= b\frac{1-2p}{\sqrt{pq}} \\ \lambda_{4_Y} &= b\frac{1-6pq}{pq} \end{aligned} \tag{A.12}$$

In the context of this thesis, b will typically be the capacity of a generator, and X will describe the randomness of the generator's operation, based on its FOR. The FOR of a generator can be viewed as the probability of failure, and thus q, as it is defined here, is equal to the FOR:

$$\begin{aligned} p &= 1 - q = 1 - \text{FOR} \\ q &= \text{FOR} \end{aligned} \tag{A.13}$$

A.1.3 Weibull distribution

The Weibull distribution is a continuous, two-parameter distribution, which will in this thesis be used to describe the wind speed at a given location. Given the Weibull parameters k and c, often referred to as the scale and shape parameter respectively, the PDF of the Weibull distributed random variable X is given by [23]

$$f_X(x) = \left(\frac{k}{c}\right)\left(\frac{x}{c}\right)^{k-1} \exp\left[-\left(\frac{x}{c}\right)^k\right], \qquad x > 0 \tag{A.14}$$

Integrating $f_X(x)$ to a value x, yields the CDF of X

$$F_X(x) = \int_{-\infty}^{x} \left(\frac{k}{c}\right)\left(\frac{x}{c}\right)^{k-1} \exp\left[-\left(\frac{x}{c}\right)^{k}\right] dx = 1 - e^{-(\frac{x}{c})^k} \tag{A.15}$$

A.2 Correlation and covariance

Correlation between random variables can be described by the means of different measures. *Covariance* and *correlation coefficient* are two measures most relevant for this thesis, hence this section will focus on providing an introduction of those concepts. For the most general case, let X and Y be two correlated random variables with joint probability distributions described by $f(x, y)$. The covariance σ_{XY} of X and Y is then: [26]

$$\begin{aligned}\sigma_{XY} &= E[(X - \mu_X)(Y - \mu_Y] \\ &= \begin{cases} \sum_x \sum_y (x - \mu_X)(y - \mu_Y) f(x, y), & X \text{ and } Y \text{ discrete} \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_X)(y - \mu_Y) f(x, y) dx dy, & X \text{ and } Y \text{ continuous} \end{cases}\end{aligned} \tag{A.16}$$

where E[·] is the expected value operator and μ_X and μ_Y are the expected values of X and Y, respectively. In accordance with the general introduction of basic statistics in section 2.4, the uppercase X denotes the random variables, while the corresponding lowercase x denotes a particular value in the range of X. Equivalently for Y and y.

The covariance σ_{XY} is a measure of the the nature of correlation between X and Y. The sign of σ_{XY} indicates whether the correlation is positive (i.e. large values of X result in large values of Y) or negative (i.e. large values of X result in small values of Y).

From the covariance σ_{XY}, the correlation coefficient ρ_{XY} can easily be obtained as

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} \tag{A.17}$$

The correlation coefficient is a dimensionless size restricted between the values -1 and 1. When the two variables are independent, $\sigma_{XY} = 0$, and consequently so is ρ_{XY}, according to eq. (A.17). A value of $\rho_{XY} = 1$ or $\rho_{XY} = -1$ corresponds to complete linear dependency, the sign dependent on the linear expression relating X and Y.

When considering a *set* of random variables, the correlation between the variables of that set can be described by the use of a correlation matrix. Let a set of n correlated random variables be represented by the vector $\mathbf{X} = [X_1, ..., X_n]$. The correlation between the variables can be described by a matrix of which the rows corresponds to $\mathbf{X}$ and the columns to $\mathbf{X}^T$.

A *covariance matrix* then describes the covariance between the variables in $\mathbf{X}$:

$$\mathbf{C}_{\mathbf{x}(\sigma)} = cov(\mathbf{X}, \mathbf{X}^T) = \begin{bmatrix} \sigma_{X_1}^2 & \sigma_{X_1X_2} & \dots & \sigma_{X_1X_n} \\ \sigma_{X_2X_1} & \sigma_{X_2}^2 & \dots & \sigma_{X_2X_n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{X_nX_1} & \sigma_{X_nX_2} & \dots & \sigma_{X_n}^2 \end{bmatrix} \tag{A.18}$$

where $cov(\cdot, \cdot)$ is the covariance operator. Notice that the diagonal element (i, i) becomes the square of the the standard deviation σ_{X_i}, also referred to as the variance ν_{X_i} (refer to section 2.4.4). This can be derived from eq. (A.17), noticing that the correlation coefficient ρ of a variable X_i with itself is equal to 1 (complete dependency). Then, according to eq. (A.17)

$$\sigma_{X_iX_i} = 1 \cdot \sigma_{X_i}\sigma_{X_i} = \sigma_{X_i}^2 \tag{A.19}$$

If it is desirable to describe the correlation between the variables of $\mathbf{X}$ by their respective correlation coefficients, the *correlation coefficient matrix* is defined in a similar way:

$$\mathbf{C}_{\mathbf{x}(\rho)} = \rho(\mathbf{X}, \mathbf{X}^T) = \begin{bmatrix} 1 & \rho_{X_1X_2} & \dots & \rho_{X_1X_n} \\ \rho_{X_2X_1} & 1 & \dots & \rho_{X_2X_n} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{X_nX_1} & \rho_{X_nX_2} & \dots & 1 \end{bmatrix} \tag{A.20}$$

where $\rho(\cdot, \cdot)$ is the correlation coefficient operator. Notice that in this case, the diagonal elements becomes unity, also in accordance with eq. (A.17)

$$\rho_{X_iX_i} = \frac{\sigma_{X_i}^2}{\sigma_{X_i}\sigma_{X_i}} = 1 \tag{A.21}$$

A.3 Orthogonal transformation

The orthogonal transformation as it is utilized in this thesis, transforms a set $\mathbf{u}$ of correlated variables to a set $\mathbf{v}$ of independent variables:

$$\mathbf{v} = \mathbf{B}\mathbf{u} \tag{A.22}$$

where $\mathbf{B}$ is the transformation matrix yet to be determined. The following derivation is mainly inspired by [7]. Let $\mathbf{C}_\mathbf{u}$ be the correlation coefficient matrix describing the correlation between the variables in $\mathbf{u}$. In most engineering applications, this matrix is positive definite,

which makes it possible to decompose it by Cholesky decomposition, obtaining:

$$\mathbf{C_u} = \mathbf{L}\mathbf{L}^T \tag{A.23}$$

where $\mathbf{L}$ is an inferior, triangular matrix. The elements of $\mathbf{L}$ can be calculated by: [7]

$$\begin{aligned} l_{kk} &= (\rho_{u_{kk}} - \sum_{m=1}^{k-1} l_{km}^2)^{1/2}, \quad k = 1, 2, ..., n \\ l_{ik} &= \frac{\rho_{u_{ik}} - \sum_{m=1}^{k-1} l_{im} l_{km}}{l_{kk}}, \quad i = k+1, k+2, ..., n \end{aligned} \tag{A.24}$$

where n is the number of variables in $\mathbf{u}$ and l_{kk}, l_{ik}, l_{km} and $l_k k$ are elements of the matrix $\mathbf{L}$. $\rho_{u_{kk}}$ and $\rho_{u_{ik}}$ are, in accordance with appendix A.2, elements of the correlation coefficient matrix $\mathbf{C_u}$.

For the set of transformed variables $\mathbf{v}$ to be independent, obviously their correlation coefficient matrix $\mathbf{C_v}$ has to be equal to the identity matrix $\mathbf{I}$. In the identity matrix, all off-diagonal elements are zero, indicating in the correlation coefficient matrix that the correlation coefficient $\rho_{ij} = 0$ for all buses i and j. Let now $\rho(\cdot, \cdot)$ denote the correlation coefficient operator as it was described in appendix A.2. Then, expressing $\mathbf{C_v}$ by the use of eqs. (A.22) and (A.23), yields

$$\begin{aligned} \mathbf{C_v} = \rho(\mathbf{v}, \mathbf{v}^T) = \rho(\mathbf{B}\mathbf{u}, \mathbf{u}^T\mathbf{B}^T) = \mathbf{B}\rho(\mathbf{u}, \mathbf{u}^T)\mathbf{B}^T = \mathbf{B}\mathbf{C_u}\mathbf{B}^T \\ = \mathbf{B}(\mathbf{L}\mathbf{L}^T)\mathbf{B}^T = (\mathbf{B}\mathbf{L})(\mathbf{B}\mathbf{L})^T = \mathbf{I} \end{aligned} \tag{A.25}$$

From eq. (A.25), an expression for the transformation matrix $\mathbf{B}$ can be found, as

$$(\mathbf{B}\mathbf{L})(\mathbf{B}\mathbf{L})^T = \mathbf{I} \quad \Rightarrow \quad \mathbf{B} = \mathbf{L}^{-1} \tag{A.26}$$

So eq. (A.22) can be expressed in terms of the decomposed correlation coefficient matrix $\mathbf{C_u}$:

$$\mathbf{v} = \mathbf{L}^{-1}\mathbf{u} \tag{A.27}$$

Taking the inverse of the transformation in eq. (A.27), allows the correlated variables $\mathbf{u}$ to be modelled as a linear combination of the independent variables in $\mathbf{v}$:

$$\mathbf{u} = \mathbf{L}\mathbf{v} \tag{A.28}$$

An equivalent derivation can be done if the covariance matrix (refer to appendix A.2)

was used instead of the correlation coefficient matrix to describe the correlation between the variables in $\mathbf{u}$. Then in eq. (A.25), the covariance operator $cov(\cdot,\cdot)$ is used instead of the correlation coefficient operator.

A.4 Linearization of power flow equations

Linearization of the line power flow equations forms the basis of the methodology of the CM. When both the state variables $\mathbf{X}$ and the line power flows $\mathbf{Z}$ are of interest, both the nodal power equations and the line power flow equations need to be linearized in order to be used in the CM. The following derivation is inspired by [28], where the nodal power equations were linearized.

Let the vector $\mathbf{W} = \mathbf{f}(\mathbf{X})$, where X are the state variables, contain the nodal power equations for all buses, including both active and reactive power. The nonlinear nodal power equations for bus i are given by

$$P_i = \sum_{j=1}^{n} |V_i||V_j||Y_{ij}| \cos(\theta_{ij} - \delta_i + \delta_j) \tag{A.29}$$

$$Q_i = -\sum_{j=1}^{n} |V_i||V_j||Y_{ij}| \sin(\theta_{ij} - \delta_i + \delta_j) \tag{A.30}$$

where the meaning of the variables are defined equivalently as in section 2.3. Let equivalently $\mathbf{Z} = \mathbf{g}(\mathbf{X})$ denote the nonlinear line power flow equations,

$$P_{ij} = V_i V_j (G_{ij} cos\theta_{ij} + B_{ij} sin\theta_{ij}) - t_{ij} G_{ij} B_{ij} V_i^2 \tag{A.31}$$

$$Q_{ij} = V_i V_j (G_{ij} sin\theta_{ij} - B_{ij} cos\theta_{ij}) + (t_{ij} B_{ij} - b_{ij0}) V_i^2 \tag{A.32}$$

where all variables are defined equivalently as in section 3.2.2.

The aim of the linearization procedure is to obtain a set of linear equations that are valid within a range of small disturbances with respect to some defined starting point. Let this point be denoted $\mathbf{X_0}$ for the state variables. In the context of this thesis, $\mathbf{X_0}$ represents the expected values of random variables contained in $\mathbf{X}$. We now want to linearize the nodal power equations around $\mathbf{W_0} = \mathbf{f}(\mathbf{X_0})$. Applying a small disturbance to the state variables $\mathbf{X}$ will result in a corresponding change in the nodal power injections, which can be described by:

$$\mathbf{W_0} + \mathbf{\Delta W} = \mathbf{f}(\mathbf{X_0} + \mathbf{\Delta X}) \tag{A.33}$$

Applying Taylor series expansion to eq. (A.33) yields:

$$\mathbf{W_0} + \mathbf{\Delta W} = \mathbf{f}(\mathbf{X_0}) + \mathbf{f}'(\mathbf{X_0})\mathbf{\Delta X} + \mathbf{f}''(\mathbf{X_0})(\mathbf{\Delta X})^{\mathbf{2}} + \mathbf{f}'''(\mathbf{X_0})(\mathbf{\Delta X})^{\mathbf{3}} + ... \tag{A.34}$$

where $\mathbf{f}'$ denotes the first-order derivative of $\mathbf{f}$. Assume now that the disturbance is sufficiently small for terms of $(\mathbf{\Delta X})^2$ or higher to safely be ignored. Then eq. (A.34) simplifies to:

$$\mathbf{W_0} + \mathbf{\Delta W} = \mathbf{f}(\mathbf{X_0}) + \mathbf{f}'(\mathbf{X_0})\mathbf{\Delta X} \tag{A.35}$$

Substituting $\mathbf{W_0} = \mathbf{f}(\mathbf{X_0})$ into eq. (A.35) yields

$$\begin{gathered} \mathbf{f}(\mathbf{X_0}) + \mathbf{\Delta W} = \mathbf{f}(\mathbf{X_0}) + \mathbf{f}'(\mathbf{X_0})\mathbf{\Delta X} \\ \Updownarrow \\ \mathbf{\Delta W} = \mathbf{f}'(\mathbf{X_0})\mathbf{\Delta X} \end{gathered} \tag{A.36}$$

In power flow analysis studies, it is usually the nodal power injections $\mathbf{W}$ that are the known input values, leaving the state variables $\mathbf{X}$ to be the output variables. Using eq. (A.36) to express the state variables yields:

$$\mathbf{\Delta X} = (\mathbf{f}'(\mathbf{X_0}))^{-1}\mathbf{\Delta W} = \mathbf{J_0}^{-1}\mathbf{\Delta W} \tag{A.37}$$

where $\mathbf{J_0}$ is the Jacobian matrix, calculated by

$$\mathbf{f}'(\mathbf{X_0}) = \left.\frac{\partial \mathbf{f}(\mathbf{X})}{\partial \mathbf{X}}\right|_{\mathbf{X}=\mathbf{X_0}} \tag{A.38}$$

and

$$\mathbf{S_0} = \mathbf{J_0}^{-1} \tag{A.39}$$

is the sensitivity matrix relating changes in the nodal power injections to the corresponding changes in the state variables. Then eq. (A.37) can be expressed as

$$\mathbf{\Delta X} = \mathbf{S_0}\mathbf{\Delta W} \tag{A.40}$$

Now consider performing the exact same procedure in eqs. (A.33)–(A.36) for the line power flow equations $\mathbf{Z} = \mathbf{g}(\mathbf{X})$. eq. (A.36) then yields, for the line power flows

$$\mathbf{\Delta Z} = \mathbf{g}'(\mathbf{X_0})\mathbf{\Delta X} \tag{A.41}$$

where

$$\mathbf{g}'(\mathbf{X_0}) = \left.\frac{\partial \mathbf{g}(\mathbf{X})}{\partial \mathbf{X}}\right|_{\mathbf{X}=\mathbf{X_0}} \tag{A.42}$$

Denoting the sensitivity matrix relating changes in the state variables to corresponding changes in the line power flows $\mathbf{D_0}$, yields

$$\mathbf{\Delta Z} = \mathbf{D_0 \Delta X} \tag{A.43}$$

B | Probabilistic modelling of wind turbines

The generated power from a wind turbine (WT) obviously highly depends on the wind speed at the site where the turbine is located. The wind power is further associated with a high level of uncertainty and needs to be modelled as a random variable. The Weibull distribution is used to model the wind speed uncertainty in the majority of the sources including wind power modelling that are referred to in this thesis [23, 18, 24, 7] and is also chosen to model the wind speed uncertainty in the case studies of chapter 4. Expressions for the PDF and CDF of the Weibull distribution are provided in appendix A.1.3.

Notice that it is the *wind speed* that is Weibull distributed, not the generator output from the wind turbine. The WT power output is dependent on several other factors that also need be taken into account. For reasons related to the construction of the turbine, the WT will not produce any power if the wind speed is below a certain minimum value v_{ci} or above a maximum value v_{co}. These speed limits are often referred to as the *cut-in* and *cut-out* speeds of the WT, depicting the wind speeds at which the WT goes in to production and falls out of production, respectively. When the cut-in speed v_{ci} is reached and increasing, the power output from the WT will generally increase linearly with the wind speed v up to a certain point v_r. v_r will be referred to as the *rated* speed, and is the maximum speed the WT is constructed to retain. This means that as the wind speed increases above the rated speed, the speed of the WT will not increase, and consequently neither will the power output of the WT. The power output is thus kept constant for increasing wind speed, until a wind speed of v_{co} is reached, and the WT is shut down completely.

The above reasoning can be modelled by a *speed-power curve* relating the wind speed v to the generator power output P. The speed-power curve will generally vary between different manufacturers of WTs [18], but an idealized curve can provide a sufficiently good model for studies such as the ones performed in chapter 4 in this thesis. A typical, idealized

speed-power curve can be written as follows [23]:

$$P = \begin{cases} 0, & \text{if } v \leq v_{ci} \\ \dfrac{v - v_{ci}}{v_r - v_{ci}} P_r, & \text{if } v_{ci} \leq v \leq v_r \\ P_r, & \text{if } v_r \leq v \leq v_{co} \\ 0, & \text{if } v \geq v_{co} \end{cases} \tag{B.1}$$

The speed-power curve of eq. (B.1) is not a continuous function of v and there is no straightforward way to easily relate the probability distribution of the wind speed to the power output of the WT. The most obvious way to handle this issue, is to simply generate a random sample from the Weibull distribution with specified parameters k and c, and then to run all values from the sample through the speed-power curve in eq. (B.1). This procedure is described in [23] and also adopted in this thesis.

When the sample from the Weibull distribution has been run through the speed-power curve, a sample of the same size N is obtained for the WT power output. By viewing this sample as a discrete distribution, raw moments, cumulants and properties like standard deviation, skewness and kurtosis can easily be found from the expressions provided in section 2.4.

C | Complete data of test systems

C.1 IEEE 14 bus test system

Base case data and probabilistic data for the IEEE 14 bus test system used for case studies in section 4.2 is presented here. All data is identical to [28]. R and X are the resistance and reactance of the line, respectively. b_{ij0} is the susceptance to ground (given in table C.1 as half the value of the full line susceptance) and t_{ij} is the transformer tap value.

Table C.1: Line data for IEEE 14 bus test system

From bus	To bus	R (p.u.)	X (p.u.)	b_{ij0} (p.u.)	t_{ij}
1	2	0.01938	0.05917	0.0528	0
1	5	0.05403	0.22304	0.0492	0
2	3	0.04699	0.19797	0.0438	0
2	4	0.05811	0.17632	0.034	0
2	5	0.05695	0.17388	0.0346	0
3	4	0.06701	0.17103	0.0128	0
4	5	0.01335	0.04211	0	0
4	7	0	0.20912	0	0.978
4	9	0	0.55618	0	0.969
5	6	0	0.25202	0	0.932
6	11	0.09498	0.1989	0	0
6	12	0.12291	0.25581	0	0
6	13	0.06615	0.13027	0	0
7	8	0	0.17615	0	0
7	9	0	0.11001	0	0
9	10	0.03181	0.0845	0	0
9	14	0.12711	0.27038	0	0
10	11	0.08205	0.19207	0	0
12	13	0.22092	0.19988	0	0
13	14	0.17093	0.34802	0	0

Probabilistic modelling of the loads is presented in tables C.2 and C.3. In table C.2 mean values μ are given in MW and MVAr (active and reactive power, respectively), and standard deviations σ in percentage of the corresponding mean value.

Data for probabilistic modelling of generation is given in table C.4.

Table C.2: Probabilistic load data for normal distributed loads of IEEE 14 bus test system

	P_i (MW)		Q_i (MVAr)	
Node	μ	σ	μ	σ
1	0.0	0.0	0.0	0.0
2	21.74	0.09	12.7	0.092
3	94.20	0.10	19.0	0.105
4	47.8	0.11	-3.9	0.097
5	7.60	0.09	1.6	0.05
6	11.20	0.06	7.5	0.063
7	0.0	0.0	0.0	0.0
8	0.0	0.0	0.0	0.0
10	9.0	0.10	5.8	0.10
11	3.5	0.095	1.8	0.095
12	6.1	0.076	1.6	0.086
13	13.5	0.105	5.8	0.095
14	14.9	0.086	5.0	0.086

Table C.3: Probabilistic distribution of discrete load at bus 9 of the IEEE 14 bus system

Active power (MW)	13.4	19.6	30.2	34.8	
Probability	0.10	0.15	0.30	0.25	0.20
Reactive power (MVAr)	7.5	11.0	17.0	19.6	21.0
Probability	0.10	0.15	0.30	0.25	0.20

Table C.4: Probabilistic modelling of generation in the IEEE 14 bus system

Bus no.	Capacity (MW)	No. of units	FOR
1	25	10	0.08
2	22	2	0.09

C.2 IEEE RTS

In this section, base case load demand data and line data for the IEEE RTS (24 bus case) is presented explicitly. For additional data, including generator data, refer to [11]. For the probabilistic modelling of the load demands, refer to the various cases in chapter 4.

Table C.5: Base case load demand data for the IEEE RTS

Bus i	P_i (MW)	Q_i (MVAr)
1	108	22
2	97	20
3	180	37
4	74	15
5	71	14
6	136	28
7	125	25
8	171	35
9	175	36
10	195	40
11	0	0
12	0	0
13	265	54
14	194	39
15	317	64
16	100	20
17	0	0
18	333	68
19	181	37
20	128	26
21	0	0
22	0	0
23	0	0
24	0	0

Table C.6: Line data for IEEE RTS

From bus	To bus	R (p.u.)	X (p.u.)	b_{ij0} (p.u.)	t_{ij}
1	2	0.0026	0.0139	0.4611	0
1	3	0.0546	0.2112	0.0572	0
1	5	0.0218	0.0845	0.0229	0
2	4	0.0328	0.1267	0.0343	0
2	6	0.0497	0.192	0.052	0
3	9	0.0308	0.119	0.0322	0

Table C.6 – continued

i	j	R	X	b_{ij0}	t_{ij}
3	24	0.0023	0.0839	0	1.03
4	9	0.0268	0.1037	0.0281	0
5	10	0.0228	0.0883	0.0239	0
6	10	0.0139	0.0605	2.459	0
7	8	0.0159	0.0614	0.0166	0
8	9	0.0427	0.1651	0.0447	0
8	10	0.0427	0.1651	0.0447	0
9	11	0.0023	0.0839	0	1.03
9	12	0.0023	0.0839	0	1.03
10	11	0.0023	0.0839	0	1.03
10	12	0.0023	0.0839	0	1.03
11	13	0.0061	0.0476	0.0999	0
11	14	0.0054	0.0418	0.0879	0
12	13	0.0061	0.0476	0.0999	0
12	23	0.0124	0.0966	0.203	0
13	23	0.0111	0.0865	0.1818	0
14	16	0.005	0.0389	0.0818	0
15	16	0.0022	0.0173	0.0364	0
15	21	0.0063	0.049	0.103	0
15	21	0.0063	0.049	0.103	0
15	24	0.0067	0.0519	0.1091	0
16	17	0.0033	0.0259	0.0545	0
16	19	0.003	0.0231	0.0485	0
17	18	0.0018	0.0144	0.0303	0
17	22	0.0135	0.1053	0.2212	0
18	21	0.0033	0.0259	0.0545	0
18	21	0.0033	0.0259	0.0545	0
19	20	0.0051	0.0396	0.0833	0
19	20	0.0051	0.0396	0.0833	0
20	23	0.0028	0.0216	0.0455	0
20	23	0.0028	0.0216	0.0455	0
21	22	0.0087	0.0678	0.1424	0

D | Additional results from case studies

D.1 3-bus test system

Results for the case studies with correlation were presented in sections 4.1.2 and 4.1.3 for PEM and CM, respectively. As a complement, the results for the special case with no correlation is presented here.

Table D.1: Results of PLF for PEM on 3-bus test system without correlation

	V_i (p.u.)		δ_i (deg)	
Bus i	μ	σ	μ	σ
1	1.0500	0.00000	0.0000	0.0000
2	1.0300	0.00000	-2.8239	0.3534
3	1.0321	0.00163	-2.1322	0.1874

Line ij		P_{ij} (MW)		Q_{ij} (MVAr)	
i	j	μ	σ	μ	σ
1	2	22.7722	2.5396	1.7151	0.7091
1	3	70.259	4.8154	9.1575	3.9431
2	3	-9.078	3.9452	18.6747	4.3283

Table D.2: Results of PLF for CM on 3-bus test system without correlation

	V_i (p.u.)		δ_i (deg)	
Bus i	μ	σ	μ	σ
1	1.05	0	0	0
2	1.03	0	-2.743	0.3509
3	1.03174	0.00162	-2.0923	0.1862

Line ij	P_{ij} (MW)	Q_{ij} (MVAr)

i	j	μ	σ	μ	σ
1	2	22.1885	2.5203	1.8701	0.7081
1	3	69.2745	4.7827	10.0677	3.9199
2	3	-8.1713	3.923	17.6634	4.2903

D.2 IEEE 14 bus

PLF results for voltages and power flow in lines from [28] are presented in tables D.3 and D.4, respectively. These values are the basis for comparison of the corresponding values obtained from the case studies in section 4.2.

Table D.3: PLF results for voltages from [28], IEEE 14 bus

	V_i (p.u.)		δ_i (deg)	
Bus i	μ	σ	μ	σ
1	1.06	0	0	0
2	1.045	0	-4.98429	0.44298
3	1.01	0	-12.73054	0.99757
4	1.01714	0.00202	-10.30872	0.68979
5	1.01873	0.00164	-8.76485	0.57883
6	1.07	0	-14.219	0.84952
7	1.06128	0.00286	-13.35621	0.97527
8	1.09	0	-13.35621	0.97527
9	1.05571	0.0051	-14.93501	1.14956
10	1.0508	0.00441	-15.09401	1.09751
11	1.05681	0.00231	-14.78788	0.97113
12	1.05517	0.00069	-15.07309	0.88307
13	1.05035	0.0012	-15.15407	0.90842
14	1.03539	0.00368	-16.03092	1.06123

Table D.4: PLF results for line power flows from [28], IEEE 14 bus

Line ij		P_{ij} (MW)		Q_{ij} (MVAr)	
i	j	μ	σ	μ	σ
1	2	156.9366	13.3943	-18.9334	3.1295

Table D.4 – continued

		P_{ij}		Q_{ij}	
i	j	μ	σ	μ	σ
1	5	75.4682	4.7871	5.502	0.5054
2	3	73.2721	5.7571	4.7525	0.5651
2	4	56.1419	3.3318	-0.4093	0.6566
2	5	41.522	2.4094	2.5914	0.5004
3	4	-23.2535	4.4619	4.5501	2.0654
4	5	-61.0946	4.4898	16.0791	1.4391
4	7	28.0606	3.5716	-9.8291	0.8584
4	9	16.0705	2.0367	-0.4891	0.7895
5	6	44.111	2.6636	12.1028	0.5586
6	11	7.3663	1.4758	3.6053	1.0306
6	12	7.789	0.4155	2.5089	0.1953
6	13	17.7556	1.234	7.2398	0.6805
7	8	0	0	-17.3021	1.674
7	9	28.0607	3.5716	5.7639	2.3431
9	10	5.215	1.5654	4.1753	1.0612
9	14	9.4161	1.2503	3.5818	0.6851
10	11	-3.7978	1.4467	-1.6587	1.0088
12	13	1.6172	0.3685	0.7594	0.1822
13	14	5.654	1.1001	1.7751	0.6725

D.3 IEEE RTS base case

In this section, results are presented for the case studies of the base case IEEE RTS. Refer to section 4.3 for a description of the probabilistic modelling.

The results for voltage magnitudes are provided in table D.5 with all values of μ and σ in p.u. In table D.6 results for voltage angles are provided, all values of μ and σ in degrees. tables D.7 and D.8 present the results for active and reactive power flows, all values given in MW and MVAr, respectively. In all the aforementioned tables, the percentage-wise deviation ϵ between the results obtained from the PEM and the CM are included in the last two columns.

Table D.5: Voltage magnitude results for IEEE RTS base case

Bus i	PEM		CM		ϵ	
	μ	σ	μ	σ	ϵ_μ	ϵ_σ
1	1.035	0	1.035	0	0.000	0.000
2	1.035	0	1.035	0	0.000	0.000
3	0.98932	0.00196	0.98938	0.00196	0.006	0.000
4	0.9979	0.00112	0.99794	0.00112	0.004	0.000
5	1.0185	0.00092	1.01853	0.00092	0.003	0.000
6	1.01234	0.00198	1.0124	0.00198	0.006	0.000
7	1.025	0	1.025	0	0.000	0.000
8	0.99261	0.00138	0.99266	0.00138	0.005	0.000
9	1.00127	0.00114	1.00133	0.00114	0.006	0.000
10	1.0284	0.00137	1.02846	0.00137	0.006	0.000
11	0.98986	0.00069	0.98989	0.0007	0.003	1.449
12	1.00249	0.00092	1.00253	0.00092	0.004	0.000
13	1.02	0	1.02	0	0.000	0.000
14	0.98	0	0.98	0	0.000	0.000
15	1.014	0	1.014	0	0.000	0.000
16	1.017	0	1.017	0	0.000	0.000
17	1.03855	0.00008	1.03855	0.00008	0.000	0.000
18	1.05	0	1.05	0	0.000	0.000
19	1.02324	0.00031	1.02325	0.00031	0.001	0.000
20	1.03849	0.00018	1.03849	0.00018	0.000	0.000
21	1.05	0	1.05	0	0.000	0.000
22	1.05	0	1.05	0	0.000	0.000
23	1.05	0	1.05	0	0.000	0.000
24	0.97783	0.00135	0.97786	0.00135	0.003	0.000

Table D.6: Voltage angle results for IEEE RTS base case

Bus i	PEM		CM		ϵ	
	μ	σ	μ	σ	ϵ_μ	ϵ_σ
1	-7.2827	0.7541	-7.2779	0.7541	0.066	0.000
2	-7.3746	0.7522	-7.3698	0.7521	0.065	0.013
3	-5.5877	0.7904	-5.5838	0.7904	0.070	0.000
4	-9.6942	0.6782	-9.6899	0.6782	0.044	0.000
5	-9.9683	0.69	-9.964	0.69	0.043	0.000
6	-12.4252	0.711	-12.4207	0.7109	0.036	0.014

Table D.6 – continued

Bus i	PEM		CM		ϵ	
	μ	σ	μ	σ	ϵ_μ	ϵ_σ
7	-7.3637	1.0743	-7.3575	1.0743	0.084	0.000
8	-11.0937	0.938	-11.0881	0.938	0.050	0.000
9	-7.4381	0.5773	-7.4349	0.5773	0.043	0.000
10	-9.5063	0.5993	-9.5028	0.5993	0.037	0.000
11	-2.1555	0.3898	-2.1541	0.3898	0.065	0.000
12	-1.5187	0.3151	-1.5175	0.3151	0.079	0.000
13	0	0	0	0	0.000	0.000
14	2.2568	0.6218	2.2584	0.6218	0.071	0.000
15	11.5641	0.8695	11.5658	0.8695	0.015	0.000
16	10.4471	0.7747	10.4487	0.7747	0.015	0.000
17	14.9297	0.8796	14.9313	0.8796	0.011	0.000
18	16.2902	0.9368	16.2919	0.9368	0.010	0.000
19	8.916	0.6586	8.9174	0.6586	0.016	0.000
20	9.5285	0.5332	9.5296	0.5332	0.012	0.000
21	17.1156	0.9099	17.1173	0.9099	0.010	0.000
22	22.7643	0.8981	22.7659	0.898	0.007	0.011
23	10.5712	0.4576	10.5723	0.4576	0.010	0.000
24	5.2972	0.8089	5.2992	0.8089	0.038	0.000

Table D.7: Active power flow results for IEEE RTS base case

P_{ij}		PEM		CM		ϵ	
i	j	μ	σ	μ	σ	ϵ_μ	ϵ_σ
1	2	11.9402	3.9042	11.9399	3.9042	0.003	0.000
1	3	-7.9648	3.0441	-7.9667	3.044	0.024	0.003
1	5	60.0246	3.1059	60.0268	3.1058	0.004	0.003
2	4	38.4351	2.6804	38.4358	2.6804	0.002	0.000
2	6	48.5011	2.493	48.5005	2.4929	0.001	0.004
3	9	22.8935	4.9078	22.898	4.9078	0.020	0.000
3	24	-211.2061	5.2481	-211.2063	5.2481	0.000	0.000
4	9	-36.155	3.2023	-36.1514	3.2023	0.010	0.000
5	10	-11.7184	3.5024	-11.7139	3.5024	0.038	0.000
6	10	-88.595	5.5946	-88.5923	5.5947	0.003	0.002
7	8	115	6.25	115	6.25	0.000	0.000
8	9	-36.9262	5.3095	-36.9233	5.3094	0.008	0.002
8	10	-21.1993	5.4466	-21.1944	5.4466	0.023	0.000

Table D.7 – continued

P_{ij}		PEM		CM		ϵ	
i	j	μ	σ	μ	σ	ϵ_μ	ϵ_σ
9	11	-105.936	4.7567	-105.9186	4.7567	0.016	0.000
9	12	-120.4846	5.4672	-120.4663	5.4672	0.015	0.000
10	11	-151.1934	5.7423	-151.1769	5.7423	0.011	0.000
10	12	-166.7559	6.1753	-166.7391	6.1753	0.010	0.000
11	13	-86.1921	14.1755	-86.1459	14.1755	0.054	0.000
11	14	-171.7625	10.7729	-171.7731	10.7729	0.006	0.000
12	13	-60.5561	11.708	-60.5143	11.708	0.069	0.000
12	23	-227.6969	4.0886	-227.7015	4.0886	0.002	0.000
13	23	-225.2697	9.3335	-225.3026	9.3335	0.015	0.000
14	16	-367.5468	9.8568	-367.551	9.8567	0.001	0.001
15	16	112.296	12.0858	112.3009	12.0858	0.004	0.000
15	21	-214.9188	3.884	-214.9193	3.8839	0.000	0.003
15	21	-214.9188	3.884	-214.9193	3.8839	0.000	0.003
15	24	215.5416	5.4765	215.5378	5.4765	0.002	0.000
16	17	-322.6704	9.3272	-322.6765	9.3272	0.002	0.000
16	19	115.0675	12.176	115.0826	12.176	0.013	0.000
17	18	-186.938	9.2541	-186.9409	9.2541	0.002	0.000
17	22	-139.088	0.3959	-139.0883	0.3959	0.000	0.000
18	21	-60.2885	3.993	-60.2892	3.993	0.001	0.000
18	21	-60.2885	3.993	-60.2892	3.993	0.001	0.000
19	20	-33.1846	6.3611	-33.175	6.3611	0.029	0.000
19	20	-33.1846	6.3611	-33.175	6.3611	0.029	0.000
20	23	-97.2998	6.692	-97.2881	6.6919	0.012	0.001
20	23	-97.2998	6.692	-97.2881	6.6919	0.012	0.001
21	22	-156.4641	0.3999	-156.4638	0.3999	0.000	0.000

Table D.8: Reactive power flow results for IEEE RTS base case

Q_{ij}		PEM		CM		ϵ	
i	j	μ	σ	μ	σ	ϵ_μ	ϵ_σ
1	2	-26.9196	0.7238	-26.9206	0.7238	0.004	0.000
1	3	21.6041	1.086	21.5654	1.0863	0.179	0.028
1	5	4.8755	1.0635	4.8289	1.0642	0.956	0.066
2	4	19.1896	0.8774	19.1509	0.8779	0.202	0.057
2	6	-1.0018	0.8897	-1.0381	0.8901	3.623	0.045
3	9	-16.9874	1.0229	-17.0094	1.0231	0.130	0.020

Table D.8 – continued

P_{ij}		PEM		CM		ϵ	
i	j	μ	σ	μ	σ	ϵ_μ	ϵ_σ
3	24	6.1091	1.5072	6.117	1.5072	0.129	0.000
4	9	5.455	0.9192	5.4277	0.9194	0.500	0.022
5	10	-9.5901	1.1069	-9.6279	1.1073	0.394	0.036
6	10	-130.2706	1.2218	-130.3052	1.222	0.027	0.016
7	8	26.9401	2.542	26.8396	2.5419	0.373	0.004
8	9	3.3929	1.5055	3.3576	1.5055	1.040	0.000
8	10	-17.9705	1.3417	-18.0058	1.3417	0.196	0.000
9	11	-12.776	0.801	-12.7652	0.801	0.085	0.000
9	12	-25.6909	1.0988	-25.6897	1.0987	0.005	0.009
10	11	36.0207	1.0329	36.0287	1.0328	0.022	0.010
10	12	23.1781	1.2746	23.1771	1.2745	0.004	0.008
11	13	-54.9832	2.116	-54.9653	2.116	0.033	0.000
11	14	48.1282	2.5177	48.193	2.5181	0.135	0.016
12	13	-33.3478	1.4787	-33.3016	1.4791	0.139	0.027
12	23	-6.0989	1.438	-6.0666	1.4382	0.530	0.014
13	23	5.1277	3.0118	5.0972	3.0119	0.595	0.003
14	16	-23.7456	2.7362	-23.7658	2.7362	0.085	0.000
15	16	-32.5902	1.2981	-32.6033	1.298	0.040	0.008
15	21	-41.9655	0.888	-41.9691	0.888	0.009	0.000
15	21	-41.9655	0.888	-41.9691	0.888	0.009	0.000
15	24	48.6557	2.5176	48.587	2.5178	0.141	0.008
16	17	-33.8432	2.2634	-33.8591	2.2633	0.047	0.004
16	19	-43.3121	1.619	-43.3542	1.6191	0.097	0.006
17	18	-58.6991	0.7999	-58.6945	0.7998	0.008	0.013
17	22	4.2772	0.077	4.2786	0.077	0.033	0.000
18	21	5.1209	0.5673	5.1191	0.5673	0.035	0.000
18	21	5.1209	0.5673	5.1191	0.5673	0.035	0.000
19	20	-39.3138	0.9706	-39.3188	0.9706	0.013	0.000
19	20	-39.3138	0.9706	-39.3188	0.9706	0.013	0.000
20	23	-44.3555	0.9981	-44.3445	0.9982	0.025	0.010
20	23	-44.3555	0.9981	-44.3445	0.9982	0.025	0.010
21	22	20.1235	0.092	20.1234	0.092	0.000	0.000

E | Matlab scripts

(Restricted public access)