EXTENDING THE HERMAN-BETA TRANSFORM FOR PROBABILISTIC LOAD FLOW ANALYSIS OF RADIAL FEEDERS

by

MUNYARADZI JUSTICE CHIHOTA

(BSc. Eng., MSc. Eng.)

Thesis presented for the degree of Doctor of Philosophy (Ph.D.)

In the Department of Electrical Engineering

University of Cape Town

September 2019

Plagiarism Declaration

Plagiarism is the use of another person's work and pretend that it is yours.

The work reported in this thesis is original apart from what has been properly referenced and cited. The IEEE convention for citation and referencing is used.

SIGNED:

Signed by candidate

Chihota Munyaradzi Justice

Abstract

The increased penetration of distributed generation (DG) mostly derived from renewable energy sources (RES) such as wind and photovoltaic (PV) systems presents two main forms of technical issues. The system is more susceptible to voltage-rise and equipment overload conditions, while the levels of uncertainty associated with the power system's parameters increase significantly. The resulting networks characterised by the bi-directional flow of power, uncertainty from stochastic customer load variations and the intermittent generation from DGs, require appropriate load flow assessment tools.

Classic deterministic load flow (DLF) approaches which assume the input parameters are fixed, or address the associated variability using empirical factors, have been shown to be inadequate. As a result, probabilistic load flow (PLF) approaches based on statistical foundations were proposed to account for the impacts of the uncertainty. For the PLF methods, the efficiency of a specific approach is influenced by two major constraints; the accuracy of the method and the associated computational effort that affects speed. For instance, numerical methods such as the renowned Monte-Carlo simulation (MCS), offer the most accuracy (within the limits of randomness) but are very computationally demanding and undesirable for practical applications. Usually, speed is coupled with loss of accuracy. Most approaches based on analytical and approximation approaches, which offer higher computational speeds, have limited accuracy due to excessive simplifications. Consequently, a trade-off between accuracy and speed is key to a robust PLF approach. Furthermore, the model solutions must be applicable to both small and large systems, consider the dependency between random variables, and avoid complex formulations which limit the practical usefulness.

This research proposes a non-iterative analytical approach referred to as the Herman-Beta extended (HBE) transform to meet the performance and scope requirements of a model PLF solution. The method is based on the beta probability density function (PDF) as a universal descriptor of inputs, and the method of moments for the computation of the output PDFs. The novel formulation of the transform with consideration of complex-type input parameters redresses the network model simplifications of unity power factor loads and resistive feeders in the original HB algorithm (HBA) and the limitation of the representation using absolute values. Further, the effects of dependency between loads and generators are incorporated directly using covariances. The proposed approach opens many possibilities for new applications, including the accurate analysis of the PLF for feeders at any voltage (LV, MV and HV), compensated feeders (shunt reactors and shunt capacitors), and systems with voltage-dependent load or DG.

The performance of the proposed technique is demonstrated using representative test feeders, modified IEEE 33, 34 and 69-bus test systems, as well as practical distribution networks. The results from several test cases demonstrate a good correlation between the LF outcomes from the proposed method and those

from the MCS method and with significant computational advantage. The performance of the method compared with its predecessors shows advanced accuracy while maintaining high computational speed.

Acknowledgements

Above all, I thank God unreservedly for his guidance and blessing on this journey; In him I live, move and have my being.

Second, I am thankful and indebted to Professor Trevor Gaunt, my supervisor and mentor, for his support and guidance throughout the period of the Ph.D. research. The extent to which you have trained me in leadership, critical thinking, and research practice is beyond imagination. Through your supervision and support, your contributions to this thesis are profoundly valuable and have helped to see this work completed.

I am also grateful to the Department of Electrical Engineering at the University of Cape Town and the Trustees of the Dutkiewicz Family for funding this work since 2016. Your unwavering commitment to this work has been encouraging and enabling.

I also appreciate the love, support and companionship of my fiancé throughout this journey. Walking along with you undoubtedly facilitated the journey and encouraged me in difficult times. To my family, the Chots Empire, words could never be sufficient to say thank you. You've been there all along, and you keep supporting me even now. Finally, I am grateful to my sister, Bridget Madzeke, and her husband for your faith, vision and support, which has made my tertiary education in South Africa possible.

Contents

List of Figures

List of Tables

List of Acronyms and Abbreviations

χ^2	Chi-square
1p-2w	Single-phase Two-wire
2p-2w	Two-phase Two-wire
2p-3w	Two-phase Three-wire (Bi-phase system)
3p-3w	Three-phase three-wire
3p-4w	Three-phase four-wire
CDF	Cumulative Density Function
DG	Distributed Generation
DLF	Deterministic Load Flow
E(x)	Expected value of x
ESS	Energy Storage System
EV	Electric Vehicle
FLF	Fuzzy Load Flow
FOSMM	Fist-Order-Second-Moment Method
GMM	Gaussian Mixture Model
HBA	Herman-Beta Algorithm
HBE	Herman-Beta Extended
HV	High Voltage
KS	Kolmogorov Smirnov
LF	Load Flow
LV	Low Voltage
MCS	Monte-Carlo Simulation
MSM	Mixed Skew Model
MV	Medium Voltage
PDF	Probability Distribution Function

PEM	Point Estimate Method
pf	Power factor
PLF	Probabilistic Load Flow
PoC	Point of Connection
PoLF	Possibilistic Load Flow
pu	Per unit
PV	Photovoltaic
PVDG	Photovoltaic Distributed Generation
RES	Renewable Energy Sources
RQ	Research Question
RV	Random variable
SD	Standard Deviation
Trfmr	Transformer
UT	Unscented Transform
Wind-DG	Wind Distributed Generation
$Y \sim Beta(\alpha, \beta, \mathrm{c}, \mathrm{d})$.	Random variable Y follows a beta PDF with shape parameters α and β, and is defined in the range [c, d]

Chapter 1: Introduction

This chapter provides a brief background to the subject of network planning under uncertainty, discussing the strengths and weaknesses of the existing methods, leading to the formulation of the research hypothesis. The anticipated contributions and the outline of the thesis are also presented.

1.1 INTRODUCTION

Load Flow (LF) analysis is a fundamental tool by which power engineers analyse the steady state conditions of critical system variables such as bus voltages, real and reactive power flows, line losses and faults. This information is essential for regulation according to the statutory limits devised to ensure the optimal performance of connected loads and that of the infrastructure. Consequently, LF studies are important tools for power system design, planning and operation.

The suitability of a LF approach, which operates as a transform, depends on how adequately it models the system inputs, how well its network model represents practical systems, and the accuracy of the LF output calculations. In these respects, two key LF methodologies have been applied; deterministic and probabilistic.

Deterministic Load Flow (DLF) approaches – with specified (fixed) values of power generation, loads and network parameters – have been used since the early 1900s but the first digital solutions appeared in the literature in 1956 [1], [2]. However, the stochasticity of the customer loads, uncertainty of the distributed power generation and the application of correction factors to the model parameters suggest the inputs are not fixed. As a result, DLF methods, despite the use of correction factors, are typically not appropriate for power system analysis under uncertainty. The failure to capture explicitly the associated uncertainty makes the results of DLF techniques unrealistic and potentially misleading. Consequently, network engineers who rely on DLF tools are predisposed to sub-optimal network design with potentially serious cost implications: increased overall costs for an over-designed system, and costly maintenance and reinforcement for an under-designed system.

In efforts to achieve more realistic load flow results enabling optimal system design and planning, new techniques based on statistical analysis, termed the Probabilistic Load Flow (PLF), were introduced in the mid-70s [3]. Unlike deterministic approaches, a PLF analysis aims to characterise the uncertainty elements of the power system adequately using samples, cumulants, moments or probability density functions (PDFs). With the inputs determined, the role of a PLF transform is to project the impacts of the input variability to the outputs, such that the full spectrum of the probable system operating states is considered in decision making processes. With more representative solution sets, a planner is well-

equipped to make informed decisions by applying a design risk or confidence limits to the range of possible outcomes. This makes PLF potentially more representative and reliable.

The use of PLF techniques has increased in popularity and application. Recently, PLF techniques have become unavoidable due to the increased penetration of Distributed Generation (DG). The reason with this is that DG elements, which are mostly based on Renewable Energy Sources (RES), complicate planning by addition of substantial level of uncertainty.

Two major constraints determine the usefulness of a specific PLF approach; the accuracy of the method and the associated computational effort which impacts speed. Usually, speed is coupled with loss of accuracy. For instance, numerical methods based on iterated deterministic calculations, such as the well-known Monte-Carlo simulation (MCS) using random number generation and sampling, offer high accuracy but are computationally demanding. A standard MCS analysis requires up to 10,000 scenarios to accomplish acceptable accuracy. The associated computational expense limits the use of the MCS to the validation of other methods although it is also used for calculations in the absence of alternatives. Several remedies for the computational burden offered more effective means to the sampling. While the solutions offer improvements, they do not fully resolve the problems.

The identified inadequacy of numerical methods opened opportunities for other PLF approaches based on analytical, approximate and fuzzy concepts. Approximate approaches, such as the Point estimate method (PEM), can be thought of as 'analytic-numerical' owing to the hybrid approach; the selection of representative samples of the inputs (which is numerical) using devised algorithms to ensure the adequate capture of the statistical properties of the inputs (a concept applied by analytical methods). While the number of trials can be reduced to only two for a network with a single variable, the computational effort increases with the number of variables. Further, the accuracy of the outputs depends on approximation series required to characterise the output solutions as PDFs or cumulative distribution functions (CDFs). The sensitivity of the accuracy of the approach to system size and the inconsistent approximation of outputs using series expansions makes the approach undesirable. To avoid the characterisation through samples, analytical approaches using statistical properties of the inputs such as moments, cumulants or transforms such as the Laplace and Fourier series, were suggested. Generally, analytical approaches offer a great deal of computational efficiency, but the formulations of the algorithms are usually complex. This constraint necessitates assumptions and simplifications that, in most cases, are the reason for the loss of accuracy in the LF results. The applied assumptions are attributed to the modelling of the inputs, the nature of the system, and the characterisation of the outputs. A wide range of simplifications has been applied:

- The modelling of inputs through standard distributions; for instance, the Gaussian model [4], [5] is often selected based on computational convenience.
- The linearization of the load flow equations around selected operation points.

- Disregarding the unbalance characteristic of the connected loads for the ease of a balanced load flow analysis.
- Reducing the scope of considered uncertainty; parameter and model uncertainties are mostly ignored.
- Assuming independence or simplified linear dependencies between input variables.

As demonstrated through practical data [5], [6], the skewness associated with load demand and, far more, the generation profiles from RES, such as wind and solar, are significant and cannot be ignored [6]. For this reason, the adoption of the normal distribution, based on simplicity instead of its fitness to practical data, raises concerns about the accuracy and reliability of the associated PLF results. Where Gaussian mixture models (GMMs) are used to cater for the skewed variability to preserve the computational convenience, complexity arises in two aspects: the selection of the optimal number of components and the extended data requirements for modelling.

On the other hand, the beta PDF is versatile; (1) it is bounded – matching the practical characteristics of load and DG magnitude occurring between the limits of zero and full capacity, such as restricted by a circuit breaker, (2) it is capable of modelling variously skewed data, including symmetrical data and, (3) it requires only two shape parameters (in addition to the two extrema bounds) to fully characterise the distribution of the uncertain variables, making it both convenient and suitable for practical applications. Based on these characteristics, and success on fitness tests such as the Chi-square (χ^2) and Kolmogorov Smirnov (KS) using practical data, the beta PDF is preferred for modelling loads [7], PV [8] and in some cases wind [9]. The application of the beta PDF model in a PLF approach based on moments, termed the Herman-Beta algorithm (HBA), demonstrated high computational efficiency and acceptable accuracy compared with model solutions from the MCS. However, the method was developed specifically for low voltage (LV) feeders and with assumptions limiting the accuracy of the approach and the extent of its applications.

1.2 PROBLEM FORMULATION

The uncertainty of input variables in power systems requires the deployment of statistical tools for voltage-drop computation on distribution feeders. Furthermore, the variability in load and generation profiles needs to be considered explicitly to minimise error in the computation of system operational states. A PLF tool must avoid costly simplifications especially the omission of sensitive variables or conditions such as feeder configuration and electrical properties, and load properties like the operating power factor, unbalance, and correlation. With all these considerations, there is also a constraint of computational time. Since some applications are real-time based, a fast algorithm allows for a quick analysis of the network enabling power system engineers to make timely grid operation and management decisions. In off-line applications, computational strain is also undesirable as it requires

too much computer memory and powerful processors that might be unnecessarily costly. In applications requiring repeated PLF calculations, for instance the assessment of the impacts of varied scenarios of DG penetration, the requirement of speed becomes more apparent. Generally, a non-iterative PLF approach which optimizes computational efficiency is desired. This has been one of the main drivers in the development of new PLF approaches in the past decade.

The original HB algorithm meets most of these specifications (speed and adequate accuracy), only it has some assumptions that limit the scope of its applications. First, the approach was formulated for LV feeders with insignificant line reactance (X/R close to zero) which allowed the assumption of purely resistive feeders. For higher voltage networks, the neglected variables are significant and their omission results in notable error [10]. Secondly, the algorithm was intended for modelling feeder conditions at maximum demand, for which loads tend to be characterised by resistance and could be modelled at unity power factor. Lastly, the HBA assumes all loads are independent of each other. As such, the practical likelihood of loads, DG and phase currents in three-phase loads being affected by a common factor is not considered.

Since the HBA 'transfer function' can only take real-type inputs, there can be two approaches to overcome this limitation. The first, which is relatively easy, is an approximation approach involving the compensation for the imaginary components of the variables using the magnitude of the complex variable as opposed to neglecting the imaginary component. However, this was found to be a limited solution with inconsistent and substantial errors [10], [11]. Alternatively, the limitation can be avoided by the explicit representation of the complex random input variables in the algorithm. This means a complete re-formulation of the network equations in the HB algorithm, and renewed statistical solutions incorporating the interdependence between loads, DG, and phases in the case of balanced loads. Moreover, to remove further limitations, the candidate PLF solution must cater for a wide range of network topologies including the three-phase three-wire (3p-3w) which is common on medium voltage (MV) and high voltage (HV) networks and also include voltage regulation elements such as shunt capacitors which are common on MV and HV networks. Lastly, considering the load can be characterised as constant-current or voltage-dependent (constant-impedance, and constant-power), catering for these models potentially expands the applicability of the approach.

1.3 RESEARCH HYPOTHESIS

The original Herman-Beta algorithm is a proven effectual method for the computation of feeder voltage-drops for radial LV systems of passive (with loads only) or active (feeders with loads and DG) configuration, 3p-4w and bi-phase feeder topologies but with limitations. At present, analysis on the MV network in South Africa remains based on deterministic approaches.

In summary, the original HB algorithm has the following limitations and their removal could lead to wider applications of the approach.

1. The assumption of unity power factor loads is limited to loads with similar behaviour and only applicable during the interval of maximum demand. As a result, the approach is limited to the analysis of passive LV feeders for voltage-drop conditions. In the presence of DG, the intervals of interest include the noon interval in which maximum DG production coincides with relatively low demand. In this interval, the load characteristics may be different. Furthermore, the loads supplied by higher voltage networks have electrical characteristics significantly different from the conditions set through the assumption. Allowing power factor less than unity power factor to be modelled will make the modelling more representative.
2. The assumption of resistive feeders may be generally applicable on low voltage feeders but is not generally applicable on higher voltage feeders where it leads to inaccuracy. Incorporating the reactive impedance in feeders will allow the approach to be extended to higher voltages.
3. The extension to higher voltage networks demands new formulations to cater for the various feeder topologies such as the 3p-3w.
4. The assumption of independency between random variables is only valid when considering customers of the same class. The possible correlation between customer classes, DGs, and for balanced loads, needs to be accounted for. New statistical solutions with this consideration are necessary.
5. The approach only caters for constant-current load models without provision for constant-impedance and constant-power models. Extension of the algorithm to allow voltage-dependent characteristics for some shunt-connected elements would extend the application to include shunt capacitors and higher voltage lines or cables.
6. The results from the study by Chihota et al. (2015) which investigated the possibility of the extension of the HB algorithm to MV using an approximate approach inspire the development of the in-full probabilistic solution for radial feeders [10], [11].

The HB algorithm, if reformulated in light of the identified opportunities, has the potential to satisfy the requirements of high computational efficiency and accuracy without the limitations of voltage level, feeder configuration or load properties. Therefore, the hypothesis of this research is:

> "Reformulating the equations in the existing Herman-Beta transform opens many possibilities for accurate feeder calculations"

From here onwards, the new transform will be referred to as the Herman-Beta extended (HBE) transform, while the predecessor formulations are denoted simply as the Herman-Beta algorithm (HBA) or alternatively, the original HB algorithm.

1.4 RESEARCH QUESTIONS

The following questions are necessary to test the validity of the research hypothesis:

RQ.1 How is the HB algorithm reformulated with complex random variables?

RQ.2 What uncertainty should be modelled, and how?

RQ.3 Is the beta PDF a suitable descriptor of the input parameters?

RQ.4 How accurate are the results of the HBE and what is the associated computational burden?

RQ.5 Apart from constant-current load and DG models, is extension to voltage dependent models possible?

RQ.6 Can the HBE approach account for the dependence between random variables? If not, what is the error associated with such an omission?

RQ.7 How does the reformulated HBE increase the scope of application?

1.5 RESEARCH SCOPE AND LIMITATIONS

The research reported in this thesis concerns the derivation of a new probabilistic load flow algorithm based on the HB approach, hence termed the Herman-Beta extended (HBE) transform, by removing the existing limiting assumptions and potentially extending its application to medium and high voltage systems. The application of the HBE is limited to radially operated feeders; either by the distribution system design or based on operating conditions as with the case of ring networks with normally open switches.

The suitability of the beta PDF for modelling the input variables and the output functions (magnitudes of bus voltages and line currents) will need to be confirmed. Though the thesis reports investigations justifying the use of the beta model as opposed to other renowned models, it does not cover the extraction of load/DG models from actual data.

Since the HBA approach is suitable for load flow analysis on three-phase four-wire (3p-4w) feeders with balanced and unbalanced loads, and without DG (termed passive feeders) and with DG (termed active feeders), it is anticipated the HBE will have similar attributes after the limiting assumptions are removed, but with greater accuracy and wider application. A formulation suitable for three-phase three-wire (3p-3w) feeders will be needed.

The anticipated HBE transform is intended for beta PDF input parameters of constant-current models of loads and DG, and its extension to voltage-dependent models will be investigated.

The test networks and input models used in this thesis are devised with technical considerations suitable for testing the developed tools.

1.6 THESIS STRUCTURE

The rest of the thesis is organised as follows:

Chapter 2 and 3 are based on extensive literature review of the existing PLF approaches and the sources and modelling of uncertainty, respectively. Chapter 4 describes the original Herman-Beta algorithm (HBA) and the underlying theoretical foundations of the approach. In Chapter 5, the validity of the beta PDF as descriptor of inputs is tested.

In Chapter 6 the HBE and its extensions are developed, and rigorous testing is presented in Chapter 7.

Chapter 8 describes a practical application using DG penetration studies as an example. Chapter 9 concludes the thesis with a summary of the answers to the research questions and assessment of the validity of the research hypothesis.

Appendices contain other relevant information that supplements the description of the research.

Chapter 2: Probabilistic Load Flow Techniques

This chapter reviews the existing probabilistic load flow approaches to examine the manner in which the solutions are derived, their strengths and limitations, and the resulting research opportunities. In the review, special attention is given to the performance of the methods in terms of the accuracy compared to model solutions, the associated computational efficiency, and complexity.

2.1 INTRODUCTION AND THE CLASSIFICATION OF APPROACHES

Probabilistic load flow (PLF) techniques were introduced in the mid-1970s by Borkowska [3] with the objective to model input uncertainties and provide a set of load flow results more representative of the range of probable network conditions. The proposed application compared with deterministic formulations demonstrated the advantage of broader, clearly defined assessment of the expected conditions of the system in the presence of various forms of uncertainty. Several applications, mostly from Leita da Silva and Allan [12]–[16], which took after the founding contribution, refined and extended the applications of PLF approaches. As such, probabilistic techniques have increased in popularity and application to decision-making. The developments in the subject area before 1999 [17], up until 2008 [4], and recently [18], [19] detailed through the referenced extensive review papers demonstrate the research activity on the subject area.

In principle, PLF approaches use system inputs defined statistically using properties such as moments, cumulants, probability density functions (PDFs) or cumulative distribution functions (CDFs) to calculate the associated properties of the outputs in a LF problem. Simplified, they operate as transforms propagating the uncertainty characteristic of the input random variables through load flow equations to the output functions.

The PLF can be considered as an approach to solving a multivariate system with a load flow function $y = f(x)$ where x is a vector of random variables $x = \{x_1, x_2, x_3, \dots, x_n\}$ of inputs such as the load demand and DG currents. Assuming the statistical properties of the inputs are known, the objective of a PLF is to determine the statistical characteristics (samples, mean, variance, cumulants, moments, PDF or CDF) of the uncertain output y. The vector y potentially represents bus voltages, line flows, losses, and so forth, depending on the specific requirements of an application. Mathematically, there are several ways in which the PLF objective can be fulfilled. Based on the manner in which the solutions are derived, PLF approaches can be classified into numeric or sampling-based, analytical, approximate and heuristic approaches.

2.2 NUMERIC APPROACHES

Numeric approaches, referred to as simulation techniques in some contexts, are based on the repeated sampling of the input PDFs, or raw sampled data represented as arrays or histograms, to generate numerous scenarios of the input states, and performing deterministic load flow calculations for each scenario. The possible analysis of a very large number of scenarios sufficient to characterise the 'randomness' associated with the LF makes numeric approaches highly accurate. Further, they are not dependent on the complexity of the load flow equations; they are applicable to any system that can be solved through deterministic approaches. The Monte-Carlo Simulation (MCS) is a renowned numerical method which has been used extensively in PLF studies [15], [20]–[28]. The key features of the MCS analysis are random number generation and random sampling. In essence, the MCS creates system operating scenarios through the random selection of samples from the uncertain inputs and computes a deterministic load flow for each sampled scenario, leading to the accumulation of output samples.

Depending on the scope of analysis, the output vector can represent any LF output elements such as voltage magnitude, angle, power flows, line currents and losses. For an MCS with N_t trials, the estimate of the statistical mean and variance of a selected output variable Z can be determined through (2.1) and (2.2) [29].

$$E(\boldsymbol{Z}) = \frac{1}{Nt}\sum_{j=1}^{Nt} Z_j \qquad (2.1)$$

$$Var(\boldsymbol{Z}) = \frac{1}{Nt-1}\sum_{j=1}^{Nt} \left(Z_j - E(Z)\right)^2 \qquad (2.2)$$

where $E(Z)$ and $Var(Z)$ are the expected value and variance of the LF outputs.

The MCS approach boasts conceptual simplicity and flexibility [29]; each trial represents a possible system state, and it is relatively straightforward to include system constraints and various network elements in the analysis. As such, the approach can avoid simplifications of input parameters and network properties which often result in errors. This makes the MCS potentially very accurate and flexible to many applications. However, its accuracy depends on the number of trials conducted which impact the convergence to a solution. Several studies have indicated that simulations higher than 10,000 simulations are sufficient to obtain a good degree of accuracy while limiting the computational burden [30], [31]. Generally, this is applicable to any system without regard to size. However, large networks may contribute to the computational complexity related to the sampling because every node introduces a new load uncertainty that must be represented in combination with all others [32]. Beyond this, further trials ensure convergence with reduced expected error.

Due to the feasible high accuracy, the MCS is often used as a validation tool in the testing of other PLF approaches [33]. However, the requirement of a large number of simulations present a huge computational burden. As a result, the MCS is unattractive for use in real applications as the

computational time is high [15]. Moreover, this issue becomes unbearable in applications like DG penetration studies which require several scenarios of PLF calculations.

The challenge of computational burden in the MCS, stemming from the repeated sampling and the use of the exact network models, stimulated research on alternative sampling approaches and linearization techniques to minimize the computational time. Allan et al. [15] demonstrated that an MCS based on a linearized load flow formulation attains acceptable accuracy with a limited number of trials. Furthermore, combining the MCS with multi-linearized equations, with different linearization points based on the total active system load, reduces the computational burden further [20]. Other approaches have focussed on optimising the sampling in systems with correlated random variables. Approaches such as the Latin hypercube sampling (LHS), with random permutations (LHS-RP) [21] and Cholesky decomposition (LHS-CD) [22], and the Latin supercube sampling (LSS) [24] have offered further computational improvements. The stratified sampling approach characteristic of the LHS offers improved sampling efficiency compared to the simple random sampling (SRS) technique used in the conventional MCS [21]. Jirutitijaroen and Singh [21] demonstrated that the LHS-RP achieves better distribution-tail behavior (upper and lower bounds) than the SRS with the same sample size. The same study also indicates that a discretised LHS-RP performs comparably with a continuous-form LHS but with reduced computational time and storage requirements. However, the accuracy of the LHS-RP is compromised in cases involving a mixture of dependent and independent input variables where the generated samples sometimes project undesired correlations [22]. The application of the LHS combined with the Cholesky decomposition (LHS-CD) solves this problem. Yu et al. [22] proposed the application of the LHS-CD approach to PLF studies and conducted performance comparative studies with other sampling approaches. Their results demonstrated that the LHS-CD required smaller number of simulations than those in the LHS-RP and SRS, to achieve the same measure of accuracy. Nonetheless, both the LHS-CD and the LHS-RP have limited accuracy when applied to systems with dependent variables. An extended LHS approach proposed by Yu and Rosehart [23] provides a more robust and effective sampling approach considering correlated renewable energy inputs. Further developments resulted in the introduction of the Latin Supercube Sampling (LSS) approach which demonstrates improved performance compared to the SRS and all the variants of the LHS [24]. However, the consistency of the method cannot be guaranteed. Cui and Franchetti [25] later proposed the Quasi-Monte Carlo Simulation (QMC) which uses low discrepancy sequence samples to achieve a uniform coverage of the high dimension random variable space. Tests indicate that the method significantly reduces the number of trials required to achieve a credible PLF solution [27].

Recently, a QMC combined with Copula functions is reported to use much fewer samples to achieve the desired accuracy, which helps to reduce the computation more significantly than possible before [27]. Apart from sampling optimisation, other approaches have considered optimisation of the software implementation of the MCS [28]. By exploiting the parallel processing architecture of the graphics

processing unit (GPU) and the use of open computing language (OpenCL), the computational time to a standard MCS was reduced by a factor of 20. Compared to sampling optimisation approaches such as the LHS-CD which improve the computational efficiency by factors up to 300 [26], the use of the OpenCL approach with the SRS offers little improvement. Nonetheless, the OpenCL approach demonstrates that the computing environment can be optimised to better the computational time associated with any load flow tool. Its computational advantage when applied with optimised sampling approaches such as the QMC could be more beneficial but has not been tested.

Despite the several advancements through a wide range of contributions, the MCS approach, is inherently challenged by computational inefficiency due to the numerical sampling such that even in the highly improved form of the solution, the computational burden remains significant. As a result, several classes of approaches avoiding the numerical approach in quest for improved computational efficiency were proposed.

2.3 ANALYTICAL APPROACHES

The formulation of analytical approaches makes use of arithmetic approaches premised on the representation of PDFs using properties such as cumulants, moments, and Fourier components to determine the corresponding properties in the output variables. The chief advantage in this class of approaches is that it limits the high computational expense associated with repeated sampling from PDFs by using only a few parameters carrying sufficient information to characterise the spread of data. However, the derivation of analytical solutions is associated with complexity, making the use of simplifications inevitable.

One of the major challenges in the derivation of analytic solutions is the non-linearity of load flow equations. As a result, the linearization of the LF equations around the expected means of the system's operating states was proposed [15]. Apart from the non-linearity problem, several challenges complicate the derivations of analytical approaches:

- dependence between input random variables
- skewness in non-symmetrical input distributions
- estimation of the output distribution functions, especially when they are not Gaussian
- variability in network or model parameters

In order to allow derivations with lower computational costs, simplifications (related to the identified challenges) are applied at the cost of accuracy. Therefore, the performance of analytical approaches is a compromise between accuracy and speed. Several methods have been proposed and will be discussed in accordance with their mathematical methodologies.

2.3.1 Convolution

In efforts to reduce the computational burden associated with numeric approaches, Allan et al. [14] proposed a PLF method based on the principles of convolution. By applying the additive and multiplicative properties of the Laplace transform, the sum and products of RVs in the LF equations are calculated leading to the construction of output density functions. While the method demonstrated improved computational efficiency which achieving acceptable accuracy, the convolution techniques applied are inherently complex. In another study, Allan et al. [34] proposed a discrete frequency-domain convolution technique based on the Fast Fourier Transform (FFT), instead of the Laplace. The comparison between the two methods for execution time and accuracy on common test networks predominantly demonstrated superior performance in the FFT solution. At this point, none of the solutions had dealt with dependency until Allan et al. [16] proposed an approach combining convolution and the MCS techniques. The approach, unlike its predecessors, could include in the calculation the impacts of economic dispatch, load shedding, and re-dispatch. However, only economic dispatch was found to have significant impacts on the PLF outcomes. An extension to include network contingency effects of uncertain network elements, thereby removing the assumption of deterministic network parameters, was later reported by da Silva et al. [12]. The method is relevant in the quantification of adequacy indices and certainly gives more representative results. However, it results in a higher computational burden than the conventional approach. In general, while convolution techniques demonstrate an improved computational efficiency compared to the MCS, they remain computationally inefficient while storage problems, especially with large systems, present great challenges [35]. Moreover, the representation of input dependence is challenging and further complicate the approach as high dimensional dependent discrete convolution (HD-DDC) approaches are required to solve the multi-variate dependent systems [36].

2.3.2 Cumulants

To avoid the computational complexity and the storage issues associated with the convolution of PDFs to derive output CDFs, cumulants were introduced [37]. Using the approach, the inputs are characterised using cumulants, and the corresponding output parameters are determined by taking advantage of the arithmetic properties of cumulants. Then, by using expansion series such as the Edgeworth (EW), Gram-Charlier (GC) and Cornish-Fisher (CF), the CDFs and PDFs of the outputs can be estimated without the need of complicated convolution computation.

The founding article published by Sanabria and Dillon [38] paved the way for cumulative principles to be exploited to solve the probabilistic current flow, taking into account continuous RVs and discrete RVs through an extension based on the Von Mises function. The approach, which was tested against existing solutions such as the FFT method, showed significant computational advantages including the reduction of storage requirements. The same work demonstrated the limitations of the Gram-Charlier (GC) expansion in the estimation of the output distributions, especially when the non-linear LF

equations are not simplified. However, Zhang and Lee [39], having linearized the equations and hence anticipating unimodal output distributions, based their work on the GC expansion series. While the accuracy of the method compared to the MCS is appreciable, it is not clear what order in the GC expansion is optimal for both accuracy and speed. In the reported study, it is demonstrated that higher expansion orders, up to the 7^{th}, provide better estimates to the tail behaviour of the output distributions. Beyond this, increasing the expansion order does not improve results but increases errors. Based on this limitation, together with the likely convergence issues especially in cases in involving discrete RVs, Hu and Wang [40] proposed a modified approach combining the concept of cumulants with convolution techniques to address systems with Gaussian (continuous) and discrete distributed uncertainties. In the unified approach, the Von Mises technique is applied to the discrete random variables to generate a few impulses characterising the RVs and then convoluting them with the continuous RVs. The results indicate improved accuracy and overcomes the inconsistencies associated with the application of the conventional GC expansion to such cases. While the approach provides means to the analysis of discrete random variables, it assumes that the continuous random variables follow Gaussian distributions. In cases of non-Gaussian continuous profiles, where high variability exists in the load or DG data, the GC expansion experiences serious convergence issues. Usaola [41] identifies the problem and proposed a combined cumulants Cornish-Fisher (CCCF) approach. Compared to the cumulants approach based on the GC expansion series, the CCCF is superior in approximating the output distributions for systems with non-Gaussian inputs, while maintaining the computational efficiency. In subsequent developments, the CCCF approach was employed to cater for dependent continuous and discrete random variables in a method termed the Enhanced Linear Method (ELM) [42]. This approach also provides a better approximation of outputs by calculating a few convolutions in the output stage, which enables a more accurate approximation for multimodal output distributions, the contribution will allow the analysis of systems with dependent, continuous, discrete, and multi-nodal output distributions. Up to this point, the CM formulations were based on a dc load flow model, Yuan et al. [43] then provided a new tool based on an ac model and using the GC expansion series. The authors claim a generalized improved accuracy with increasing number of terms in the GC expansion. While this can be true for the orders tested (up to the 6^{th} term), as prior studies with more extensive tests indicated that the expansion series lacks consistency beyond the 6^{th} order [39]. In addition, the use of the GC series with skewed inputs is prone to errors since the approximation technique works optimally with symmetrical inputs. Fan et al. [44] conducted comparative studies on the performance of the Cumulants Method combined with the EW, GC and CF approximations. Their findings were as follows:

- Increasing the order of truncation of the expansion series does not guarantee accuracy or convergence (across all the three series).
- There is no guarantee of convergence with any of the approaches.

- The CF and EW series have poor performances at the extremities (or tails) of a CDF (probabilities close to zero and those close to one).
- The CF and the EW are generally better than the GC but demonstrate poor accuracy when estimating the distribution tails.

As part of the study, the use of joint cumulants to address the interdependency between input variables was addressed. Notwithstanding the improvement in accuracy, computational burden is increased. Subsequent developments have attempted to address the issues of accuracy [30], computational efficiency [45], and convergence [30], but with limited success.

Le et al. [45] suggested a new way to combine the GC-series with the Von Mises technique to cater for both continuous and discrete random variables resulting in an improved computational efficiency compared to the method earlier suggested by Hu and Wang [40]. In another study [30], it was shown that the CCGC combined with the multiple integrated method (MIM), when applied to systems with high uncertainty coefficients (tested by increased PV penetration), and with varying input correlation levels, improves computational efficiency and accuracy.

In general, the performance of the combined cumulants with expansion series is optimal in cases where the outputs can be estimated by Gaussian distributions [45]. In cases where this is not valid, convolution is considered, but with an increase in computational effort. Furthermore, the issue of convergence is not completely addressed in any of the developments. The approach is therefore susceptible to failure under unfavourable operating conditions. Moreover, the selection of an expansion series and the level of truncation necessary for optimal results is not straightforward; depends on the application, sensitivity parameters in the LF, and the nature of the anticipated output distributions.

Driven by the demonstrated challenges associated with the approximation of the output density functions using expansion series such as the GC, CF and EW, recent research has been focused on alternative methods to estimate the output functions. Methods such as the Maximum Entropy (ME) [46] and the La Place Transform (LT) [47] have been suggested. The ME, compared with the GC, achieves improved computational efficiency and more accurate estimates of the output functions [46]. Another method, termed the combined cumulants and Laplace Transform (CCLT) [47], makes use of the Laplace transform to derive the output functions. The achieved results demonstrate a slight improvement in the computational speed while the gain in accuracy is insignificant.

2.3.3 First Order Second Moments

Where input variability does not show significant skew and the number of random variables along with the characteristics of the sensitive coefficients dictating the behaviour of the system is such that the outputs can be approximated by a Gaussian PDF based on the CLT, a PLF can be solved based on two parameters; the mean and variance. The First-Order-Second-Moment method (FOSMM) is such a

method. The approach is based on solving the PLF using the first (mean) and second (variance) central moments of the inputs to determine the corresponding characteristics of the output functions. Where dependent variables exist, the covariance properties of the input variables are directly incorporated into the load flow equations without the need of any transformations. To minimize computational expense, the load flow equations are linearized around the mean using a truncated Taylor series up to the first order. Summarized, the approach utilizes the first order terms of the Taylor's expansion for linearization, and the first two moments to characterise the output functions, hence the nomenclature 'First-Order-Second-Moment Method'.

In the first application to power systems, Madrigal et al. [48] applied a combination of approaches to compute the LF considering interdependencies and objective functions for optimal power flow computation. The formulation combines the standard FOSMM approach to solve a dependent load flow system, while the Newton's method and Lagrangian functions are used to handle the inequality conditions for optimal power flow assessment. Conducted validity tests demonstrate immense computational efficiency with good accuracy when compared with the conventional MCS. Nonetheless, the use of the mean and second-order central moment limits its application to Gaussian-distributed random variables. In the application reported by Wan et al. [49], claimed to be the first proposal of the FOSMM for PLF analysis, the method appears similar to the one reported by Madrigal et al. [48] save for the differences in application; a standard load flow in Wan et al. [49], and an optimal power flow in Madrigal et al. [48]. The method tested against the MCS with 10,000 iterations is 180 to 190 times faster and yields comparable results.

There are two other moment-based approaches [50], [51]. The first one was introduced in Herman [50] and is based on Beta PDF and is known as the Herman-Beta algorithm (HBA). The approach uses the inputs ' first two raw moments to calculate the respective output parameters. Based on the versatility of the beta PDF, the approach is capable of capturing different input distribution shapes and estimating different output distribution shapes as well. Chapter 4 will discuss the HBA approach in detail.

The second moment-based approach is reported by Celli et al. [51] and is more similar to the FOSMM because it is based on mean and variance and assumes the symmetrical distribution of the inputs and outputs. Besides this assumption that limits its application, the approach demonstrates high computational efficiency.

In general moment-based approaches demonstrate the computational convenience of using a single standard distribution as the input model to a linearized PLF; the PLF can be computed in a single-pass, the correlation between input RVs can be directly incorporated and, the output distribution functions are easily determined without the need for transformation functions. However, the neglect for skewness disrupts the advantage.

2.4 APPROXIMATE APPROACHES

Approximate techniques are based on the analytical selection of sample points from input distributions, the analysis of the PLF by analytical means for each selected set of points, and the approximation of the output distributions using expansion series. Therefore, they are similar to numeric approaches on the basis they use selected points as representation of the inputs, yet they differ in the manner in which the sampling is done, and the means to calculate the PLF. The heart of these approaches is in generating the optimal number of samples that maintain sufficient information about the input variables' statistical characteristics. The discussion of the proposed approaches will give more attention to the widely applied methods such as the Point Estimate Method (PEM) and the Unscented Transform (UT) method.

2.4.1 Point Estimate Method

The Point Estimate Method (PEM) is the most applied approximate approach. The approach employs the selection of inputs samples with assigned weights or concentrations to determine the output moments of output variables. A number of approaches [52]–[63] employing different numbers of sample have been proposed and are referred to as PEM schemes. Depending on the scheme, the mean, variance, skewness, and kurtosis for the outputs can be calculated.

Mathematical Background for the PEM schemes

The foundations of the PEM are based on the statistical work by Rosenblueth [52], Harr [53], Li [56], and Hong [57]. The first PEM scheme introduced by Rosenblueth [52] and later revised by the same author in 1981 [58] utilises 2^m representative points (where m is the number of random variables in a system), and, can handle skewed, as well as correlated random variables. While the required representative points appear small, the fact that the computational effort increases exponentially with the number of RVs makes the application only beneficial for small systems. For large systems, for instance $m = 20$, Rosenblueth's PEM requires ten times more samples than an MCS with 100,000 trials. The subsequent advancement by Harr [53] introduced a $2m$ PEM scheme which substantially reduced the computational burden, especially for large systems. However, the problem with this approach is that it can only handle symmetrical random variables, as it is limited to modelling the variance of the inputs. The development by Li [56] successfully caters for skewed and correlated random variables. However, the improvement is coupled with the loss of computational efficiency; the method requires $0.5(m^2 + 3m + 2)$ simulations. As a result, Hong [57] introduced the $mn + k$ scheme (where n and k are integers) to address the limitation of symmetry and speed together. The simplest case of the scheme uses $2m$ concentrations and considers the skewness and correlation of the random variables while the extra concentration in the $2m + 1$ approach considers kurtosis. The PEM has been applied in several PLF applications.

Application of PEM schemes in PLF studies

The first application of the PEM to power system analysis was reported by Su [59]. The proposed method is based on Hong's *2m* scheme and formulated from non-linear load flow equations considering skewed distributions and correlated inputs. Compared to the MCS, the approach performs with great speed but with limited accuracy. In order to improve the performance, the author recommended higher PEM schemes. Generally for *2m* scheme, the accuracy in the solutions is sensitive to system size, performing unsatisfactorily with a high number of random variables [60]. This is because as the random variables increase, the selected concentration points may fall at points where the probability of the distribution is not well-known or worse outside the permissible range of the PDF. To make suitable the application of the PEM to networks involving a large number of RVs, Morales and Pérez-Ruiz [60] proposed a PLF method based on another variant of Hong's PEM, the *2m+1* scheme. The approach solves the issue of outlier point by implementing a different selection algorithm which is independent of the total number of RVs in the system. Furthermore, the additional order in the *2m + 1* scheme allows the consideration of kurtosis thus providing more information for the construction of the output functions. Unfortunately, the processes necessary to construct the output functions, given the advanced information, are complicated. As with the cumulants method, expansion series such as the Gram-Charlier have to be applied to derive the distributions of the outputs [60].

Higher orders of the PEM such as the *4m+1* have been proposed with the anticipation of achieving enhanced performances. However, comparative tests between the *2m*, *2m+1* and *4m+1* scheme conducted on the IEEE 14-bus and 118-bus feeders, and using the MCS as a benchmark, indicate the *2m+1* scheme delivers the best overall performance (applying a trade-off between computational efficiency and accuracy) with both discrete and continuous inputs. As anticipated based on the higher orders of moments used, the *4m+1* scheme achieved better accuracy, but inconsistently and also at additional computational expense.

Caramia et al. [61] extended the application of the PEM to unbalanced three-phase systems with dependent inputs. In the study, three Hong's PEM schemes (*2m*, *2m+1* and *4m+1*) with a modified approach to the modelling of dependency, are studied. The approach highlights the errors associated with the application of the conventional PEM approach to systems with high levels of interdependency between inputs. To address the identified limitation, a rotational transformational matrix based on eigenvectors of the correlation matrix was proposed. The published results demonstrate the new approach achieves significant improvement in accuracy, while the comparative studies between Hong's schemes revealed the 2m + 1 scheme to have the best performance. However, the application only considers Gaussian correlated inputs. A subsequent development [64] applied the Third-Order Polynomial Normal Transformation (TPNT) to solve systems with other PDFs (Yang and Zou, 2012). In other developments dealing with the same issue of correlation, the Cholesky decomposition [62], [63] and the Nataf transformation [54] have been suggested.

Overall, the developments in the PEM are well appreciated as the computational burden is hugely reduced compared to numerical approaches such as the MCS. The shortcomings of the PEM are mainly linked to accuracy and the trade-off with computational effort [60]; PEM schemes with a higher number of estimating points can offer improved accuracy but then, the computational time increases. A performance comparison between the 3m, 5m and 7m PEM schemes conducted by Gupta & Daratha [63] demonstrated that significant computational expense is incurred with higher schemes. For instance, increasing the order of representative points from 3m to 7m nearly doubles the computational time. Furthermore, convergence issues and invalid outcomes are possible; the chance for flawed, non-real outcomes (negative CDFs or infinite sets of solutions due to the failure of convergence) increases as estimation points may fall outside of the regions in which distributions are defined [55], [60]. Mohammadi et al. [55] attempted to address the related issue using the Rosenblatt transformation to extend the domain of the random variables to the infinity range. However, the approach's use of joint PDFs to model the inputs, which must be known a priori, is not always practical [54].

In addition, the use of CDF/PDF approximation series such as the CF, GC and EW brings along the inherent accuracy limitations in the approximation of the output functions [41], [44]. Furthermore, the accuracy in the PEM is sensitive to the number of input random variables and the levels of correlation between them [62]. This is because as the number of RVs increases, the concentrations of the estimate points are bounded to their limits, resulting in error.

Several other approaches [65]–[73] employing different solutions to the selection of estimate points, the handling of correlation, and the estimation of the output distributions have been suggested and are discussed further.

2.4.2 Generalized Polynomial Chaos

A generalized polynomial chaos (gPC) is a sum of a system of orthogonal polynomials of random variables approximating a random variable function [74]. The orthogonal polynomials used in the approximation depend on the type of random variables in the problem. For example, for Gaussian type, the Hermite polynomial is recommended, while for beta-distributed variables, a Jacobi is more appropriate. Using the polynomial functions with known means and variances, the corresponding properties for the outputs can be calculated by applying linear properties. While the approach achieves comparable results to the MCS, it suffers computational burden with large systems [74]. The multi-variate form of the gPC approach with *n* random variables and using a *p-th* order polynomial requires up to $(n + p)! / (n! \cdot p!)$ deterministic load flow calculations. For instance, a 3rd order gPC analysis for a system with 50 prosumers all with installed PVs (n = 50 + 50) requires up to 176,851 DLFs. For such system scale, the gPC is more demanding than the conventional MCS with 10,000 trials. In another application of the gPC [65], a Galerkin projection is used to compute the polynomial coefficients in a single run, as opposed to other approaches which depend on iterations. However, the polynomial

coefficients for each random variable still need to be calculated, which leads to higher-dimensional problems. This problem is only remedied is cases where the grouped customers share common load/DG models. Apart from this special case, the computational burden increases significantly with differentiated customer load or DG models. Further, the method performs poorly at the tails due to polynomial expansion truncation errors.

2.4.3 Stochastic Response Surface Method (SRSM)

Ren et al. [66] having highlighted the limited performance of the PEM as a result of the approximation at the output stage proposed a PLF based on the stochastic response surface method (SRSM) and polynomial chaos expansions to solve the PLF. The basic idea of the SRSM is to model the standard input PDFs using inverse quantile functions, determine the power flow response of the system using a Hermite polynomial chaos expansion, compute the unknown polynomial expansion coefficients, and approximate the output PDFs using a kernel density estimation method. Similar to the methodology in the PEM, the solution of the unknown coefficients is derived using a selection of representative points, termed collation points, used in repeated deterministic load flow calculations.

The performance of the proposed method tested against the MCS demonstrates a good level of accuracy. Compared to the PEM, the SRSM generally achieves better accuracies at the tails, and relieves the issue of impermissible probability outcomes from the use of expansion series such as the GC, CF and the EW. However, in cases involving dependence, the advantage over accuracy is lost, especially on predicting the tails. Withal, the SRSM incurs a higher computational burden compared to the PEM schemes. Also, similar to the PEM, computational efficiency declines with system scale. Hence the authors recommend the method only for small to medium scale systems.

2.4.4 Unscented Transformation Method

The unscented transformation method (UT) is another approximate PLF approach that has recently received attention [67]–[70].

In principle, the UT method is a transform capable of mapping the mean and covariances of a given set of inputs to the outputs through a non-linear function. It's capability to deal with non-linear functions makes it suitable to solve the non-linear LF equations without simplifications. Further, the direct transformation of covariances is computationally convenient. It avoids the computational expense and complexity associated with the indirect handling of interdependence through transforms such as the Cholesky Decomposition and the Nataf transformation, especially considering the non-linear space.

Theoretically, the UT is similar in principle to the PEM in that it uses a set of selected points to capture the probabilistic information of the inputs. However, in the UT, the input random variables are discretised and then approximated using a set of carefully selected points called sigma points. As a result, the performance (accuracy and computational efficiency) of the method depends on the manner

in which the sigma points are selected. The differentiated approaches to the selection of the sigma points underpin the formulations of several UT schemes reported in literature.

In the first application reported by Oke et al. [67], the selection of the representative points and the corresponding weights is handled using a Gaussian Quadrature (GQ) and Stieltjes procedure. With the selected sigma points and weights, a recursive Newton-Raphson (NR) LF is computed leading to the calculation of the expected value, normalized cumulants and moments of the outputs. With these, the output distribution is estimated using expansion series. As with the demerits of the PEM, the UT's accuracy is sensitive to the number of sample points used, which improves the estimates of the higher order moments of the outputs. However, increasing the sample points attracts significant computational cost and complexity; the method, referred to as the simple UT, uses $(N + 1)^m$ representative points (where N is the order of the approximation scheme and m is the total number of random variables in the system). In this first application, the reported tests were limited to a 3-bus system, and interdependence was not addressed. The MCS validated the approach for accuracy, but it is difficult to determine the advantage of the approach due to the limited tests carried out. Moreover, due to the exponential relationship between the number of RVs and the required representative points for PLF calculation, the computational burden is expected to rise dramatically with system size.

In another application of the UT [68], a different sample point selection method requiring only $2m + 1$ points is proposed. The approach employs a modified technique for the calculation of the weights; the weights can be negative, as long as the sum of the weights for that system equals unity. As a result, the approach avoids the convergence issues and errors resulting from out-of-range selection of sample points in the PEM and simple UT approaches, where the weights are confined to the closed interval (0, 1). The advancements of this approach are the direct handling of correlation using co-variance matrices and the avoidance of errors associated with the location of sampled points. Compared with the $2m$ PEM, the approach has a slightly higher computational burden but is significantly less accurate. The only advantage over the $2m$ PEM is in the handling of correlation. However, the comparison with the $2m$ scheme is misplaced and potentially misleading, the method should have been compared to the $2m + 1$ scheme which uses the same number of representative points.

Baghaee et al. [69] introduced a $2m + 1$ UT-based PLF method exploiting the characteristics of radial basis function neural networks (RBFNN) for non-linear mapping, to solve exactly the non-linear power flow equations. Using the approach, the need to calculate the partial derivatives and inverse Jacobian matrix in the power flows is avoided, equipping the algorithm with computational advantage. Further, a direct computation of the output means and covariances is possible using the RBFNN as a transform mapping the inputs to the outputs. This equips the approach with the advantage of accuracy and computational efficiency as approximate transforms and expansion series are avoided. Based on its capability to address systems with non-linear equations, the method is applicable to meshed networks. Conducted tests show better performance compared with other analytical approaches ($2m$, $2m + 1$ PEM;

the conventional $2m + 1$ UT) demonstrates. Notwithstanding the enhanced accuracy and convenience for non-linear LF calculations, the use of neural networks, and the associated training, are complex, and the accuracy of the method depends on the effective design of the neural network on which the input-output transformations depends.

Canon and Jafarzadeh [70] carried out a comparative analysis on the performances of the general ($2n + 1$), basic ($2n$), simplex ($n + 1$), and spherical ($n + 2$) schemes. Their results, summarised in Table 2.1, provide more detail to the performance of the UT schemes.

The analysis demonstrated that the spherical and general schemes are superior in accuracy and computational efficiency over both the Simplex and Basic UT schemes. The Simplex has the worst performance, with inconsistent computation time and mediocre accuracy.

Table 2.1: Various schemes of the Unscented Transform method

PERFORMANCE INDICATORS	**UT SCHEMES**			
	General	***Basic***	***Simplex***	***Spherical***
Sample Points	$2n + 1$	$2n$	$n + 1$	$n + 2$
Computational Effort	High	Average	Low	Low
Accuracy	High	Average	Low	High[1] Low[2]

The diversity of tests performed also demonstrate that the performance of the UT schemes is dependent on system size and the weighting approach. Solutions such as the scaling of weighting parameters by system size and the careful selection of weighting parameters are believed to improve the approach [70].

2.4.5 Taguchi

Another approximate PLF approach which approaches the sampling of representative points by using Taguchi orthogonal arrays was recently proposed [71]. The method utilizes a few deterministic load flow calculations to estimate the mean and standard deviation of the outputs. The number of the required scenarios for a given number of random variables is prescribed from the properties of the Taguchi arrays. As an example, for a 118-bus feeder with 105 statistical loads in service, the Taguchi array $L_{128}2^{127}$which can handle up to 127 RVs prescribes 128 load flow calculations. For the same system, a 2m-PEM requires 210 calculations. Furthermore, tested against the MCS, the Taguchi-based approach achieves better accuracies than the 2m-PEM. However, the computational expense increases rapidly with the number of random variables to extents where the MCS with limited trials (e.g. 10,000) is less

[1] For small systems
[2] For large systems

computationally demanding. Furthermore, the approach for the selection of orthogonal arrays is not clear when dealing with skewed random variables.

2.4.6 Others

Gupta [72] suggested a Gauss Quadrature (GQ) based PLF methodology. Similar to the UT approach, the GQ method sis used to determine the location and concentrations of sample points, which are of the order *3m* in this approach. The work was aimed at improving the performance issues, particularly sensitive accuracy, characteristic of most approximate techniques, and demonstrating a hybrid input application involving PV, wind and EVs. Further, the approach allows the characterisation of voltage-dependent load models. Compared with a *7m*-PEM, a *3m*-GQ scheme offers more accuracy especially in cases where the load demonstrates dominant multimodality. However, the computational time is reported to be approximately only four times faster the MCS, while no information is provided for the corresponding comparison with the *7m*-PEM.

Recently, a new PLF approach employing the mean-value-first-order saddle-point approximation (MVFOSPA) technique was proposed by Mohammadi et al. [73]. The method requires only $m + 1$ simulations, similar to the UT's Simplex scheme. Its performance compared with the $2m$-PEM and $2m+1$ UT schemes, on the basis of the MCS demonstrates superiority in both computational efficiency and accuracy.

2.5 HEURISTIC METHODS

Heuristic approaches have also been used to solve the uncertain load flow problem. Methods such as those based on fuzzy logic[3], Genetic algorithms (GA) and the imperialist competitive algorithm (ICA) have been reported.

2.5.1 Fuzzy Logic

While the advancement in the field of probabilistic analysis of power systems is well appreciated and gained popularity, some scholars argued that the PLF techniques were limited to cases where statistical historical information or the variability of the input variables were fully known [75]–[77]. As a result, a new class of approaches for solving the uncertain load flow, termed Possibilistic load flow (PoLF), was proposed to deal with cases where the variability of the inputs was vaguely known and could not be defined using probabilistic functions except through concepts of 'limitations' and 'restrictions.'

Using 'linguistically' defined expected operating conditions, boundaries derived from expert knowledge or the available limited data sets, fuzzy numbers (FNs) are used to depict 'possibilistic' models of the uncertain random variables. The computation of a fuzzy load flow (FLF), based on these

[3] While the approaches based on fuzzy logic fall outside the scope for probabilistic approaches, they are considered here due to the attention given in the literature, and the close relationship between the concepts of probability and possibility.

foundations, was first reported by Miranda and Matos [75]. In the FLF, fuzzy numbers are used to compute the power flow problem in a deterministic routine based on an incremental DC power flow approach [78]. The outputs of the computation are fuzzy numbers which can take a variety of shapes, but usually the trapezoidal and triangular representations for simplicity. Since its first application, several formulations of the FLF theory has been reported in the existing literature [76]–[86]. The developments provided extension of the original approach in the following ways; the inclusion of both uncertain loads and DG in the FLF studies [76], generator uncertainty optimization [78], AC power flow [77] and synchronous generator uncertainties [80], transmission network analysis [81], the conversion of actual wind power outputs PDFs into fuzzy possibilistic models [82], and the inclusion of dependency [83].

With the advancement of metering applications, the statistical data required for probabilistic models has been increasingly available. In addition, it is highly unlikely to have a system where all the inputs are not fully known. Usually, the practical data necessary for the characterisation of loads, DG, and EVs is commonly found. However, model parameters, based on practical values, for variables such as resistance, reactance, and shunt capacitance (for both line models and voltage regulation devices) are vaguely known. Consequently, the consideration of hybrid solutions incorporating both probabilistic and possibilistic techniques were proposed [79], [86]. Soroudi and Ehsan [86] proposed a combined approach involving the probabilistic representation of wind DG using Weibull PDFs and fuzzy sets for the load demand. While the concept demonstrated is worthwhile, the approach by Aien et al. [84] is more practical as the selection of the models reflects the availability of practical data; the approach represents loads, wind, and PV power probabilistically while electric vehicles and DG based on gas turbines are modelled possibilistically.

FLF approaches have limited applications in power system analysis as a result of a couple of drawbacks. Firstly, the methods do not demonstrate adequate accuracy and computational efficiency compared to model solutions. Attempts to address these challenges through a linearized systemic fuzzy load flow (LSFLF) improves the performance and provides the possibility of application to large systems but still remains inadequate [85]. Secondly, with the increasing availability of data, a FLF becomes unimportant. Thirdly, the application of the hybrid possibilistic-probabilistic approach, which is potentially useful, has not been fully developed to overcome the issues of inaccuracy and computational inefficiency.

2.5.2 Imperialist Competitive Algorithm

Nikmehr and Ravadanegh [87] proposed another heuristic approach based on the Imperialist Competitive Algorithm (ICA). The advantage of the approach is that it can be applied to weakly meshed networks. However, the approach is computationally inefficient; in the tests conducted, at least 100 iterations for each of 5,000 scenarios of input samples were required to converge to a solution.

Moreover, the computational expense is expected to rise with network size. This makes the approach undesirable and not suitable for large power systems.

2.5.3 Particle Swarm Optimization

Recently, an application using Particle Swarm Optimization (PSO) algorithm for probabilistic load flow calculation was proposed [88]. In this method, samples from random variables are clustered into several groups based on similarities, then a cluster representative point (cluster centre) is determined using the PSO algorithm, followed by the calculation of the deterministic load flow based on the scenarios of representative cluster centres. From the publication, it is not clear how the selection (prior clustering) of samples from density functions is done, but it seems a random selection is applied. Also, there is no information on how the optimal number of clusters is selected, rather a fixed number of five clusters is selected. From the five DLFs calculated, five output quantities are achieved. Then, the output distribution function is estimated by sorting the output values by magnitude and assigning each value a probability of occurrence in increments of 0.2. The result is a stair-case CDF. As expected, based on the crude approximation of the CDF, the error in the variance is very significant, up to 50% in one of the tested cases. While the authors claim an overall good accuracy based on the results for the means, the error in the spread of the distribution impacts the percentile values (extracted at selected design risk) which are more useful to the planner for decision making.

This concludes the review of the formulation of probabilistic load flow solutions. The next section discusses the scope and interpretation of probabilistic outputs.

2.6 VARIABLES OF INTEREST AND COMPLIANCE CRITERIA

The outputs from a PLF analysis results in probabilistic outcomes with characteristic distributions for the selected variables of interest, depending on a specific application. For network design, particularly feeder design, PLF approaches must determine the following:

- The probability of nodal voltages being outside acceptable levels
- The probability of exceeding the thermal rating of conductors

In the presence of active components such as DG and ESS (including EVs discharging into the grid), two voltage extremes must be considered; voltage-rise, and voltage-drop. Many countries, including South Africa, allow a $\pm$ 10% deviation on LV bus voltages and $\pm$ 5% for MV networks [89]. On the other hand, some countries, such as New Zealand, comply with the IEC 60038 voltage regulation standard [90] which specifies a more stringent upper limit of +6% for LV (the lower limit being -10%). The South African grid code classifies these regulation limits for voltage as compatibility levels.

Given the variability in the load flow inputs, a network planner cannot guarantee that the statutory limits will not be violated. Rather, with the application of a design risk or planning confidence level, the

uncertainty can be carefully managed. Regarding the statistical compliance for voltage regulation, the latest South African national electricity supply guideline [89] states:

> "The highest and lowest of the assessed 95% weekly values over the full measurement period shall not be outside the compatibility levels given in Table 1, or as otherwise contracted (see note 1);"

In the quote, the reference "Table 1" refers to the aforementioned compatibility levels on voltage variations and "note 1" caters for cases involving interruptions. The meaning of '95% weekly values' can be interpreted statistically as 95% confidence or conversely, a design risk of only up to 5%; which translates to 2.5% per limit for variables regulated on both tails of the distribution [91]. While this means there is a 5% allowance for values to violate compatibility levels, extreme limits, at an additional 5% tolerance, are enforced. These are termed 'limits' according to the nomenclature of the South African grid code.

While the regulation of voltage is explicitly enforced, the code does not specify the thermal limits for distribution conductors nor those for transformers. As a result, it is common to assume the rated capacity as the compatibility level [92] and applying the same design risk as for voltage regulation, except that it is single tailed (without the likelihood of underloading) in this case.

The plot of Figure 2.1 illustrates the concept of permissible limits and the applied design risk for a parameter such as voltage magnitude which has both upper and lower limits.

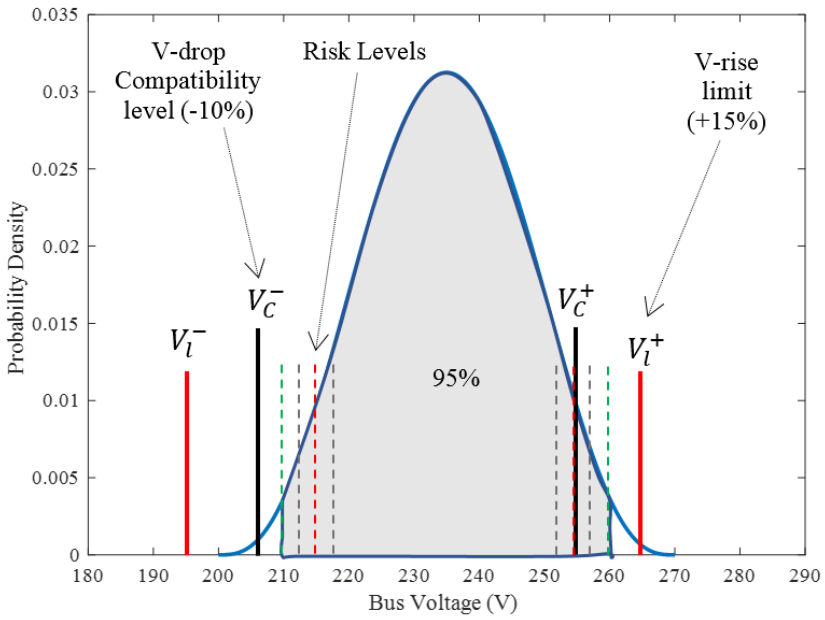

Figure 2.1: Compliance assessment of uncertain outputs

The figure also illustrates that for each parameter, the compliance depends on a selected design risk. For instance, adopting a 5% risk (broken green traces) in the design meets the voltage-drop compatibility level but violates the voltage-rise level (V_c^+). However, if a more relaxed design risk, such

as the one indicated by the broken red traces, is used, both compatibility levels are met. From this illustration, it is clear that the selection of a design risk and the knowledge of the relevant grid code requirements are of critical importance in the analysis of probabilistic outcomes. In most cases [93][37] [39] the level of design risk (in the form of confidence levels or percentiles) is adopted without practical justification.

From the probabilistic voltages and line currents, system variables such as active and reactive power flows, losses, and unbalance can be determined easily.

2.7 IMPLICATION ON ONGOING WORK

This chapter has discussed in detail the various approaches to solving the uncertain load flow and the associated performances in terms of accuracy, computational speed and complexity in the formulations.

The conducted review indicates that two critical performance metrics, accuracy and computational efficiency, separate the proposed techniques. Ideally, high computational efficiency and precision are both desired. However, speed is associated with loss of accuracy. Approaches aiming for highly accurate solutions (compared to model solutions such as the MCS) at high computational speed usually result in complex formulations that are unfit for practical applications.

While accurate and fast, a PLF approach must demonstrate consistency and scalability; the accuracy and computational efficiency of the approach must be consistent over various system conditions and system size. As been demonstrated through the review, most PLF solutions have sensitive accuracy influenced by the number of RVs in a system, conditions of dependence, and the shape of the input parameters. Furthermore, the computational efficiency of most approaches, particularly iterative methods based on repeated LF calculations, is largely influenced by system size. On the other hand, non-iterative approaches, particularly those based on a standard PDF input models and the method of moments, are less susceptible to scale-related computational issues.

2.8 CONCLUSION

This extensive review of many approaches to the PLF suggests that analytical approaches have the best potential to provide accuracy and computational efficiency. A PLF based on a standard distribution, whose parameters can be solved without significant computational burden nor complexity, has potential to fulfil the performance requirements. Among the analytical approaches, the moments approach offers ways of single-pass calculation, direct incorporation of correlation, and determining the output distributions without the use of expansion series. This suggests the Herman-Beta approach based on the beta PDF is a fitting candidate. It was formulated for LV feeders and has not been developed to meet all the identified requirements, especially for application on MV and HV networks. The detailed review of the approach in terms of the mathematical formulation, scope and its performance and the changes

needed to remove the assumptions will be discussed in Chapter 4. Before then, it is necessary to review the characteristics of the input models to the PLF, in the next chapter.

Chapter 3: Input Parameters and Models

The previous chapter reviewed the existing approaches to the probabilistic load flow. This chapter identifies the inputs required for probabilistic load flow analysis and reviews the approaches to the modelling. By discussing the sources of uncertainty in power systems and their characterisation using statistical approaches, and interdependence between random variables, the chapter provides a review of probabilistic load modelling.

3.1 IDENTIFICATION OF SYSTEM INPUTS

The identification and modelling of power system inputs forms a critical part of the PLF analysis. The validity of the results from a PLF study are dependent on the validity of input models; when the input models do not adequately capture the properties of the input states, the calculated output states will likewise inadequately represent the system behaviour despite the accuracy of the PLF input-output transformation.

The inputs and their characterization are application-specific and are influenced by system properties such as network topology and connected technologies. As such, the suitability of a PLF approach to a specific application is determined by how well it captures the system's characteristics; its inputs and the system's behaviour which determines the relationship between the input and the output states. The system's input parameters may include feeder models and parameters, connected technologies (such as loads, DG and capacitance), and other application-specific parameters.

3.1.1 Feeder Models

The input parameters related to the feeder are influenced by the topology of the network of analysis, whether high voltage (HV), medium voltage (MV) or low voltage (LV). The characteristics of these networks differ in several ways. The first apparent difference is in the nominal voltage. In South Africa, HV networks are operated at 44 kV and higher [94], usually run for long distances, and are mostly designed with overhead lines as opposed to underground cabling. Due to the large spacing between phase conductors, HV line are mostly inductive, with reactance-to-resistance (X/R) ratios larger than unity. Transmission lines can be classified according to length. Where the feeders are short (<80 km), the electrical representation of the feeders through a series impedance (resistor and inductor) is sufficient [95]. However, the charging effects of the lines by shunt capacitance must be included for longer feeders. The representation of the resulting models is possible through the nominal T- and π-models. However, the representation using the π-model is convenient for power flow analysis. Figure 3.1 shows a single-node representation of a feeder using a π-model, including how the technologies such as loads, and DG are interconnected.

As shown in the diagram, the π-model lumps the shunt capacitance (represented as admittances Y/2) into two equal parts at either ends of the line; the sending and receiving ends.

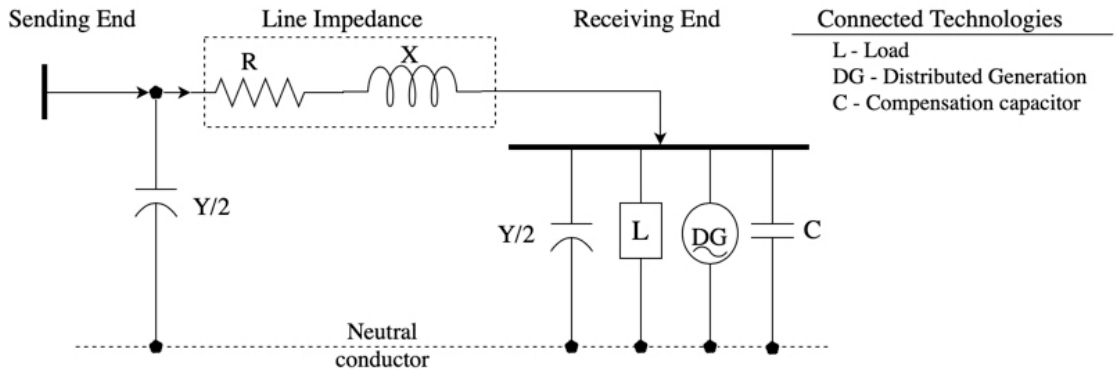

Figure 3.1: Pi-model representation on a long feeder with various connected technologies

Distribution networks can be separated into medium voltage (MV) and low voltage (LV) depending on the operating voltage levels specified by the relevant design guidelines. In South Africa, LV networks are classified for voltage levels up to and including 1 kV, while MV networks have nominal voltage levels greater than 1 kV and less than 44 kV. MV conductors are usually characterised by X/R ratios close to unity [96] whereas LV conductors are typically more resistive resulting in much lower X/R ratios. Due to the relatively shorter length of the feeders, both MV and LV feeders can be represented electrically using a series resistance and inductance, without shunt admittance.

Apart from the classification based on feeder length, feeders can be further separated based on phase technology or feeder layout [97]. Two main design practices exist; sometimes termed the North American and European. On the primary terminals (MV side) of the transformers, the North American layout is predominantly 3p-4w, three phase conductors and a multi-grounded neutral conductor, while the European is mostly 3p-3w, without neutral conductor. On the secondaries (LV terminals), the North America layout often has single-phase two-wire (1p-2w) laterals while the European has 3p-4w layouts. Other two-wire (2w) topologies are possible. On 3p-4w systems, single-phase loads are connected between a phase conductor and the neutral (1p-2w), whereas two phase conductors (2p-2w) are used in a 3p-3w system. Bi-phase systems with two phase conductors and a neutral (2p-3w) are also not uncommon.

Power system circuits can be operated in several ways including radial, ring and mesh configuration. The radial configuration can be common on HV and MV networks due to low cost, easier fault current protection and simpler voltage control amongst other benefits. However, ring and mesh configurations are used in cases requiring high reliability. Nevertheless, in some cases the tie points are normally open such that the circuit is open-loop. In this configuration, operation is essentially radial. For LV networks, the radial configuration is predominant, but meshed networks are also not uncommon, especially in metropolitan areas. These configurations and their operation obviously differ worldwide.

Table 3.1 provides a summary of the characteristics of the networks according to the classification by voltage.

Table 3.1: Characteristics of power system network components

VOLTAGE LEVEL	NETWORK CHARACTERISTICS					
	Voltage Level[4]	***Topologies***	***Conductors***	***Operation***	***Loads***	***Technologies***
HV	$V_{HV} > 44\ kV$	3p-3w	high X/R, Y/2 shunts	Radial	Balanced	PV-DG, W-DG, Shunt Caps.
MV	$1\ kV < V_{MV} \leq 44\ kV$	3p-3w 3p-4w[5] 1p-2w 2p-2w SWER[6]	X/R ~ 1	Radial, Open/Closed ring	Moderately Unbalanced	PV-DG, W-DG, Shunt Caps, ESS.
LV	$V_{LV} \leq 1\ kV$	3p-3w[7] 3p-4w[8] 1p-2w 2p-2w 2p-3w	Low X/R	Radial, Mesh	Highly Unbalanced	PV-DG, EVs, ESS

As indicated in Table 3.1, MV networks are manly operated in radial, open loop and closed loop ring configurations. However, the open-loop ring configuration, which has 'normally-open' network switches, is in essence operated radially. Hence, the application of a radial load flow can be suitable but with consideration of the resulting different network layouts from the switch configurations. LV networks, on the other hand, are sometimes operated in mesh configuration, but simpler and less costly radial systems are not uncommon.

3.1.2 Connected Technologies

Apart from the load demand, several technologies are interconnected to the power system. Compensation devices such as shunt capacitors are commonly connected to the power system to control the reactive power thereby reducing system losses and improving the quality of supply [98]. In addition, distributed generation from PV (PV-DG) [99], that from wind (W-DG) [100] and recently, electric vehicles (EVs) [101] and energy storage systems (ESSs) [102] form the components of the modern power system. The uptake and integration of electric vehicles (EVs) to power systems, particularly the LV network, has proliferated. Electric vehicles can operate as both consumers and producers; through EV charging and battery discharge as an ESS, respectively. However, due to the high costs of EVs, uptake in developing countries such as South Africa is still stunted and not expected to change significantly in the near future [103]. In such contexts, the assessment of EV loads may not be necessary.

[4] Values are based on the South African design guideline [89]
[5] Common in North America
[6] Single wire earth return
[7] Commonly found in Central, North and South America
[8] Commonly found in Africa, Asia, and Europe

Another feature of the inputs that affects how the load flow is performed is unbalance. HV networks are considerably balanced as both the loads [95] (which is mostly through transformers supplying underlying networks) and utility scale DG [104] are balanced through three-phase connections. As a result, a method developed to solve transmission networks can apply single-phase analysis on the basis that unbalance is insignificant. Unfortunately, the assumption cannot be made from MV and LV networks.

On MV networks, the load demand from the direct industrial, commercial and agricultural customers, especially when supplied through single-phase connections, can exhibit significant unbalance [105], [106]. As a result, load flow analysis on MV systems must consider unbalance. On LV networks, the levels of unbalance are much higher due to the prominent single-phase loads and small DG systems usually regulated for single-phase configuration (for systems less than 4.6 kVA in South Africa [104]) [107].

3.1.3 Electrical Behaviour of Connected Technologies

Electrical loads can be modelled as constant power (P), constant-current (I), constant-impedance (Z), or in composite form, depending on the response of the load power to the system voltage. Constant-power models assume independence from voltage variations, while constant-current and impedance models assume a linear and quadratic relationship, respectively. (3.1) provides the mathematical model representation of these relationships [108], [109].

$$P = P_o \cdot \left(\frac{V}{V_o}\right)^{\alpha};\ Q = Q_o \cdot \left(\frac{V}{V_o}\right)^{\beta} \qquad (3.1)$$

where, P_o and, Q_o are the values of the active and reactive load powers at nominal voltage, respectively. The parameters V and V_o refer to the actual bus voltage and nominal system voltage, respectively. The exponential α, β = 0, 1, and 2 for constant power, current and impedance load models respectively.

The electrical characteristic of the load has significant effects to the outcomes of a load flow and its convergence [110], [111]. Nonetheless, in most cases the load model is assumed to be constant-power (real and reactive) without practical justification. Such assumptions may lead to inconsistent and misleading results regarding the voltage and loading conditions on a passive network [110] or the impacts of DG which significantly affects decision making in penetration studies [111].

To understand the response of loads to system variations, load modelling studies are essential [112]–[116], whereas the assumption of a model without practical justification constrains the accurate and realistic assessment of a network. In South Africa, findings from a conducted load research on residential loads led to the adoption of the constant-current model [117]. Only a few PLF approaches have incorporated the voltage-dependence characteristics of loads in the formulation [72], [118] .

3.2 INPUT UNCERTAINTY

Generally, uncertainty can be regarded as the limited knowledge of a variable's state or value. The variable, instead of having a single 'deterministic' or a known singular state, multiple states are possible. This possibility of multiple states at a given instant, and the failure to determine, with certainty, which of the possible states the variable takes introduces the concept of 'uncertainty'. In statistical studies, mainly those in the field of risk assessment, the dominant taxonomy categorises uncertainty into two groups, aleatory and epistemic [119][120]. Aleatory uncertainty, also referred to as irreducible uncertainty, is due to the inherent variation or randomness characteristic of a variable occurring among members of a population. The variability arising from characteristics such as location (spatial) and time (temporal) are classified under aleatory uncertainty. On the other hand, epistemic uncertainty arises from simplifications in models, assumptions about the system inputs, and approximations made due to the lack of detailed information or limited knowledge about a system under analysis. The differences in the fundamental characterisation separates the representation and modelling of the two types of uncertainty. For aleatory uncertainty, the probability or chance of occurrence of the variable can be modelled using a probability density function (PDF) or a cumulative distribution function (CDF), whereas intervals and fuzzy numbers are often employed to model epistemic uncertainties.

Characterization and representation by mathematical models enable uncertain systems to be analysed. And by using computational tools that enable input uncertainty to be propagated in the analysis, the outputs are uncertain, reflecting the uncertainty in the system inputs. Given the representative results, it is possible to extract a singular result for decision-making by applying a risk or confidence level.

Briefly, the analysis of uncertain systems requires a three-fold framework, which includes:

- Data assimilation:
    - the identification of the uncertain inputs
    - characterisation and modelling
- Uncertainty propagation:
    - deriving solutions accounting for the identified uncertainties
    - characterisation and modelling of the uncertain outputs
- Certification:
    - establish acceptance criteria for the probable outputs based on design risk and regulations

Items (2) and (3) were discussed in the preceding chapter. This leaves item (1) to be discussed here.

The first part of the discussion identifies the sources of uncertainty in power systems, which can be categorized as input uncertainty and network model uncertainty.

3.2.1 Network Model Uncertainty

Network model uncertainty concerns two aspects; the limited knowledge regarding the system model, and that for its parameters.

Topological Uncertainty

Topological uncertainty arises from the inaccuracies in the representation of a studied network; the topology of a network, what is included or ignored and the random outage of network elements such as transmission lines [121], generators [40], and transformers [122]. With unlimited knowledge of the network model and its elements, the probability that the system operates in its basic configuration as designed is not unity [12]. The probability of losing network elements, or operation outside the normal conditions is not null. The possible changes in the network configuration influence the system states (such as line flows and bus voltages) and impact network operation.

Leite da Silva [12] reported the first PLF application that considered the uncertainty in the network configuration due to network outages. Since then, several approaches addressing topological uncertainty arising from line outages [40], [121], [123], [124] were developed. Most approaches use the concept of failure rate (or outage rate) to model the random operational states of the network elements. With the failure rates known, the effects on system operation and the resulting impact on the outputs states can be determined using appropriate load flow techniques.

Parameter Uncertainty

In reality, the true values of model parameters for power systems are in most cases not fully known [125]. There are several factors that contribute to the disparity between the practical values of the parameters and those reflected in the system design. Variations in the system's operating conditions such as temperature, the proximity of adjacent overhead conductors and the magnitude of line flows make some model parameters, such as resistance and reactance, uncertain. Further, where measurements are involved, the associated errors present additional uncertainty [12].

The characterisation of parameter uncertainty is uncommon due to the lack of extensive data sets for modelling. Apart from this, it appears the uncertainty is simply ignored with assumption that is does not significantly impact power system applications, yet it probably contributes more than we understand to failure by overloading. In the few cases where it is considered, the rated values (or value under anticipated operating conditions) and limits in the form of extremum values are used for characterisation. As a result, the modelling is commonly achieved using uniform distribution functions or interval analysis [126].

3.2.2 Load Demand

In addition to the electrical characteristic of the load discussed earlier, there are two further key attributes of the load that influence its characterisation and modelling; the variation with time (the

stochasticity element), and the diversity of the load for a population of grouped customers (the variability element). The stochasticity refers to the load changes with time of day, type of day, seasons and weather, and, the variability in the demand is as a result of the diversity in the load demand within a group of customers. These two elements are further discussed in relation to the sources of uncertainty for load demand.

A customer's behaviour: The stochasticity of the load

For a single customer, the variation in electricity consumption is dependent on the 'electricity behaviour' of the customer [127]–[130]. The behaviour can be characterised in terms of the demand as a function of the 'time-of-use' (TOU) of electricity as influenced by habits such as times present at home, cooking times, and bathing times, which in turn determine the selection of appliances used [127], [129]. These variations are responsible for the intra-diurnal load patterns. Beyond that, the load demand varies from day-to-day depending on customer activities. However, a customer tends to have more or less consistent weekday patterns (Monday-Friday) which can be differentiated from weekends (Saturday & Sunday) [130]. Furthermore, seasonal variations introduce yet another layer of load variability [128].

While the variations between months in the same season can be considered relatively constant, inter-seasonal variability is significant; the winter load demand is anticipated to be generally higher due to the increased use of heating devices, whereas a summer month often has lower demands unless air-conditioning is heavily utilised. Consequently, the identification of load model parameters adequately capturing seasonal effects, daily-weekly periodicity and diurnal variations is vital [131]. Figure 3.2 demonstrates load model segmentation based on TOU, season and day-type.

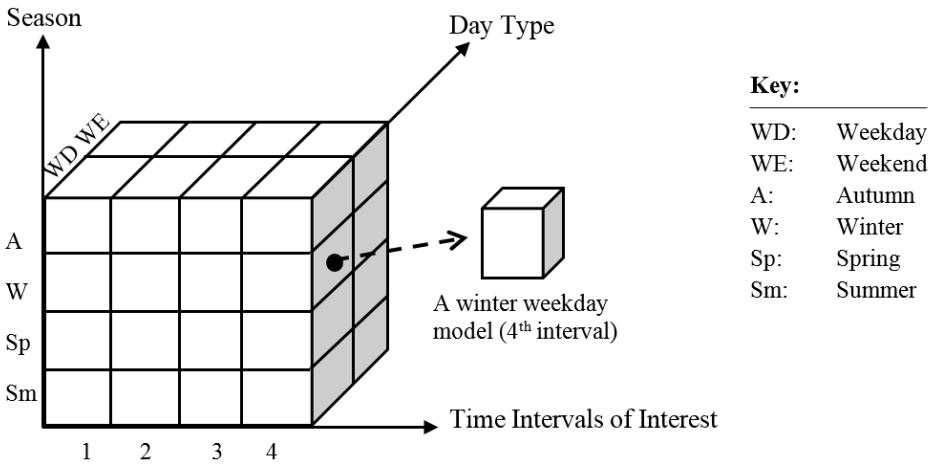

Figure 3.2: Load Model Segmentation for Uncertainty Analysis

The model shows a 3-dimensional analysis of load variations across seasons, day-type and diurnal intervals as identified in the review. In the representation of Figure 3.2, only four intervals are used for demonstration purposes. The selection and representation of intervals of analysis depend on the requirements of a specific application. For example, for commercial customers, the load demand is

almost constant throughout the day, except for the lunch period. Accordingly, only two intervals might be of interest, whereas for an industrial customer, several intervals connected to work shifts may exhibit significant differences. Likewise, the intervals around noon are of interest for renewable energy integration studies involving photovoltaic systems [132].

A group's behaviour: Load diversification

Utility companies often perform customer classification to capture customers with similar consumption profiles using clustering approaches [133]–[140]. This reduces the volume of network data associated with classifying each customer with a profile and also relieves the utilities from continual load research projects for new customers. When customers are grouped, they are identified with a single representative profile [136]. However, as is expected, the customers within a class or load group are not exclusively homogenous, diversity in terms of consumption exists. As a result, where a representative load profile is used, the spread of the individual customer's profiles around the selected representative profile demonstrates the associated uncertainty. Several load research projects [141], [142] involving the clustering of electricity customers have concluded that social factors, economical and lifestyle are some of the reasons for the intra-class variability.

3.2.3 Distributed Generation

The introduction of distributed generation from renewable energy sources has seen significant changes in the design and operation of the power system. Firstly, the formerly passive configured system where power flowed unidirectionally and downstream from centrally located generators becomes active, with bi-directional flow of power [143]. Secondly, the generation capacity becomes uncertain as a result of the intermittent power production from DG plants [144]. The cause for the intermittency is due to the dependence of the outputs of the DG plants on meteorological conditions such as wind and the sun's irradiance. As a result, the stochasticity in the weather patterns is reflected in the DG power output. This introduces a significant degree of uncertainty in the power system which in turn presents operational challenges [145]–[147]. The uncertainty related to the most common renewable energy sources, solar and wind, is discussed henceforward.

Wind Power Generation

The output from wind power plants is dependent on the characteristics of the wind it receives, the land terrain at the site of installation, and the technical specifications such as the height and type of wind turbines [148], [149]. Of these variables, wind speed exhibits an intermittent nature owing to the dependence on weather conditions which are stochastic in nature. As a result, the output power from wind turbines is uncertain and exhibits diurnal and seasonal variability.

The relationship between the power output, P, derived from wind with speed, v, and swept across an area, A, where the air has density ρ, is non-linear according to (3.2) [150].

$$P = \frac{1}{2}\rho A v^3 \qquad (3.2)$$

To avoid the non-linear formulation and reduce the computational burden (especially when several wind turbines are considered), a linear formulation based on the manufacturer's power curve for a specific turbine is commonly used [149]–[151]. Equation (3.3) gives the output power as scaled form of the wind speed according to a three-part piece-wise function.

$$P = \begin{cases} 0, & v < v_{in}^c; v > v_{out}^c \\ \frac{v - v_{in}^c}{v_R - v_{in}^c} \cdot P_R, & v_{in}^c \le v \le v_R \\ P_R, & v_R < v \le v_{out}^c \end{cases} \qquad (3.3)$$

Where P_R and P_i are the rated power of the wind turbine and the power generated at bus *i*, respectively, whereas v_{in}^c, v_{out}^c, and v_R refer to the cut-in, cut-out and rated wind turbine speed, respectively.

Considering the wind speed as a RV, the active power output from the wind turbine is a RV scaled and truncated according to (3.3).

Photovoltaic Systems (solar irradiance, PV arrays)

Generally, without consideration of the technical variables, the performance of solar arrays in harvesting electric power from the sun relies on the solar radiation incident on the panels and the ambient temperature at the site of installation [8], [149], [152]. This radiation is dependent on the weather conditions attributed to the location of installation and its surroundings which can cause shading and potentially limit the impact of the incident radiation. As expected, due to the varied irradiance of the sun from sunrise, through noon, to sunset, photovoltaic (PV) generation exhibits strong diurnal variability. Further, seasonal variability is significantly distinct between summer and winter seasons; summer days having more sun hours and irradiance than winter days.

Considering solar irradiance (I_S), as a random variable, the power output from a solar panel with surface area A and efficiency η, the power output (P_{PV}), assuming the system includes a maximum power point tracking (MPPT) system, is given by [152]:

$$P_{PV} = A \cdot I_s \cdot \eta \qquad (3.4)$$

The model can be expanded to include the effects of cloudiness or clearness [151], [153], [154] but retaining the linearity of the transform. Accordingly, the uncertainty in PV power generation is a scaled RV of the uncertain solar irradiance.

3.2.4 Scope for Uncertainty Analysis

While it might be of interest to a planner to consider the full range of uncertainties, there are challenges. The modelling of uncertainty demands the collection of historical data and supporting load research to determine the suitable models for characterisation. Furthermore, the consideration of a great number of random variables impacts the computational efficiency of most power flow programs. Consequently,

the decision whether to include or neglect a particular source of uncertainty depends on the system's sensitivity to that uncertainty. Accordingly, it may not be necessary to model or accurately model all uncertainties especially when they have insignificant impact on the system's variables of interest, despite the addition of considerable computational inefficiency and data requirements [155].

In the light of these constraints, most approaches neglect epistemic uncertainties (model and parameter uncertainty) [156]. These types of uncertainties are less significant than those associated with input active and reactive power variations.

3.3 PROBABILISTIC MODELLING

Uncertainties (as identified in the previous section) must be taken into account in power system analysis in order to achieve effective network designs that take into account the full spectrum of input possibilities. For the relevant analysis, mathematical models characterizing the uncertainty are needed [157].

Probabilistic load modelling is founded on the principle of expressing the likelihood or chance of occurrence of an event as a numeric probability. The probabilities of event occurrences are modelled based on hypothetical concepts or historic data [157]. Either way, the characterisation of the probability is through a probability density function (PDF) which when integrated over the domain of the random variable leads to a cumulative distribution function (CDF). However, in some cases, uncertain variables are characterised with a mean, which is usually the deterministic model of the variable, and hypothetically assigned a variance or standard deviation to cater for the anticipated variability. In such cases, the assumption of the normal PDF is nearly automatic. The processes underlying the task of probabilistic characterisation are two-fold; selecting a suitable distribution and determining its parameters.

While the models differ, the acceptability of any probabilistic assessment is determined by the validity of its input models. As such, the review of the various modelling approaches, the related models, and the analysis of their suitability (accuracy, data requirements and computational implications for load flow programs) is critical.

3.3.1 Hypothetical Models

Though the development of statistical load flow approaches using diverse load models has been progressive and widely reported, the quantification of the required input parameters has mostly not been explicitly described. In a large span of publications [51], [68], [154], [156], [158]–[162], input random variable parameters such as load variance, standard deviation, and correlation factors are given either as arbitrary values or devised without practical justification. For load and DG modelling, coefficients of variation (CV) between 0 and 15% are commonly applied to deterministic models in order to facilitate a PLF analysis. This lack of clarification of how statistical input parameters can be derived for

a probabilistic load flow assessment potentially limits the usefulness of PLF methods to system operators, planners and design engineers.

The limitation of hypothetical modelling is that the input models are not validated with the practical behaviour of the input variables. As a result, the results from such analysis cannot be expected to be more meaningful than its input models [120].

To derive more meaningful models reflecting the practical variability observed in the data, detailed load research is required. Such studies involve a great deal of data collection from load monitoring and data-logging. The general scope of the modelling process, regardless of the final selected statistical model, comprises the following steps [13], [163]:

- data collection
- data processing to clean-up bad data
- clustering of data according to design characteristics (customer class, voltage level, etc.) and for a desired period depending on the scope of uncertainty planning, whether short, medium or long-term
- evaluation of goodness of fit (GoF) tests to determine the representative statistical distribution functions for the grouped data
- evaluation of dependency between system inputs

Out of the curve fitting exercise, two forms of models are possible; standard distributions or mixture models. These are discussed in more details in the subsequent sub-sections while the subject on dependency is dealt with later.

3.3.2 Standard PDFs

Load Demand Models

There have been significant research contributions on the statistical description of loads [7], [13], [15], [127], [133], [134], [139], [164]–[171] and DG [6], [8], [9], [12], [23], [24], [170], [172]–[176] using standard parametric distributions. There are a couple of factors influencing the choice of a distribution over another. For loads, the Gaussian distribution is widely adopted due to its simplicity [13], [51], [126], [177]; the mean and variance are easily computable, and the outputs are characteristically Gaussian due to the linear property. Nonetheless, the distribution has been shown to be limited and inadequate as the load data demonstrates significant skewness [15], [117], [167], [178]. Owing to the identified limitations, Sirisena et al. [170] suggested the use of multiple Gaussian distributions to the model the skewness. However, the modelling application further constrained the pre-existing computational inefficiency associated with the load flow tool.

Herman and Kritzinger [7] reported on a load research in which various distribution functions such as the Weibull, normal, and beta were fitted to grouped residential loads in South Africa. The outcomes of the fitness tests evaluated using χ^2 and K-S statics, and the consideration of flexibility in modelling variously skewed data distributions, including symmetrical data, led to the recommendation of the beta PDF. Further, the beta model is convenient for the computation of the uncertain load flow using an analytical approach based on moments.

A similar load research is reported by Ghosh et al. [165] but for an application on distribution system state estimation and considering commercial load data. Conducted χ^2 fitness tests on models such as the normal, log-normal, and beta distributions revealed the log-normal as the best fit and the normal as the worst. However, considering the comparable accuracy between the beta and log-normal PDFs for modelling left-skewed load data, and the advantage over the log-normal in modelling right-skewed data, the authors recommended the beta PDF. In other cases, the log-normal [166], lambda [179], gamma [180] and Weibull [135] distributions have been applied as alternatives to the modelling of skewed load data.

The models discussed thus far are mostly applicable to LV distribution systems at the customer level. For higher voltage levels such as the MV and transmission systems, the load and generation from lower voltage levels are commonly aggregated at the supplying transformer terminals unless the feeder is analysed as a composite system. For aggregated loads, the model of the compound load is dependent on the respective constituent load models until a sufficient number of loads is reached to satisfy the conditions of the central limit theory (CLT). When the CLT takes effect, the aggregate load models approach normality [13], [180].

Filho et al. [13] conducted Chi-square (χ^2) and Kolmogorov-Smirnov (K-S) tests to evaluate the goodness of fit (GoF) of normal distributions on substation level aggregated load data on Brazilian networks. Their findings led to the conclusion that the assumption of the Gaussian model is valid for both aggregated active and reactive power.

The question that arises regards the fulfilment of the CLT conditions to justify the assumption of normality. Related, Carpaneto and Chicco [180] conducted a study aimed at the statistical characterisation of aggregated residential customers and assessing the application of the CLT with various sizes of aggregated loads. Extensive K-S goodness-of-fit tests rejected the normal distribution for aggregated loads for groups with less than 20 customers due to significant skewness. However, as the number of customers increases, the aggregated loads become acceptably Gaussian. For a group of 40 customers, about 94% of the histograms from various time intervals passed the normality test.

However, where direct MV or HV customers exists, the characteristics of the load is anticipated to be different. For these, there has not been substantial studies for the statistical characterisation of the loads. However, studies involving clustering [181]–[186] present promising approaches to the derivation of

the relevant models. In particular, the approach reported by Nassar and Salama [185] provides a potential framework for deriving statistical load properties for grouped MV clients using clustering algorithms such as the k-means. The GoF tests (Anderson-Darling and K-S tests) performed on the segmented data clusters showed that the Weibull and beta PDFs provide the best fits for the data, whereas the Gaussian PDF was inadequate in modelling the skewness. The drawback in the modelling approach is that the load is represented using a mean (deterministic) value and a probabilistic error, which is not convenient for the derivation of PLF solutions.

Generation from Non-renewable-energy Plants

Binomial distributions [12], [23], [24], [170], [187] are commonly used to model the power production from non-renewable-energy power plants (as are coal and nuclear power plants that serve as base load units and operate independently of network load variations). This representation is adequate considering no intermediate generation states exists. In actual fact, without anticipation for outage, a deterministic model is acceptable. This however changes when the generation is dependent on a renewable energy source whose availability is uncertain. In such cases, as with DG from RES, the applied models reflect the variability associated with the RES.

Wind Power Generation

The modelling of the expected wind turbine power generation can be achieved through statistical models of wind speed and the application of the characteristic function of the specific wind turbine (3.2). Since the 1940s, a large number of research studies have proposed a range of statistical models to describe the variability in wind speed data. An extensive review of the body of literature from the 1940s until 2009 is provided by Carta et al. [188].

Distributions such as the Gaussian, Gamma, Rayleigh, Log-normal, Beta and Weibull have been considered. However, the majority of the publications demonstrate that the Weibull provides the best model unless the data contains dominant multi-modes. The investigations by Carta et al. [188] also confirmed that in addition to good performances on GoF tests, the Weibull was preferred on the basis of (i) its flexibility, (ii) the limited number of parameters characterising the PDF, (iii) simplicity in solving for the distribution parameters and (iv) its closed-form (bounded interval) which matches practical expectations of bounded wind speeds.

Although generally applicable to most cases, the Weibull is not always the most appropriate model. In some instances, the extreme-value [172], [173], beta [9], [172], Gamma [189] and Burr [189] PDFs are better.

Photovoltaic (PV) generation

The statistical models for power generation from PV arrays depend on the models for solar irradiance and the characteristic transforms relating the irradiance and the generated power in the solar cells (3.3). Like the variability associated with wind speed, that of solar irradiance is also highly skewed [190].

Graham and Hollands [175] reported one of the earliest work on the statistical modelling of the measure of irradiance. In their work, irradiance is not modelled directly but through the atmospheric transmittance or clearness index. Analyses of the historical meteorological data revealed that hourly atmospheric transmittance histograms were closely modelled by beta PDFs. Taking a slightly different approach, Salameh et al. [8] conducted statistical modelling tests using solar irradiance data measured on a daily basis for about 30 years in hourly intervals. The observed frequency histograms tested using $\chi^2$and K-S GoF tests revealed the beta PDF as the most suitable distribution. On the other hand, the Weibull and Log-Normal were found to be inconsistent in modelling the variously skewed histograms for different time intervals throughout the day. In another research, Youcef Ettoumi et al. [176] showed that the statistical features of solar measurements collected from various locations in Algeria were acceptably modelled by the beta PDF. Further research has shown the suitability of beta PDF for modelling irradiance (or the indices for clearness and transmittance) using 5-minute averaged data [191], segmented data according to seasons [192], and for long-term planning [193].

3.3.3 Mixture Models

Though the aforementioned standard parametric distributions have been suggested for various load and DG sources, it has been argued and demonstrated that the variation in the practical load data cannot be perfectly fit by any single unimodal distribution [167]. As a result, mixed models, which approach modelling by employing a weighted sum of a number of parametric distributions of the same kind, to estimate the shape of the load distributions, have been suggested [163], [167], [178]. In this regard, the application of the Gaussian mixture model (GMM) has been proven to be an adequate approach to deal with the heterogeneity associated with grouped customers and the multi-modality that arises in the data.

While representation by GMMs achieves a high level of modelling accuracy, there are challenges. The optimal number of components required to obtain a desired accuracy in the developed models is difficult to determine [136]. Accordingly, iterative techniques such as the Bayesian Information Criterion (BIC) [194] and the Expectation Maximization (EM) [167] based on the maximum likelihood or maximum *a-posteriori* (MAP) estimation have been applied to determine the optimal number of components and their estimates, respectively.

The accuracy of the GMM depends on the number of components in the mixture; one study suggests that the optimal number, considering accuracy and computational efficiency, resulting in a satisfactory approximation of the PDF is up to five [195]. However, the number of components used in the GMM may present a strong limitation when dealing with large systems with a lot of random variables. This issue is two-dimensional:

- The impracticality of load modelling using many parameters (a 5-component GMM requires a minimum of 15 parameters and a 5-by-5 covariance matrix),

- The computational complexity associated with the PLF; in most cases, the PLF solution, though computable analytically for a scenario in which every input has a single component, requires an iterative approach to assess the impact of the combination of the components [195].

Similar in principle to the GMM, Huang et al. [178] proposed a mixed skewness model (MSM) based on combinations of skewed distributions as opposed to the symmetrical Gaussian components in the GMM. In the application, focused on modelling the long-tailed conditional forecast errors in wind power, the authors conclude that an optimal number of up to 3 components is sufficient for most engineering applications; beyond this number, the modelling accuracy improves but the computational expense increases significantly.

Clearly, standard PDFs cannot perfectly model the variability and the heterogeneity associated with practical load or generation data. On the other hand, mixture models have the ability to accurately model the data using summated, weighted standard PDFs. However, there are several computational issues associated with mixture models, while the use of standard PDFs is computationally convenient for load flow analysis tools.

3.4 DEPENDENCE BETWEEN INPUTS

From a statistical point of view, dependency in a multi-variate system refers to the relationships between random variables in terms of their probability of occurrence. This interdependency can either be positive or negative such that a high chance of occurrence of one variable is linked to a high, or low chance of occurrence of another, respectively [196]. For instance, extreme heat or cold weather conditions are likely to be coupled with high electricity demand as a result of the increased use of air conditioning and heating systems respectively [197]. Likewise, the power outputs from wind generators in a common geographic location tend to rise and fall together [198]. In both these cases, a positive correlation exists between the random variables. As such, the existence of correlation potentially has significant impacts on the occurrence of probabilistic events and the resulting probable set of input states. Since the output states of a system depend on the range of inputs, the outcomes of a probabilistic analysis are influenced by the dependencies between random variables. For that reason, the appropriate representation of interdependencies between input parameters is of critical importance [120].

PLF studies incorporating correlation between inputs are extensively reported [16], [35], [68], [72], [159], [163], [169], [187], [199]–[204], however only a few publications report on modelling from practical data [198], [205]–[211]. The review of the relevant literature points to four main forms of interdependence which form the discussion that proceeds.

3.4.1 Load-Load Correlation

The assumption of independence between loads has been contested right from the onset of the development of probabilistic load flow methods [187]. The consistent argument in many publications is that a group of customer loads from the same community or residential area will tend to behave in a like manner due to demographic and environmental factors [16], [142]. As a result, some form of dependence between them is inevitable. In the founding publication, Allan at al. [187] demonstrated the correlation in load demand between aggregated loads using practical load data. The conducted sensitivity tests showed that neglecting dependency results in misleading results where the output density functions are either broader or narrower than the true outputs. Despite the means (μ) being unaffected, the differences in dispersion affect risk-based decisions which rely on percentile values. Stemming from this practical demonstration, subsequent publications adopted the modelling of load dependence as linear correlation [16], [169].

Filho et al. [13], [203] tested the correlations between loads on a Brazilian network in Rio de Janeiro. From the calculation of correlation coefficients from the load samples, loads in one area presented high levels of correlation (more than 0.7) while the rest were weakly dependent. However, in the PLF analysis also reported in the same paper, only two dependence states, either completely dependent or independent, are considered. Humeau et al. [206] later demonstrated the correlation between customer loads, dependence between loads over time, and the dependence of the load demand on temperature.

Generally, a few PLF approaches have demonstrated the impact of correlated loads [68], [159], [163], [201], [212]. In some cases, if customer classifications are based on similarities in electrical behaviour, customers belonging to the same class and sharing a common load model may be considered independent. However, this is only valid for the specified interval in which the load model is defined. [142].

3.4.2 Generator-Generator Correlation

Allan et. al [169] extended their previous work [187] which introduced the impacts of load-load dependence to include correlations between generators. The paper takes on a perspective inclined to the application of generation dispatch. It suggests that generator control for generators serving a common load area are raised and lowered in a correlated fashion to meet the requirements of power in that area. As a result, the output of one generator is dependent on the output of another. Several publications have also modelled this type of generator-generator correlation until distributed generation became prominent.

With distributed generation from renewable energy sources, the correlation between random variables becomes more significant. The main reason for this is that the power output from DG plants is dependent on weather conditions which are in turn stochastic. As a result, generators in the same geographic area and experiencing common meteorological conditions, such as wind and solar irradiance, are likely to

exhibit a positive dependence between their power outputs. Several studies have demonstrated the interdependence between wind farms [100], [198], [210], [213]–[216] and that between PV systems [211], [217], [218]. The studies generally indicate that the correlations in wind speed and solar irradiance is strong (up to 1) within the same area, and gradually falls with distance. A study by Ekstrom et al. [211] is discussed in more detail, providing a comparative analysis of the characteristics of DG - DG relations between PV and wind power plants.

The correlation test results presented by the authors show that PV plants in the same area exhibit correlations up to full-dependence ($\rho = 1$), whereas wind turbines under the same conditions are only partly dependent ($\rho < 0.8$). Between the PV and wind plants in the same location, negative correlation coefficients of up to 0.25 were observed. Their analysis of the trend of correlation with distance indicates that for PV and wind plants within a radius of 180 km and 100 km, respectively, the correlation remains significant (up to $\rho \cong 0.5$, reduction by $\cong 50\%$). On the other hand, the correlation between PV and wind generation reduces to about –0.2 (reduction by 25%) in a distance of 100 km.

There are limited publications focusing on the derivation of correlation factors from practical information. In addition, most published test networks are intended for deterministic analysis and therefore do not provide any statistical parameters. The performance of PLF applications with dependent RVs is thus mostly tested using arbitrarily selected coefficients of correlation [23], [24].

3.4.3 Load-Generator Correlation

Allan et al. [169] argued the existence of another form of input correlation; that between loads and generators. For non-renewable generation plants, the load-generator correlation is as a result of the coupled control of dispatched generation according to variations in the load demand. A few publications [205], [207], [208], [210] have investigated the relationship between the load demand and power production from DG. Ziser et al. [205] demonstrated the existence of correlation between the demand diversity and weather diversity in Queensland, Australia. The dependence of the load demand to weather conditions such as temperature, wind speed, and cloud cover, was measured to be between 32% and 86%. Since DG production depends on weather conditions, the achieved results suggest the existence of load-DG correlations. Bell et al. [208] investigated the relationship between wind velocity and electricity demand in order to assess the capacity of wind turbine generators to meet the demand for electricity without additional power. Through the investigation, which was performed for numerous sites, it was discovered that both positive and negative correlations were possible, depending on location and its characteristic wind patterns. The study reported by Thornton et al. [210] offers the characterisation of the relationship between electricity demand and wind power supply and is potentially useful for both short and long term power system planning. However, the analysis does not lead to the derivation of correlation coefficients usable for PLF analysis. Unlike the earlier two forms of dependency, load-generator uncertainty is not commonly included in PLF input models.

3.4.4 Other Forms of Interdependence

From the onset of the development of PLF approaches, Leita da Silva et al. [16] suggested that the linear model for correlation as applied for the dependence between nodal powers could be used to account for the relationship between the active and reactive components of the load. However, according to the authors, the correlation is likely to depend on the type of loads and the sensitivity factors of the load flow. As a result, the assumption of dependence between the quadrature components of the powers should not be taken as a general conclusion. The paper makes recommendation for detailed study of the load characteristics if the conditions of correlations were to be considered. Due to the complexities anticipated with the in introduction of this type of correlation, including the unavailability of detailed studies on the load behaviour, the authors assumed the components were independent.

Since this article, there has not been, to the knowledge of the writer, any research directly addressing the identified gap in the body of knowledge. As a result, it remains that the subject area is poorly understood. The majority of PLF approaches assume one of three modelling approaches:

- The power factor is constant, and deterministic, such that the variability in the active power is the same as that in the reactive power [6], [30], [151], [158], [200], [219], [220].
- The active and reactive powers are linearly correlated [62], [201]
- The active and reactive component are described using different PDFs and treated as independent RVs [23], [40], [49], [51], [154], [221], [222].

Obviously, the models are expected to be application specific. Due to the heavy use of lighting and resistance heating, particularly during evenings, the residential load can be predominantly resistive with the power factor close to unity and with little power factor variation between grouped customers. However, where appliances such as computers and motors form part of the load, the variation in power factor changes. As such, commercial and industrial loads are expected to demonstrate considerable power factor variations. Further, the operation of inverters, particularly at MV and HV level and sometimes at LV, where DG uptake regulation may be implemented using power factor control, introduces another aspect to the modelling of power factor.

Delgado and Dominguez-Navarro [62] investigated the effects of dependency between the active and reactive components of the load demand. The achieved results demonstrated that where the load components were fully correlation, the dispersion in the voltage distributions is higher. Then recently, a study reported by Bingyan et al. [223] demonstrated similar effects but for wind-DG inputs. While the outcomes from both investigations appear compelling, there is still no indication from practical data what the anticipated power factor variations would be. Accordingly, the true implications of assuming a particular model for the power factor is not certain. Nevertheless, the results inspire load research focussed on the modelling of power factor variations.

Another form of dependence exists for asynchronized statistical profiles. Graham and Hollands [175] in their investigations on the characterisation of the variability of solar irradiance using atmospheric transmittance discovered a correlation coefficient of about 0.54 ± 0.14 between successive transmittance hourly values. Studies conducted by Chen et al. [224] led to the recommendation of joint-normal distributions to model load time series for use in time-sequential simulations for distribution system planning. Bernard et al. [163] further investigated 15-minute residential load data for correlation between consecutive time interval input models for a time series probabilistic analysis application. Subsequent quarterly load measurements demonstrated correlation factors of up to 0.92. As anticipated, the degree of correlation decreased by the time interval separation so that correlation factors for intervals one hour apart decreased by as much as 40%, while those more than 2 hours apart had little correlation and can therefore be considered independent.

In any multi-variate system (the power system in our case), several forms of correlations between RVs potentially exist. However, if all interdependencies are taken into account, without justification, the dimensions of dependence become very high leading to complex load flow formulations [36], let alone the high volumes of data required to characterise the dependencies. As a result, the consideration of correlation in PLF studies must follow practical justification and consider the computational cost over accuracy the inclusion of the dependencies makes.

3.5 IMPLICATION ON ONGOING WORK

This chapter has dealt with the characterisation of the inputs which can be separated into feeder components and connected technologies.

For a PLF approach to be widely applicable without specificity to a voltage level, it must carefully consider the feeder configuration, layout and model parameters which impact the LF analysis. Across all networks, at different voltage levels, the 3p-3w, 3p-4w, 1p-2w, 2p-2w and bi-phase (2p-3w) are common configurations used in power system design around the world. The electrical modelling of the feeders using only a series impedance (series resistance and reactance) model is sufficient for feeders of short length (<80 km), while the π-model (or T-model) with shunt admittance is required to simulate line charging effects for longer feeders. All feeders are commonly operated in radial configuration, including open-loop ring networks. However, closed-loop ring and mesh layouts are not uncommon. Notwithstanding the benefits of a PLF approach applicable to all the feeder layouts, the scope for this work is limited to radial-operated feeders. For such systems, the non-linear load flow equations can be avoided using simpler voltage-drop calculations based on Ohms Laws and the application of the superposition theory.

Apart from feeder properties, the characterisation of the various technologies such as voltage regulation shunt capacitors, the load demand, PV-DG, wind-DG, EVs, and ESS connected to the modern power system is essential. The connection configurations at the respective PoCs, whether single-phase or

three-phase, affects the balance across the feeder phases, which significantly impacts the LF outcomes in turn. Further, the sensitivities of the LF (its outcomes and convergence) to the electrical behaviour of the inputs necessitates the consideration of voltage-dependence in the models.

The quantification of the inputs at a given instant is complicated by uncertainty. Approaches such as possibilistic and interval analysis offer means to characterise the associated uncertainty in the absence of detailed data. However, in the presence of data, probabilistic approaches are necessary to explicitly model the characteristic variability.

Tests with practical data demonstrate the variability in the input data can result in symmetrical or unsymmetrical (skewed) distributions. Further, the heterogeneity in the data components can be significant to cause multimodality which cannot be exactly modelled by standard distributions. For such cases, mixture models are more appropriate. The Gaussian model is the most applied model for the load demand. In most cases, the selection of the model is based on its simplicity. However, there is substantial indication that the practical data demonstrates significant skewness unless the demand or generation is aggregated. As a result, several non-symmetric distributions with the ability to model skewed data such as the Log-Normal, Weibull, and the Beta were suggested and are more appropriate. For DG, the Weibull is dominantly recommended for wind generation while the beta PDF is popular for PV generation. However, the beta PDF also acceptably models the variability associated with wind generation and in some cases is the preferred distribution.

Three forms of input dependence have been identified; load-load, load-DG and DG-DG. The active and reactive components of the load are also most likely correlated. However, there has not been sufficient research using practical data to quantitatively describe the supposed dependencies. Apart from the majority of publication which do not comment on the representation of the dependence, a few approaches are based on marginal distributions while others have assumed the reactive component is a projection of the variability in the active power through a constant power factor. The approach in this thesis follows after the latter.

3.6 CONCLUSION

This extensive review of the many aspects and approaches to input modelling suggests that a widely applicable PLF approach must consider symmetric and unsymmetrical input distributions, include the effects of voltage-dependence, model the correlation between RVs, and model feeders accurately considering complex line impedance and shunt elements for voltage regulation (shunt capacitance) or line regulation (shunt admittance). Furthermore, the approach must be based on an unbalanced load flow. These findings provide the features to be covered by the HBA algorithm reformulation. However, it is necessary to test the validity of beta PDF as a descriptor of the different inputs in advance.

The review reveals that the beta PDF is validated for modelling both the variability associated with the load demand and PV-DG. However, in some cases, the Log - Normal provides a better model for positive - skewed data, whereas the Gaussian can provide the best distribution approximation to aggregated load data. For wind, the Weibull is widely applied. Where the data demonstrates multi-modality, mixture models with the optimal number of components offer more accurate means to model uncertain variables. However, the use of many parameters and the complexity in the computation of the associated load flow present significant challenges. The following questions arise: (1) how accurate should input models be, and (2) where the beta PDF does not provide the best-fit to the practical data, what is the error in assuming the beta PDF as a model. Hence, the next chapter tests the validity of the beta PDF for probabilistic load modelling by carrying out an investigation on the impact of input model accuracy on the PLF outputs.

Chapter 4: The Original Herman-Beta Algorithm

In this chapter, an alternative method to probabilistic load flow analysis known as the Herman-Beta Algorithm (HBA) is discussed. The theoretical foundations of the approach and the progression of the developments are detailed. The scope of its applications is assessed leading to the identifications of its strengths, weaknesses, and limitations. The performance evaluation is carried out through several case studies on representative test networks and validated using the Monte-Carlo Simulation tool.

4.1 HISTORY AND PROGRESSIVE DEVELOPMENTS

A load research project (LRP) conducted in the early 90s on residential customers in South Africa indicated that the load demand of grouped customers was acceptably modelled by right-skewed beta distributions [7], [117], [225]. Further, the power factors for most loads, analysed during the interval of maximum demand (relevant for passive network design), were found to be very close to unity, while the feeders supplying them were mostly resistive. These factors, including the versatility of the beta PDF and the relative ease to manipulate it for PLF calculations, led to the development of the initial Herman Beta Algorithm (HBA). The proposal, made in 1993, was intended for statistical voltage-drop analysis of passive, radially operated, low voltage (LV) feeders [50].

After several refinements and extensive testing, the method was later adopted as a national LV feeder design standard in South Africa [94]. At that time, distributed generation was not as popular such that national utility companies had no provisions for their integration. However, with time, various factors including the advancement of renewable energy and the deregulation of electricity markets, opened doors for DG penetration. The newly actively configured power system required a new approach allowing for the representation of non-centrally located generators and the resulting bi-directional power flow. For that reason, the HBA was extended to active topology [226]. In the quest to make the tool more deployable to practical applications than previously possible through the Excel-spreadsheet format [227], a C# software program was made [228]. The software was developed for the analysis of networks with up to 1,000 nodes and many different loads and DG models. The application of the method, without modifications, on MV networks in Brazil [229] inspired further research on the possibility of voltage-drop calculations on MV networks [10], [11].

To provide more details, brief discussions on the key underlying developments of the algorithm are presented using a timeline as follows:

- 1993: Ron Herman identifies the beta PDF as a descriptor of currents, develops a transform based on moments, which was later named after him [50], [230]. The transform was derived

to compute voltage-drop on low voltage feeders for two configurations: three phase systems with a neutral conductor, and bi-phase systems [231]. However, to reduce the computational complexity, certain assumptions taking advantage of the findings gathered from a conducted LRP were made. The feeders were considered purely resistive, and the loads operated at unity power factor. To limit the scope further, the proposed algorithm was derived to solve only two nodes, as demonstration of the methodology. At this time, no distributed generation (DG) was considered. Accordingly, the algorithm was only suitable for passive LV systems with insignificant reactance and unity power factor loads and was not fully developed to be applied to sizeable practical networks.

- 1998: Herman and Heunis took the work further by testing the extension of the transform beyond the two nodes for which it was initially formulated [227]. The research discovered that using the original approach to summate the voltage-drop effects of successive nodes resulted in significant errors. To solve the identified issues, the superposition theory was suggested to summate the effects of multiple nodes. Its success meant the transform could be applied to practical systems without the loss of accuracy in the solutions.
- 1999: After validation tests on the accuracy and consistency of the Herman-Beta Algorithm across numerous benchmark networks, and parallel work on load modelling in South Africa, the HBA was adopted as a national standard for LV feeder design [94]. The guideline, which is also adopted in Kenya [232], caters for three LV distribution topologies;3p-4w, single phase (1p-2w), and bi-phase (2p-3w).
- 2008: Herman and Gaunt [142] further extended the application of the HBA by investigating an approximate approach for estimating load parameters in countries where a large load database is not available. The approach uses the commonly known after-diversity-maximum-demand (ADMD) parameters for an identified target group and estimates the corresponding Beta parameters from established variance-ADMD curves. These curves, which depict the correlation between variance (dispersion) and ADMD values according to customer class, were derived through the statistical analysis of the large datasets available from conducted load research.
- 2010/11: Gaunt et al. [233] extended the approach to active systems with distributed generation . The extension was premised on modelling DG as negative loads, also described as beta currents. To maintain the algebraic identity, the DG nodes were separated from the load nodes by ‘voltage-drop-insignificant’ feeder sections of about 0.5 metres. The details of this work are developed in a thesis by Namanya [234] which led to two journal publications: one describing the PLF algorithm for active systems [226], and another demonstrating its use in the analysis of the limits of PV-DG penetration [132]. The extended

transform was later deployed on a C# software platform capable of handling feeders with up to a 1,000 nodes and many different load and DG models [228].

- 2011: Gaunt et al. [235] studied the impacts of design risk on the outcomes of probabilistic network design for LV systems. The objectives of the work, which is detailed in a thesis by Kadada [91], were two-fold: the derivation of design parameters aligned with the quality of supply (QoS) standards [89] and building a universal framework by which network designers are able to assess the appropriate design parameters specified to customer classes. The research provided the acceptance criterion for probabilistic voltages in line the South African design standards. The compliance criterion stipulates a 95% confidence or 5% design risk for voltage regulation on LV, MV, and HV networks.
- 2013: Siebert et al. [229] applied the HBA algorithm to support decision making processes for self-healing in smart grids, aiming to improve reliability indexes. In the application, two real composite MV-LV networks in Rio de Janeiro were used to assess the performance of the HBA in comparison to the MCS and a deterministic approach. The results demonstrated the HBA was more suitable for fast decision making despite the associated errors (which measured up to 6% for bus voltages at two standard deviations). As expected, the errors emanate from the neglect of reactance and reactive power which were significant on the tested networks.
- 2015: Inspired by the earlier reported research by Siebert et al. [229] , Chihota et al. [10], [11] investigated the extension of the approach to medium voltage feeders with significant reactance and reactive power flow. The suggested approach was based on an approximate technique which modelled the complex impedance through its absolute value, and the complex power likewise. The approach was meant to only compensate for the reactance and reactive power flow initially neglected. While the method performed significantly better than the original approach, it results in significant (up to 6% of the true voltage values) and inconsistent errors.

4.2 THE MATHEMATICAL FOUNDATIONS

The Herman-Beta method for calculating voltage-drop operates as a transform, manipulating beta-distributed input variables of current into beta-distributed output variables of voltage. The transform is based on the method of moments which solves for the distribution parameter of the outputs based on unbalanced LF equations. Using the derived parameters, the output functions can be determined directly, including the interpretation using risk margins or confidence intervals. The subsequent sections detail the components of the HB algorithm and their mathematical basis.

4.2.1 The Modelling of Inputs

The key characteristic of the HB approach is that it is based on the beta probability density function as the sole descriptor of inputs. The selection of the beta PDF over other distributions, which competitively fit the load data, is premised mainly on the versatility of the distribution. The characteristics of the beta distribution are discussed hereafter.

Properties of the Beta Distribution

The beta distribution is a family of probability curves, in continuous mode, bounded exclusively between two defined intervals [a, b]. The shape of the distribution is controlled by two positive shape parameters, denoted by α and β, while its range is dependent on the parameters *a* and *b*. For the general form of the distribution, the parameters *a* and *b* are fixed to 0 and 1, respectively. In this form, the distribution is called a two-parameter beta distribution. While its applications are numerous, for cases involving variables confined to that interval, the set limits restrict its applications. The four-parameter variant which can be applied to any closed interval $[a, b]$ is extensible and more useful.

The characteristic function for a random variable *x* which follows a beta distribution defined over the range $[a, b]$, and with shape parameters $\alpha, \beta > 0$, is given by:

$$f(x) = \frac{(x-a)^{\alpha-1}(b-x)^{\beta-1}}{(b-a)^{\alpha+\beta-2}} \frac{1}{B(\alpha,\beta)} \qquad (4.1)$$

where $B(\alpha, \beta)$ signifies the beta function whose value is equivalent to the gamma function $\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}$.

For such a distribution $X \sim \text{Beta}(\alpha, \beta, a, b)$ defined over a domain $[a, b]$, the variables for the measure of central tendency, dispersion, and moments, are calculated as follows:

- Expected value (mean),

$$E(X) = a + \frac{\alpha}{\alpha+\beta}(b-a) \qquad (4.2)$$

- Variance,

$$var(X) = E[(X-\mu)^2] = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}(b-a)^2 \qquad (4.3)$$

- Standard Deviation,

$$\sigma(X) = \sqrt{var(X)} \qquad (4.4)$$

- Skewness,

$$\gamma = \frac{2(\beta-\alpha)}{\alpha+\beta+2}\sqrt{\frac{\alpha+\beta+1}{\alpha\beta}} \qquad (4.5)$$

The versatility of the Beta distribution

The beta distribution is capable of modelling variously distributed data defined over any two limits. The limit parameters a and b offer flexibility in terms of the applicable domain, while the shape parameters α and β offer further flexibility on dispersion. Depending on the comparative values between the two shape parameters, the beta PDF can assume a right (or positively) skewed, symmetrical, or left (or negatively) skewed distribution shapes. Furthermore, the distributions can either be unimodal or bi-modal.

In statistical terms, positively-skewed distributions are those whose mean is higher than the median, while for negatively-skewed distributions, the median is higher than the mean. For a symmetrical distribution, the median and the mean are equal. Figure 4.1 shows a positively-skewed distribution typical of an urban residential customer load (parameters: a = 1.23, b = 5.56; over a circuit breaker of 60 amps) according to the South African design standards [94].

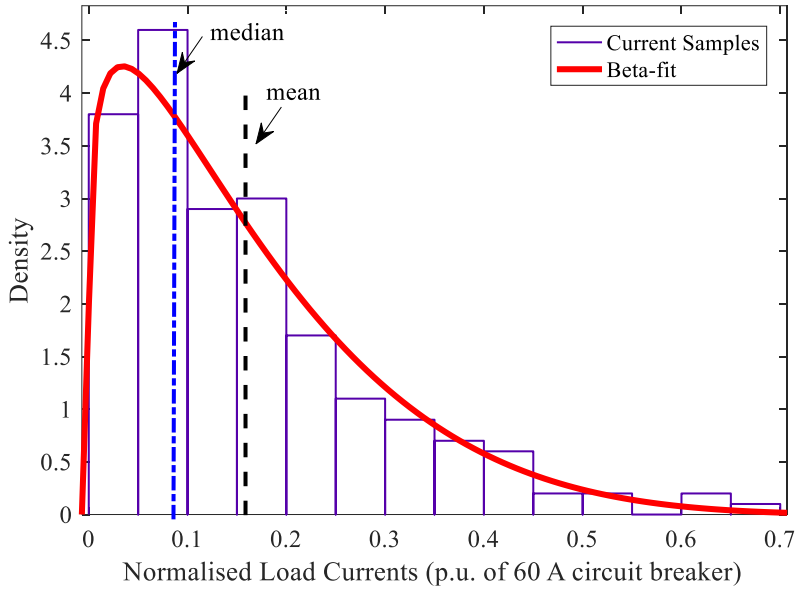

Figure 4.1: Right (positive) skewed distribution

To demonstrate the extent to which the beta PDF is flexible, various cases covering the full spectrum of the possible relationships between the two shape parameters are discussed:

Case 1: $\alpha \neq \beta$

- Unequal shape parameters result in skewed distributions. The functions are left-skewed when $\alpha > \beta$ and right-skewed when $\beta > \alpha$. Further, the magnitudes of the shape parameters determine the specific shape the skewed distributions can assume. If both parameters are less than 1, the distributions are U-shaped and bimodal (a mode on either limit). When both are greater than 1, the distributions are unimodal (solid red and broken green traces of Figure 4.2). Then, when only one of the parameters is less than one, as with the cases ($\alpha \geq 1, \beta <$

1) and $(\alpha < 1, \beta \geq 1)$, J-shaped distributions result. Depending on the relational magnitudes of the parameters under these conditions, left and right-triangular distributions, as well as the power distribution can be modelled.

- The various forms of these distributions are convenient for the modelling of practical variables which demonstrate significant skew. For instance, residential loads can be modelled by unimodal positive-skewed distributions, while the power output from PV-DG plants during noontime can be highly negative-skewed. On the other hand, U-shaped distributions are often characteristic of the wind power output from turbines facing highly variable wind speeds [236]. Likewise, rapid cloud transients on PV arrays introduce high variability in the irradiance incident on the panels. As a result, the power output fluctuates between 'highs' and 'lows' such that the distributions are U-shaped.

Case 2: $\alpha = \beta; (\alpha, \beta) < 1$

- An equality condition between the shape parameters results in symmetrical distributions. And, if both parameters are less than one, the symmetrical distributions are U-shaped. The special case ($\alpha = \beta = 0.5$) yields an arcsine distribution (broken blue trace of Figure 4.2). If the shape parameters are reduced further such that $(\alpha, \beta) \rightarrow 0$, the distribution takes the form of a two-point Bernoulli distribution. The density function has two Dirac functions at both extremities (x = a and x = b), while zero elsewhere. Figure 4.3 (b) shows such a distribution. Bernoulli distributions are often used to model discrete RVs which can take either of two forms, a zero or a one. Accordingly, they can be manipulated to model power system variables such as the failure rates of transformers and the probability of line outages.

Case 3: $\alpha = \beta = 1$

- Increasing the shape parameters towards unity diminishes the modes and raises the probabilities values between the two extremities, until the distribution is flat. The resulting mode-less distribution (depicted by the black dotted trace of Figure 4.2), with an equal chance of occurrence for all x, equivalently models a uniform distribution with variance $var(X) = \frac{1}{12}(b-a)^2$ defined over a domain $[a, b]$. As discussed in Chapter 2, uniform distributions are relevant for the characterisation of model parameters such as line resistance and reactance whose uncertainty is commonly modelled using intervals.

Case 4: $\alpha = \beta; (\alpha, \beta) > 1$

- Beyond unity, the equality condition results in unimodal symmetrical distributions whose means equal both the mode and median. Moving from 1 to infinity, the distributions start off dome-shaped, turn parabolic as the parameters are increased, and eventually resemble truncated Gaussian distributions. When tending to infinity ($\alpha = \beta \rightarrow \infty$), the dispersion

diminishes until the distributions resemble a Dirac function located at $x = 0.5$. For this distribution, shown in Figure 4.3(a), the probability of occurrence is concentrated at the mid-point and is zero elsewhere. The resulting invariable function offers means to model deterministic elements such as shunt capacitance.

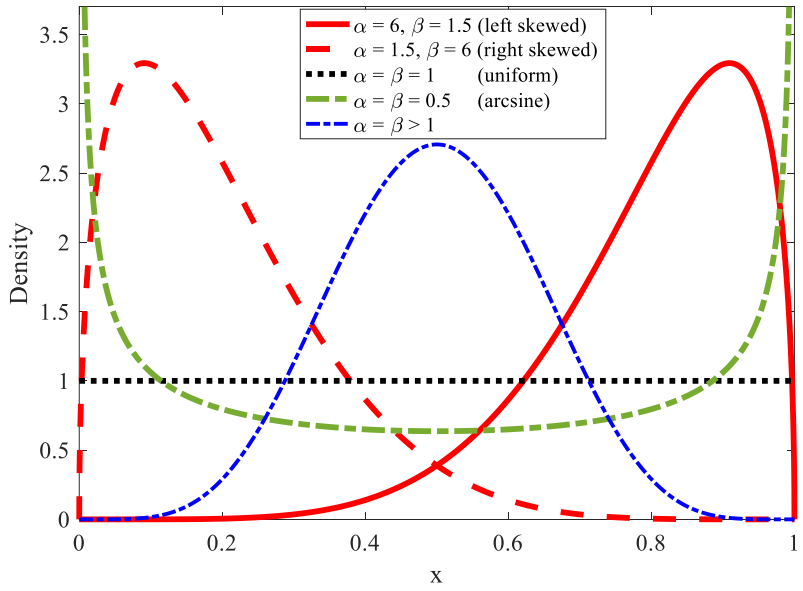

Figure 4.2: Beta distributions for variables with significant variability

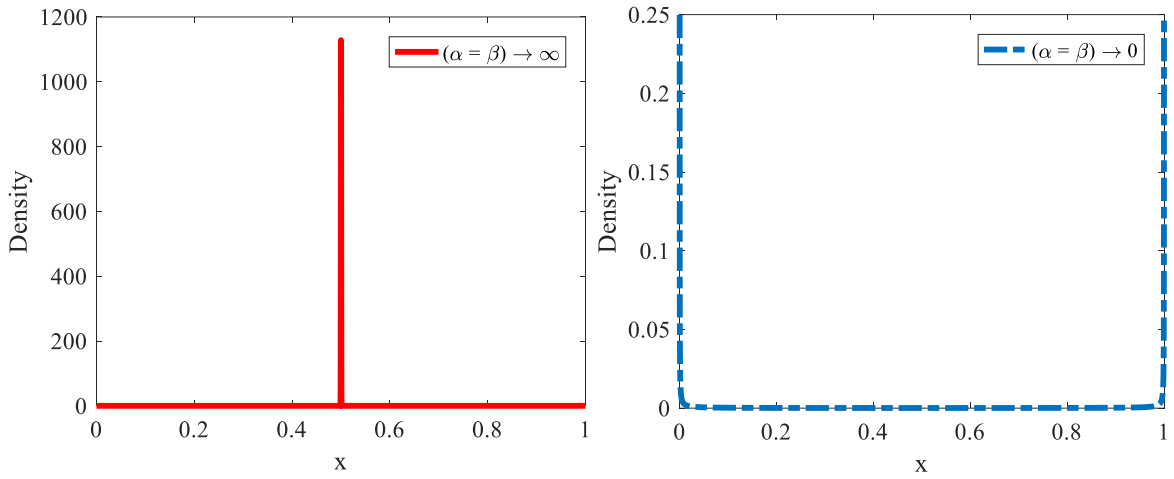

Figure 4.3: Beta distributions for (a) deterministic variables, and (b) discrete variables

The basis for choosing the beta PDF

Considering the discussed attributes of the beta PDF, the motivations for choosing the distribution for the application of load modelling and probabilistic load flow are summarised below:

1. The beta parametric distribution is bounded within selected limits a and b; matching the practical expectation of loads falling between zero (no load) and a maximum (full-load) most likely restricted by the size of the customers' circuit breaker. In contrast, the use of other distributions, especially those with infinite extremum limits, can unnecessarily increase the computational burden in some PLFs [51], [167] as unlikely and impractical load conditions are also assessed. Where the issue is addressed, the process of truncation is often used to make practical the abstract models [148]. However, this comes with computational expense.

2. The versatility of the beta distribution to model variously skewed data, including symmetrical data, gives it a great advantage. The distribution can take various shapes which can cater for a wide range of variables: (i) symmetrical and positively-skewed load demand, (ii) skewed and U-shaped (considering cloud transients) PV generation distributions, (iii) skewed unimodal and bimodal U-shaped wind distributions (iv) uniformly distributed model parameters such as resistance and reactance, (v) failure or outage rates for transformers and line sections using equivalent models to discrete functions, and (vi) deterministic models for insignificantly variable quantities such as shunt capacitance. The demonstrated flexibility makes possible the use of a universal distribution to model a wide range of RVs. This is powerful in avoiding the computational complexity associated with the use of several distribution families in a single PLF tool.

3. The statistical properties (such as the mean, mode, skewness, moments, and kurtosis) of a beta PDF are fully captured by only two parameters α and β in addition to the limit parameters a and b which allow for scalability. The low number of parameters is convenient for load modelling and the derivation of the statistical solutions.

Based on these arguments, the HB transform takes as inputs, beta distributed DG and load currents at nominal voltage and unity power factor. The description is through a 4-parameter beta distribution, $X \sim Beta(\alpha, \beta, 0, C)$, where 0 and C are the limit parameters. The parameter C is usually taken as the circuit breaker rating suppling the customer or otherwise as a general scaling constant. With the parameters to the input functions known, it is possible to solve the PLF using the method of moments.

4.2.2 Solving the PLF Problem Using Moments

The theoretical development of the Herman-Beta algorithm is detailed in several published articles [50], [142], [227], [230], [231]. Here, the detailed statistical and algebraic operations are intentionally left out to only demonstrate the theoretical framework of the approach. However, in the subsequent chapters dealing with the formulation of the extended algorithm, the finer details of the derivation are exposed.

While the HBA has been formulated for various feeder topologies, the demonstration here will be limited to the 3-ph, 4-wr model depicted in Figure 4.4.

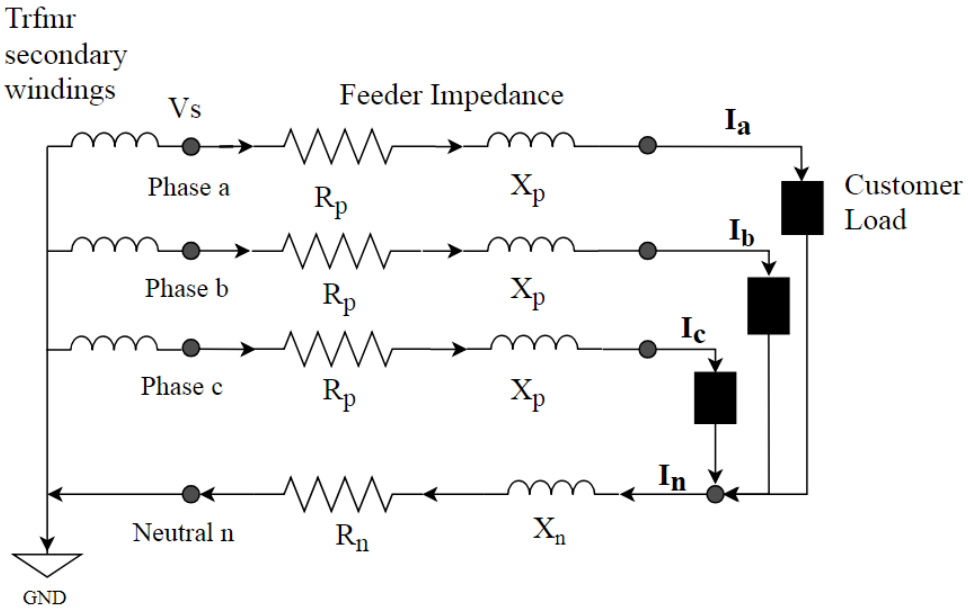

Figure 4.4: Three-phase, four-wire LV feeder model

As a base case, the feeder is simplified to a single node, and without DG. Thereafter, the formulation is extended to successive nodes, and the interconnection of DG. In the mathematical representations, the bold-face is used to separate complex variables from absolute quantities. This notation is consistent throughout this report.

On a radial feeder, voltage change (drop or rise) can be calculated simply by considering the product of the current (line) flow and the impedance of a section under analysis according to (4.6).

$$\Delta \boldsymbol{V} = \boldsymbol{IZ} \qquad (4.6)$$

Where $\boldsymbol{Z}$ denotes the impedance of the feeder segment while $\boldsymbol{I}$ is the summated load current from the individual customer currents connected to that node. Suppose there are m_a, m_b and m_c customers connected to the respective phases a, b and c, then the sum of the currents in each phase may be expressed by (4.7) – (4.9):

$$\boldsymbol{I_a} = \boldsymbol{I}_1 + \boldsymbol{I}_2 + \cdots + \boldsymbol{I}_{ma} \qquad (4.7)$$

$$\boldsymbol{I_b} = \boldsymbol{I}_1 + \boldsymbol{I}_2 + \cdots + \boldsymbol{I}_{mb} \qquad (4.8)$$

$$\boldsymbol{I_c} = \boldsymbol{I}_1 + \boldsymbol{I}_2 + \cdots + \boldsymbol{I}_{mc} \qquad (4.9)$$

Where $I_i \sim Beta(\alpha, \beta, 0, C)$.

The neutral conductor current is expressed in phasor terms as:

$$\boldsymbol{I_n} = \boldsymbol{I_a} + \boldsymbol{I_b} + \boldsymbol{I_c} \qquad (4.10)$$

Applying the analysis of voltage-drop on the red phase (phase A) according to (4.6) results in the following expression:

$$\Delta \boldsymbol{V} = \boldsymbol{I_a}(R_p + jX_p) + \boldsymbol{I_n}(R_n + jX_n) \quad (4.11)$$

Where $\boldsymbol{I_a}$, $\boldsymbol{I_n}$, R_a, R_n, X_a, X_n denote conductor currents, resistance/wire and reactance/wire for phase A and the neutral, respectively.

And, where the cross-section of the phase and neutral wires are different, it is useful to define scaling factors:

$$k_r = \frac{R_n}{R_p}, and\, k_x = \frac{X_n}{X_p} \quad (4.12)$$

With unbalanced loads on the feeder, a current in the neutral conductor is associated with either a voltage-drop or rise. Allowing for the neutral current and the angular shifts between the phase currents, (4.6) can be extended and separated into quadrature components, ΔV_{i_x} and ΔV_{i_y} given by:

$$\Delta V_x = \boldsymbol{I_a} R_p(1 + k_r) - \frac{1}{2}(\boldsymbol{I_b} + \boldsymbol{I_c}).k_r.R_p - \frac{\sqrt{3}}{2}(\boldsymbol{I_b} - \boldsymbol{I_c}).k_x.X_p \quad (4.13)$$

$$\Delta V_y = \boldsymbol{I_a} X_p(1 + k_x) - \frac{1}{2}(\boldsymbol{I_b} + \boldsymbol{I_c}).k_x.X_p + \frac{\sqrt{3}}{2}(\boldsymbol{I_b} - \boldsymbol{I_c}).k_r.R_p \quad (4.14)$$

Suppose the loads connected to each phase operate at a power factor $cos\,\phi_{(a,b,c)}$, the phase currents become complex according to (4.15).

$$\boldsymbol{I}_{(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})} = I_{(a,b,c)}\, cos\,\phi_{(a,b,c)} \pm j\, I_{(a,b,c)} sin\,\phi_{(a,b,c)} \quad (4.15)$$

Once the voltage change is determined, the bus voltage at the receiving end or consumer point of connection, V_{con} can be calculated as follows:

$$\boldsymbol{V_{con}} = \boldsymbol{V_S} - \boldsymbol{\Delta V} \quad (4.16)$$

Where, V_S is the sending end voltage, usually taken to be the secondary side transformer bus voltage.

Considering the sending end voltage as the reference voltage, and using the complex components of voltage-drop from (4.13) and (4.14) leads to the calculation of the absolute voltage through the following:

$$\boldsymbol{V_{con}} = (V_s - \Delta V_x) - j(\Delta V_y) \quad (4.17)$$

$$|\boldsymbol{V_{con}}| = V_{con} = \sqrt{(V_s - \Delta V_x)^2 + (\Delta V_x)^2} \quad (4.18)$$

The probabilistic voltage calculation requires the computation of the distribution parameters $\alpha_{V_{con}}$ and $\beta_{V_{con}}$, and its limits. Accordingly, the HBA approach is premised on solving for the four statistical parameters of the absolute value of the consumer voltage given by (4.18).

Based on the properties of the beta PDF discussed in the preceding sections, the shape parameters are computable from any of the characteristic equations (4.2) – (4.5). Assuming the limits can be estimated, solving for the two unknown shape parameters requires a system of two or more simultaneous

equations. For algebraic simplicity, avoiding terms of high orders, the first two moments of V_{con} are chosen; the mean and the second raw moment. The selected moments can be determined by computing the expected value of the functions of V_{con} and V_{con}^2. However, the non-linearity of (4.18) introduces computational burden. As a result, linearization using a Taylor's expansion truncated at the second term is applied to get (4.19). The square of the consumer voltage is computed exactly from (4.19) and is given by (4.20).

$$V_{con} \approx V_s \left(1 - \frac{\Delta V_x}{V_s} + 0.5 \left(\frac{\Delta V_y}{V_s}\right)^2\right) \quad (4.19)$$

$$V_{con}^2 = V_S^2 - 2.V_s.\Delta V_x + \Delta V_x^2 + \Delta V_y^2 \quad (4.20)$$

The expected value of each of these equations ((4.19) and (4.20)) lead to the first two moments for the consumer voltage given by (4.21) and (4.22).

$$E(V_{con}) \approx V_s \left(1 - \frac{E(\Delta V_x)}{V_s} + 0.5 \frac{E(\Delta V_y^2)}{V_S^2}\right) \quad (4.21)$$

$$E(V_{con}^2) = V_S^2 - 2.V_s.E(\Delta V_x) + E(\Delta V_x^2) + E\left(\Delta V_y^2\right) \quad (4.22)$$

Equations (4.21) and (4.22) indicate that the statistical moments of the outputs are dependent on the expected values of voltage-drop, which are in turn dependent on the input RVs as depicted by (4.13) and (4.14). The first moment can be achieved fairly easily by calculating the expected moment of the quadrature components of voltage-drop. However, the second moment requires the square of both (4.13) and (4.14).

4.2.3 Simplifications in the HBA

With the complex currents of (4.15) substituted into (4.13) and (4.14) the equations become much more complex and challenging. As a result, the scale of the problem was reduced by applying the following assumptions:

1. The loads are assumed to operate at unity power factor thereby not capable of injection or absorption of reactive power. The result is that the complex currents of (4.15) become real currents $I_{(a,b,c)}$.
2. The impedance of the lines is assumed to be completely resistive at a specified temperature. Consequently, the terms involving X_p are dropped.
3. The random variables are assumed to be totally independent. As a result, no correlation is built into the statistical analysis of the equations.

Applying these assumptions, the voltage-drop equations are simplified as follows:

$$\Delta V_x = I_a R_p (1 + k_r) - \frac{1}{2}(I_b + I_c).k_r.R_p \quad (4.23)$$

$$\Delta V_y = \frac{\sqrt{3}}{2}(I_b - I_c).k_r.R_p \qquad (4.24)$$

After following a fairly tedious analysis applying principles of algebra and parametric statistics, the moments of the quadrature components of the voltage-drop equations are determined.

$$E(\Delta V_x) = C_1.R_p.C.G \qquad (4.25)$$

$$E(\Delta V_y) = C_2.R_p.C.G \qquad (4.26)$$

$$E(\Delta V_x^2) = R_p^2 C^2(C_3.H + C_4.G^2) \qquad (4.27)$$

$$E(\Delta V_y^2) = R_p^2 C^2(C_5.H + C_6.G^2) \qquad (4.28)$$

Where the constants C_1 to C_6 (provided in Appendix B1) depend on the feeder properties, the numbers of customers connected to the phases and the normalised moments of the input currents, G and H.

Finally, the solutions to (4.25) – (4.28) are plugged back into the expressions (4.21) and (4.22) to determine the first two moments of the outputs.

4.2.4 Construction of Outputs

The anticipated output functions are beta-distributed voltages characterised by four parameters according to (4.29):

$$V_{con} \sim Beta(\alpha_{V_{con}}, \beta_{V_{con}}, V_{min}, V_{max}) \qquad (4.29)$$

where $\alpha_{V_{con}}$ and $\beta_{V_{con}}$ are shape parameters and V_{min} and V_{max} are the limits of the beta functions.

The functions are transformed versions of the standardised beta function, $Beta(\alpha_{V_{con}}, \beta_{V_{con}}, 0,1)$ according to:

$$V_{con}^* = \frac{V_{con} - V_{min}}{V_{max} - V_{min}} \qquad (4.30)$$

From (4.30) the moments of the normalised voltage, V_{con}^*, can be determined:

$$E(V_{con}^*) = \frac{E(V_{con}) - V_{min}}{V_{max} - V_{min}} = \frac{\alpha_{V_{con}}}{\alpha_{V_{con}} + \beta_{V_{con}}} \qquad (4.31)$$

$$E({V_{con}^*}^2) = \frac{E(V_{con}^2) - 2.V_{min}.E(V_{con}) + V_{min}^2}{(V_{max} - V_{min})^2} = \frac{\alpha_{V_{con}}(\alpha_{V_{con}} + 1)}{(\alpha_{V_{con}} + \beta_{V_{con}})(\alpha_{V_{con}} + \beta_{V_{con}} + 1)} \qquad (4.32)$$

The simultaneous system of equations indicates the shape parameters are a function of the already solved moments for the actual voltage and the unknown limit parameters. Accordingly, the subsequent problem in the construction of the outputs is determining the extremum voltages (limits) of the output functions, V_{min} and V_{max}.

The calculation of Extremum voltages

The expected extremum voltages are limits to the consumer voltage equation in (4.19).

$$V_{min} = min(V_{con}) = min\left(\sqrt{(V_s - \Delta V_x)^2 + \left(\Delta V_y\right)^2}\right) \quad (4.33)$$

$$V_{max} = max(V_{con}) = max\left(\sqrt{(V_s - \Delta V_x)^2 + \left(\Delta V_y\right)^2}\right) \quad (4.34)$$

Analysis of (4.33) and (4.34) using the expanded expressions of (4.23) and (4.24) sets up conditions for minima and maxima of the expected voltages. These limits depend on the system's nominal voltage, and the related voltage changes which in turn depend on the sizes of the loads and levels of unbalance (due to phase loading).

The minimum voltage in each phase, is met when the phase under study is fully loaded whilst the others are unloaded. Expressed mathematically, this is a case when $I_b, I_c = 0$ while I_a= 1, for normalized phase currents. The inverse applies for the computation of maximum voltage; it occurs when the phase under study is unloaded, while the other two phases are fully loaded. These conditions applied to (4.33) and (4.34) leads to the solution of the limit parameters.

$$V_{min} = V_s - (1 + k_r).R_p.C.m_a \quad (4.35)$$

$$V_{max} = \sqrt{\left(V_s + \frac{1}{2}k_r R_p.C.(m_b + m_c)\right)^2 + \left(\frac{\sqrt{3}}{2}.k_r.R_p.C.(m_b - m_c)\right)^2} \quad (4.36)$$

With the limit parameters determined, the shape parameters are easily computable.

Determining the distribution shape parameters

The final step in the statistical characterisation of the outputs involves solving for the shape parameters $\alpha_{V_{con}}$and $\beta_{V_{con}}$. The substitution of the extremum voltages calculated in the preceding steps allows the derivation of simultaneous solution for $\alpha_{V_{con}}$and $\beta_{V_{con}}$ in (4.31) and (4.32).

$$\alpha_{V_{con}} = \frac{E\left(V_{con}^{*2}\right) - E(V_{con}^{*})}{E(V_{con}^{*}) - \frac{E(V_{con}^{*2})}{E(V_{con}^{*})}} \quad (4.37)$$

$$\beta_{V_{con}} = \frac{\alpha_{V_{con}}}{E(V_{con}^{*})} - \alpha_{V_{con}} \quad (4.38)$$

Statistical interpretation

Having solved for the distribution parameters, the probability density functions, or cumulative distribution functions, can be plotted directly. Further, analysis of the plots to evaluate percentile values based on risk is possible using the beta inverse function.

$$V_{con\%}^{*} = betainv\left(\alpha_{V_{con}}, \beta_{V_{con}}, r\right) \quad (4.39)$$

To rescale the normalised voltage, the following transformation is used:

$$V_{con\%} = V_{con\%}^{*} \cdot (V_{max} - V_{min}) + V_{min} \quad (4.40)$$

where $V_{con_{\%}}$ is the percentile consumer voltage value considering a risk of value $r\%$ as applied in (4.40).

4.2.5 Extension to Successive Nodes

Up to this point, the analysis has been limited to feeders with a single node. To extend the approach to multi-nodes, the principle of superposition is applied. The theory implies that if the voltage at a certain point in a circuit is the function of more than one voltage- or current source, the voltage may be written as the sum of the voltages due to each source.

Applied in context, the total voltage-drop on the feeder with N loads (or random variables) located on independent nodes is determined by the summation of the voltage-drops due to the individual loads. To avoid confusion, a notation separating these quantities is adopted; the subscript "i" for variables associated with the individual loads and "t" for the summated quantities.

As a result, the voltage-drop calculations in (4.23) and (4.24), and, the resulting extremum voltages (4.33) and (4.34) are calculated separately for each node. In the respective calculations, the resistance R_p and R_n are substituted with the summated resistances of the feeder sections leading to a selected node i.

$$R_{p_i} = \sum_{k=1}^{i} R_{p_k} \;\; ; \; R_{n_i} = \sum_{k=1}^{i} R_{n_k} \qquad (4.41)$$

With the independent voltage-drop components for each load calculated, the summated voltage-drops, and their moments can be determined using Riemann Sums as follows:

$$\Delta V_{x_t} = \sum_{i=1}^{N} \Delta V_{x_i} \;\; and \; \Delta V_{y_t} = \sum_{i=1}^{N} \Delta V_{y_i} \qquad (4.42)$$

$$E\left(\Delta V_{x_t}\right) = \sum_{i=1}^{N} E(\Delta V_{x_i}) \qquad (4.43)$$

$$E\left(\Delta V_{y_t}\right) = \sum_{i=1}^{N} E(\Delta V_{y_i}) \qquad (4.44)$$

The expected values of the squares of the functions in (4.42) lead to the second moments of the total voltage-drops:

$$E\left(\Delta V_{x_t}^2\right) = \sum_{i=1}^{N} E\left(\Delta V_{x_i}^2\right) + \sum_{j=1}^{N} \sum_{\substack{k=1 \\ k \neq j}}^{N} E\left(\Delta V_{x_k}\right) E\left(\Delta V_{x_j}\right) \qquad (4.45)$$

$$E\left(\Delta V_{y_t}^2\right) = \sum_{i=1}^{N} E\left(\Delta V_{y_i}^2\right) + \sum_{j=1}^{N} \sum_{\substack{k=1 \\ k \neq j}}^{N} E\left(\Delta V_{y_k}\right) E\left(\Delta V_{y_j}\right) \qquad (4.46)$$

The substitution of the summations represented by (4.41) – (4.46) in the computation of the statistical properties of the outputs given in (4.21) and (4.22), and summation of the extremum voltages due to each node, completes the solution for the multi-node feeder.

From the mathematical illustrations presented in (4.41) – (4.46), it is apparent that the singular node case that has been used to illustrate the properties of the HBA is a special case $N = 1$. As such, besides

the changes discussed here, the process of calculations detailed in the preceding Sections 4.2.2– 4.2.4 remain the same.

4.2.6 Solving Feeders with DG

The analysis thus far has been on passive systems without distributed generation. The introduction of DG poses two challenges to the conventional way of calculating voltage-drop, particularly the way the HBA algorithm is formulated:

1. DG currents flow in the opposite direction to that of the loads. As a result, the voltage-changes in reverse cause voltage-rise as opposed to voltage-drop. Consequently, the polarity of voltage-drops is negative in the case of DG.
2. The load parameters for DG are expected to be different due to the distinct load behaviour different from that of the loads. Thus, the effect of loads and DGs cannot be readily summed. To include the voltage-rise effects of DG, the application of superposition is extended to DG nodes.

In the light of these considerations, DG injections are modelled as negative beta currents (negative loads). And, with the anticipated differences in the voltage-change impacts, DG nodes are separated from load nodes using voltage-change insignificant spurs of length 0.1 m. This helps maintain algebraic identity in the voltage-drop elements, as well as allowing separate models for the two. The outcomes of voltage-rise calculations from DG nodes are included in the superposition calculations for whole feeder. However, a few modifications in the HBA are needed:

- Generators connected to a node are represented as 'negative customers.'
- In the calculation of the second moments using (4.27) and (4.28), the expectation of the squared currents is used. As a result, the polarity of the customer variables m_{ai}, m_{bi} and m_{ci} must be positive. To achieve this, the negative variables are replaced with their absolute values for the calculation of constants C_3 to C_6.
- As a result of the reverse power flow, the conditions for the calculation of the extrema voltages are reversed. Minimum voltages are likely to occur on a phase with no generation, while the other phases are at peak generation. Conversely, the maximum voltage occurs when a phase is generating at its peak, while the other phases have no generation. Equations (4.48) and (4.49) accommodate the required changes.

$$V_{min} = \sqrt{\left(V_s + \frac{1}{2}k_r R_p . C . (m_b + m_c)\right)^2 + \left(\frac{\sqrt{3}}{2} . k_r . R_p . C . (m_b - m_c)\right)^2} \qquad (4.47)$$

$$V_{max} = V_s - (1 + k_r) . R_p . C . m_a \qquad (4.48)$$

The description of the HBA provided in this chapter served the purpose of exposing the theoretical framework underlying the approach. A comprehensive guideline to the calculations of voltage-change on LV feeders for three-phase and bi-phase systems is provided in [11].

4.3 SIMULATION CASE STUDIES AND DISCUSSION

This section aims at assessing the performance of the HBA algorithm, and demonstrate its practical strengths, weaknesses, and limitations.

Case studies are conducted to demonstrate the performance firstly under the conditions for which it was developed (Cases 1 and 2), and then extended to other practical conditions which challenge the key assumptions in the method (Cases 3 and 4). A Monte-Carlo simulation with 15,000 trials is used to validate the outcomes of the algorithms, tested under a common software environment in MATLAB.

Three test statistics, the mean, standard deviation and percentile voltage at 5% risk, expressed on the base of the slack bus are used to assess the accuracy of the method. To quantify the accuracy, a relative percentage error index, ε_{HBA}^{t} given by (4.48) is used. The statistic measures the relative difference between the outcomes of the HBA, t_{HBA}, and those from the MCS, t_{MCS}.

$$\varepsilon_{HBA}^{t} = \left|\frac{t_{MCS}-t_{HBA}}{t_{MCS}}\right| \times 100\ \% \qquad (4.49)$$

It is important to indicate the anticipated error margin in the repeated calculations of the Monte-Carlo simulation. In that way, it gives a clear perception of the accuracy of the method under test.

For a MCS with *N* samples, the expected error associated with the estimation of the variable *t* is proportional to the standard deviation of the samples and the square root of the conducted trials as follows:

$$E(\varepsilon_{MCS}) = \frac{\sigma_{MCS}^{t}}{\sqrt{N}} \qquad (4.50)$$

Using (4.50), it follows that for an outcome with a 5% standard deviation, an MCS with 15,000 samples has a mean error of approximately 0.04%.

For all cases, a representative 8-bus radial test system with no laterals is used. Table 4.1 provides the original properties of the test system according to voltage levels. However, where necessary, modifications are applied to suit the conditions of the test cases.

Table 4.1: Feeder and load properties for test cases

TEST SYS.	FEEDER PROPERTIES							LOAD PROPERTIES				
	Line Voltage (kV)	Cond. Type	R (Ω)	X (Ω)	X/R ratio	A	Branch Length (m)	α	β	C	ADMD (kVA)	pf
LV	0.4	Al35	0.944	0.0895	0.0949	108	30	1.50	4.00	60	3.764	1.00
MV	19	Al300	0.1187	0.0825	0.6953	200	2000	1.50	4.00	200	600	0.80

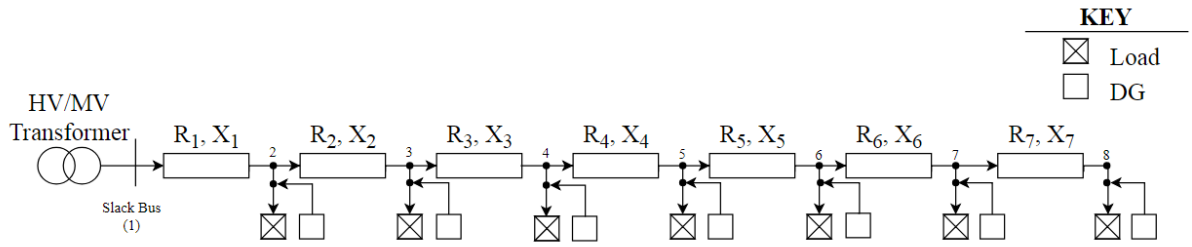

Figure 4.5: 8-bus representative test network

For each of the network models, the customers on the test feeder are assumed to belong to the same load group such their statistical load parameters are shared. However, the behaviour of each customer in the group is assumed independent. Where DGs exists, the statistical parameters are differentiated from the loads, but common between the DGs. The assumption of independency is extended and applies to both load-DG and DG-DG relationships.

Across all tests, the connected loads or DG are unbalanced. The specific phase distribution of the loads, including the properties provided in Table 4.1 are detailed in Appendix B1.

4.3.1 Base Case: Passive LV Feeder ($X = 0$, $Q = 0$) with Unbalance

The aim of this investigation is to evaluate the performance of the HBA under the conditions for which it was initially developed; resistive feeders, unity power factor loads, and no DG.

The properties of the test system are provided in Table 4.1 under the 'LV' row. Aluminium cables with a cross-section are of 35 mm^2 and low X/R ratio are used in the model. The loads are separated by short distances of 30 metres to model distribution feeders in low income communities which are commonly densely populated. The results of test compared to outcomes from a MCS with no assumptions are discussed hereafter.

The plot of Figure 4.6 shows the trend of mean voltages on the feeder. As expected on a passive radial feeder, the bus voltages fall with electrical distance from the source. The trend of voltages meets the theory expectations and confirms the correct implementation of the network theory. The comparison of the voltage traces between the MCS (blue trace) and the HBA (red trace) demonstrates the accuracy of the HBA. The voltage plots are nearly indistinct, especially for bus voltages closer to the source. However, moving away from the source toward the terminal node, little mismatches are observed. This is due to the amounting impact of reactance, assumed insignificant in the HBA, as the electrical distance from the source increases. Nonetheless, the observed error is very small Table 4.2 provides a quantitative analysis of the results. For mean voltages, the maximum error, expected to occur at the terminal node due to the linear nature of the error, is only up to 0.032%.

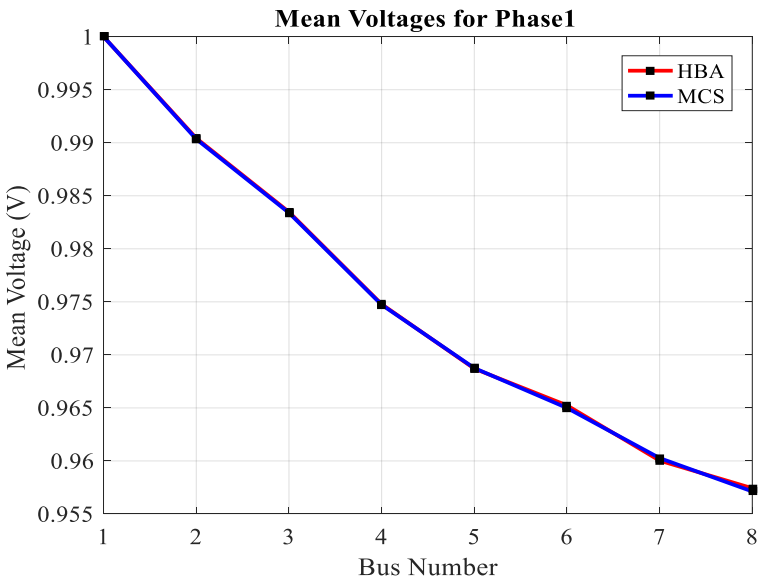

Figure 4.6: Mean bus voltages (Case 1, Red phase)

Table 4.2: Comparison of voltage outcomes between the MCS and HBA methods – Case 1

Bus ID	MCS Outcomes			HBA Outcomes			Relative Error (%)		
	E(V)	σ(V)	Vp%	E(V)	σ(V)	Vp%	E(V)	σ(V)	Vp%
1	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	0.000	0.000	0.000
2	0.9903	0.0053	0.9812	0.9904	0.0053	0.9818	0.008	0.093	0.057
3	0.9834	0.0097	0.9664	0.9835	0.0097	0.9674	0.008	0.265	0.108
4	0.9747	0.0143	0.9500	0.9748	0.0143	0.9513	0.005	0.342	0.136
5	0.9688	0.0179	0.9378	0.9687	0.0179	0.9393	0.005	0.353	0.163
6	0.9650	0.0206	0.9296	0.9652	0.0205	0.9315	0.025	0.350	0.201
7	0.9603	0.0230	0.9206	0.9600	0.0232	0.9219	0.025	0.895	0.135
8	0.9571	0.0247	0.9132	0.9574	0.0246	0.9169	0.032	0.575	0.410

The comparison of the standard deviation is illustrated in Figure 4.7. The MCS and the HBA traces are comparable in a similar manner to what was observed for mean voltages in Figure 4.6. However, the trends of the plots differ. As illustrated in Figure 4.7, the standard deviation increases with electrical distance. This is consistent with the theoretical expectations for a direct proportion between the standard deviation of a RV and its scaling coefficients. In this case, the scaling coefficients are the impedance properties of the feeder.

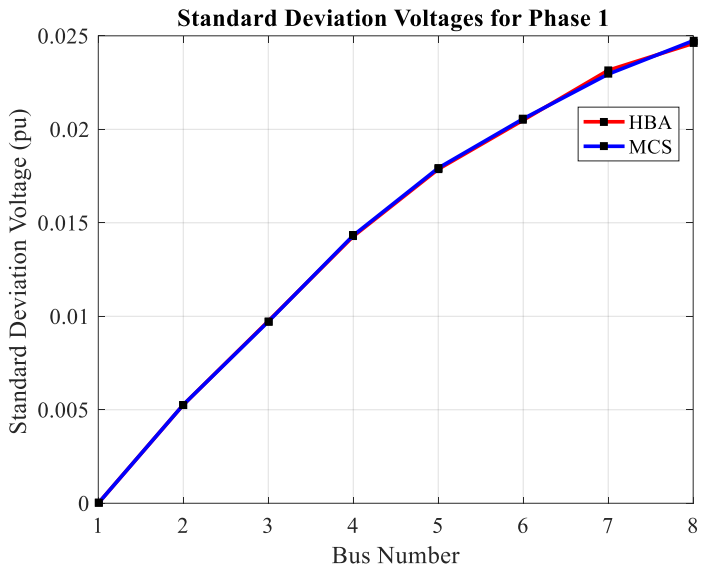

Figure 4.7: Standard deviation of bus voltages (Case 1, Red phase)

The errors associated with the standard deviation outcomes are also provided in Table 4.2. The maximum error recorded is under 1%. The accuracy in the SD values suggests the HBA-PDFs closely model the output data. The plot of Figure 4.8 shows the PDF of bus voltages for bus 8, plotted on the same axes with the histogram of voltages from the MCS. The beta PDFs fit the MCS histograms fairly well. However, similar to the deductions in Chapter 3 on input models, fitting a parametric PDF to the output data will result is some error. To test the fitness of the output PDFs, a quantitative assessment of the outputs (mean, variance and percentile values), as performed in Table 4.2, can be used. Figure 4.8 validates two aspects of the HBA approach. The differentiated PDFs for the phase voltages demonstrate the correct implementation of the unbalanced load flow approach while the good fit of the HBA PDF to the MCS data confirms the accuracy.

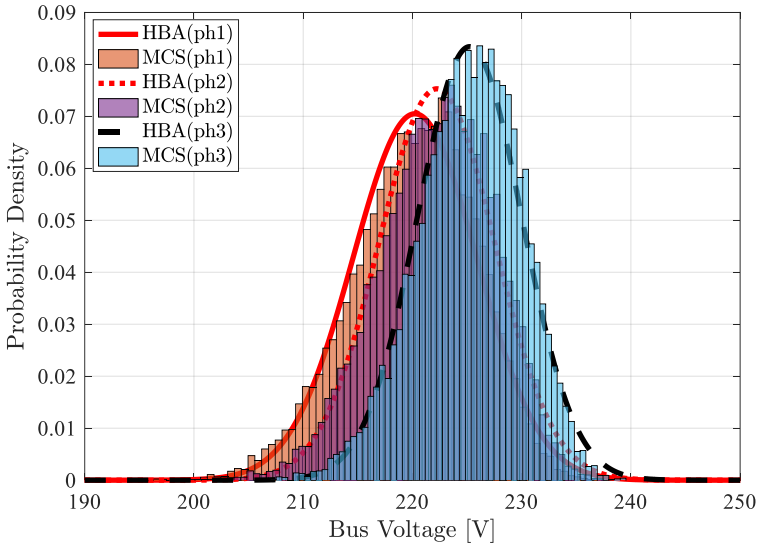

Figure 4.8: Distribution of bus voltages (Case 1, bus 8, All phases)

Further tests on the accuracy of the HBA-PDFs to fit the MCS data is offered through the comparison of the percentile voltages in Table 4.2. While the observed errors are higher than those for mean voltages, they are still considerably low. The highest error is 0.4% which indicates a relative difference of about 0.0038 pu.

The test results demonstrate that the HBA performs reasonably well when applied to LV feeders with negligible reactance and unity pf loads. The test is extended to active feeders in the consecutive section.

4.3.2 Case 2: Active LV Feeder (X = 0, *Q* = 0, DG)

This case assesses the validity of the extension of the HBA approach to active networks. In the test, the network properties as used in the base case are maintained while DG allocations, each per connected customer, are made. The parameters of the connected DG are provided in Table 4.3.

Table 4.3: Statistical DG parameters for Test Network 1

TEST SYSTEM	DG PROPERTIES				
	α	β	C	ADMD (kVA)	pf
LV	7.00	3.50	30	4.600	1.00

The selection of the DG parameters for is such that the distributions are left-skewed (similar to PV-DG generation profiles) and have the maximum possible capacity for single phase connections in South Africa. The PLF analysis is undertaken in the same way as the base case save for the design risk which when split between the voltage-drop and rise limits becomes 2.5% for each limit.

Figure 4.9 shows the trend of bus voltages for the system with DG.

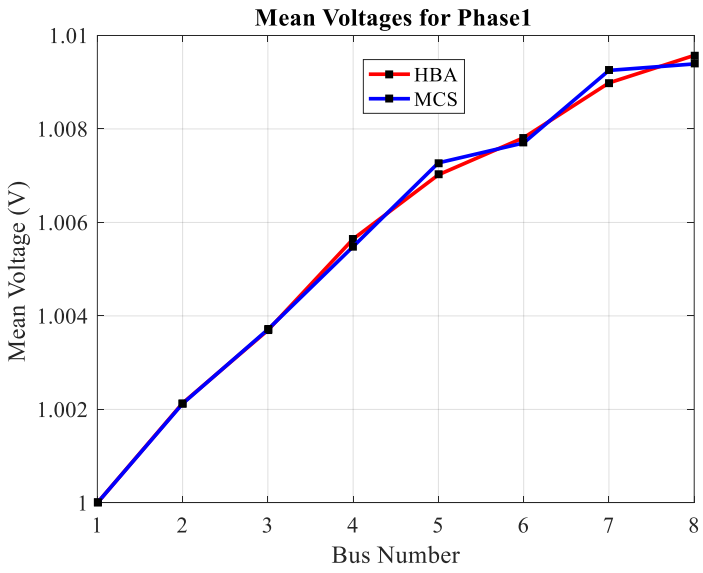

Figure 4.9: Mean bus voltages (Case 2, phase 1)

The trend of voltages observed in Figure 4.9 differs from that of Figure 4.6 in that instead of voltage-drop, voltage-rise occurs. Since the capacity of generation (4.6 kVA) exceeds the load demand (3.8 kVA) at each node, the net load is negative. As a result, the feeder operates similar to a purely active feeder. The consistent voltage-rise profile observed meets this expectation. Further, the aligned traces from the two methods and the quantitative analysis provided in Table 4.4 demonstrate comparable accuracy to that achieved for the passive case.

Table 4.4: Comparison of voltage outcomes between the MCS and HBA methods – Case 2

Bus ID	MCS Outcomes			HBA Outcomes			Relative Error (%)		
	E(V)	σ(V)	Vp%	E(V)	σ(V)	Vp%	E(V)	σ(V)	Vp%
1	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	0.000	0.000	0.000
2	1.0021	0.0057	1.0126	1.0021	0.0057	1.0132	0.001	0.193	0.060
3	1.0037	0.0105	1.0233	1.0037	0.0105	1.0242	0.002	0.135	0.091
4	1.0055	0.0154	1.0339	1.0056	0.0154	1.0357	0.016	0.382	0.171
5	1.0073	0.0191	1.0424	1.0070	0.0192	1.0446	0.025	0.482	0.216
6	1.0077	0.0219	1.0480	1.0078	0.0221	1.0509	0.011	0.640	0.280
7	1.0093	0.0248	1.0548	1.0090	0.0249	1.0577	0.026	0.636	0.269
8	1.0094	0.0266	1.0576	1.0096	0.0265	1.0613	0.018	0.316	0.350

The discussion in Section 4.3.1 on the trend of differences between the results of the MCS and those of the HBA also applies here. Then, the series of voltage underestimations and overestimations noted are due to the randomness of the Monte-Carlo simulations. This is tested further in the development of the HBE in Chapter 6. The comparable order of error between those recorded in Table 4.3 and Table 4.4 indicates the accuracy is maintained despite the addition of DG. The plot of Figure 4.10 shows the trend of the standard deviation. Compared to that for the passive case (Figure 4.7), the standard deviation in the active case is slightly higher.

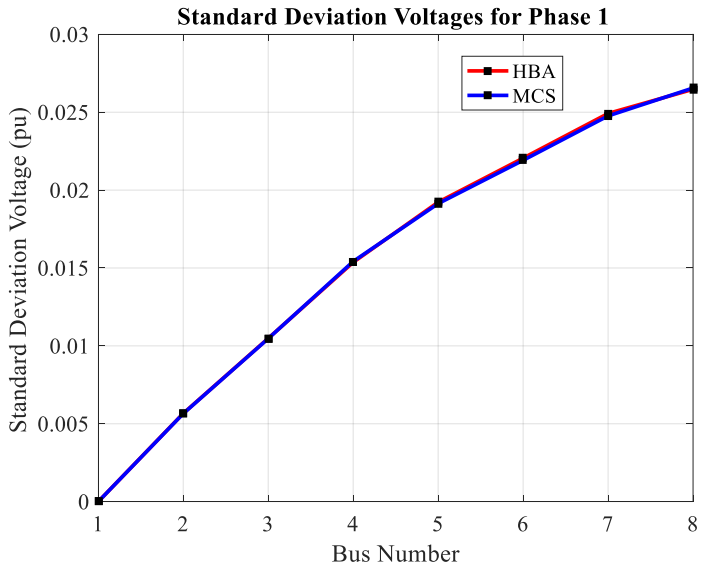

Figure 4.10: Standard deviation of voltages (Case 2, Red phase)

This meets the statistical expectation of increased variability with the number of uncertain variables; the active case has increased uncertainty owing to the addition of DG RVs.

Figure 4.11 shows the PDFs from the HBA fitted to the MCS voltage histograms. The plots are relevant to demonstrate two things; the impact of unbalanced DG and the accuracy of the output distributions.

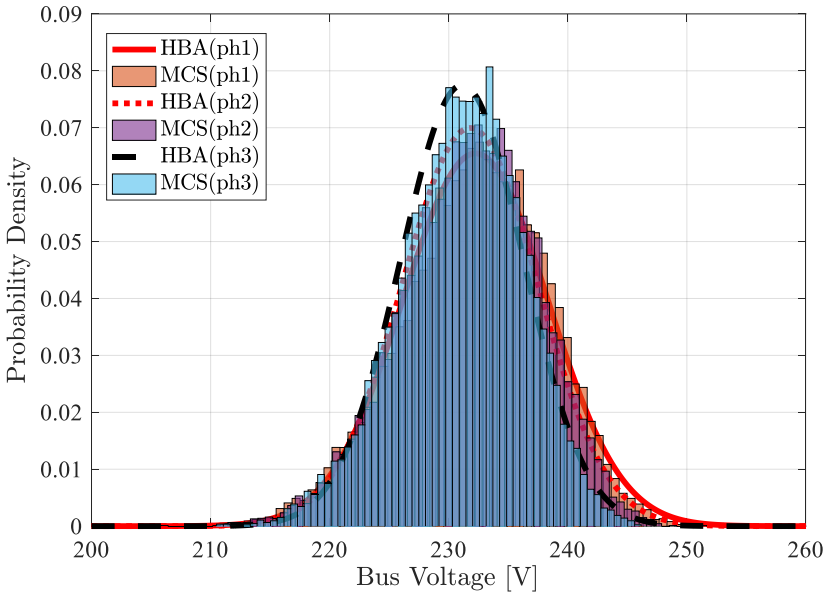

Figure 4.11: Distribution of bus voltages (Case 2, bus 8, all phases)

The PDFs shown in the plot are differentiated across the phases as expected for an unbalanced feeder. However, opposed to the observations in the passive case, the red phase has the highest voltages while the yellow phase has the least. This is explained by the fact that DG connected to the same nodes and phases causes reversed unbalance to that due to loads. Where both types of loads exist, the net-unbalance follows the polarity of the net-generation while the level of unbalance is influenced by its magnitude. The same plot also demonstrates the accuracy of the HBA PDFs in predicting the distribution of the output functions. The PDFs sit well on the histograms as seen with the yellow phase and is reasonably representative of the results for the other phases not clearly visible in the plot.

Overall, the discussed results demonstrate a plausible implementation of DG in concept and accuracy. The outcomes instil confidence in the consistency of the HBA accuracy with the introduction of DG. The conducted analysis concludes the conditions for which the HBA was formulated. The subsequent case tests the performance of the approach outside these conditions.

4.3.3 Case 3: Active MV Feeder ($X \neq 0$, $Q \neq 0$, passive)

In this study, the suitability of the HBA, with its inherent assumptions and simplifications, for load flow computation on MV feeders is tested. To make the test network compliant with practical models for MV feeders, both conductor and load parameters in the passive case are modified such that the line

reactance and reactive power flow become significant. The changes are reflected in the entries under the 'MV' row of Table 4.1. The feeder is tested in passive configuration, without DG.

The plot of Figure 4.12 shows the trend of voltages on the MV feeder. The dispersion of the HBA from the MCS trace demonstrates the error resulting from the neglect of feeder reactance and reactive power flow. The HBA under-estimates voltage-drops resulting in a higher profile of voltages. The magnitude of the error is a function of the electrical distance from the source through Ohms Law. Hence, closer to the source, where the electrical distance is small, the impact of the inaccurate input models is also small. However, as the distance increases, the error in the models increase, and the impact on the outputs becomes more significant.

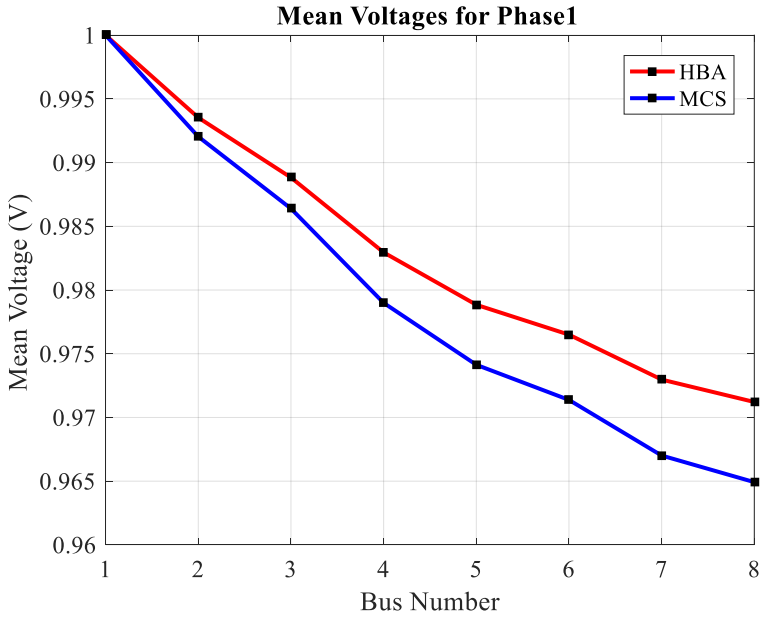

Figure 4.12: Mean bus voltages (Case 3, bus 8, Red phase)

Table 4.5 indicates the error in the mean voltages is up to 0.7%. The plot of Figure 4.13 compares the trends for standard deviation. The plot also reflects significant error margins which vary with electrical distance from the source.

Table 4.5: Comparison of voltage outcomes between the MCS and HBA methods – Case 3

Bus ID	MCS Outcomes			HBA Outcomes			Relative Error (%)		
	E(V)	Var(V)	Vp%	E(V)	Var(V)	Vp%	E(V)	Var(V)	Vp%
1	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	0.000	0.000	0.000
2	0.9921	0.0043	0.9846	0.9935	0.0036	0.9877	0.145	17.991	0.317
3	0.9864	0.0080	0.9728	0.9888	0.0066	0.9780	0.245	17.329	0.536
4	0.9793	0.0118	0.9589	0.9830	0.0096	0.9671	0.377	18.143	0.848
5	0.9739	0.0148	0.9484	0.9788	0.0121	0.9590	0.505	18.260	1.114
6	0.9714	0.0168	0.9422	0.9765	0.0138	0.9537	0.525	17.575	1.224
7	0.9672	0.0190	0.9343	0.9730	0.0156	0.9472	0.596	17.758	1.382
8	0.9646	0.0203	0.9288	0.9712	0.0166	0.9438	0.684	18.230	1.621

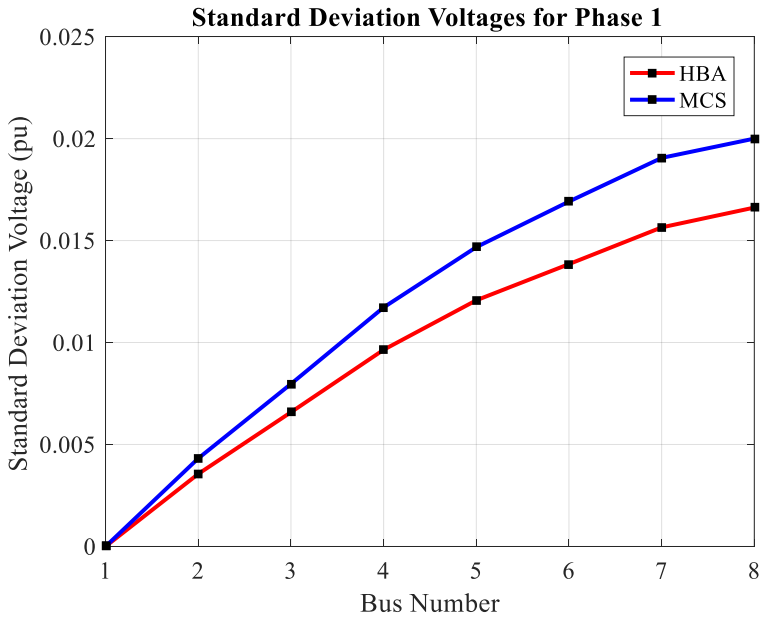

Figure 4.13: Standard deviation of voltages (Case 3, Red phase)

The under-estimated standard deviation in the HBA results, as noted in the differences between the red and blue trace of Figure 4.13, reflects reduced variability in the inputs. The omission of the uncertain quadrature components of the inputs leads to the under-estimation of the standard deviation. The error associated with both the mean and standard deviation is reason for the significant difference in the output distributions in Figure 4.14. The insert shows the differences in the outputs for phase 1.

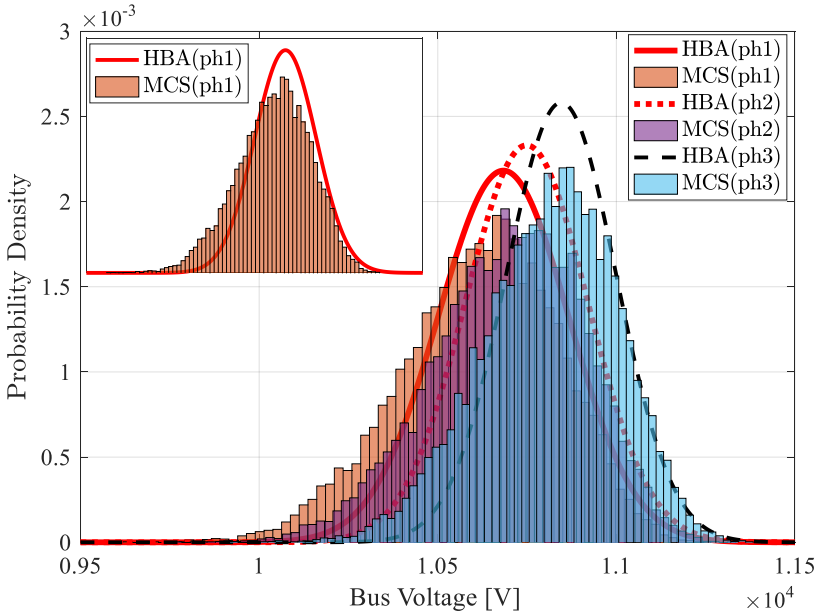

Figure 4.14: Distribution of bus voltages (Case 3, bus 11, phase 1)

Apart from the differences between the mean and standard deviation, the extrema voltages differ. As expected from the voltage-drop optimistic results of the HBA, the error in extrema voltages is more significant on the lower tail. The neglected parameters are liable for the excess voltage-drop observed

on the MCS histogram. The magnitude of the errors in the percentile voltages (up to 1.6%) and standard deviation (up to 18%) reflected in Table 4.5 confirms the extent of the mismatch between the output functions.

The results demonstrate the unsuitability of the HBA method for voltage-drop calculation on distribution systems with substantial line reactance and reactive power flow.

4.4 CONCLUSION

This chapter introduced the HBA, discussing the underlying theoretical foundations and the process for deriving the PLF solution. The HBA approach is a viable tool for the calculation of probabilistic voltages on radial distribution feeders. However, its application is only suitable to LV feeders with insignificant line reactance and unity power factor loads and DGs. When applied to MV feeders with significant reactive power flows, substantial errors result. The assumptions related to the representation of line impedance and load currents must be removed.

Apart from the assumptions applied for simplification, the HBA and its extension to multi-nodes with loads and DG retained the assumption that the beta PDF was the appropriate model of the random input variables. Since that decision was based on load studies over 25 years ago, and several load models have been published since, it is necessary to re-assess the suitability of the beta-BDF. This forms the subject of the chapter that follows.

Chapter 5: The Validity of the Beta PDF as a Descriptor of Inputs

The extension of the HBA to several applications is based on the validity of two hypotheses. The one tests the suitability of the beta PDF to model various inputs such as wind, PV and the load demand. The other concerns the feasibility of deriving the statistical solutions to LF equations considering the full characteristics of the concerned random variables. This chapter tests the first hypothesis; the validity of the beta PDF as a descriptor of inputs. The beta PDF is fitted to various practical and synthetic load models to test its fitness, while the load flow based on Monte-Carlo simulations is performed to assess the impact of the input model accuracy. The MCS based on the actual data is used as a validation tool. The sensitivity of the outputs to the input errors directs the acceptability and limitations of the beta model.

5.1 INTRODUCTION

The literature review presented in Chapter 3 unpacked the contributions on the characterisation of loads and distributed generation (DG) using statistical distributions. The selection of a particular input distribution over another is commonly made on the basis of two factors; the measure of the model accuracy assessed through goodness of fit (GoF) tests and the convenience for PLF analysis. Standard parametric distributions such as the log-normal [133], [166], beta [7], [166], [173], [176], Gaussian, and Weibull [135], [171], [174], [237] are widely adopted based on their computationally convenience for PLF algorithms. On the other hand, mixture models such as the Gaussian mixture model (GMM) [167], [238] and the mixed skew model (MSM) [178] have been proposed to enhance the accuracy in the input models. The motivation is based on the argument that the input data often demonstrates multi-modality which cannot be perfectly fitted using a single standard distribution. However, the use of mixture models may result in computational complexity, inefficiency, and impractical results. Consequently, the selection of a modelling approach must not only consider the model accuracy at the inputs. A trade-off between the impact of the model on the accuracy of the PLF outputs and the associated computational efficiency (in deriving the load models and also the use of the models in a PLF study) must be applied.

The beta PDF has been validated to model skewed residential customer loads [7], [166], PV [176] and wind power [173]. However, in some cases it is chosen for its versatility over other distributions with better GoF outcomes. For instance, the Weibull, log-normal, and normal distributions are widely acceptable models for wind power, the load demand (in the UK and Finland), and aggregated loads (and DG). This implies that the use of the beta model in any of these cases only provides approximations, with error, to the actual distributions. Notwithstanding the possible implications on accuracy, there is

no record of research conducted to test the sensitivity of PLF outputs to the input models. This chapter tests the acceptability of the beta PDF to model load data accurately characterised by the Weibull, log-normal, Gaussian, deterministic, and mixture models. This is achieved by comparing the outcomes from an MCS-based PLF using exact models and beta-PDF approximate models. The GMM, log-normal, Weibull, Gaussian and deterministic models are considered as the 'exact models' on the basis of the assumption that they are the acceptable models. The validity of the beta model is evaluated by analysing the magnitude of error in the output mean, variance, and more significantly, the risk-based percentile values.

5.2 INPUT MODELS

This section details the characterisation of the inputs used in the PLF studies. For easy reference, the models are classified into mixture models and standard distributions. It must be noted that all distributions serve as validation models to the beta distribution.

5.2.1 Mixture Models

Based on its reputed modelling accuracy, the GMM is used as a validation tool to assess the accuracy of the beta PDF in modelling multimodal data and the associated impacts on the PLF outputs.

A GMM distribution is defined as a weighted sum of a finite number (N) of component Gaussian distributions (i), each with a characteristic mean (μ_i), standard deviation (σ_i) and an assigned weight (ω_i). The function of such a distribution is given by (5.1).

$$f(x) = \sum_{i=1}^{N} \omega_i \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{(x-\mu_i)^2}{2\sigma_i^2}} \qquad (5.1)$$

where $0 \leq \omega_i \leq 1$ and $\sum \omega_i = 1$

The assignment of the weight parameters ω_i should be in such a way that the second Kolmogorov axiom is observed; ensuring the sum of all probabilities on the GMM is unity.

5.2.2 Standard Probability Distribution Functions

Based on the literature survey of Chapter 2, four distributions are of most interest: Gaussian, log-normal, and Weibull. In addition, deterministic models may be required for variable with little to no variability. The characteristic functions for these distributions are provided in (5.2) – (5.5):

$$\textit{Gaussian:} \quad f(x) = \frac{1}{(\sigma\sqrt{2\pi})} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad (5.2)$$

$$\textit{Log-normal:}\ f(x) = \frac{1}{(x\sigma\sqrt{2\pi})} e^{-\frac{(\ln x-\mu)^2}{2\sigma^2}} \qquad (5.3)$$

$$\textit{Weibull:}\quad f(x)=\begin{cases}\frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1}e^{-\left(\frac{k}{\lambda}\right)^{k}}, & x\geq 0\\ 0, & x<0\end{cases} \qquad (5.4)$$

$$\textit{Deterministic:}\ f(x)=1,\ x\in[-\infty,+\infty] \qquad (5.5)$$

The provided distribution functions are used to derive synthetic input data. The data constructed from many random samples are used for validating the fitted beta PDF models in the simulation studies.

5.3 SIMULATION CASE STUDIES AND DISCUSSION

The aim of the conducted investigation is to establish the validity of the beta PDF as a versatile descriptor of inputs in the way suggested in the previous chapter. Its accuracy when applied to input data best modelled by log-normal, Gaussian, Weibull, deterministic and mixture models, is assessed. To achieve this, five case studies where each for the identified models are employed are considered. The five identified input models are used as validation models (VM) and the corresponding MCS-PLF is recognized with ' -VM, ' while the identifier ' -beta ' is used for the MCS based on the respective beta-PDF models.

As in the previous chapter, a simple 8-bus feeder (Figure 4.1) is used with the MV properties detailed in Case 4.3 and Table 4.2. The mean demand or generation is kept constant across all cases, save for the shape of the distributions which is dictated by the relevant input models. All of the connected buses are loaded equally with loads of the same model parameters (according to the PDF), with a constant power factor of 0.8 lagging, and balanced across the phases (unity correlation). In all cases except for the Weibull case which involves wind generation, the networks are passive configured.

The test protocol used is discussed hereafter and is common to all cases.

5.3.1 Test Protocol

The flow chart for assessing the sensitivity of the PLF outputs to the various input load models is shown in Fig. 1. The procedure has four main stages:

1. **Generation of load samples**: An array of customer loads is constructed by random sampling from the respective statistical distributions of the validation models defined in Sections 5.3.2 – 5.3.5. Where possible, practical models based on real load data are used. Otherwise, synthetic models are used. The data vectors are 500 × 1 in size, resembling load data samples from a group of 500 customers.
2. **Fitting the beta PDF to the load data**: In this stage, the parameters of the beta models to the synthetic load data are determined. The approach uses a selected normalisation factor *C* (which is taken to be any value equal or greater than the maximum data point) and the expressions of mean and variance for the beta PDF to derive the shape parameters.

The relevant expressions are [142]:

$$\alpha = \frac{\mu(C\mu - \mu^2 - \sigma^2)}{C\sigma^2} \quad (5.6)$$

$$\beta = \frac{(\sigma - \mu)(C\mu - \mu^2 - \sigma^2)}{C\sigma^2} \quad (5.7)$$

3. **Computation of a probabilistic load flow**: The parameters of the beta PDF derived in Step 2 are used as input models in a PLF evaluation on a simple 8-bus representative feeder. The PLF is based on a 15,000-trial MCS. Another MCS-PLF using the synthetic load data (derived in Step 1) is performed and used as the validation tool. The MCS-PLFs are denoted MCS-VM and MCS-beta, respectively.

4. **Comparison of outcomes**: Three performance metrics are used. The mean, variance, and percentile voltages assessed at 5% design risk are computed for each PLF of Step 3. Comparison of these values leads to analysis on the impacts of model 'accuracy' on PLF output accuracy.

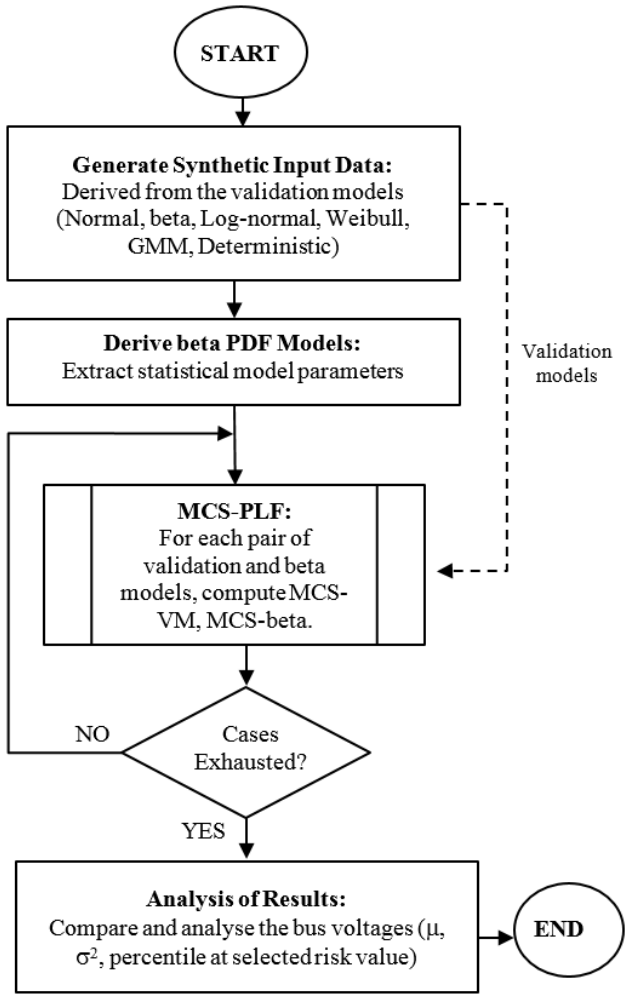

Figure 5.1: Flow chart for the sensitivity analysis of PLF outcomes to estimate beta input models

5.3.2 Case 5.1: Aggregated Gaussian Loads

This case study assesses the performance of the beta models in system involving aggregated data, which are commonly well characterised by Gaussian functions. To facilitate the analysis, the load properties provided in Table 5.1 are adopted. The mean is selected such that the ADMD is about 550 kVA while the standard deviation is set to 20% of the calculated mean.

Table 5.1: Aggregated load Gaussian parameters, Case 5.1

LOAD DEMAND CHARACTERISTICS			
μ	σ	C	ADMD
0.25	0.05	200	550 kVA

Figure 5.2 shows the beta-fit with parameters Beta (11.98, 11.99, 0, 0.5) and the underlying histogram of the synthetic load data. The plot demonstrates the accuracy of the beta PDF in modelling symmetrical, aggregated load currents estimated through normal distributions.

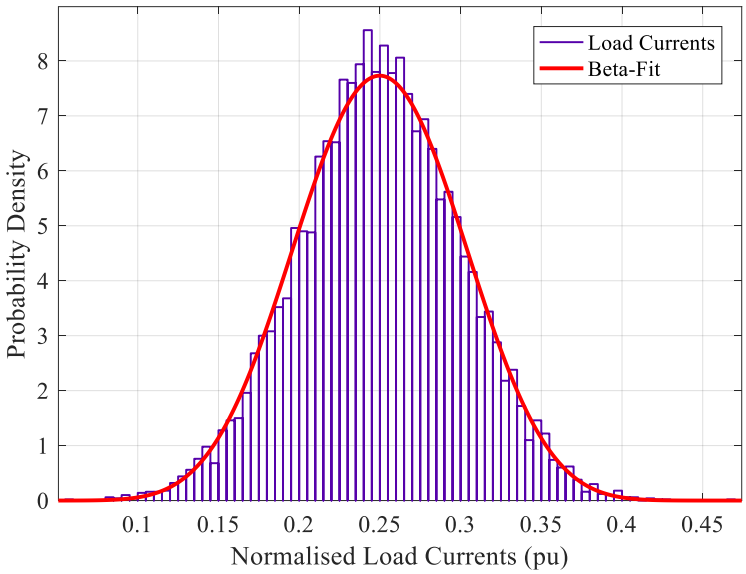

Figure 5.2: Beta-Fit to aggregated load data

Considering the accuracy in the input model, as reflected by a low χ^2 fitness statistic of 0.1255, the output functions from the conducted PLFs are anticipated to be undistinguishable within tolerance of the random MCS errors. Figure 5.3 confirms the anticipated accuracy. The MCS-beta trace accurately follows the cumulative distribution function derived from the MCS with the validation model. The outcomes demonstrate the characteristic accuracy of the beta PDF in modelling symmetrical RVs.

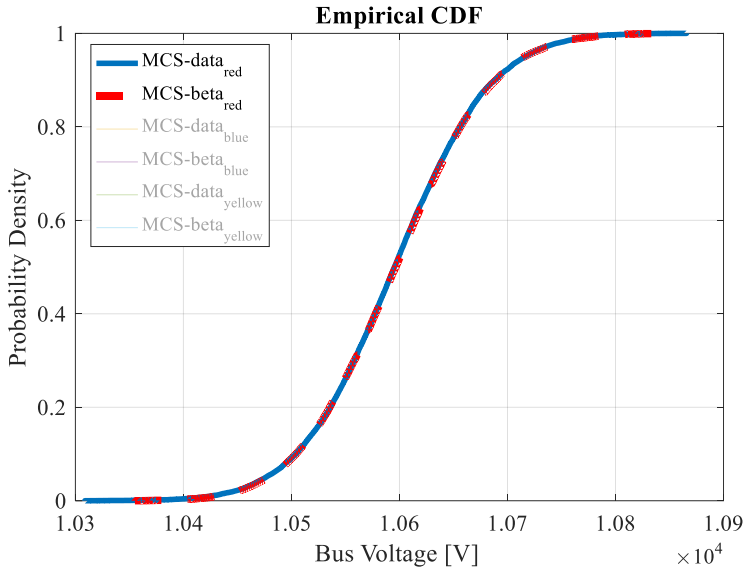

Figure 5.3: Comparison of CDF plots of voltage magnitude at bus 8, Case 5.1

The quantitative analysis of the outcomes will be provided collectively, across all conducted cases, later in the discussion.

5.3.3 Case 5.2: Log-normal Distributed Loads

In some counties, such as the UK [136], Finland [127], and Sweden [130], statistical load models have already been developed. For these countries and possibly others, the log-normal distribution has been identified as the best descriptor of the load demand. The simulation conducted here serves to demonstrate the practical extension of the beta PDF to cases where the variability in the load data is modelled by other skewed distributions. The log-normal is used as an example.

By using the mean and variance of a log-normally distributed RV, the corresponding parameters of the distribution can be determined as previously established in (5.6) and (5.7). Table 5.2 provides the parameters of the synthetic load model used.

Table 5.2: Log-Normal distributed load parameters, Case 5.1

LOAD DEMAND CHARACTERISTICS			
μ	σ	C	ADMD
-1.72076	0.592333	200	469 kVA

Figure 5.4 demonstrates that while the beta-PDF is flexible, it fails to accurately mimic the sharply skewed log-normal distribution. This is also reflected by a high χ^2 fitness statistic of 98.425; the approximate beta model has a lower mode and a broader distribution. If accuracy is determined only at the input stage, one would conclude the errors are significant, making the beta model unacceptable. However, looking at the outputs CDFs of Figure 5.5, the order of error does not compare to that observed in the inputs.

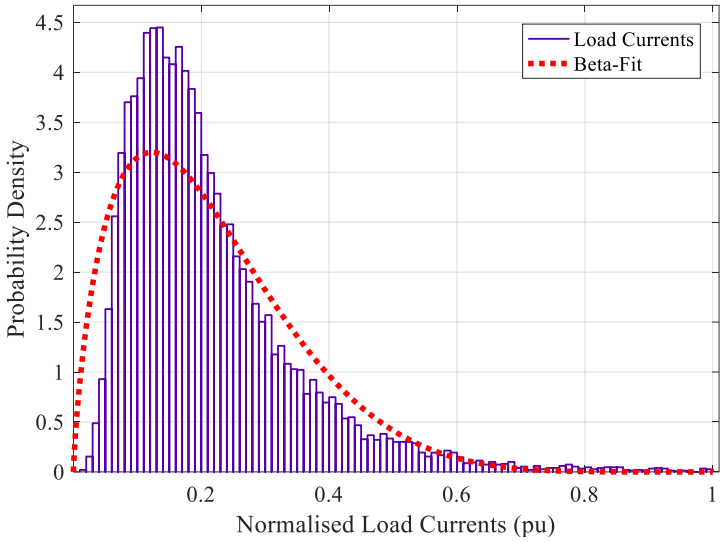

Figure 5.4: Beta-Fit to log-normal distributed load data

On analysis of the output plots, the CDFs appear generally in sync with the exception of slight mismatches at either extremity. These observations are substantiated by the fact the model errors reflected in Figure 5.4 are mostly noted at the tails. Even so, the order of mismatch is insubstantial.

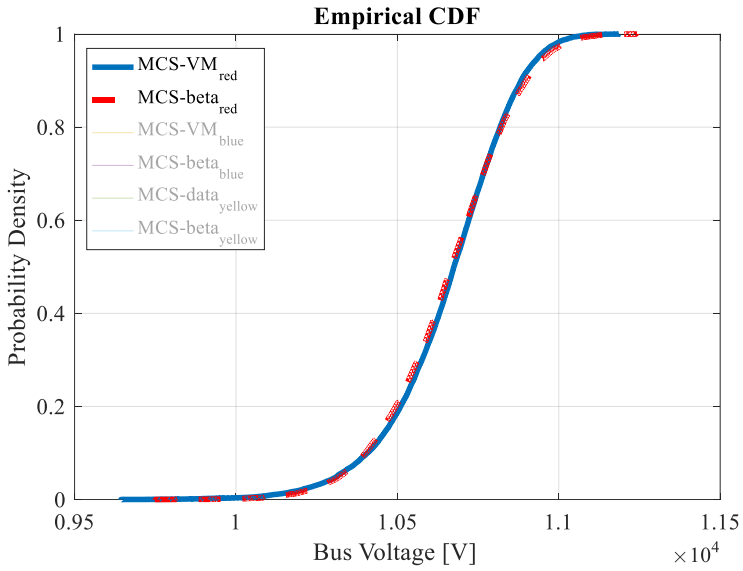

Figure 5.5: Comparison of CDF plots of voltage magnitude at bus 8, Case 5.2

Considering the outcomes, it can be generalised that the beta PDF is not limited to load data which is perfectly characterised by its function. The approximation of the beta PDF to other distributions, the log-normal in this case, results in acceptable accuracies at the output stages. A similar analysis is extended to the Weibull distribution commonly applied to wind speed data and in some cases, the load demand.

5.3.4 Case 5.3: Weibull-distributed Wind Speed

In this study, the capacity of the beta PDF to model wind power distributions is evaluated. The test system properties are modified such that the loads are beta-distributed with parameters $\alpha = 1.5$ and $\beta = 4.00$. Then, on each bus and phase, wind farms with a rated capacity of 2.2 MW are connected.

The load parameters are common in both MCS simulations in order to ensure the studied impacts are solely due to the DG models. The characteristics of the wind generator and the Weibull distribution parameters for wind speed are adopted from practical data [237], [239]. The assumed models allow the analysis of the fitness of the beta PDF to model highly variable [8] and high [14] wind power outputs. The tests related to these distributions of wind speed are denoted Case 5(a) and 5(b) respectively. Table 5.3 provides the model parameters.

Table 5.3: Wind turbine and wind speed parameters

CASES	WIND GENERATOR CHARACTERISTICS				DISTRIBUTION PARAMETERS	
	Cut-in Speed	Rated Speed	Cut-out Speed	Rated Power	Scale	Shape
Highly variable wind (5.3a) [237]	4.00 ms^{-1}	13.61 ms^{-1}	25.00 ms^{-1}	1 MVA	11.0086	1.9622
Medium-High Wind (5.3b) [239]	4.00 ms^{-1}	13.61 ms^{-1}	25.00 ms^{-1}	1 MVA	10.8000	6.3000

The means to derive power output distributions from wind speed data has already been covered in Chapter 2. Using the characteristic function provided in Section 2.2.1 through (2.2), the speed data arrays are converted into power-output arrays. The beta PDFs fitted to the histograms of the derived power data are shown in Figure 5.6 and Figure 5.8.

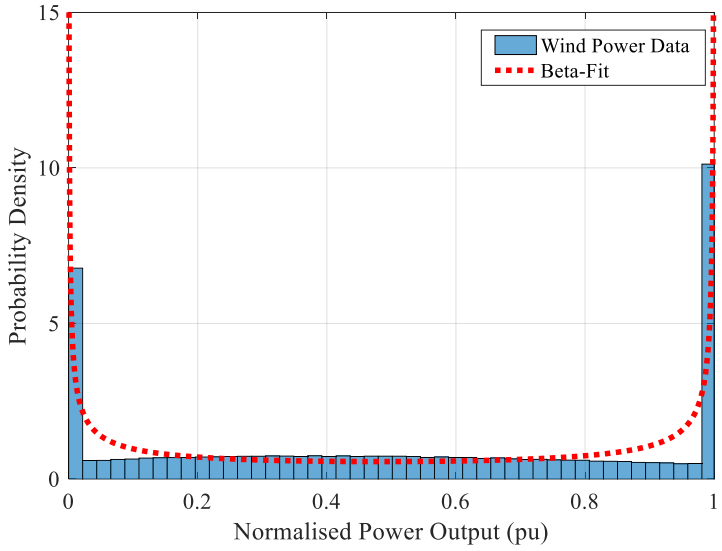

Figure 5.6: Beta-Fit to highly variable wind power output

Figure 5.6 depicts the distribution of the injected currents from the wind generators. The distribution is tri-modal, with the dominant modes occurring at the extreme ends of the distribution where power

output is truncated based on the wind speed. The comparison the output functions demonstrates the capability of the beta PDF to model the dominant modes. However, the densities of these modes are much higher than those observed in the validation data. Furthermore, the beta distribution has an anti-mode between the dominant modes. As a result, it fails to accurately follow the sharp cornered variations at either extremity, and the concave variation in between them. This is also reflected by a high χ^2 fitness statistic of 2868.2.

Despite the identified model errors, the output functions shown in Figure 5.7 are aligned with great accuracy. The smooth CDFs demonstrate that the multimodality in the input models is not reflected in the outputs.

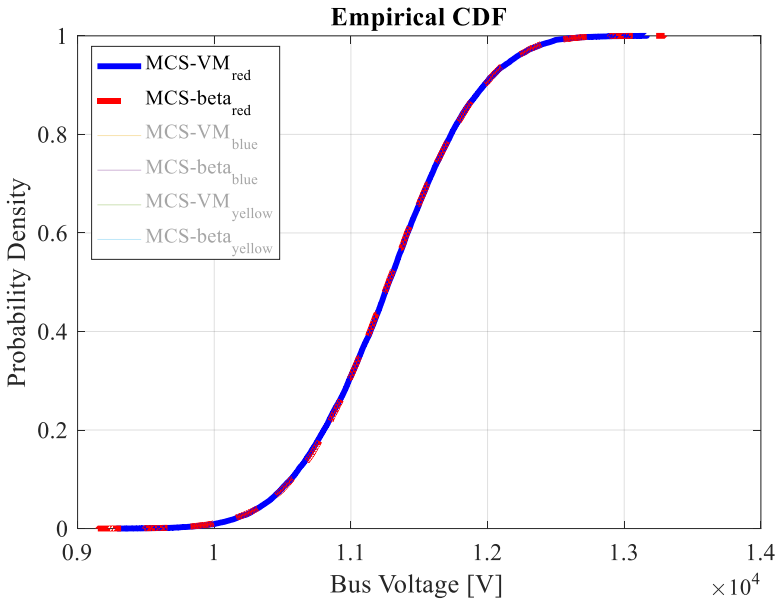

Figure 5.7: Comparison of CDF plots of voltage magnitude at bus 8, Case 5.3a

When higher wind speeds are considered, a larger number of data points fall between the cut-in and cut-off speeds. As a result, the output distributions are much denser between the extremum output states. While the resulting distribution remains trimodal, the density of the modes at the extremities are minor. Accordingly, the beta PDF approximates the tri-modal distribution using a unimodal distribution as shown in Figure 5.8. While the distribution does not perfectly model the underlying data, the error is limited (χ^2 fitness statistic of 28.482). The beta PDF is slightly broader and has a lower mode due to the compensation of the probabilities of the two unrepresented modes.

Nonetheless, the outcomes depicted in Figure 5.9 validate the adequacy of the approximate model. Similar to the outcomes in Case 5(a) involving highly variable wind, the probabilistic outputs are coincident. The results demonstrated in both cases provide confidence in the use of the beta PDF for in PLF studies involving wind distributions.

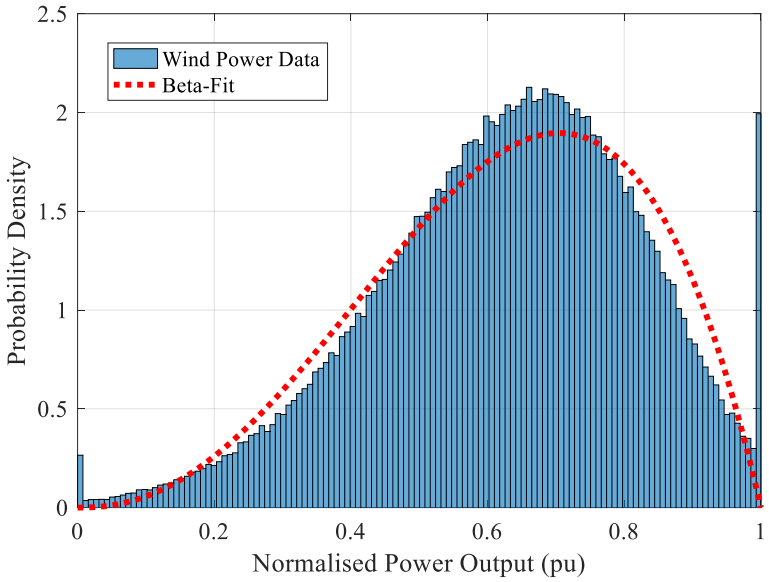

Figure 5.8: Beta-Fit to high wind-speed power output

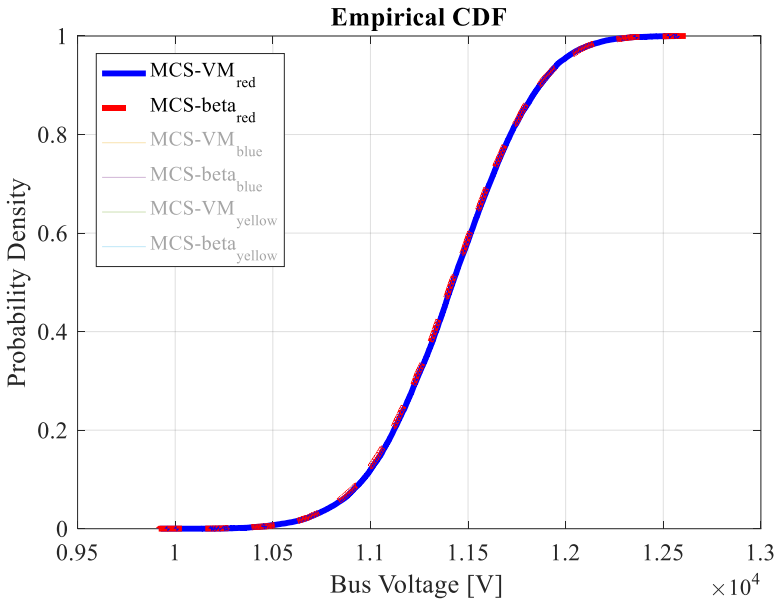

Figure 5.9: Comparison of CDF plots of voltage magnitude at bus 8, Case 5.3b

Apart from the multimodality caused by the truncation of output power as a result of wind speed falling below and above the threshold levels characteristic of the wind turbines, the distributions discussed thus far are related to standard distributions. Next, the beta PDF is tested against mixture models.

5.3.5 Case 5.4: Mixture Models

Standard distributions are most favourable for the characterisation of unimodal data or at the most U-shaped bi-modal distributions. This test case investigates the accuracy of the beta PDF when applied to data demonstrating heterogeneity resulting in multi-modes or irregular distributions.

Based on the accuracy associated with the Gaussian Mixture Model (GMM), it is used here to regenerate the load data with the characteristics sought for this investigation. Table 5.4 presents the GMM parameters used to model the nodal loads characterized as currents. The shape parameters were derived from published models based on real load data [238], converted to current at nominal voltage and scaled down to suit the selected test system.

Table 5.4: Gaussian Mixture Model Parameters, Case 5.5a

GAUSSIAN COMPONENT	DISTRIBUTION PROPERTIES		
	Mean (A)	Std. dev (A)	Weight (%)
1	32.4068	9.2641	17.90
2	27.6210	2.0460	19.56
3	34.4864	6.8014	46.49
4	24.1748	4.5416	16.05

Figure 5.10 illustrates the flexibility of the GMM distribution in characterising the irregular load data. On the other hand, the regular beta distribution only provides an approximation to the histogram of currents; it follows the nearly symmetrical distributed data but averages the irregularities in the distribution. A χ^2 fitness statistic of 462 explains the discrepancies on the mode and the bins around it.

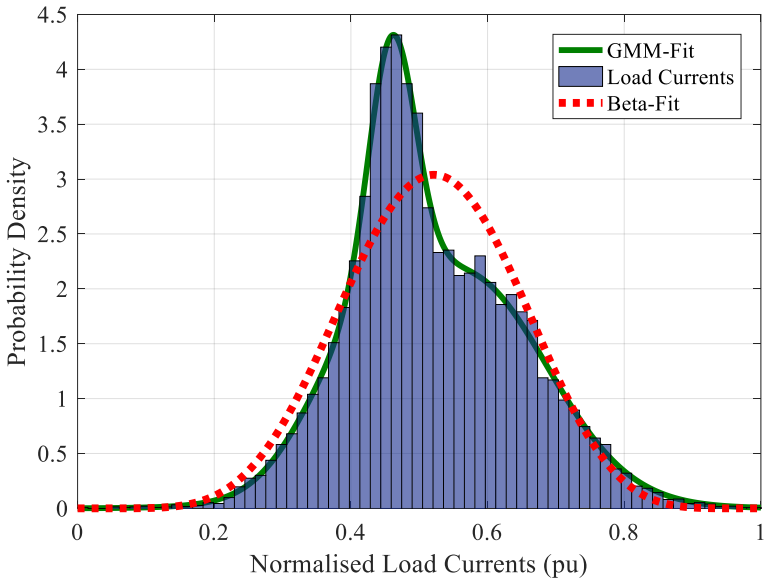

Figure 5.10: Beta-Fit to a dominant mode GMM distribution

On performing the PLF, the outcomes associated with either load models are closely similar as demonstrated in Figure 5.11. On looking closer into the tails of the distributions, only slight misalignments are observed particularly at the lower tail.

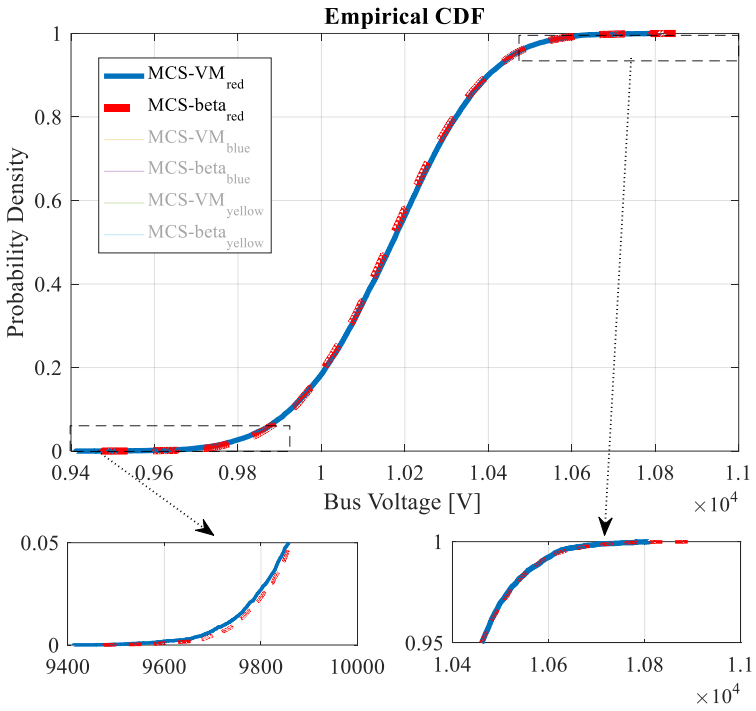

Figure 5.11: Comparison of CDF plots of voltage magnitude at bus 8, Case 5.4a

Similar to the previous cases involving multimodal wind distributions, the shape of the inputs is not reflected in the outputs. However, it can be argued that the margin of error in the outputs is dependent on the level of data heterogeneity. The GMM model, derived from reported load data and on which this case was based, does not have multiple dominant modes; the PDF of the load data is nearly unimodal. To investigate the behaviour of the outputs with increased multi-modality and skewness in the load data, the GMM model is 'stretched' to take on a more extreme multi-modal profile. Although this model may not be representative of practical loads, the procedure allows further investigation of the behaviour of the system in terms of input-output accuracy relationships. Table 5.5 provides the parameters of the modified GMM.

Table 5.5: Gaussian Mixture Model Parameters, Case 5.5b

GAUSSIAN COMPONENT	DISTRIBUTION PROPERTIES		
	Mean (A)	Std. dev (A)	Weight (%)
1	6.52	1.45	35
2	12.35	2.046	30
3	20.5	4.85	25
4	30.65	4.542	10

Figure 5.12 shows the approximate model offered by the beta PDF. It follows the skew in the data, tracks the mean, and averages the effect of modes resulting in a regular, unimodal distribution. Judging

from the visual appearance of the differences between the input models and a high χ^2 fitness statistic of 1951.49, an anticipation for significant errors in the outputs is almost obvious.

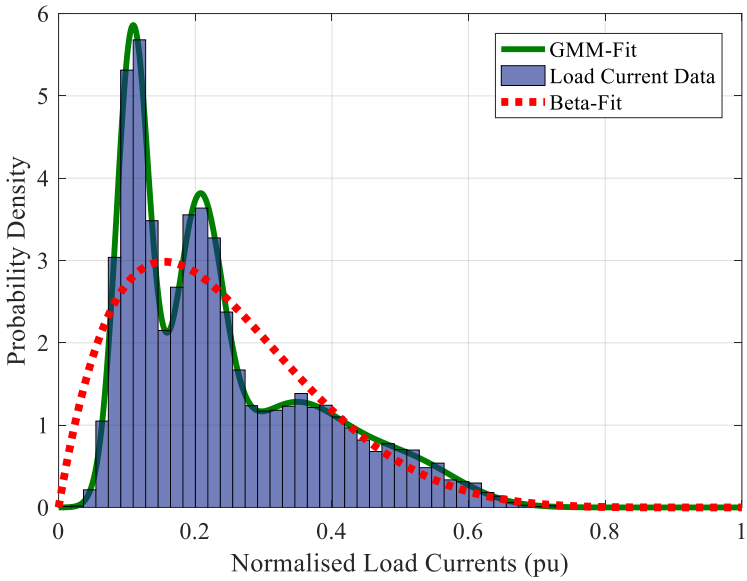

Figure 5.12: Beta-Fit to a dominant multi-mode GMM distribution

On the contrary, the outcomes of the PLF studies suggest otherwise. The voltage distributions at the terminal node of the feeder (bus 8), using each of the distributions in the MCS, are shown in Figure 5.13. The plotted CDFs show that the load models derived from the beta PDF and the GMM data result in similar voltage distributions.

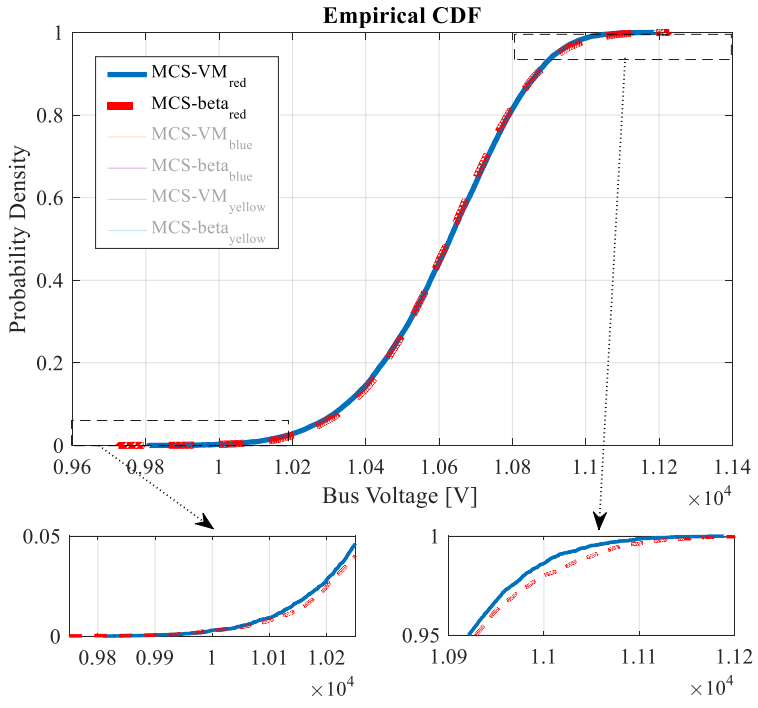

Figure 5.13: Comparison of CDF plots of voltage magnitude at bus 8, Case 5.4b

Looking closely at the upper and lower 5% intervals, the MCS-beta achieves the best approximate to the lower tails, while a slightly wider mismatch is noted at the upper tails of the voltage profile. While the graphical differences might appear large, the quantitative difference in the voltages at the tails is less than 0.5% of the rated voltage (0.005 pu).

Up to this point, it appears the outcomes of the PLF are not significantly swayed by the input model accuracy as long as the mean, variance, and general skewness of the load data is captured in the approximate model. A plausible explanation to these findings is related to the central limit theory (CLT) and the conditions of interdependency between RVs.

The RVs in all the tested cases were assumed independent. Under such conditions, the sum of a sizeable number of the input RVs, scaled by the voltage-drop sensitivity factors, is governed by the central limit theory (CLT). As the output functions gradually tend to normality, the error in the input models tends to zero. This is an expected result considering the beta PDF follows the skewness of the distribution, and, has the same mean and variance as the load data. However, the application of the CLT and the preservation of the input model errors is disrupted when the conditions of interdependence, or the sensitivity factors are changed. For a system with fully dependent RVs, the outputs are a scaled function of the inputs such that the distribution shape is maintained. Moreover, the output distributions of a multivariate system mainly depend on the RVs with the dominating sensitivity factors. The result is that the multi-modality in the inputs is projected to the outputs. Accordingly, the worst-case error observed at the outputs is expected when the loads are fully correlated, and the sensitivity coefficients are large.

To facilitate the related investigation, the properties of the test network are modified by considering the loads to be perfectly correlated and extending the feeder by two nodes such that the voltage-drops are increased. Figure 5.14 depicts the CDF of bus 10 voltage according to the two input load models. By comparison with the GMM CDF in Figure 5.13, it is clear that input multi-modality is reflected significantly in the several ‘inflection points’ of the output, and the fat tail of the GMM load model in the extended tail of the voltage-drop. The shape of the CDF confirms the anticipated inapplicability of the CLT and the conformance of the output distributions to the shapes of the inputs.

Since the beta PDF is limited to the estimation of unimodal and U-shaped bimodal distributions, its approximation of the multi-modal output distribution is prone to error. From the analysis of the plots of Figure 5.14, the MCS-beta CDF differs from that based on the GMM by up to 0.015 p.u. of the rated voltage. Considering the design limits of permissible voltage variation on MV feeders (0.05 p.u. variation), this level of difference could be significant, potentially leading to under- or over-design. Nonetheless, with the anticipation of weak correlations between loads, and that between DGs in separate locations, the expected error margins would fall.

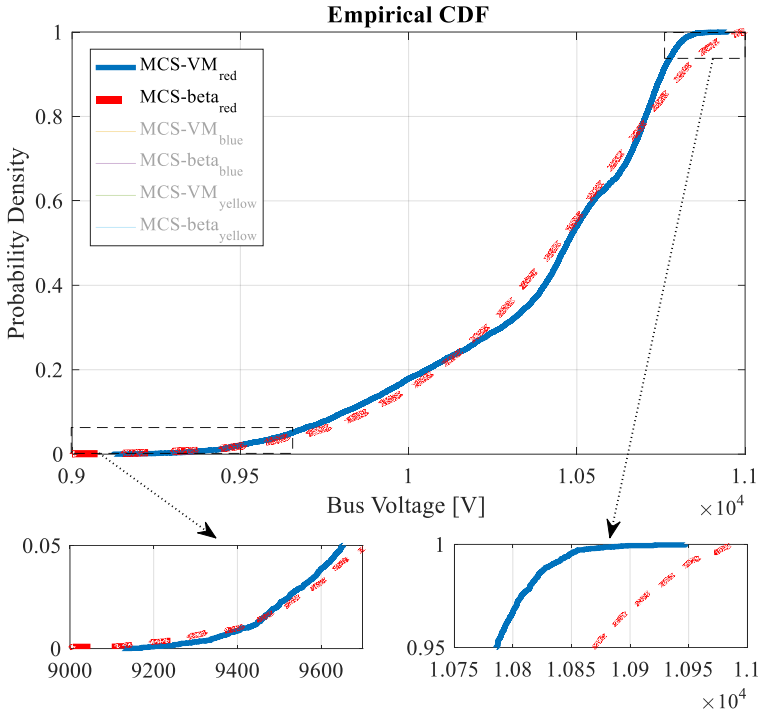

Figure 5.14: Comparison of CDF plots of voltage magnitude at bus 8, Case 5.4c

We have used the MCS-based PLF to test the sensitivity of the PLF outputs to beta-modelled inputs of various load models. The obtained results suggest the beta PDF is fit to model variously skewed load demand, various levels of wind power, and multimodal distributions for PLF analysis. Figure 5.15 demonstrates the error margin in the percentile voltage for the tested cases across all buses.

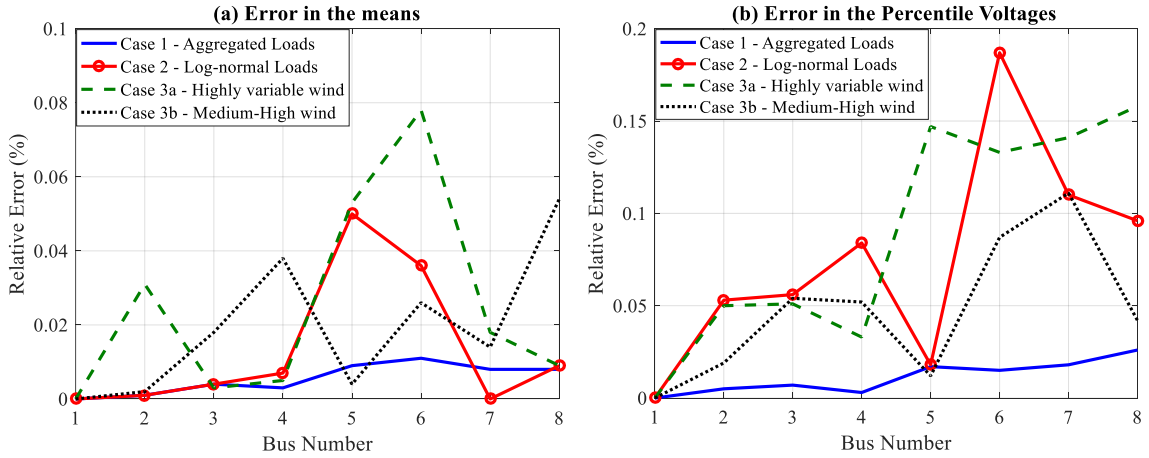

Figure 5.15: Error Analysis on the performance of the beta-based PLF on standard distributions

However, the error is limited. When the input models are independent, the application of the CLT forces the outputs into unimodal distributions which can be well approximated by then sum of beta RVs. For such cases, the suitability of the beta model is only determined by how well it captures the mean and variance of the load data such that the corresponding properties for the outputs are accurate.

However, where the inputs are correlated, the accuracy at the output stage is more sensitive to the input models. Figure 5.16 demonstrates the increased error margins between the uncorrelated and correlated GMM cases.

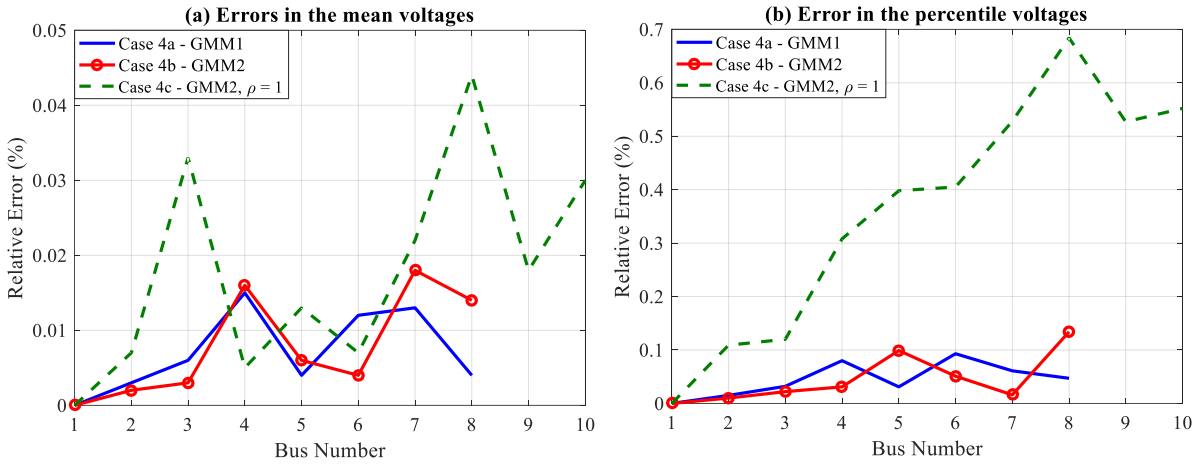

Figure 5.16: Error Analysis on the performance of the beta-based PLF on multimodal distributions

In cases where the input is nearly unimodal (unimodal or with low-weighted multi-modes) or bi-modal (U-shaped), without fat tails, the assumption of a standard beta model only leads to insignificant errors in the outputs. However, multimodal and fat-tailed load data inappropriately modelled by unimodal standard distributions may introduce significant error in PLF outputs.

5.4 IMPLICATIONS ON ONGOING WORK

This chapter investigated the effects of input model accuracy on PLF outputs, comparing mixed models with standard unimodal parametric distributions. Simulations reveal that although the beta distribution, compared to other models such as the GMM, does not always provide the best fit to various load data, the differences in PLF outputs are not always significant. Where the actual input has a dominant mode (with other low-weighted multi-modes) and is not fat-tailed, the errors in assuming a unimodal, standard distribution are insignificant. And, where the inputs have significant multi-modes and fat-tails, the errors depend on the sensitivity factors and the conditions of interdependence. If the RVs are independent, errors are also limited as the output functions approach normality on the basis of the CLT. The extreme case of error occurs when the inputs are fully dependent and is proportionate to changes in voltage (drop or rise). However, such levels of dependency between input variables are unlikely in practical networks. These errors are anticipated to fall with levels of input interdependence or sensitivity factors in a PLF.

Considering the outcomes from this chapter including the likely conditions of network operation discussed in Chapter 3, the beta PDF is a valid descriptor of probabilistic load flow inputs. It offers an important advantage because it is suitable for the modelling of various power system parameters and supports highly efficient PLF approaches. Moreover, clustering methods for statistical load modelling

can be used when input data are highly heterogeneous, to separate the data and reduce complications of dominant multi-modes [181], [183], [186], [240], [241].

Having established the validity of the beta PDF as a descriptor of inputs, the development of the candidate PLF approach based on the beta PDF is detailed in the chapter that follows.

Chapter 6: The Extended Herman-Beta Transform (HBE)

Based on the initially formulated Herman Beta algorithm (HBA) for low-voltage feeders only, this chapter develops a new probabilistic transform suitable for feeders of all voltages. The new approach removes the model simplifications of unity power factor loads and resistive feeders that were suitable only for selected LV systems. The approach is then extended to include dependence between random variables, voltage-dependent input models, and the calculation of branch currents.

6.1 INTRODUCTION

To facilitate the extension of the Herman-Beta algorithm to higher voltages, two issues have been identified. The first, related to the validity of the beta PDF as the sole descriptor of currents, was tested in Chapter 5. The work in this chapter derives and solves the load flow with the initial assumptions removed. Furthermore, the normalization technique for the consumer voltages, also based on the initial assumptions, is replaced by a new approach based on a per-unit system. The approaches combined enable the analysis of feeders of any X/R ratio and load or DG at any power factor.

While the new approach removes the limiting assumptions, the foundations of the proposed approach remain:

- All input random variables are characterised by the beta distribution and are valid in the same time interval of analysis
- The statistical input parameters describe the variability of the active power. The reactive power component is derived from the power factor which is assumed to be constant and deterministic for a given load class.
- Consequently, the extension of the algorithm is premised on three reformulation processes:

1. The reformulation of the voltage-change equations
2. The reformulation of the statistical solutions
3. The reformulation of the normalization technique
    - Then, the sign correction distinguishing DG from loads needs to be applied to the reformulated equations. These processes are detailed separately in the subsequent sections.

6.2 REFORMULATION OF SYSTEM EQUATIONS

The consideration of complex impedance and nodal powers requires the reformulation of the 3p-4w equations and a new formulation for the 3p-3w topology. These are discussed hereafter.

6.2.1 Three-phase Four-wire Model

The network model of Chapter 4, including some of the voltage-drop equations, is recalled here. The electrical representation used was depicted in Figure 4.4 with complex impedances consistent with most distribution lines parameters.

As detailed in Section 4.2.2, the voltage-drop for phase A (used to demonstrate the analysis throughout this chapter) given by (4.11), is:

$$\boldsymbol{\Delta V} = \boldsymbol{I_A}(R_p + jX_p) + \boldsymbol{I_n}(R_n + jX_n) \qquad (6.1)$$

The complex phase currents according to (4.15) become:

$$\boldsymbol{I_{A,B,C}} = I_{A,B,C}\, cos\, \phi_p \pm j\, I_{A,B,C} sin\, \phi_p \qquad (6.2)$$

Equation (6.2) can be represented in quadrature components of currents I_r and I_x as follows:

$$\boldsymbol{I_{A,B,C}} = I_r \pm jI_x \qquad (6.3)$$

Assuming the complex currents are modelled as beta currents of the active power (I_r) at a deterministic power factor ($cos\, \phi_p$) common to all the loads connected at the node of analysis, the distribution of the reactive current component (I_x) is a projection of the active component through $\tan \phi_p$ [6], [219]. For simplicity, the notation I_a, I_b and I_c is used to denote the currents of the active power in phases *a*, *b* and *c* respectively. Accordingly, the complex phase currents become:

$$\boldsymbol{I_{A,B,C}} = I_{a,b,c} \cdot (1 \pm j\, tan\, \phi_p) \qquad (6.4)$$

The polarity between the quadrature components of (6.4) is influenced by the power factor; it assumes a negative or positive state when the load power factor is lagging (inductive loads) or leading (capacitive loads), respectively. Substituting for the phase currents and allowing for the neutral current and the angular shifts between the phase currents results in the quadrature components given by:

$$\Delta V_x = k_3 \cdot I_a - \frac{1}{2} \cdot k_1 \cdot (I_b + I_c) - \frac{\sqrt{3}}{2} \cdot k_2 \cdot (I_b - I_c) \qquad (6.5)$$

$$\Delta V_y = k_4 \cdot I_a - \frac{1}{2} \cdot k_2 \cdot (I_b + I_c) + \frac{\sqrt{3}}{2} \cdot k_1 \cdot (I_b - I_c) \qquad (6.6)$$

Where,

$$k_1 = k_r R_p \pm k_x X_p tan\, \phi_p \qquad (6.7)$$

$$k_2 = k_x X_p \mp k_r R_p tan\, \phi_p \qquad (6.8)$$

$$k_3 = (1 + k_r) \cdot R_p \pm (1 + k_x) \cdot X_p tan\, \phi_p \qquad (6.9)$$

$$k_4 = (1 + k_x) \cdot X_p \mp (1 + k_r) \cdot R_p tan\, \phi_p \qquad (6.10)$$

The constant k_r is a proportion of the phase wire and the neutral wire's resistances, while k_x is a proportion of their reactance according to (4.12). The polarity modes in (6.7) – (6.10) follow the inductive or capacitive nature of the load, respectively.

With the components of voltage-drop determined, the magnitude of the consumer voltage, V_{con}, considering only a single node supplied by a sending-end voltage V_s is:

$$V_{con} = \sqrt{(V_s - \Delta V_x)^2 + \Delta V_y^2} \qquad (6.11)$$

The statistical solutions to the derived system of equations will be developed later.

6.2.2 Three-phase Three-wire Model

The literature review conducted in Chapter 3 identified the need to define the proposed PLF approach for 3p-3w systems. This section provides the mathematical framework underlying the algorithm's development.

Figure 6.1 depicts the network model for a 3p-3w system with delta-connected loads supplied by a delta source.

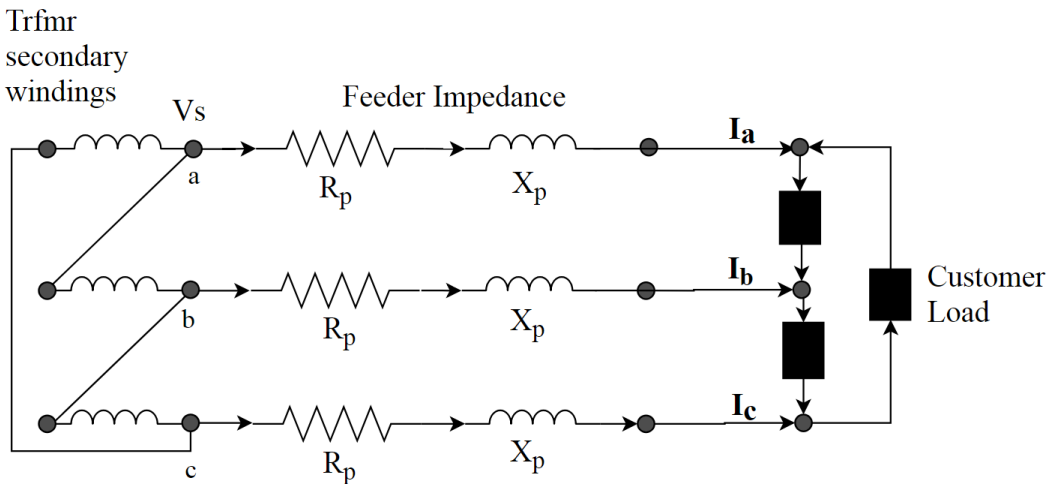

Figure 6.1: Three phase, three wire network model

Where different source configurations such as a grounded wye exist, transformations can be applied to derive the sending-end line voltages. Similar to the previous analysis, the red phase is used to demonstrate the derivation of the load flow equations and can easily be extended to the other two phases by changing the reference phase.

Applying Kirchhoff's Voltage Law (KVL) to Loop 1, which includes the load connected across phase *A* and *B*, results in:

$$\Delta \boldsymbol{V} = \boldsymbol{Z_p}(\boldsymbol{I_{A_L}} - \boldsymbol{I_{B_L}}) \qquad (6.12)$$

Where the subscript 'L' denotes variables related to line quantities.

The line currents can be expressed in terms of the load currents by applying Kirchhoff's Current Law (KCL) to the respective load nodes:

$$\boldsymbol{I_{A_L} = I_{AB} - I_{CA}}$$

$$\boldsymbol{I_{B_L} = I_{BC} - I_{AB}} \qquad (6.13)$$

$$\boldsymbol{I_{C_L} = I_{CA} - I_{BC}}$$

Where $\boldsymbol{I_{AB}}$, $\boldsymbol{I_{BC}}$, and $\boldsymbol{I_{CA}}$ are the complex load currents. For simplicity, the notation $\boldsymbol{I_A}$, $\boldsymbol{I_B}$ and $\boldsymbol{I_C}$ as previously used, is adopted. Re-writing (6.12) in terms of the load currents results in the expression of (6.14).

$$\boldsymbol{\Delta V} = (R_p + jX_p) \cdot (2\boldsymbol{I_A} - \boldsymbol{I_B} - \boldsymbol{I_C}) \qquad (6.14)$$

Then, by applying (6.4), substituting for the complex load currents, and allowing for the angular shifts between the phase currents, the quadrature components of the voltage-drop are determined as follows:

$$\Delta V_x = w_1 \left(2 \cdot I_a + \tfrac{1}{2} \cdot (I_b + I_c)\right) + \tfrac{\sqrt{3}}{2} \cdot w_2 \cdot (I_b - I_c) \qquad (6.15)$$

$$\Delta V_y = w_2 \left(2 \cdot I_a + \tfrac{1}{2} \cdot (I_b + I_c)\right) - \tfrac{\sqrt{3}}{2} \cdot w_1 \cdot (I_b - I_c) \qquad (6.16)$$

Where,

$$w_1 = R_p \pm X_p tan\, \phi_p \qquad (6.17)$$

$$w_2 = X_p \mp R_p tan\, \phi_p \qquad (6.18)$$

The polarity modes in (6.17) and (6.18) are related to the nature of the load power factor in the same way as (6.7) – (6.10). Also consistent with the approach for the 3p-4w model, (6.11) is applied to determine the consumer voltage.

The similarity (in structure and actual terms) between the voltage-drop equations of (6.15) and (6.16) and those in (6.5) and (6.6) for the 3p-4w model is due to the fact that in both cases, the KVL loops involve two conductors; a phase conductor and a neutral conductor in the 3p-4w model, and two phases in the 3p-3w model. Then, since the properties of the neutral conductor are expressed in terms of those of the phase conductor through k_r and k_x, the special condition ($k_r = k_x = 1$) makes the models nearly equivalent, except for the difference in the direction of current between the neutral conductor (3p-4w model) and phase conductor (3p-3w model).

6.3 REFORMULATED STATISTICAL SOLUTIONS

To determine the voltage output distribution, four parameters must be calculated; the shape parameters α and β, and the limit parameters dictating its range. The solution to the unknown parameters would require a system of four or more simultaneous equations. It is theoretically possible to use the first four

moments. However, practically, the dimension of the associated algebra increases the computational burden significantly [50]. Consequently, a normalization technique (reported in the consecutive sections) is applied to solve the limits of the outputs, leaving only two parameters unsolved. For these, the first and second raw moments are used.

The expression for the consumer voltage magnitude given in (6.11) presents a non-linearity problem in the computation of $E(V_c)$. To address this, a linearization technique based on Taylor's expansion is used. Re-writing the function of (6.11) according to Taylor's formula gives:

$$V_c = V_s\left(1 + \left(\frac{\Delta V_x^2 + \Delta V_y^2}{V_s^2} - \frac{2\Delta V_x}{V_s}\right)\right)^{\frac{1}{2}} \qquad (6.19)$$

And, since we intend to keep the order of moments only up to the second order, the approximate consumer voltage is determined according to Taylor's series truncated to terms of the first order.

$$E(V_c) \approx V_s\left(1 - \frac{E(\Delta V_x)}{V_s} + 0.5\frac{E(\Delta V_y^2)}{V_S^2}\right) \qquad (6.20)$$

The calculation of the second moment is based on the exact square of the consumer voltage by squaring (6.11) leading to (6.21).

$$E(V_c^2) = V_S^2 - 2.V_s.E(\Delta V_x) + E(\Delta V_x^2) + E\left(\Delta V_y^2\right) \qquad (6.21)$$

To solve (6.20) and (6.21), the moments for the voltage-drop equations, (6.15) and (6.16), must be determined. The calculation of the moments of these expressions is achieved through the application of the additive and multiplicative properties of moments. The analysis is now separated based on topology and is discussed next, first for the 3p-4w system, then the 3p-3w system.

6.3.1 Three-phase Four-wire System

The moments are derived based on the voltage-drop equations achieved in (6.5) and (6.6). The first moments are achieved by the direct expectations of the respective equations. On the other hand, the second moments are obtained by computing the expectations of the squared voltage-drops.

1st raw moments

The expected voltage-drops depend on the expected phase currents according to:

$$E(\Delta V_x) = k_3 \cdot E(I_a) - \frac{1}{2}\cdot k_1 \cdot (E(I_b) + E(I_c)) - \frac{\sqrt{3}}{2}\cdot k_2 \cdot (E(I_b) - E(I_c)) \qquad (6.22)$$

$$E(\Delta V_y) = k_4 \cdot E(I_a) - \frac{1}{2}\cdot k_2 \cdot (E(I_b) + E(I_c)) + \frac{\sqrt{3}}{2}\cdot k_1 \cdot (E(I_b) - E(I_c)) \qquad (6.23)$$

For each of the phases, the currents I_a, I_b and I_c, generalised to I_p for illustration, are a sum of the individual customers' random active power currents Y_i connected to that phase, according to:

$$I_p = C_b\left(Y_1 + Y_2 + Y_3 + \cdots + Y_{m_p}\right) = C_b\sum_{j=1}^{m_p} Y_j \qquad (6.24)$$

Assuming the active power currents for all customers at the node of analysis have a common distribution $I_i \sim \mathrm{Be}(\alpha, \beta, 0, C_b)$, the expectation of the currents is given by:

$$E(I_p) = m_p \cdot G \cdot C_b \qquad (6.25)$$

where p denotes the phase identity *a*, *b*, or *c*, while m_p is total number of customers connected to phase p. *G* is the expected current for the individual beta currents.

The results of (6.25) substituted for the expected currents in (6.22) and (6.23) concludes the computation of the first moments for the voltage-drops.

$$E(\Delta V_x) = (k_3 \cdot m_a - k_1 \cdot C_1 - k_2 \cdot C_2) \cdot G \cdot C_b \qquad (6.26)$$

$$E(\Delta V_y) = (k_4 \cdot m_a - k_2 \cdot C_1 + k_1 \cdot C_2) \cdot G \cdot C_b \qquad (6.27)$$

where the constants $k_1 - k_4$ are the same as those provided in (6.7) to (6.10) while C_1 and C_2 are given by:

$$\begin{bmatrix} C_1 \\ C_2 \end{bmatrix} = \begin{bmatrix} 0.5(m_b + m_c) \\ 0.5\sqrt{3}(m_b - m_c) \end{bmatrix} \qquad (6.28)$$

2nd raw moment

The second moments can be interpreted as the expected value of the squared voltage-drop components in (6.5) and (6.6), and are presented as sums of element-wise (Hadamard matrix [242]) products as follows:

$$E(\Delta V_x^2) = \sum \left\{ \begin{bmatrix} k_3^2 \\ \frac{\sqrt{3}}{2} k_1 k_2 + 0.75 k_2^2 + 0.25 k_1^2 \\ -\frac{\sqrt{3}}{2} k_1 k_2 + 0.75 k_2^2 + 0.25 k_1^2 \end{bmatrix} \circ E \begin{bmatrix} I_a^2 \\ I_b^2 \\ I_c^2 \end{bmatrix} + \begin{bmatrix} -k_3 k_2 \sqrt{3} - k_1 k_3 \\ k_3 k_2 \sqrt{3} - k_1 k_3 \\ -\frac{3}{2} k_2^2 + \frac{k_1^2}{2} \end{bmatrix} \circ E \begin{bmatrix} I_a I_b \\ I_a I_c \\ I_b I_c \end{bmatrix} \right\}$$

$$(6.29)$$

$$E(\Delta V_y^2) = \sum \left\{ \begin{bmatrix} k_4^2 \\ -\frac{\sqrt{3}}{2} k_1 k_2 + 0.75 k_1^2 + 0.25 k_2^2 \\ \frac{\sqrt{3}}{2} k_1 k_2 + 0.75 k_1^2 + 0.25 k_2^2 \end{bmatrix} \circ E \begin{bmatrix} I_a^2 \\ I_b^2 \\ I_c^2 \end{bmatrix} + \begin{bmatrix} k_4 k_1 \sqrt{3} - k_2 k_4 \\ -k_4 k_1 \sqrt{3} - k_2 k_4 \\ -\frac{3}{2} k_1^2 + \frac{k_2^2}{2} \end{bmatrix} \circ E \begin{bmatrix} I_a I_b \\ I_a I_c \\ I_b I_c \end{bmatrix} \right\}$$

$$(6.30)$$

The expressions for the second moments given in (6.29) and (6.30) depend on the moments of the currents, I_a^2, I_b^2 and I_c^2, and those to the products between currents, $I_a I_b$, $I_a I_c$ and $I_b I_c$.

Extending the inference of (6.24) to squared currents and applying some algebra operations results in (6.31). Then, by computing the expected value of the expression, the second moments of the phase currents are determined as given in (6.32).

$$I_p^2 = \left(Y_1 + Y_2 + Y_3 + \cdots + Y_{m_p}\right)^2 = m_p \sum_{j=1}^{m_p} Y_j^2 + m_p(m_p - 1) \sum_{j=1}^{m_p} \sum_{\substack{k=1 \\ k \neq j}}^{m_p} Y_j \cdot Y_k \quad (6.31)$$

$$E\left(I_p^2\right) = m_p \cdot H \cdot C_b^2 + m_p(m_p - 1) \cdot G^2 \cdot C_b^2 \quad (6.32)$$

where H denotes the second moment of the individual load currents with parameters Beta$(\alpha, \beta, 0, C_b)$.

To determine the expectation of the current products, the current RVs are assumed independent such that the multiplicative property is applicable. Systems with dependent RVs will be addressed later on.

The multiplicative property for multivariate problems is provided in (6.33) and its application to the load currents differentiated by subscripts *m* and *n* is shown in (6.34).

$$E(I_n I_m) = E(I_n)E(I_m) \quad (6.33)$$

$$E(I_n I_m) = m_n m_m \cdot G^2 \cdot C_b^2 \quad (6.34)$$

By applying the outcomes from (6.32) and (6.34) back into the founding equations (6.29) and (6.30) the computation of the second moments is resolved. The expectations for the quadrature squares are:

$$E(\Delta V_x^2) = C_b^2(C_3 \cdot H + C_4 \cdot G^2) \quad (6.35)$$

$$E\left(\Delta V_y^2\right) = C_b^2(C_5 \cdot H + C_6 \cdot G^2) \quad (6.36)$$

where the constants $C_3 - C_6$ are given by:

$$\begin{bmatrix} C_3 \\ C_4 \\ C_5 \\ C_6 \end{bmatrix} = \begin{bmatrix} k_3^2 m_a + c_1 m_b + c_2 m_c \\ c_3 + c_4 \\ k_4^2 m_a + d_1 m_b + d_2 m_c \\ d_3 + d_4 \end{bmatrix} \quad (6.37)$$

The constants $k_1 - k_4$ are the same as those provided in (6.7) to (6.10) while $c_1 - c_4$ and $d_1 - d_4$are derived from (6.38) and (6.39), respectively:

$$\begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix} = \begin{bmatrix} 0.5k_1k_2\sqrt{3} + 0.75k_2^2 + 0.25k_1^2 \\ -0.5k_1k_2\sqrt{3} + 0.75k_2^2 + 0.25k_1^2 \\ k_3^2 m_a(m_a - 1) + c_1 m_b(m_b - 1) + c_2 m_c(m_c - 1) \\ m_a m_b(-k_3k_2\sqrt{3} - k_1k_3) + m_a m_c(k_3k_2\sqrt{3} - k_1k_3) + m_b m_c(-1.5k_2^2 + 0.5k_1^2) \end{bmatrix} \quad (6.38)$$

$$\begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \end{bmatrix} = \begin{bmatrix} -0.5k_1k_2\sqrt{3} + 0.75k_1^2 + 0.25k_2^2 \\ 0.5k_1k_2\sqrt{3} + 0.75k_1^2 + 0.25k_2^2 \\ k_4^2 m_a(m_a - 1) + d_1 m_b(m_b - 1) + d_2 m_c(m_c - 1) \\ m_a m_b(k_4k_1\sqrt{3} - k_2k_4) + m_a m_c(-k_4k_1\sqrt{3} - k_2k_4) + m_b m_c(-1.5k_1^2 + 0.5k_2^2) \end{bmatrix} \quad (6.39)$$

6.3.2 Three-phase Three-wire System

To avoid unnecessary repetition, the theory development discussed in Section 6.3.1 is not covered here. The derivation herein is restricted to demonstrating differentiated results between the two topologies. The derivations are based on the voltage-drop equations in (6.15) and (6.16).

1st raw moments

The expectation of (6.15) and (6.16) considering the results of (6.25) results in:

$$E(\Delta V_x) = (F_1 \cdot w_1 + F_2 \cdot w_2) \cdot G \cdot C_b \tag{6.40}$$

$$E(\Delta V_y) = (F_2 \cdot w_1 - F_1 \cdot w_2) \cdot G \cdot C_b \tag{6.41}$$

where the constants w_1 and w_2 are provided in (6.17) to (6.18) while F_1 and F_2 are given by:

$$\begin{bmatrix} F_1 \\ F_2 \end{bmatrix} = \begin{bmatrix} \left(2 \cdot m_a + 0.5 \cdot (m_b + m_c)\right) \\ \frac{\sqrt{3}}{2} \cdot (m_b - m_c) \end{bmatrix} \tag{6.42}$$

2nd raw moments

As applied in Section 6.3.1, the squared voltage-drops in (6.15) and (6.16) are used to derive the moment of the second order as follows:

$$E(\Delta V_x^2) = \sum \left\{ \begin{bmatrix} 4w_1^2 \\ \frac{\sqrt{3}}{2} w_1 w_2 + 0.75 w_2^2 + 0.25 w_1^2 \\ -\frac{\sqrt{3}}{2} w_1 w_2 + 0.75 w_2^2 + 0.25 w_1^2 \end{bmatrix} \circ E \begin{bmatrix} I_a^2 \\ I_b^2 \\ I_c^2 \end{bmatrix} + \begin{bmatrix} 2w_1 w_2 \sqrt{3} - 2w_1^2 \\ -2w_1 w_2 \sqrt{3} + 2w_1^2 \\ -\frac{3}{2} w_2^2 + \frac{w_1^2}{2} \end{bmatrix} \circ E \begin{bmatrix} I_a I_b \\ I_a I_c \\ I_b I_c \end{bmatrix} \right\} \tag{6.43}$$

$$E(\Delta V_y^2) = \sum \left\{ \begin{bmatrix} 4w_2^2 \\ -\frac{\sqrt{3}}{2} w_1 w_2 + 0.75 w_1^2 + 0.25 w_2^2 \\ \frac{\sqrt{3}}{2} w_1 w_2 + 0.75 w_1^2 + 0.25 w_2^2 \end{bmatrix} \circ E \begin{bmatrix} I_a^2 \\ I_b^2 \\ I_c^2 \end{bmatrix} + \begin{bmatrix} -2w_2 w_1 \sqrt{3} + 2w_2^2 \\ 2w_2 w_1 \sqrt{3} + 2w_2^2 \\ -\frac{3}{2} w_1^2 + \frac{w_2^2}{2} \end{bmatrix} \circ E \begin{bmatrix} I_a I_b \\ I_a I_c \\ I_b I_c \end{bmatrix} \right\} \tag{6.44}$$

Then, by applying the theory developed in (6.31) – (6.34), the second moments are solved. The outcomes are given by:

$$E(\Delta V_x^2) = C_b^2 (F_3 \cdot H + F_4 \cdot G^2) \tag{6.45}$$

$$E\left(\Delta V_y^2\right) = C_b^2 (F_5 \cdot H + F_6 \cdot G^2) \tag{6.46}$$

where the constants $F_3 - F_6$ are given by:

$$\begin{bmatrix} F_3 \\ F_4 \\ F_5 \\ F_6 \end{bmatrix} = \begin{bmatrix} 4k_1^2 m_a + e_1 m_b + e_2 m_c \\ e_3 + e_4 \\ 4k_2^2 m_a + f_1 m_b + f_2 m_c \\ f_3 + f_4 \end{bmatrix} \tag{6.47}$$

The constants w_1 and w_2 are provided in (6.17) to (6.18) while $e_1 - e_4$ and $f_1 - f_4$ are derived from (6.48) and (6.49), respectively:

$$\begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \end{bmatrix} = \begin{bmatrix} 0.5_1 w_2\sqrt{3} + 0.75w_2^2 + 0.25w_1^2 \\ -0.5w_1w_2\sqrt{3} + 0.75w_2^2 + 0.25w_1^2 \\ 4w_1^2 m_a(m_a - 1) + e_1 m_b(m_b - 1) + e_2 m_c(m_c - 1) \\ (2w_1w_2\sqrt{3} - 2w_1^2)m_a m_b + (-2w_1w_2\sqrt{3} + 2w_1^2)m_a m_c + (-1.5w_2^2 + 0.5w_1^2)m_b m_c \end{bmatrix} \quad (6.48)$$

$$\begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \end{bmatrix} = \begin{bmatrix} -0.5w_1w_2\sqrt{3} + 0.75w_1^2 + 0.25w_2^2 \\ 0.5w_1w_2\sqrt{3} + 0.75w_1^2 + 0.25w_2^2 \\ 4w_2^2 m_a(m_a - 1) + d_1 m_b(m_b - 1) + d_2 m_c(m_c - 1) \\ (-2w_2w_1\sqrt{3} + 2w_2^2)m_a m_b + (2w_2w_1\sqrt{3} + 2w_2^2)m_a m_c + (-1.5w_1^2 + 0.5w_2^2)m_b m_c \end{bmatrix} \quad (6.49)$$

The substitution for the expected voltage-drops into (6.20) and (6.21) completes the computation of the first and second moments of the consumer voltage for both three-phase topologies. However, to complete the statistical solutions, the limits to the expected customer voltages must be determined.

6.4 REFORMULATED NORMALIZATION TECHNIQUE

The initial HBA makes use of worst-case loading scenarios to determine the extreme voltage-drop conditions on a feeder. Using the approach, the occurrence of the extremum voltages (limits) is characteristic of specific phase-loading conditions; (011)[9] and (100) for the maximum and minimum voltages, respectively. While these were sufficient for LV feeders under the applied assumptions, rigorous tests indicate they are invalid with varied load power factor and feeder X/R ratios.

6.4.1 Testing the Validity of the HBA's Normalization Technique Under Varied Conditions of Power Factor and Line Reactance

By using the network equations in (6.1) – (6.11) and applying phase loading combinations to (000) - (111) in binary counting, to various cases of load power factor and feeder X/R ratio, the occurrence of the extrema voltages can be tested. In the investigation, the power factor was varied from 0.5 to 1, and the X/R ratio from 0.5 to 2, creating 30 test scenarios. These conditions, consistent with practical expectations and conditions reflected on most test feeders including the IEEE 33- and 69-bus, are appropriate for the rigorous testing of the performance of the normalization approach.

Results from the conducted scenarios, shown in Figure 6.2, demonstrate that the maximum voltage condition does not only occur when the phase under study (in this case, the red phase) is unloaded while the others are loaded (011); the loading states (001) and (010) also exhibits maxima voltage.

[9] (*xyz*) denotes the loading conditions at a node, such that *x*, *y*, and *z* customers are assigned to the phases *a*, *b*, and *c*, respectively.

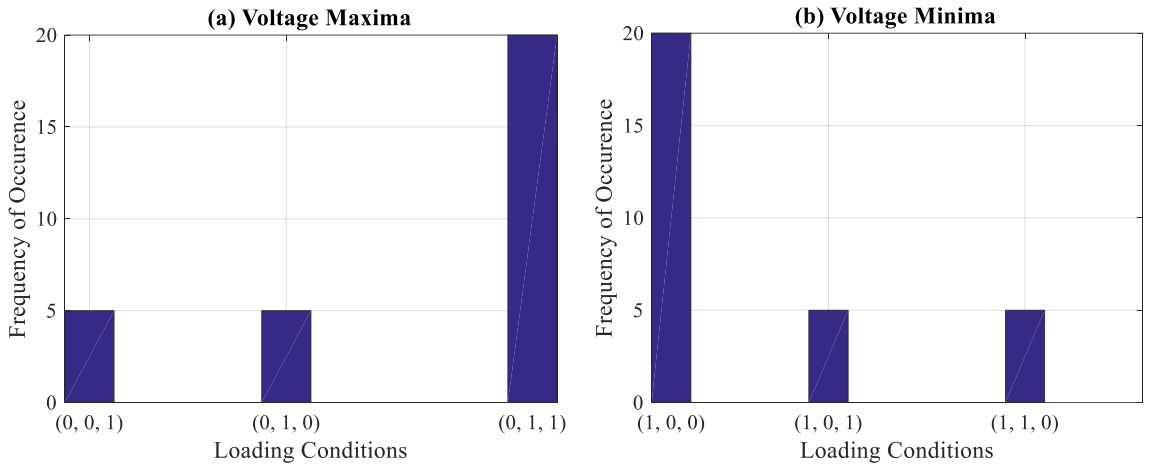

Figure 6.2: The occurrence of extremum voltages based on loading conditions

Similarly, the minimum voltage condition takes several loading states (100), (101) or (110). Adoption of the un-modified approach would result in inconsistent results and error. Hence, a new normalization technique is introduced.

6.4.2 Introducing a Per-unit System Normalisation Technique

The use of the per-unit system in power systems is common. According to this system, all input quantities such as voltage, apparent power, and impedance are specified as fractions of reference quantities or base levels. This is useful when working with composite systems with multiple nominal voltages via voltage transformers. The new approach takes advantage of this technique to normalize the load flow variables to the beta-acceptable range [0:1]. This is achieved by selecting the base quantities in such a way that the entire spectrum of probable consumer voltages falls within this range.

The statutory limits for voltage variations can guide the selection of the simulation voltage base, V_B which is a scaled version of the nominal voltage through a constant γ.

$$V_B = \gamma V_{nom} \qquad (6.50)$$

The system apparent power base, S_B, can be selected in the conventional way. With the two quantities determined, the calculation of the other base quantities can be achieved through the well-known transformations given by:

$$Z_B = \frac{V_B^2}{S_B}, \qquad I_B = \frac{S_B}{V_B} \qquad (6.51)$$

With all the base quantities determined, the scaled variables (denoted by the asterisk (*)) for the sending-end voltage V_s, the feeder impedance, Z, and the scaling current constant C_b are determined as follows:

$$V_s^* = \frac{V_s}{V_B}; \quad Z^* = \frac{Z}{Z_B}; \; C_b^* = \frac{C_b}{I_B} \qquad (6.52)$$

Performing the PLF in per-unit with V_B selected to ensure it is the absolute maximum, even with voltage-rise, the moments to the voltage-drop, $E(V_c)$ in (6.20) and $E(V_c^2)$ in (6.21), require no additional scaling. Hence, the shape parameters of the scaled voltage α_V^* and β_V^* can be determined directly. These parameters, including the limit parameters (0, V_B), model a 4-parameter beta function Beta(α_V^*, $\beta_V^*, 0, V_B$) given by:

$$V_c = f(V_c^*) = \frac{1}{b(\alpha_V^*, \beta_V^*)} \cdot \frac{{V_c^*}^{\alpha_V^*-1}(V_B - V_c^*)^{\beta_V^*-1}}{V_B^{\alpha_V^*+\beta_V^*-1}} \qquad (6.53)$$

From this distribution, the actual mean, variance, and percentile voltages can be determined through the properties of a 4-parameter distribution already discussed in Chapter 4.

6.5 EXTENSION TO DG

DGs are treated as negative loads and are specified by representing connected DGs as negative customers consistent with the original HB approach (described in Section 4.2.6), except for the normalisation technique as discussed in the previous section.

6.6 EXTENSION TO SUCCESSIVE NODES

The extension to successive nodes (up to *N* nodes) as discussed in Chapter 4, is implemented. The equations for the moments in (4.43) – (4.46) are recited here for easy reference:

$$E(\Delta V_{x_t}) = \sum_{i=1}^{N} E(\Delta V_{x_i}) \qquad (6.54)$$

$$E(\Delta V_{x_t}^2) = \sum_{i=1}^{N} E(\Delta V_{x_i}^2) + \sum_{j=1}^{N} \sum_{\substack{k=1 \\ k \neq j}}^{N} E\left(\Delta V_{x_j}\right) E(\Delta V_{x_k}) \qquad (6.55)$$

$$E(\Delta V_{y_t}) = \sum_{i=1}^{N} E(\Delta V_{y_i}) \qquad (6.56)$$

$$E(\Delta V_{y_t}^2) = \sum_{i=1}^{N} E(\Delta V_{y_i}^2) + \sum_{j=1}^{N} \sum_{\substack{k=1 \\ k \neq j}}^{N} E\left(\Delta V_{y_j}\right) E(\Delta V_{y_k}) \qquad (6.57)$$

6.7 DEPENDENCE AND CORRELATION

The original HBA assumed independence between input random variables. However, Section 3.4 in Chapter 3 identified the need to include three forms of correlation; load-load, DG-load, and DG-DG. For the sake of separation, the initial formulation of Sections 6.3.1 and 6.3.2 assumed independency between random variables. When dependence is considered, the calculation for the first moment (mean) is not affected, but the second moment is.

Looking at the equations for the second moments of voltage-drop in (6.29) and (6.30) [(6.43) and (6.44) for the 3p-3w model], the calculation of the expectation of the squared currents, $E(I_p^2)$, shows correlation between intra-nodal inputs while inter-phase interdependencies are linked to the products

between phases $E(I_n I_m)$. For the dependence between nodes (inter-nodal correlations), the calculation of the summated second moments in (6.55) and (6.57) are modified.

In general, for a bivariate system involving dependent random variables *X* and *Y*, the expected value of the product *XY* is given by:

$$E(XY) = E(X)E(Y) + cov(X,Y) \qquad (6.58)$$

where $cov(X,Y)$ is the covariance between the variables *X* and *Y*.

Under the assumption of independence, the covariance term equals zero, which is consistent with the formulation of (6.33) and (6.34). However, when dependence is factored in, the covariance term is non-zero. Assuming linear correlation modelled through a correlation factor $\rho_{X,Y}$, the covariance between the two variables can be expressed in terms of the standard deviations as follows:

$$Cov(X,Y) = \sigma_X \cdot \sigma_Y \cdot \rho_{X,Y} \qquad (6.59)$$

Then, by substituting for the standard deviations with moments, the (6.60) results.

$$Cov(X,Y) = \sqrt{E(X^2) - E(X)^2} \cdot \sqrt{E(Y^2) - E(Y)^2} \cdot \rho_{X,Y} \qquad (6.60)$$

This result put back into (6.58) achieves the following:

$$E(XY) = E(X)E(Y) + \sqrt{E(X^2) - E(X)^2} \cdot \sqrt{E(Y^2) - E(Y)^2} \cdot \rho_{X,Y} \qquad (6.61)$$

This result is now used to modify the formulation of the transform considering the identified interdependencies. The modification of the calculations is performed separately as follows.

6.7.1 Intra-phase Correlations

Correlation between currents at the same node affect the sums of the squared currents I_a^2, I_b^2, and I_c^2. The expression initially given in (6.32) with the assumption of dependency, now becomes,

$$E(I_p^2) = \sum_{j=1}^{m_p} E(Y_j^2) + m_p(m_p - 1) \sum_{j=1}^{m_p} \sum_{\substack{k=1 \\ k \neq j}}^{m_p} (E(Y_j \cdot Y_k) + \sqrt{E\left(Y_j^2\right) - E\left(Y_j\right)^2} \cdot \sqrt{E(Y_k^2) - E(Y_k)^2} \cdot \rho_{j,k} \qquad (6.62)$$

Applying the earlier assumption that the customers connected to a common node (which includes common phase connections) have the same statistical distributions of currents, and that a common correlation factor $\rho_{p,p}$ applies to all the customers in phase *p*, (6.62) is simplified to:

$$E\left(I_p^2\right) = m_p \cdot H \cdot C_b^2 + m_p G^2 C_b^2 \left(m_p - 1\right) + m_p \rho_{p,p} C_b^2 \left(m_p - 1\right)(H - G^2) \qquad (6.63)$$

The result of (6.63), applied to all phases and in both calculation for the moments $E(\Delta V_x^2)$ and $E(\Delta V_y^2)$, results in changes in the constants $C_3 - C_6$. The summary of the changes will be presented collectively at the end, in Section 6.7.4.

6.7.2 Inter-phase Correlation

The correlation between phase currents affects the calculation of the expectation of the products I_aI_b, I_aI_c and I_bI_c. Applying the outcome of (6.61) to the calculation of the moments involving two phases, a and b, whose currents are correlated according to $\rho_{a,b}$ gives:

$$E(I_aI_b) = E(I_a)E(I_b) + \sqrt{E(I_a^2) - E(I_a)^2} \cdot \sqrt{E(I_b^2) - E(I_b)^2} \cdot \rho_{a,b} \qquad (6.64)$$

With the phase currents sharing the same distribution, (6.62) can be simplified as follows:

$$E(I_aI_b) = m_am_bE(Y)^2C_b^2 + \rho_{a,b}\sigma_Y^2C_b^2 \qquad (6.65)$$

$$E(I_aI_b) = m_am_bE(Y)^2C_b^2 + \rho_{a,b}C_b^2(E(Y^2) - E(Y)^2) \qquad (6.66)$$

$$E(I_aI_b) = m_am_bC_b^2G^2 + \rho_{a,b}C_b^2(H - G^2) \qquad (6.67)$$

The result of (6.67) is applied to the other pairs of currents, I_aI_c and I_bI_c, and the application of the changes to the equations of $E(\Delta V_x^2)$ and $E(\Delta V_y^2)$ also impact the constants $C_3 - C_6$.

6.7.3 Inter-node Correlations

The dependence addressed up to this point relates to variables at a single node. To cater for the dependence occurring between node inputs, the second moments of the summated voltage-drops in (6.55) and (6.57) need reformulation. Using the result of (6.61), the changes are as follows:

$$E\left(\Delta V_{x_t}^2\right) = \sum_{i=1}^{N} E\left(\Delta V_{x_i}^2\right) + \sum_{r=1}^{N}\sum_{\substack{s=1 \\ s\neq r}}^{N}\left[E\left(\Delta V_{x_r}\right)E\left(\Delta V_{x_s}\right) + \left(\sigma_{V_{x_r}} \cdot \sigma_{V_{x_s}} \cdot \rho_{r,s}\right)\right] \qquad (6.68)$$

$$E\left(\Delta V_{y_t}^2\right) = \sum_{i=1}^{N} E\left(\Delta V_{y_i}^2\right) + \sum_{r=1}^{N}\sum_{\substack{s=1 \\ s\neq r}}^{N}\left[E\left(\Delta V_{y_r}\right)E\left(\Delta V_{y_s}\right) + \left(\sigma_{V_{y_r}} \cdot \sigma_{V_{y_s}} \cdot \rho_{r,s}\right)\right] \qquad (6.69)$$

Where,

$$\sigma_{V_{x_r}} = \sqrt{\left(E\left(V_{x_r}^2\right) - E\left(V_{x_r}\right)^2\right)};\quad \sigma_{V_{y_r}} = \sqrt{\left(E\left(V_{y_r}^2\right) - E\left(V_{y_r}\right)^2\right)} \qquad (6.70)$$

$$\sigma_{V_{x_s}} = \sqrt{E\left(V_{x_s}^2\right) - E\left(V_{x_s}\right)^2};\ \sigma_{Vy_s} = \sqrt{E\left(V_{y_s}^2\right) - E\left(V_{y_s}\right)^2} \qquad (6.71)$$

Where $\rho_{r,s}$ denotes the correlation coefficient relating the inputs at the r-th and s-th nodes.

6.7.4 Changes to the Algorithms

The approaches discussed can be applied to any node in such a way that all the three forms of dependence – load-load, load-DG and DG-DG – are covered.

To implement changes to the algorithm in accordance with the refined analysis discussed, new constants c_5, c_6, d_5 and d_6, are introduced. The first two, based on 'c' are used in the calculation of $E(\Delta V_x^2)$,

while the other two, based on 'd', for $E(\Delta V_y^2)$. The constants are given in (6.72) and (6.73), respectively.

$$\begin{bmatrix} c_5 \\ c_6 \end{bmatrix} = \begin{bmatrix} \rho_{a,a} k_3^2 m_a (m_a - 1) + \rho_{b,b} c_1 m_b (m_b - 1) + \rho_{c,c} c_2 m_c (m_c - 1) \\ \rho_{a,b} m_a m_b \left(-k_3 k_2 \sqrt{3} - k_1 k_3\right) + \rho_{a,c} m_a m_c \left(k_3 k_2 \sqrt{3} - k_1 k_3\right) + \rho_{b,c} m_b m_c (-1.5 k_2^2 + 0.5 k_1^2) \end{bmatrix} \quad (6.72)$$

$$\begin{bmatrix} d_5 \\ d_6 \end{bmatrix} = \begin{bmatrix} \rho_{a,a} k_4^2 m_a (m_a - 1) + \rho_{b,b} d_1 m_b (m_b - 1) + \rho_{c,c} d_2 m_c (m_c - 1) \\ \rho_{a,b} \left(k_4 k_1 \sqrt{3} - k_2 k_4\right) m_a m_b + \rho_{a,c} \left(-k_4 k_1 \sqrt{3} - k_2 k_4\right) m_a m_c + \rho_{b,c} (-1.5 k_1^2 + 0.5 k_2^2) m_b m_c \end{bmatrix} \quad (6.73)$$

While maintaining the structure of the equations in (6.35) and (6.36), the changes in Table 6.1 are made to the constants $C_3 - C_6$ for the computation of voltage-drops for individual nodes, and to the feeder quantities in (6.65) and (6.67).

Although this extension specifically addressed the 3p-4w model, the same approach applies to the 3p-3w model because the solutions for both systems have similar structures. The changes made in (6.72) and (6.73) are extended to the 3p-3w model according to (6.74) and (6.75) now based on constants *e* and *f*. The corresponding changes to the second moments are detailed in Table 6.1.

$$\begin{bmatrix} e_5 \\ e_6 \end{bmatrix} = \begin{bmatrix} \rho_{a,a} 4 w_1^2 m_a (m_a - 1) + \rho_{b,b} e_1 m_b (m_b - 1) + \rho_{c,c} e_2 m_c (m_c - 1) \\ \rho_{a,b} m_a m_b \left(2 w_1 w_2 \sqrt{3} - 2 w_1^2\right) + \rho_{a,c} m_a m_c \left(-2 w_1 w_2 \sqrt{3} + 2 w_1^2\right) + \rho_{b,c} m_b m_c (-1.5 w_2^2 + 0.5 w_1^2) \end{bmatrix} \quad (6.74)$$

$$\begin{bmatrix} f_5 \\ f_6 \end{bmatrix} = \begin{bmatrix} \rho_{a,a} 4 w_2^2 m_a (m_a - 1) + \rho_{b,b} f_1 m_b (m_b - 1) + \rho_{c,c} f_2 m_c (m_c - 1) \\ \rho_{a,b} m_a m_b (-2 w_2 w_1 \sqrt{3} + 2 w_2^2) + \rho_{a,c} m_a m_c (2 w_2 w_1 \sqrt{3} + 2 w_2^2) + \rho_{b,c} m_b m_c (-1.5 w_1^2 + 0.5 w_2^2) \end{bmatrix} \quad (6.75)$$

Table 6.1: Modifications to include dependence in the HBE algorithm

AFFECTED VARIABLES	**SYSTEM MODEL AND CONDITION OF DEPENDENCE**			
	3p-4w Model		3p-3w Model	
	Uncorrelated	Correlated	Uncorrelated	Correlated
$E(\Delta V_{x_i}^2)$	C_3	$C_3 + c_5 + c_6$	F_3	$F_3 + e_5 + e_6$
	C_4	$C_4 - c_5 - c_6$	F_4	$F_4 - e_5 - e_6$
$E(\Delta V_{y_i}^2)$	C_5	$C_5 + d_5 + d_6$	F_5	$F_5 + f_5 + f_6$
	C_6	$C_6 - d_5 - d_6$	F_6	$F_6 - f_5 - f_6$
$E(\Delta V_{x_t}^2)$	According to (6.55)	See changes in (6.68)	According to (6.45)	See changes in (6.68)
$E(\Delta V_{y_t}^2)$	According to (6.57)	See changes in (6.69)	According to (6.46)	See changes in (6.69)

The correlation factors required in the calculations of (6.72) – (6.75) can be derived from provided correlation matrices. For a system with *N* buses, an $N \times N$ correlation matrix is required to characterize the dependence between nodal inputs:

$$\rho = \begin{bmatrix} \rho_{1,1} & \rho_{1,2} & \rho_{1,3} & \cdots & \rho_{1,N} \\ & \rho_{2,2} & \rho_{2,2} & \cdots & \rho_{2,N} \\ & & \cdots & \cdots & \cdots \\ & & & \cdots & \rho_{N-1,N} \\ & & & & \rho_{N,N} \end{bmatrix} \quad (6.76)$$

For inter-phase dependency, the three coefficients $\rho_{a,b}$, $\rho_{a,c}$ and $\rho_{b,c}$ need to be specified per node. The same applies to intra-phase correlation factors $\rho_{a,a}$, $\rho_{b,b}$ and $\rho_{c,c}$.

6.8 COMPENSATED FEEDERS

6.8.1 Characterising Line Admittance, Shunt Capacitors and Reactors

Chapter 3, which dealt with input models for PLF analysis, indicated that the shunt capacitance of medium-length and long feeders becomes significant to affect voltage and current conditions; therefore, their representation using the π-model is essential. The model consists of a series impedance and line charging effects divided between the two shunt arms, each with an admittance of Y/2. The line thus acts as a reactive power source by means of the shunt admittances. The reactive power generated by the line is determined by its admittance Y and the voltage conditions of the line as described below [243]:

$$Q_i = \left|V_i^2\right| \cdot \frac{Y_i}{2} \quad (6.77)$$

This injected reactive power is associated with a current given by:

$$I_i = \frac{Y_{l,i}}{2} V_i \quad (6.78)$$

These formulations can be generally applied to power factor correction capacitors and rotating synchronous condensers characterised as a constant-impedance $Y_{c,i}$ as follows:

$$I_{c,i} = Y_{c,i} V_i \quad (6.79)$$

Further, where shunt reactors are used for voltage regulation, and modelled as constant-impedance $Y_{r,i}$, (6.79) can be used to characterise the current drawn by the inductors. The statistical models to these variables form the subject of the next section.

6.8.2 Statistical Modelling and Analytical Means

In Figure 6.3, L1 and L2 represent local loads of type 1 and 2 which might be defined with any particular voltage dependence. These also include downstream loads collected into the respective load groups.

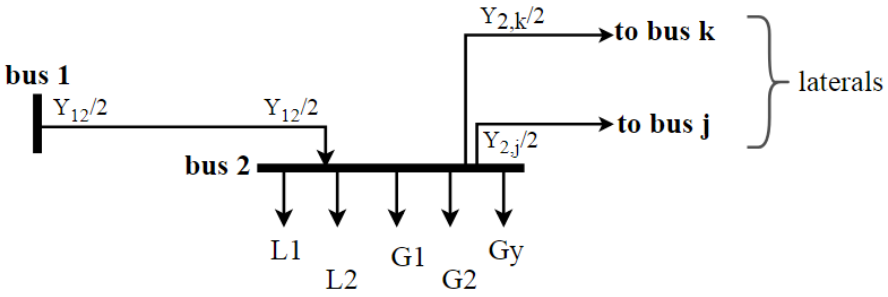

Figure 6.3: Including branch shunt capacitance in connection to node

Likewise, G1 and G2 represent the sum of local and downstream DG of type 1 and 2, respectively. Gy represents all shunt connected capacitance allocated to the node, including:

- Locally connected branches $Y_{1,2}/2$, $Y_{2,j}/2$, $Y_{2,k}/2$
- All other downstream admittances Y/2 collected to this bus
- All shunt connected capacitors connected at this node and downstream buses

All components are modelled as voltage-dependant currents of constant-impedance type. And since this thesis assumes the network parameters are fixed, the capacitors are modelled as deterministic variables. As demonstrated in Chapter 4, a beta PDF with large shape parameters ($(\alpha = \beta) \to \infty$) estimates the deterministic value x = 0.5 and can be scaled to any other value using C_b. In this work, the parameters $\alpha = \beta = 5 \times 10^5$ are adopted to limit the variability to a standard deviation of 0.1% of the nominal value. Figure 6.4 shows the distribution and limits of the adopted model with C_b = 2.

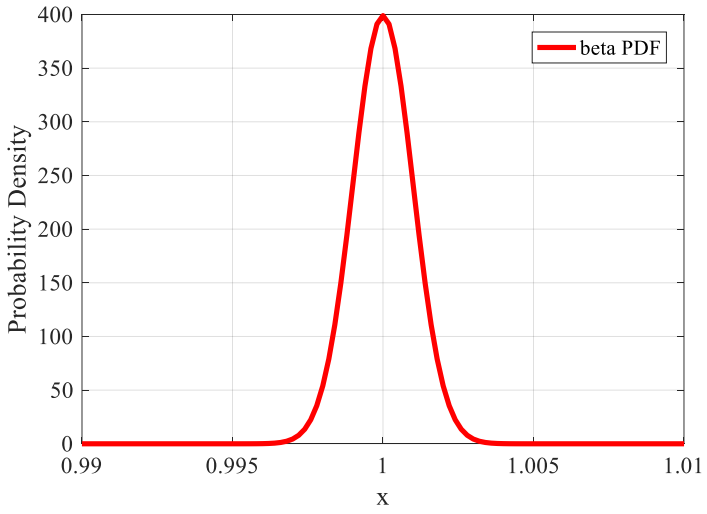

Figure 6.4: Beta PDF approximate to deterministic models of capacitance

The distribution parameters may also be used, if desired, to represent uncertainty in the capacitance. In this study, however, the model is adopted only as an approximation of the deterministic model. The

following section considers extending the HBE for the accurate analysis of feeders with voltage-dependent inputs, such as the ones identified here.

6.9 VOLTAGE-DEPENDENT LOAD MODELS

The developed approach, up to this point, is valid for inputs defined as constant-currents. To extend the approach to voltage-dependent input models, an approach based on repeated HBE calculations is proposed. Initially, all loads are assumed to operate at nominal voltage, say 1.0 pu. With this assumption, all load-types are indifferent, in other words they become constant-current models. Table 6.3 provides the calculation methodology for input currents for the various technologies.

Table 6.2: Determining load currents for various load models

NETWORK ELEMENT	**LOAD CHARACTERISTIC**	
	Voltage-dependency	***Current Equation***
Load or DG	Constant-Z	$I_k = V_k / Z_k$
	Constant-I	$I_k = I_k$
	Constant-PQ	$I_k = \left(\frac{S_k}{V_k}\right)^*$
Shunt Capacitor	Constant-Z	$I_k = Y_{c,k} V_k$
Shunt Reactor	Constant-Z	$I_k = Y_{r,k} V_k$
Shunt Admittance	Constant-Z	$I_k = \frac{Y_{l,k}}{2} V_k$

With the currents determined, an HBE is then applied to determine the bus voltages. Thereafter, the assigned current at each node is adjusted according to the calculated voltage in the previous iteration of the HBE, effectively adopting the respective voltage-dependant load model. Then, the HBE transform is computed iteratively with the updated input currents at each node.

The process is terminated when the maximum difference of voltage magnitudes in successive iterations falls below a selected tolerance value, used as a convergence criterion. The procedure can be summarised in a few steps as follows:

1. Set iteration count to $k = 1$
2. Assign currents per node by assuming the nominal voltage at each node
3. Compute the HBE based on the input currents. Store bus voltages, $V^{(k)}$
4. Increment iteration count, $k = k + 1$
5. Update the assigned currents at each node using the voltage outcomes from the previous iteration $V^{(k-1)}$
6. Repeat an HBE calculation using the updated input currents. Store bus voltages, $V^{(k)}$

7. Check convergence by testing the statistic $\left|V^{(k)} - V^{(k-1)}\right| < \varepsilon$ for each node.

8. End iterations if all buses meet the convergence criteria, otherwise repeat steps (4) – (7).

Assessing the effect of applying the voltage-dependent load models in the HBE transform PLF evaluates its extension to simulating feeders with constant-power load models and π-modelled lines with shunt capacitance.

6.10 CALCULATION OF BRANCH CURRENTS

The calculation of branch currents is important to determine the thermal loading conditions for conductors in a given network. The approach for the calculation of branch currents is similar to the calculation of voltage-drops. The total branch currents are dependent on the sum of currents locally at the node (as is the case with voltage-drops due to an individual node) and the amount of downstream currents connected to the node (as is the case with summated voltage-drops but considering upstream voltage-drops).

6.10.1 Deriving the Statistical Solution to Branch Currents

Using the pictorial of Figure 6.3 to illustrate the summation of currents, the branch current between buses 1 and 2, $I_{1,2}$, is derived from the sum of two sources of currents,

(a) Locally connected loads, DG, and shunt components (both capacitors and admittances)

(b) All downstream currents linked to the receiving bus (bus 2 in this case)

Currents from locally connected inputs

The summation of intranodal localized input currents at a single node is similar to the calculation of the voltage-drops using superposition. Using phase-A for illustration, the calculation is as follows:

$$I_a = C_b \sum_{j=1}^{m_a} Y_j \qquad (6.80)$$

Where Y_j is a random current drawn from the distribution of the j-th input current, and m_a is the total number of customers in that phase.

Assuming all customer currents at the node of analysis have a common distribution $I_i \sim \mathrm{Be}(\alpha, \beta, 0, C_b)$, the expectation of the phase currents is given by:

$$E(I_a) = m_p \cdot G \cdot C_b \qquad (6.81)$$

The second moments, considering correlation between the inputs in that phase ($\rho_{a,a}$) are given by:

$$E(I_a^2) = m_a \cdot H \cdot C_b^2 + m_a G^2 C_b^2 (m_a - 1) + m_a \rho_{a,a} C_b^2 (m_a - 1)(H - G^2) \qquad (6.82)$$

Since I_a models the active component of the load currents, the expected absolute currents are scaled quantities of the expected currents of (6.81) and (6.82) as follows:

$$E(I_A) = \frac{E(I_a)}{cos\,\theta_a} \qquad (6.83)$$

$$E(I_A^2) = \frac{E(I_a^2)}{cos\,\theta_a^2} \qquad (6.84)$$

These calculations are performed at every node on the network, which allows the separation of loads, DG, and shunt components accordingly. This completes the calculation of currents from locally connected inputs.

Total currents including those from downstream inputs

On a feeder with *N* nodes, the current in the branch leading to bus k from bus j ($I_{j,k}$), will include its local currents and downstream currents as follows:

$$I_{j,k} = \sum_{q=k}^{N} I_{A_q} \qquad (6.85)$$

The moments of these currents considering correlation between currents at nodes *r* and *s* ($\rho_{r,s}$) are given by:

$$E(I_{j,k}) = \sum_{q=k}^{N} E(I_{A_q}) \qquad (6.86)$$

$$E(I_{j,k}^2) = \sum_{q=k}^{N} E\left(I_{A_q}\right) + \sum_{r=k}^{N} \sum_{\substack{s=k \\ s\neq r}}^{N} \left[E(I_{A_r})E(I_{A_s}) + (\sigma_{I_r} \cdot \sigma_{I_s} \cdot \rho_{r,s})\right] \qquad (6.87)$$

With the moments determined, the next step involves the calculation of the beta-PDF parameters of the branch current. For this, the range of the distribution must be determined to allow normalization.

6.10.2Normalization of Branch Currents

For a passive radial network, it is expected that the bus voltage V_j at a node j closer to the source is always greater than that at the next downstream bus, V_k, unless unbalance conditions cause voltage-rise. However, on active networks, the extent of voltage-rise can result in the condition $V_k > V_j$ in which reverse power flow results. As such, the expected current flow is positive when $V_j > V_k$ and negative (reverse flow) when $V_k > V_j$. Accordingly, the distribution of branch currents potentially falls between a negative and positive limit.

These extrema limit to the branch currents can be quantified in terms of the thermal rating of the conductor and applying the normalization technique introduced in Section 6.4.2.

Taking the thermal rating of the branch to be the nominal current I_{nom}, a scalar constant γ can be applied to determine the base branch current. This quantity is denoted I_{bB} to differentiate it from the system base current I_B used in the normalization of load currents.

$$I_{bB} = \gamma I_{nom} \tag{6.88}$$

Accordingly, the extrema limits are given by:

$$I_{j,k}^{min} = -I_{bB} \tag{6.89}$$

$$I_{j,k}^{max} = I_{bB} \tag{6.90}$$

It must be noted that although 'the extrema limit to the branch currents' are the base, they are not the range within which the calculated currents will be restricted. In other words, it is possible for the calculated current to exceed 1 or -1 pu of I_{bB}. The range is defined in (6.91) and (6.92).

$$E\left(I_{j,k}^{*}\right) = \frac{E(I_{A\,j,k}) - I_{j,k}^{min}}{I_{j,k}^{max} - I_{j,k}^{min}} \tag{6.91}$$

$$E\left(I_{j,k}^{*\,2}\right) = \frac{E(I_{A\,j,k}^{2}) - 2E(I_{A\,j,k}) \cdot I_{j,k}^{min} + I_{j,k}^{min\,2}}{\left(I_{j,k}^{max} - I_{j,k}^{min}\right)^{2}} \tag{6.92}$$

As already demonstrated, the parameters of the scaled quantities, α_I^* and β_I^* can be calculated directly from the first and second scaled moments. Resultantly, the branch currents follow a 4-parameter beta function Beta$\left(\alpha_I^*, \beta_I^*, I_{j,k}^{min}, I_{j,k}^{max}\right)$. From this distribution, the actual mean, variance, and percentile values can be determined.

For the 3p-3w system, the same approach is applied but the final outcomes have to be modified. In the three-wire model, the branch currents are not equal to the load (phase) currents. The transformation in (6.93) and (6.94) is required to derive the moments of the line quantities from the corresponding phase quantities in (6.86) and (6.87).

$$E\left(I_{A_{L_{(j,k)}}}\right) = \sqrt{3}E\left(I_{A\,j,k}\right) \tag{6.93}$$

$$E\left(I_{A_{L_{(j,k)}}}^{2}\right) = 3E\left(I_{A\,j,k}^{2}\right) \tag{6.94}$$

After this modification, (6.88) – (6.94) are applicable. This completes the derivation of the approach to branch currents.

6.11 INPUT REQUIREMENTS OF THE HBE TRANSFORM

The list of equations for the calculation of bus voltages on 3p-4w and 3p-3w systems using the HBE approach are detailed in Appendix A. The input requirements necessary for the execution of the PLF analysis are listed according to parameter type as follows:

- Simulation Parameters (per PLF analysis)
- System nominal voltage, V_{nom} (V)
- Scaling constant γ to determine V_B (V).

- Base power S_B (MVA)
- Feeder Properties (per branch)
- Feeder electrical properties at rated temperature and with temperature coefficient, allowing conductor temperature to be modelled; R_i/km, X_i/km, Y_i/km and line capacity [A] and conductor length to determine the specific quantities
- For 3p-4w model, specify the constants k_x and k_y
- Input Properties (per node)
- Number of customers according to phase allocation; +ve for loads (and shunt reactors), -ve for DG (and shunt capacitors)
- Statistical load properties of the active power currents, Beta $(\alpha, \beta, 0, C_b)$
- Power factor; denoted by positive sign for inductive inputs and a negative sign for capacitive ones.
- Correlation matrix for dependent input variables
- Assessment criteria
- Risk margins V_{risk} and I_{risk} (%)

6.12 CONCLUSION

This chapter has presented the development of the HBE transform and its extensions. The developed approach overcomes the limitations associated with the original HBA algorithm by removing the following assumptions:

1. The modelling of feeders as purely resistive lines
2. The modelling of load currents at unity power factor
3. The characterization of the relationship between random variables as independent

Without the assumptions and based on a new system standardization technique per unit, the new approach allowed several extensions to new applications. The following extensions were covered:

1. PLF analysis on MV and HV feeders
2. The analysis of systems with correlated random variables, including balanced 3-phase loads.
3. The analysis of systems with voltage-dependent load or DG models
4. The analysis of compensated feeders with shunt capacitors (including line charging through shunt admittances) and shunt reactors.

5. The computation of statistical line currents

The developed approach and its extensions are tested in the next chapter.

Chapter 7: Testing the Herman-Beta Extended (HBE) Transform

This chapter tests the validity of the developed HBE transform against the Monte-Carlo simulation. Several test scenarios are conducted to ensure the approach is tested rigorously. The accuracy of the HBE is assessed using three indicators, the mean, variance and percentile values at selected risk margins.

7.1 INTRODUCTION

The new HBE transforms developed in Chapter 6 for application to radial networks are tested here to assess the validity, accuracy and computational efficiency. The assessment compares the outcomes from the HBE with outcomes from validation tools. Two tools are used; the MCS is the principal validation tool to the statistical solutions, while deterministic outcomes from published results are used to assess the accuracy of the means.

The MCS tool is based on the deterministic voltage-drop equations as developed in-part in Chapter 4 (for 3p-4w models) and 6 (for 3p-3w models). The statistical analysis is possible through simple random sampling (SRS) for systems without interdependence, and copulas where correlations exist. The approach uses 15,000 DLF scenarios to produce corresponding output scenarios which are represented as histograms and used to assess the fitness of the PDFs from the HBE. The optimal numbers of bins for the representation of the MCS histograms are determined using the Freedman-Diaconis rule [244].

For the quantitative analysis of the outcomes, three test statistics are used; the means, standard deviation and percentiles at selected risk margins. The risk levels are selected in accordance with the South African quality of supply (QoS) guideline which stipulates a 95% confidence interval for voltage measurements [89]. This is interpreted as 5% confidence interval for passive feeders (with loads only) and two-sided limits of 2.5% (2.5% and 97.5% confidence intervals) for active feeders (with loads and DG) [91]. In the results presented, the assessment of active feeders is focussed on voltage-rise, hence only the 97.5% confidence interval is considered. In the absence of guidelines on limits of thermal loading during short periods, a nominal value of 100% loading assessed at 95% confidence interval is used to assess the thermal loading conditions of line sections.

All simulations, except the referenced deterministic results, are performed under a common software environment in MATLAB (2016a version) on a machine with the following specifications: Intel ® Core™ i5, 3 GHz, 16 GB RAM and 64-bit Operating System. This information is relevant to contextualise the outcomes of computational performance tests which are also covered in this chapter.

7.2 TESTING THE REFORMULATED TRANSFORM WITH *X* AND *Q*

The objective of this test is to assess the validity of the reformulated HBE transform with the full representation of line reactance and load reactive power. A representative radial feeder of 3p-4w topology and operated at 11 kV is used. Figure 7.1 shows the one-line diagram of the 12-bus feeder with no laterals. The feeder is made up of line sections with X/R ratios between 0.5 and 3.5 and comprises a total of 22 loads with power factors ranging between 0.5 and 0.95. It is assumed that the neutral conductor has the same electrical properties as the phase conductors ($k_r = k_x = 1$).

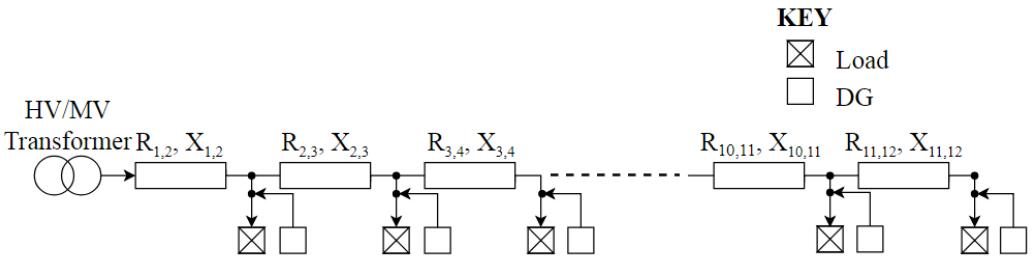

Figure 7.1: Representative 12-bus feeders with unbalanced loads

The load demand for each customer is statistically characterised by a skewed beta PDF with parameters $\alpha = 1.5, \beta = 5$ while the scaling factor, *C*, differs from node to node. The full load and line properties per node for the passively-configured feeder is provided in Appendix B1. The distribution shape parameters for DG are derived from [245] for noon-time profiles of PV power generation. Table 7.1 provides these parameters together with the selected rated power and operating power factors.

Table 7.1: Photovoltaic generation beta distribution parameters on the 12-bus representative feeder

BUS ID.	DG PARAMETERS			
	α	β	Power (kVA)	pf
2	12.62	2.21	500	0.900
6	12.62	2.21	300	1.000
8	12.62	2.21	600	0.950
12	12.62	2.21	700	0.975

The diverse conditions of power factor and X/R ratio are suitable for extensively testing the validity of the HBE's reformulation to include the reactive input components. Two test cases are performed; under passive and active configurations.

7.2.1 Testing the Validity of the HBE Formulation to Include *X* and *Q*

The PLF is conducted on a passive network (without DG) to test the validity of the reformulation before its extension to DG. By comparing the outcomes from the HBE to those from the MCS, its accuracy is assessed.

Figure 7.2 compares the output functions for a selected bus. The differentiated profiles between the phases confirms the unbalanced loading conditions on the feeder. As seen from the plots, the HBE-transform closely approximates the distribution of bus voltages from the MCS in all phases.

The distributions are reasonably representative of the results at all nodes. The full spectrum of voltage means and standard deviations (SDs) depicted in Figure 7.3 demonstrates the consistent outcomes across the nodes and phases.

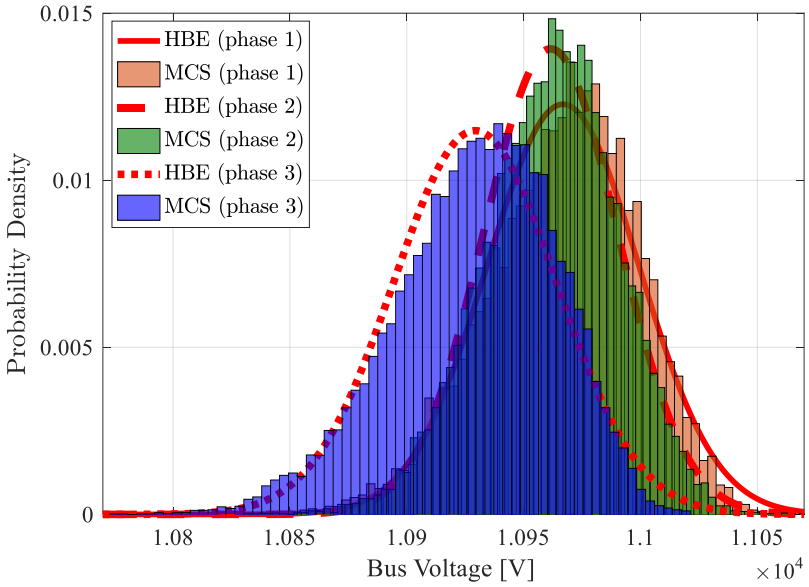

Figure 7.2: MCS data vs. HBE profile: Bus 3 voltages on passive 12-bus feeder

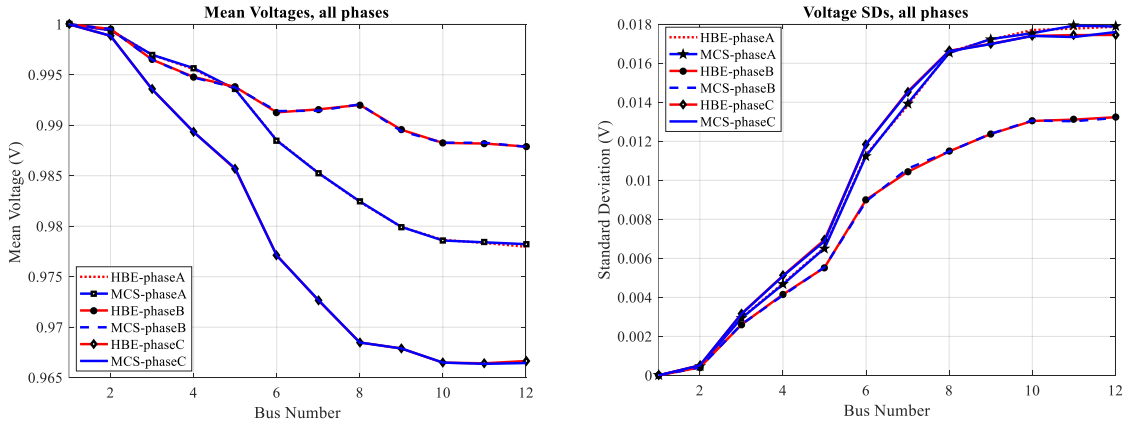

Figure 7.3: Means and SDs for all phases on the passive 12-bus feeder

Further, given the expectation of convergence to normality according to the CLT, the expected approximation errors are also reduced. The results of the many case studies carried out in this chapter test this intuition.

Table 7.2 quantitatively compares the means, standard deviation, and percentile voltage for selected buses and phases. The results show that the MCS and HBE voltages differ only slightly; the relative

percentage differences are lower than 0.02%, 0.85% and 0.04% in the mean, variance and 95th-percentile respectively. The differences are primarily due to two factors: the expected variations in the MCS outcomes due to random sampling and the approximation of the output distribution by the beta PDF.

The differences related to the approximation of the beta PDF mainly affects the calculation of percentile voltages and is dependent on how the beta-PDF fits the output data. As noted in this case study, this error is not significant because it is offset by the flexibility of the beta PDF to model differently distributed data. Further, given the expectation of convergence to normality according to the CLT, the expected approximation errors are also reduced. The results of the many case studies carried out in this chapter test this intuition.

Table 7.2: Sample voltage outcomes for the passive 12-bus feeder

BUS (PHASE)	**STATISTICAL CHARACTERISTICS OF BUS VOLTAGES**					
	MCS			**HBE**		
	$\boldsymbol{\mu_V}$ (pu)	$\boldsymbol{\sigma_V}$ (pu)	$\boldsymbol{p_{r\%}}$ (pu)	$\boldsymbol{\mu_V}$ (pu)	$\boldsymbol{\sigma_V}$ (pu)	$\boldsymbol{p_{r\%}}$ (pu)
2(A)	0.9994	4.83e-04	0.9986	0.9994	4.850e-4	0.9986
6(C)	0.9773	0.0117	0.9572	0.9771	0.0118	0.9576
9(B)	0.9896	0.0124	0.9689	0.9896	0.0124	0.9692
12(A)	0.9781	0.0178	0.9484	0.9780	0.0179	0.9482

To demonstrate the influence of the variability in the MCS outcomes on depicting the accuracy of the HBE, mean voltages and SDs for bus 3 from 500 repeated MCS PLF trials with 500 runs each are plotted in Figure 7.4 and compared with singular results from the HBE transform.

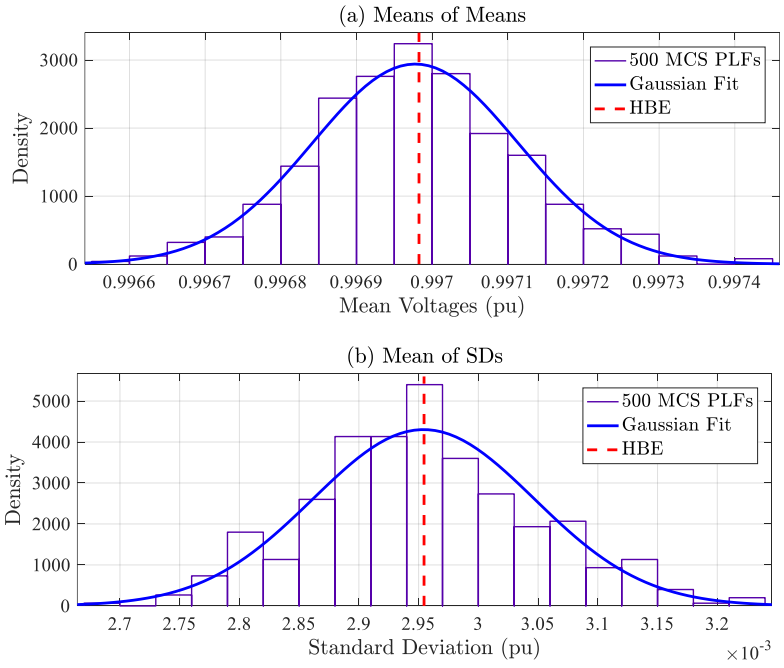

Figure 7.4: Repeated MCS LFs vs HBE singular solutions: Bus 3

The plot shows the HBE singular results are central in both MCS distributions. This implies the transform is accurate since the 'mean of the means' (Figure 7.4a) and the average of the SDs (Figure 7.4b) are the expected 'true values' of the MCS.

The results from this case study demonstrate the HBE approach appropriately accounts for the reactive components of line impedance and load power factor in passively configured feeders. The next investigation tests the validity of the approach to active feeders.

7.2.2 Testing the Validity of the Extension to Active Feeders

In this investigation, the extension of the HBE to DG nodes is evaluated. To do this, the PV-DG connected to buses 2, 5, 7, and 11 are now in service. The comparison of the bus voltages and SDs between the MCS and the HBE depicted in Figure 7.5 demonstrates the validity of the extension of the HBE to active feeders. The trend of voltages, as expected, reflects the impact of voltage-rise as a result of current injections from the DG units.

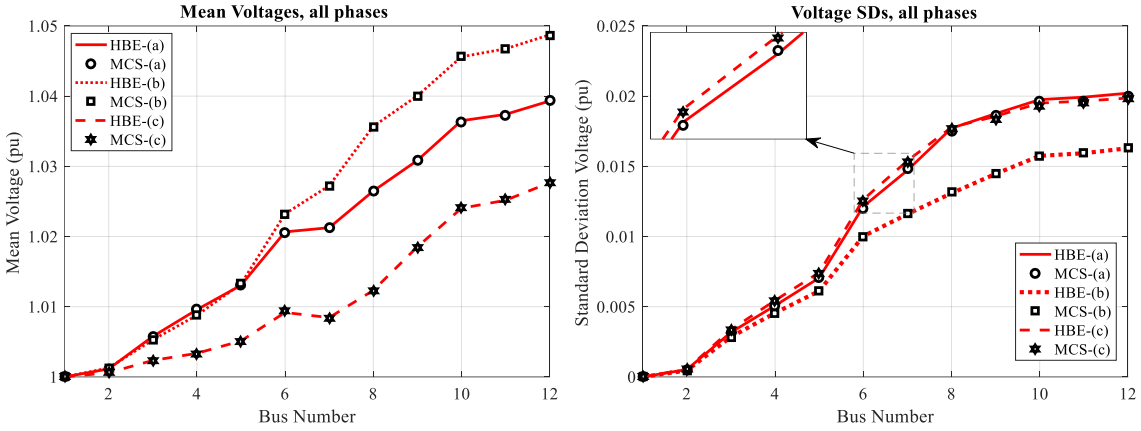

Figure 7.5: Bus Voltage standard deviations, all phases, active 12-bus feeder

On the other hand, the slight difference between the SD plots of Figure 7.3 for the passive case and those in Figure 7.5 for the active is due to the differentiated sensitivity factors (the number of RVs, the magnitude of voltage-change) in the respective load flows. Figure 7.6 shows that the HBE voltage PDFs acceptably fit the MCS data across all phases. Compared to Figure 7.2, the fitting is similar.

The conducted tests and the respective outcomes demonstrate the credibility of the HBE approach in removing the assumption of resistive feeder and unity power factor loads. These tests were conducted using a scaling factor on 1.8 in the normalization of voltages (a base voltage equal to 1.8 times the nominal voltage). The next investigation tests the impact of selected scaling factors on the HBE outcomes.

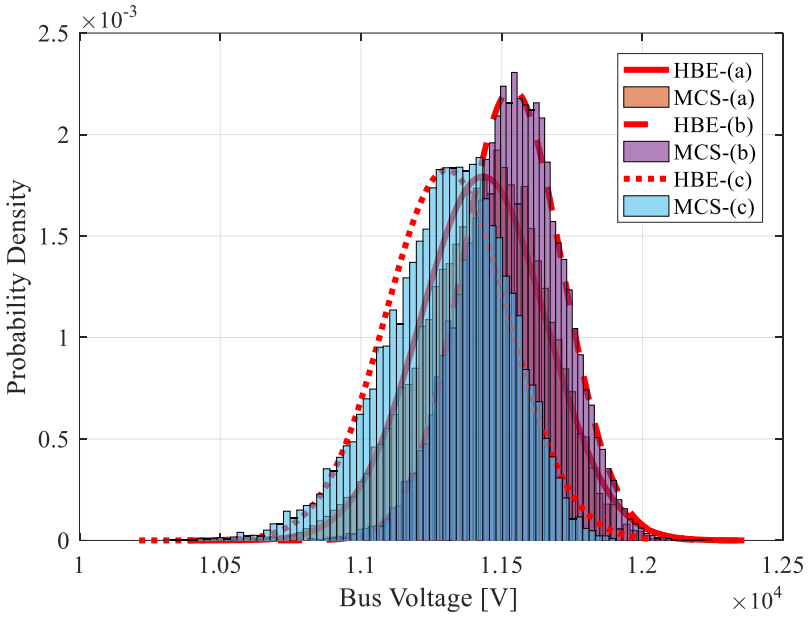

Figure 7.6: MCS data vs. HBE profile: Bus 3 voltages on active 12-bus feeder

7.2.3 Testing the Validity of the Normalization Approach

The reformulation of the HBE relies on a new normalization technique introduced to overcome the limitations of the original approach, which were linked to the key assumptions in the HBA. This case study investigates the impact of a selected scaling factor on the HBE outputs. The test is based on the previously used DG test feeder. The test involves repeated HBE calculations using scaling factors ranging from 1 to 4 in steps of 0.5. Figure 7.7 shows that the HBE transforms with different scaling factors result in similar profiles of mean voltages and standard deviations.

However, this is only valid if the extremum voltages on the tested feeder are a subset of the interval [0, V_B] (where V_B is a scaled nominal voltage by means of the selected scaling factor). Otherwise, the standardized voltages fall outside the permissible intervals of the standard beta PDF.

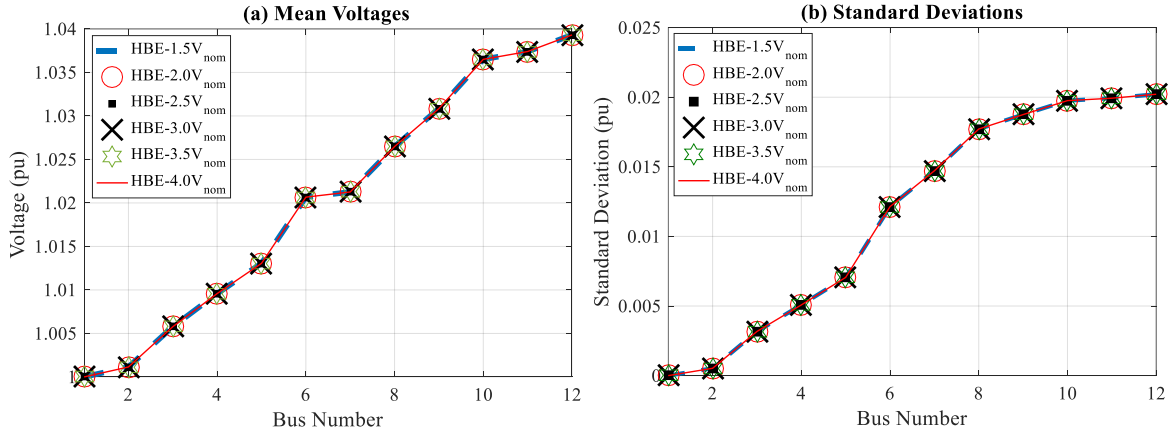

Figure 7.7: HBE outcomes based on scaling factor

7.2.4 Application to Three-wire Systems

The aim of this case study is to assess the performance of the HBE approach applied to 3p-3w systems, using the 12-bus representative feeder used in the previous case study, but with a few modifications:

1. the loads and DGs are now connected in delta configuration
2. the neutral conductor is removed
3. the input currents I_A, I_B and I_C are taken as the models for the load currents I_{AB}, I_{BC} and I_{CA}, respectively.
4. the secondary side of the supplying transformer is configured as a delta source with a sending-end voltage of 33 kV.

Having applied the modifications, the HBE and MCS software based on the 3p-3w load flow equations are applied and outcomes of voltages compared.

Figure 7.8 compares the mean bus voltages from the two PLF approaches. The general trend of voltages is generally similar to that for the four-wire system but with different voltage magnitudes as expected.

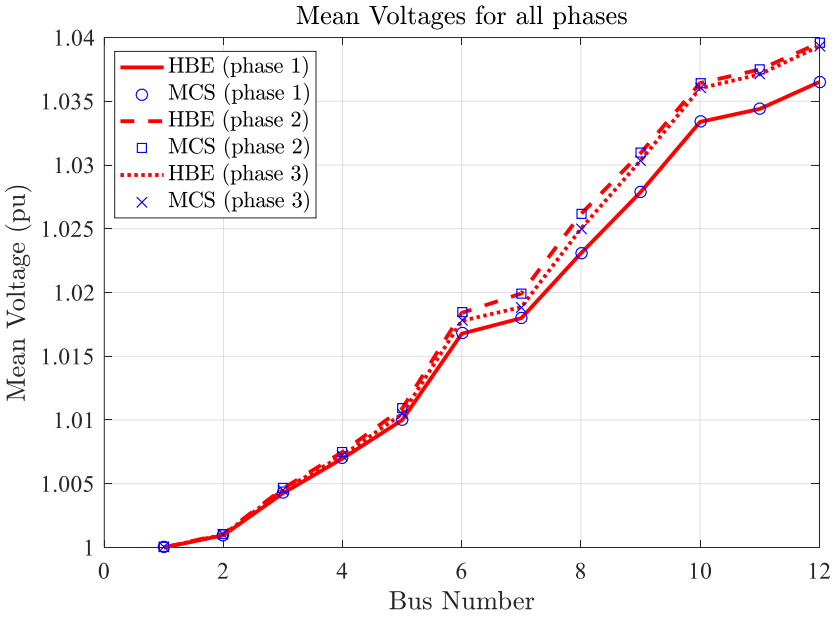

Figure 7.8: Profiles of mean voltages for all phases on the 12-bus feeder, 3p-3w system

The HBE profile of bus voltages corresponds with that of the MCS. Figure 7.9 also shows a close similarity in the profiles of standard deviation. The two outcomes demonstrate the correct implementation of the HBE approach to 3p-3w systems.

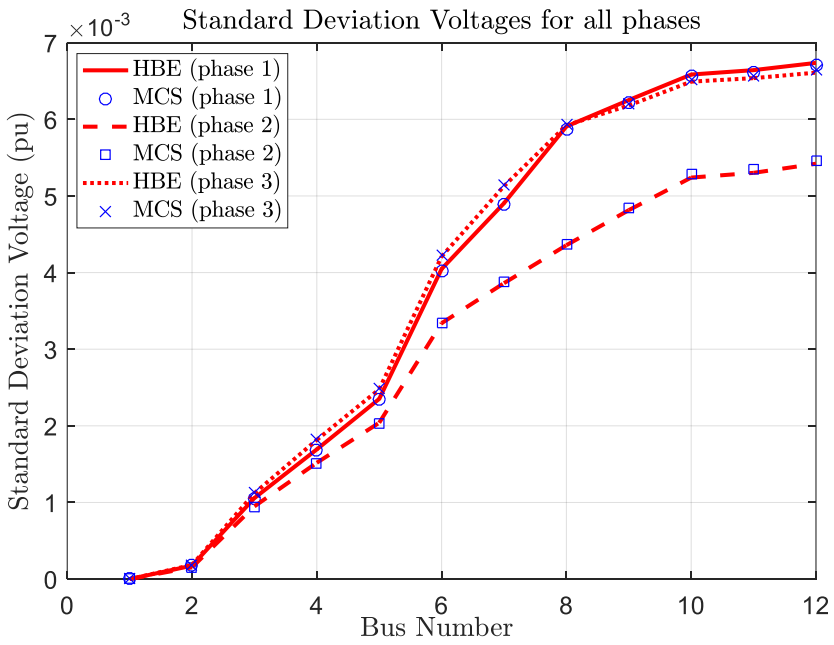

Figure 7.9: Profiles of standard deviation voltages for all phases on the 12-bus feeder, 3p-3w system

Figure 7.10 illustrates the accuracy of the HBE approach in determining the voltage output distributions. As with the PDFs for the four-wire system in Figure 7.6, the plots show similarly good fit.

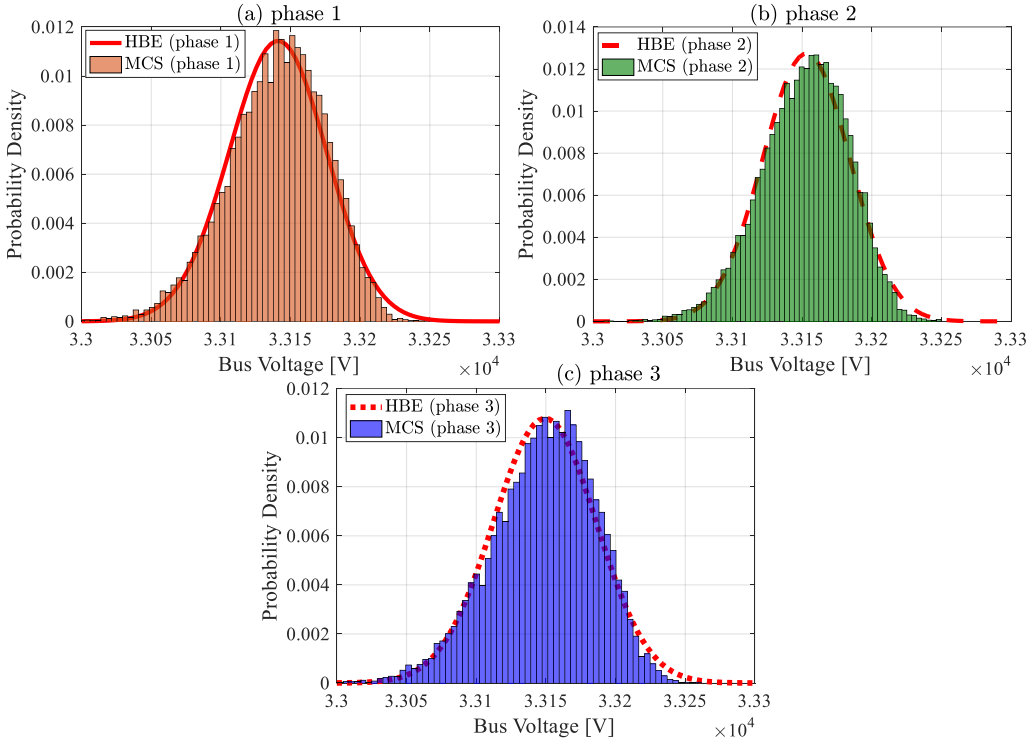

Figure 7.10: MCS data vs. HBE profile: Bus 3 voltages on active 12-bus feeder, 3p-3w

This study has validated the HBE statistical solution to the 3p-3w system of equations.

Based on the formulations of the previous chapter, the equations of voltage-change in the 3p-3w and 3p-4w model differ only in structure; the load being shared between the phase and neutral conductors in the 3p-4w model, and between two phases in the 3p-3w model. Apart from this, the HBE transforms for both systems are based on the same principles. On this premise, the 3p-4w model is used in the remainder of case studies demonstrating the validity of the HBE extensions.

7.3 TESTING THE EXTENSION TO VOLTAGE-DEPENDENT LOAD MODELS

The validity of the reformulations applied in HBE was evaluated by previous case studies. This and subsequent test studies aim to demonstrate the possibility of new applications.

The case study conducted here demonstrates the capacity of the HBE transform to solve the PLF problem with constant-power models for loads and DG. Assessing the effect of applying the constant-power load models in the HBE transform PLF evaluates its extension to simulating feeders with voltage-dependent load and DG models.

The investigation is carried out on an IEEE 33-bus with 3 laterals, operated in radial configuration at a nominal voltage of 12.66 kV. Figure 7.11 shows the one-line diagram of the feeder.

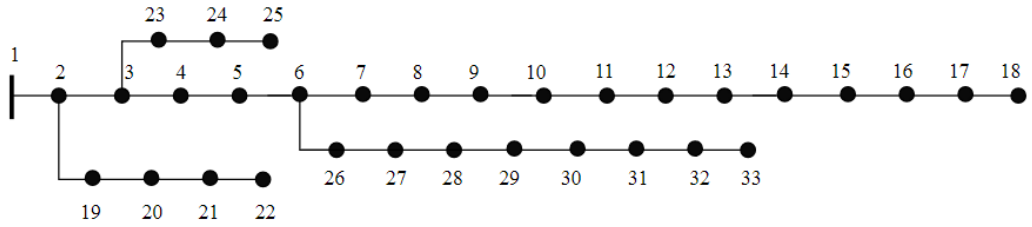

Figure 7.11: IEEE 33-bus test network

The feeder has a total of 32 constant-power loads at power factors ranging between 0.3 and 0.99 and is made up of line sections with X/R ratios between 0.3 and 3.4. The full properties of this feeder derived from published parameters [246] are provided in Appendix B2.

The test model was published with deterministic load data, taken as the mean values for a PLF, and to which a coefficient of variation (σ/μ) of 10% is assigned to model variability. Two scenarios are investigated.

7.3.1 Passive Network

The PLF is conducted on a passive feeder (without DG). Under the conditions of symmetrically distributed loads, with the same measure of variance, the published deterministic results can be compared with the means of the PLF results.

Initially, the constant-power loads are assumed to operate at nominal voltage from which equivalent current models are derived. The HBE and MCS transforms based on the current models, denoted by "I-

model", are then applied to the test system. As a result, the MCS (I-model) validates the outcomes from the HBE (I-model). After the initial execution with powers at nominal voltage, the HBE transform is applied repeatedly, after adjusting the assigned currents at each node according to the previous outcomes of voltage, effectively adopting constant-PQ load models. The updated nodal voltages are compared with the DLF solution to the feeder [247].

This comparison is possible because the mean values from a PLF can be matched with DLF results if the ADMD used in the DLF corresponds to the means of the input PDFs. In such a case, the overall trend of voltages or line currents from the PLF investigation, including the deductions thereof, will be similar to those possible using a DLF. The difference, however, is brought about by the application of risk to the spectrum of outputs, and the impact it makes on the magnitude of the output variables used for planning.

In view of this, Figure 7.5 compares the deterministic results based on constant-PQ models with those from the HBE (I-model) and the HBE with a single iteration (HBE-repeated). The similar outcomes between the MCS (I-model) and the HBE (I-model) validate the HBE (I-model). On the other hand, the HBE (I-model) and the DLF (PQ-model) approach result in differentiated results. As expected, the difference in the nodal voltages increases with voltage-drop.

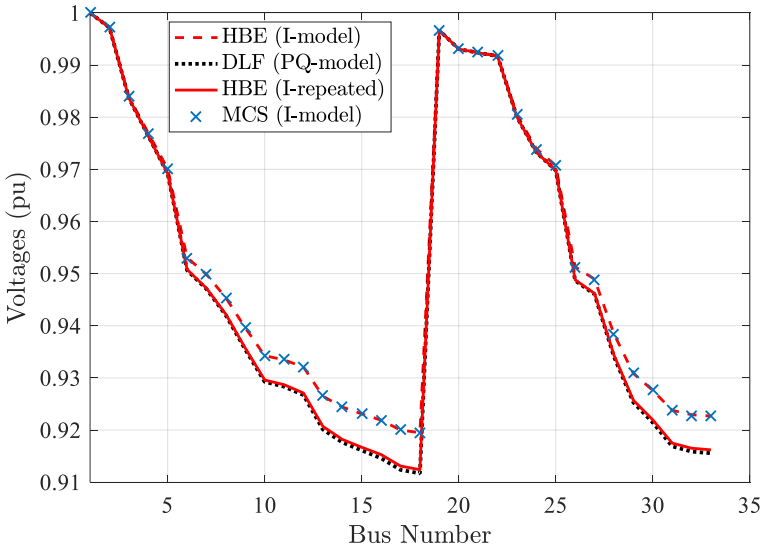

Figure 7.12: Performance of the HBE with voltage-dependent load models: mean voltages.

However, a single repetition of the HBE calculation with updated nodal voltages significantly reduces differences, as can be seen from the comparison between the DLF (PQ-model) and HBE (I-repeated) profiles. This shows that the HBE transform can accommodate constant-power models at very low computational cost. To explore its performance further, 20 iterations were conducted, and the characteristic of convergence studied. Figure 7.13 depicts the convergence of the HBE voltage at bus 11.

The study demonstrates that convergence is reached only with a few iterations; from 6 iterations and higher, the relative error between the HBE (I-repeated) and DLF (PQ model) is constant at 0.003%. However, just the single iteration reduces the difference from 0.556% in the HBE (I-model) to only about 0.046%.

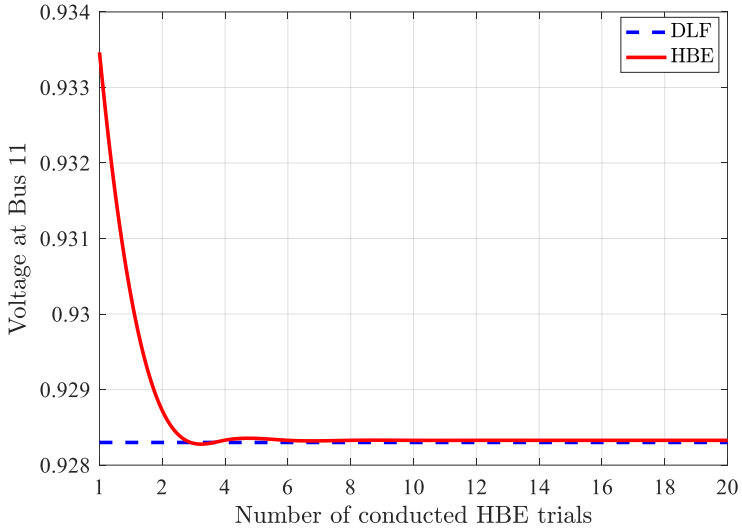

Figure 7.13: Convergence of repeated HBE calculation to PQ-model outcomes: Bus 11

This result lies within a tolerance of 0.005. It is worth mentioning that the convergence, hence the number of trials sufficient to estimate the fully converged result, is a function of the anticipated voltage deviations on a tested feeder.

Nonetheless, considering the lowest voltage on the tested feeder is 0.9131 while the normal operating conditions of MV feeders is regulated at $\pm 5\%$ of the nominal voltage [94], a single iteration is most likely sufficient on most feeders operating close to the regulated conditions.

7.3.2 Active Network

To evaluate the extension of the HBE to simulate a feeder with voltage-dependent DG models, the test system in the previous case is modified to include PV systems connected to buses 14, 18, 28 and 33, all in balanced configuration. The statistical properties of the PV-DGs are the same as those in Section 7.22.

The allocation of the generators, and the corresponding capacity and power factors are given in Table 7.3.

Table 7.3: DG properties on the modified IEEE 33-bus feeder

BUS ID.	DG PARAMETERS				BUS ID.	DG PARAMETERS			
	α	β	Power (kVA)	pf		α	β	Power (kVA)	pf
14	12.62	2.21	500	0.900	28	12.62	2.21	1000	0.91
18	12.62	2.21	750	0.950	33	12.62	2.21	750	0.94

As implemented in the previous case study, the result from the MCS is used to validate the HBE current model (HBE I-model) while the second calculation (HBE I-repeated) is compared with a converged result (HBE-PQ model) with numerous iterations.

Figure 7.14 shows the impact of the solar farms on the profile of bus voltages. A close look at the voltage plots in the insert shows the trace of the HBE (I-model) is consistent with that for the MCS (I-model), while the repeated HBE matches the profile of the converged HBE (I-converged) which simulates the PQ-model.

The reduced differences between the constant current and constant-power models observed by comparing the plots in Figure 7.12 and Figure 7.14 are because of smaller voltage deviations in the active case; the passive case has voltages between 0.91 and 1 pu, while the active case has voltages between 0.98 and 1.025 pu. This reinforces that for feeders with low voltage deviations from the nominal values, as in the case of lightly loaded and some active feeders, the impacts of assuming a current model where the actual loads behave with voltage-dependency, are minimal.

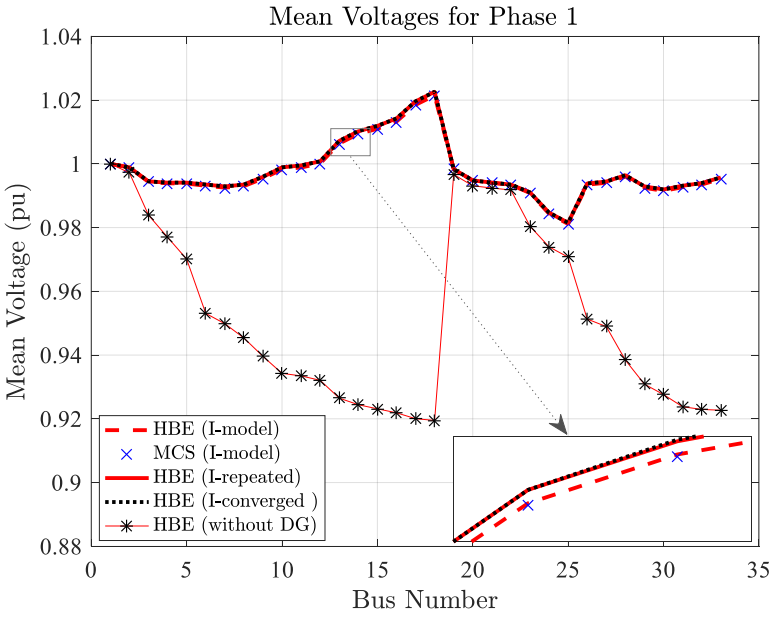

Figure 7.14: Comparison of HBE I-model and PQ-model deterministic values: Active feeder

The close similarity between the HBE (I-repeated) and the converged result (which approximates the deterministic PQ-model as demonstrated in the previous case study) shows the extension of the transform to constant-power DG models is valid. The results that follow discuss the impacts on dispersion.

First, the accuracy of the HBE (I-model) from which the HBE (I-repeated) and the converged (PQ-model) are based is discussed in more detail.

Table 7.4 presents selected results for voltage means and SDs, obtained by the tested models. The close similarity in the results between the MCS and the HBE (I-model) validates the application of the algorithm on active feeders with laterals.

Table 7.4: Sample voltage outcomes for the modified IEEE 33-bus with DG

BUS (PHASE)	**MCS (I-model)**			**HBE (I-model)**		
	μ_V (pu)	σ_V (pu)	$p_{r\%}$ (pu)	μ_V (pu)	σ_V (pu)	$p_{r\%}$ (pu)
5	0.9938	3.151E-03	0.9996	0.9938	3.158E-03	1.0000
12	1.0000	8.890E-03	1.0160	1.0001	8.853E-03	1.0174
18	1.0213	1.469E-02	1.0468	1.0215	1.456E-02	1.0500
30	0.9914	9.584E-03	1.0091	0.9915	9.620E-03	1.0103

Figure 7.15 shows a sample of bus voltage PDFs at bus 7 and bus 25. The plot generally demonstrates a good fit of the HBE PDF to the MCS data.

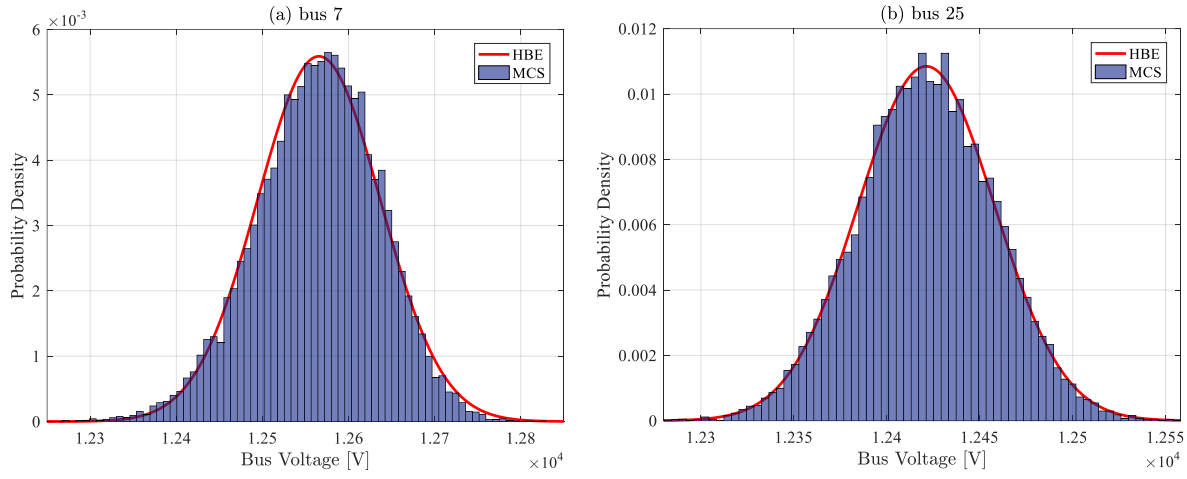

Figure 7.15: MCS data vs. HBE voltage profiles on the modified IEEE 33-bus

Figure 7.16 compares the voltage SDs between the MCS (I-model), HBE (I-model), and the iterated HBE methods. From the insert, it can be observed the constant-current models for the MCS and the HBE have similar profiles of voltage SDs. Similarly, the single iteration of the HBE transform obtains nearly the same profile of SDs as the converged model. The MCS validates the HBE I-model from which the power model is derived, while the comparison between HBE-I repeated and the converged HBE validates the HBE's approximation using a single iteration.

The conducted tests on the IEEE 33-bus network under passive and active configurations has shown that a single repeated calculation of the HBE with updated nodal voltages can offer an acceptable approximation to the converged PLF representing a constant-PQ-model based calculation. The merits of the approach in this regard validates the HBE for simulating feeder with voltage-dependent load and DG models.

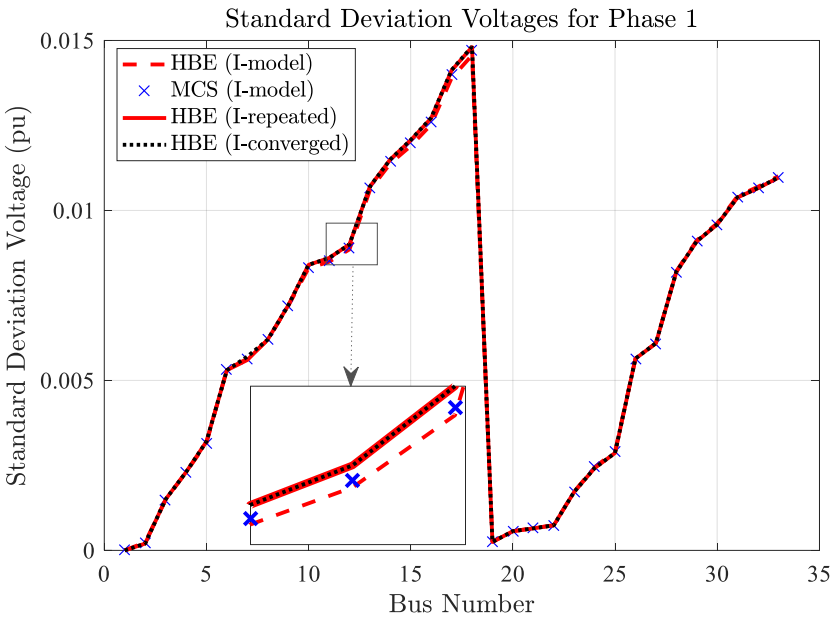

Figure 7.16: Standard deviations of the voltages for different load models: Active feeder

7.3.3 Validating the Extension to Compensated Feeders and Long Transmission Lines

This test case is aimed at evaluating the performance of the HBE in applications involving shunt capacitance, either from line charging (in medium-to-long distribution feeders) or from compensation elements. The extension to such applications has been afforded through the capacity to represent reactive components, and the extension to voltage-dependent load models tested in the preceding section. To test the validity of the HBE approach to compensated feeders, the IEEE 34-bus feeder is used along with the modifications applied in [248] to focus the testing to the effects of capacitance;

1. The in-line transformer is removed and replaced by a conductor section such that the whole network has a single reference voltage
2. The voltage regulator is omitted
3. The spot loads and distributed loads are replaced by balanced loads.

The parameters of the modified feeder, which include shunt admittances, are adopted from [248] and are also provided in Appendix B3. The investigation is conducted in three parts; the initial analysis neglects shunt admittances, which are included in the second analysis, and then adding shunt capacitors of capacity 100 kVAr (C1) and 150 kVAr (C2) to buses 30 and 32, respectively, in the final simulation. The shunt capacitors are balanced single-phase elements connected in wye configuration and modelled as constant-impedance. The one-line diagram of Figure 7.17 shows the distribution network and the placement of the compensation elements.

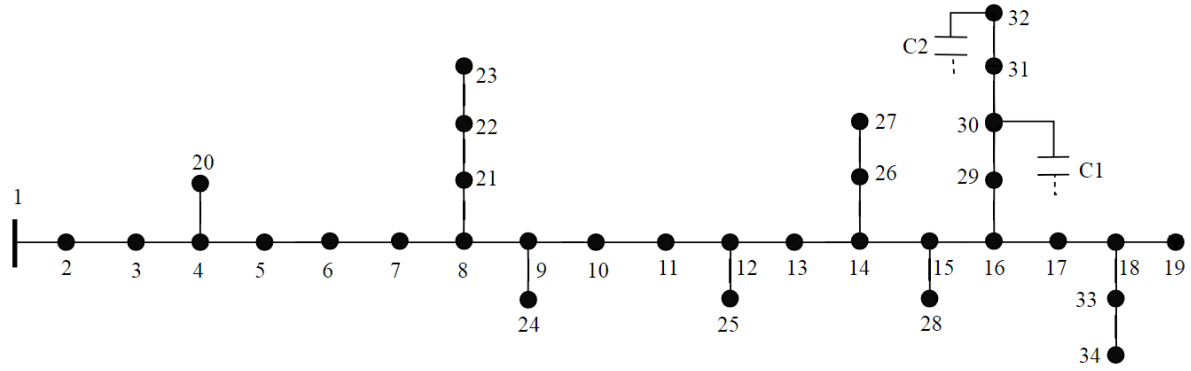

Figure 7.17: The modified IEEE 34-bus test network

Figure 7.18 shows HBE plots of mean voltages for the three tested conditions. The insert demonstrates the voltage improvement effect of line shunt capacitors. However, the difference in voltages is not significant due to the relatively short length of the feeder (about 84 km). It is anticipated that for longer feeders (> 200 km) the charging effects of the line shunt capacitance become significant. To show the likely impacts of admittance on very long feeders, the shunt admittances were scaled by 200. The plots with scaled admittance compared to the passive feeder demonstrate significant voltage improvement.

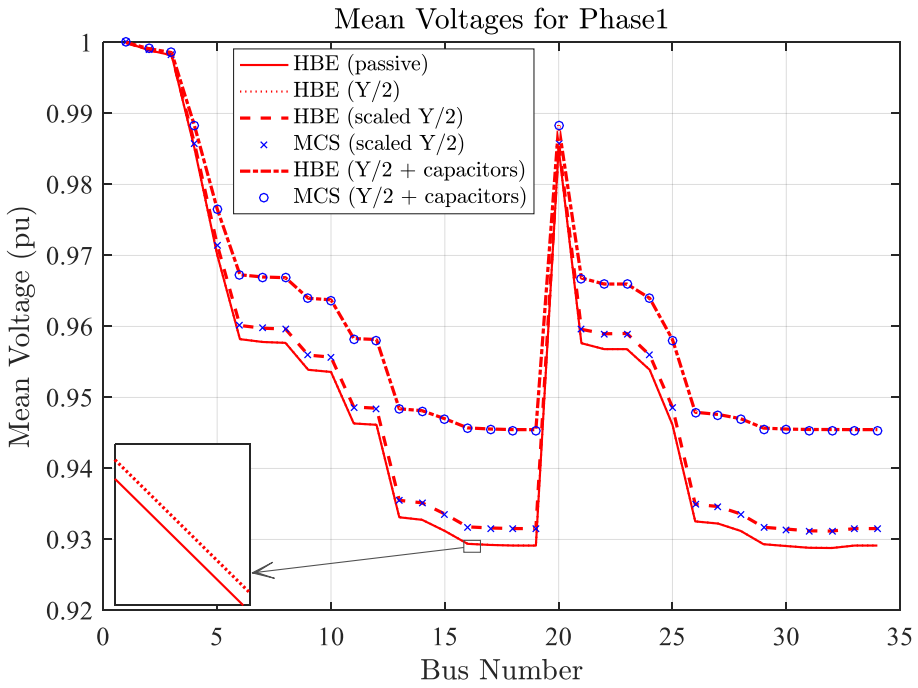

Figure 7.18: Effects of shunt capacitor elements on mean bus voltages

Likewise, compensation using capacitor banks offers significant voltage regulation. Based on the MCS results, the HBE correctly accounts for the effects of shunt capacitance in the calculation of the mean voltages.

The corresponding plots of standard deviations in Figure 7.19 show that the inclusion of line admittance and compensation devices has limited impact on dispersion.

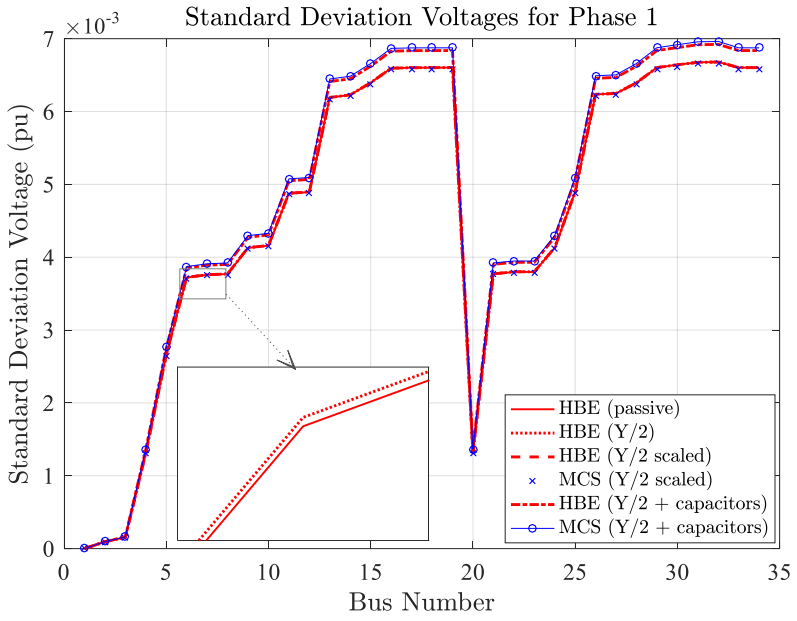

Figure 7.19: Effects of shunt capacitor elements on standard deviation of bus voltages

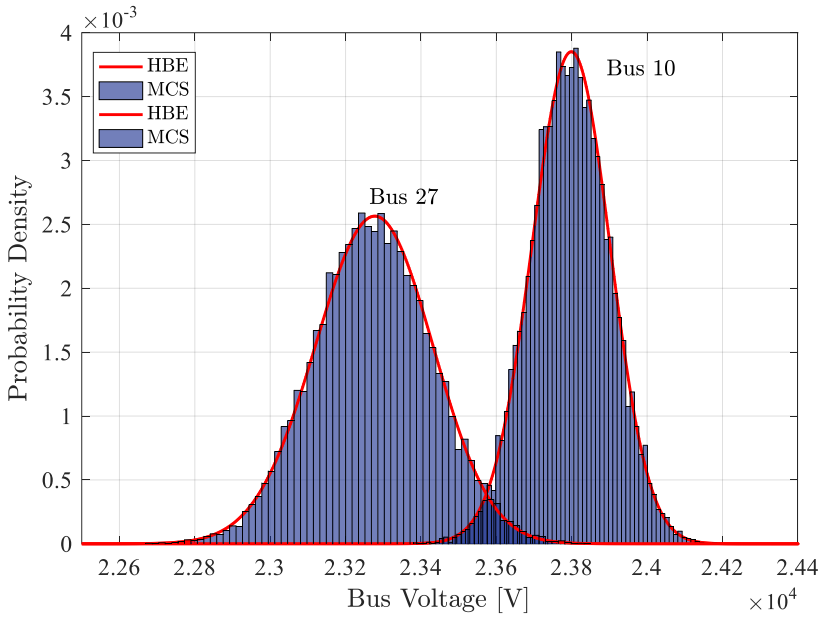

Figure 7.20: Sample bus voltage PDFs on IEEE 34-bus with Y/2 and shunt capacitors

This result is because shunt capacitor currents are modelled deterministically (mean with no variance) in the MCS and approximated by a tall spike-shaped distribution with limited variability in the HBE. However, despite the insignificant changes to the overall uncertainty in the system, the dispersion of the voltage at each node is a function of the expected voltage-drops (as depicted by the equations for the second moments). This is reason for the increased SD in the case involving compensation capacitors.

Figure 7.20 compares the distribution of voltages at buses 10 and 27 on the feeder with both line shunts and compensation capacitors. The voltage PDFs together with feeder profiles for mean and standard deviation voltages validate the HBE approach for compensated feeders and long feeders (typically HV networks).

7.4 TESTING THE APPROACH FOR SYSTEMS WITH DEPENDENCE

In this section the performance of the HBE approach with correlated random variables is evaluated using four case studies carried out on the modified IEEE 33-bus feeder as used in Section 7.3.2. To facilitate the assessment of the impacts of various forms of input dependence, the feeder is partitioned into four service areas (Areas I-V as indicated in Table 7.5); two of which are residential, one industrial and one commercial. The feeder zones according to buses are identified in Table 7.5. According to the allocation of DG in Table 7.3, the DG plants are in areas I and IV.

The correlation factors between input variables are assigned based on practical expectations. Grouped residential loads are anticipated to demonstrate low dependency as customer load behaviors are highly differentiated. On the other hand, grouped industrial loads tend to rise and fall together especially if classification is based on industrial activity. The same is expected for commercial loads which are usually strongly influenced by lighting, heating and air-conditioning over common working hours. In areas with DG, correlation between loads and DGs is considered to model the likely dependency between air-conditioning use and PV production with heat conditions. This is likely to be higher in the commercial area than the residential. The dependence between DGs in each area simulates the similarities in PV-DG productions based on installation characteristics such as tilt angle, and orientation. Considering these characteristics, the correlations in Table 7.5 are applied.

Table 7.5: Correlation conditions on the modified IEEE 33-bus

AREA	FEEDER ZONING		DEPENDENCY		
	Load Class	Buses	Load-Load	DG-DG	Load-DG
I	Residential I	2-18	0.25	0.80	0.20
II	Residential II	19-22	0.35	No DG	--
III	Industrial	23-25	0.95	No DG	--
IV	Commercial	26-33	0.85	0.90	0.50

From the indices in Table 7.5, correlation matrices are developed and entered as inputs to the LF approaches. The correlation matrix for the loads in Area II is provided in (7.1) as an example.

$$\rho_{A-II} = \begin{bmatrix} 1 & 0.35 & 0.35 & 0.35 \\ 0.35 & 1 & 0.35 & 0.35 \\ 0.35 & 0.35 & 1 & 0.35 \\ 0.35 & 0.35 & 0.35 & 1 \end{bmatrix} \quad (7.1)$$

From the various load and DG dependency conditions, three test scenarios are developed; Case 1 considers load-load dependence while Case 2 extends the dependence to DG-DG interactions, and Case 3 introduces correlation between loads and DG. The final case study (Case 4) demonstrates the application of the HBE approach in cases involving balanced three-phase loads with correlation between phase currents.

Across all cases, the Monte-Carlo simulation based on copulas is used as a validation tool.

7.4.1 Case 1: Dependence Between Loads: Gaussian (symmetrical) Distributed Inputs

In this case study, the test feeder is operated in passive configuration and load-load correlations are considered according to the properties in Table 7.5. Since the statistical load models for the loads are symmetrical distributions around the deterministic means, the test is limited to dependency between symmetrical, Gaussian-type distributions.

The comparison of the profiles of mean voltages shown in Figure 7.21 demonstrates that correlation between loads does not impact mean voltages. This meets the theoretical expectation that the calculation of mean voltages in a system with interdependent RVs does not depend on correlation factors.

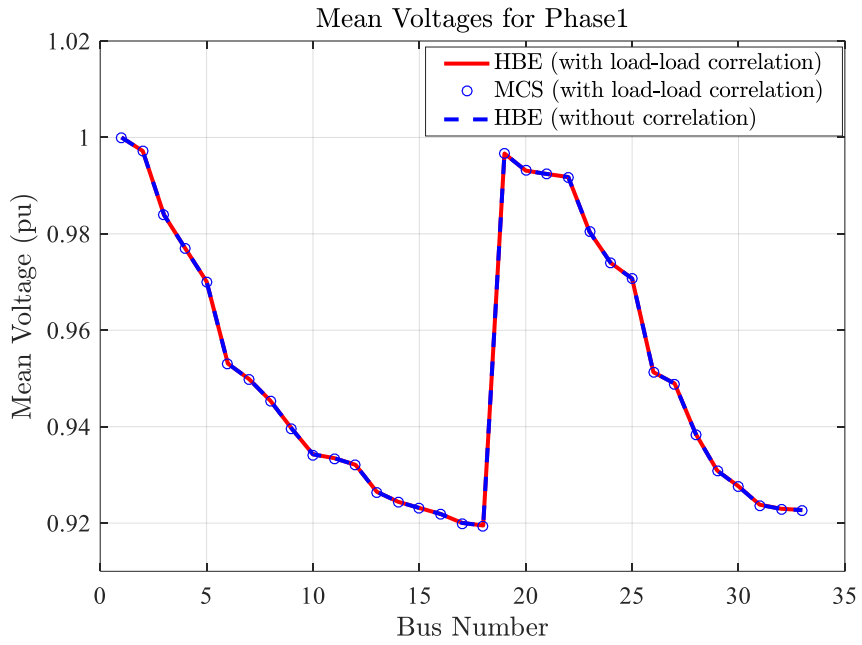

Figure 7.21: Impact of load-load correlation on mean voltages

However, the second moment and variance are affected as they depend on the measure of covariance between variables. This is confirmed in plot of Figure 7.22 which shows a significant increase in the variance when positive correlation between the loads is considered.

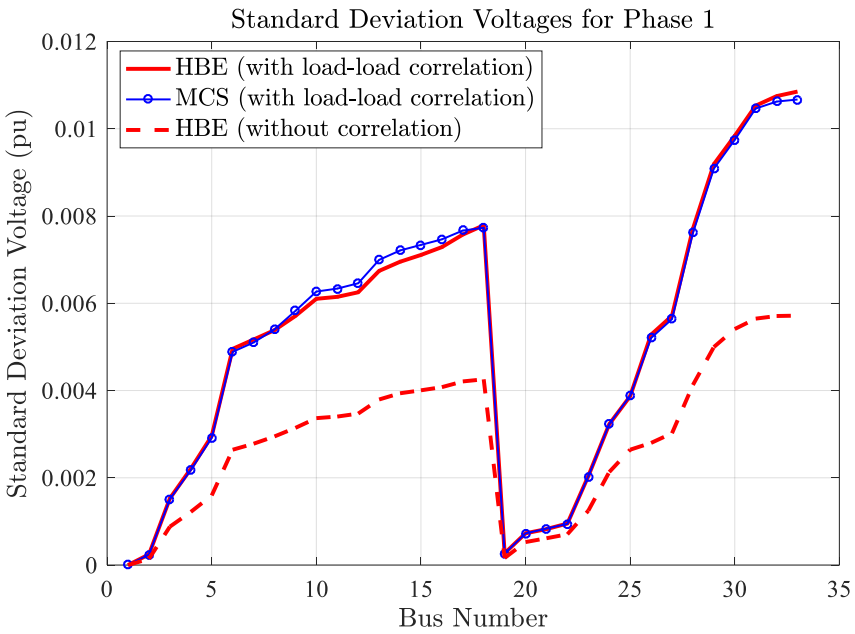

Figure 7.22: Impact of load-load correlation on standard deviations

The HBE trace closely follows the MCS profile but with some noticeable differences in some sections. The differences are related to the variance in the sampling of correlated RVs in the MCS. Figure 7.18 shows a scatter of correlated samples between two Gaussian RVs produced by the MCS sampling.

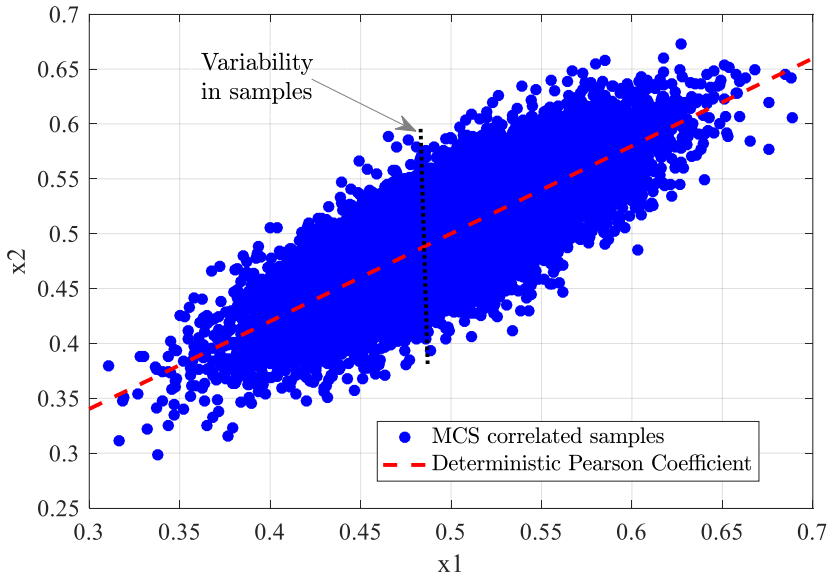

Figure 7.23: MCS correlated load samples against the assumed Pearson coefficient

Based on the sampling approach, which is random and with limited scenarios, errors in modelling the covariance between two RVs is likely. Nevertheless, the scale of differences in standard deviation does not significantly affect the output distributions. Figure 7.24 compares the voltage PDFs for buses 14 and 25 between the MCS and two HBE approaches; one with correlation and the other without. The HBE based on independent inputs while correctly tracking the mean of the distribution leads to

differently distributed outputs due to variance errors (shown in Figure 7.22). The comparison demonstrates that disregarding load-load dependence where it is significant potentially misguides a network planner's decisions; in this case underestimating the variability of the voltage conditions.

On the other hand, the HBE based on dependent inputs provides a good fit to the histogram of MCS voltages in both buses. The fit achieved on bus 14 which appears to have significant standard deviation error in Figure 7.22 confirms that the errors are not significant. Accordingly, the MCS validates the application of the HBE approach in cases involving load interdependences.

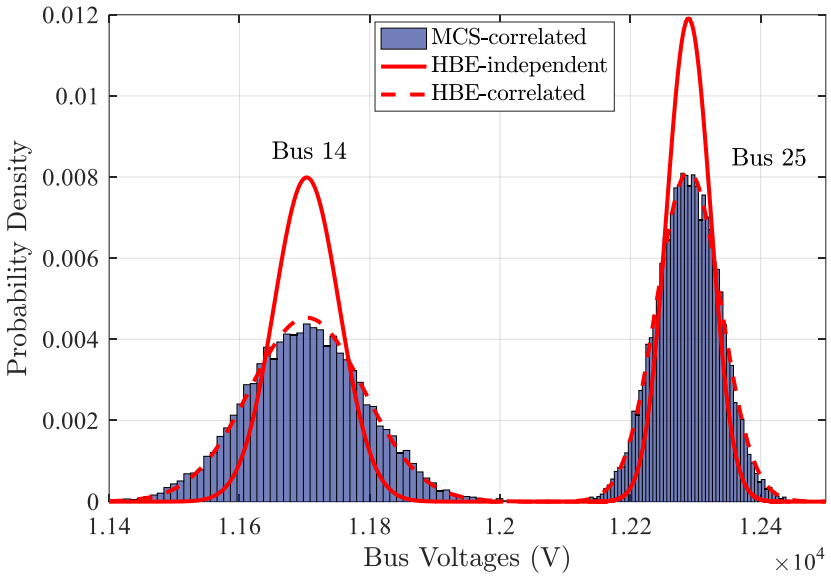

Figure 7.24: Impact of load-load correlation on bus voltage distributions

7.4.2 Case 2: Accounting for Correlation between DGs: Skewed Distributed Inputs

To test the performance of the HBE with skewed distributions, the test feeder is operated in active configuration. The correlations in Table 7.5 are applied and simulations conducted using both PLF approaches, HBE and MCS.

The plot of Figure 7.25 demonstrates the combined effects of correlated loads and DGs on standard deviation. The HBE approach based on independent inputs results in significantly lower SDs arising from the neglect of covariances between random variables. On the other hand, the outcomes of the HBE with correlated RVs matches those from the MCS with only slight differences.

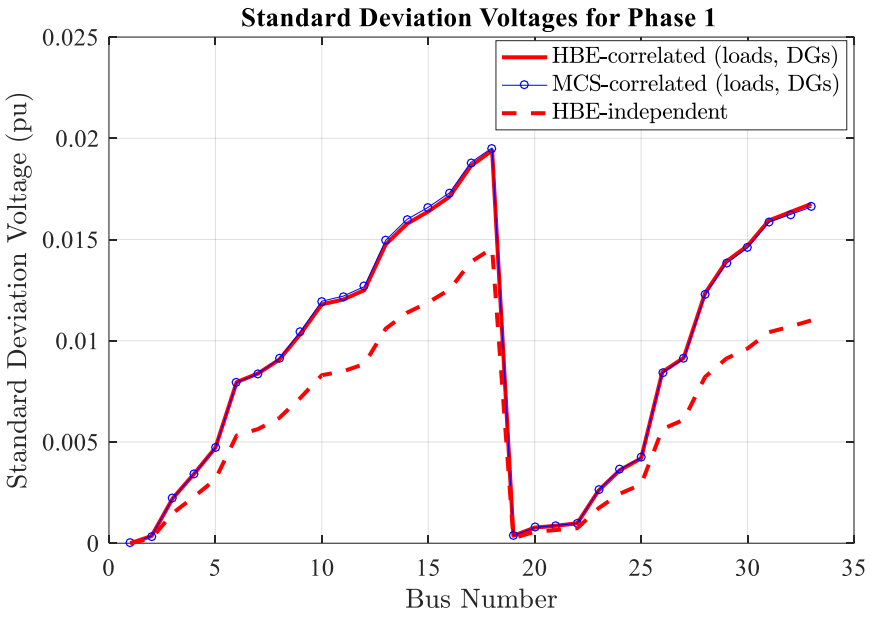

Figure 7.25: The combined effects of load-load and DG-DG correlations on standard deviations

In this case, the reduction in error compared to Case 1 is because sampling errors in the MCS decrease with increasing correlation factors. In addition, due to the higher sensitivity factors of the load flow to DG (based on its capacity), the errors initially observed in the sampling of correlated loads are offset. Figure 7.26 compares the voltage PDFs between the applied PLF approaches.

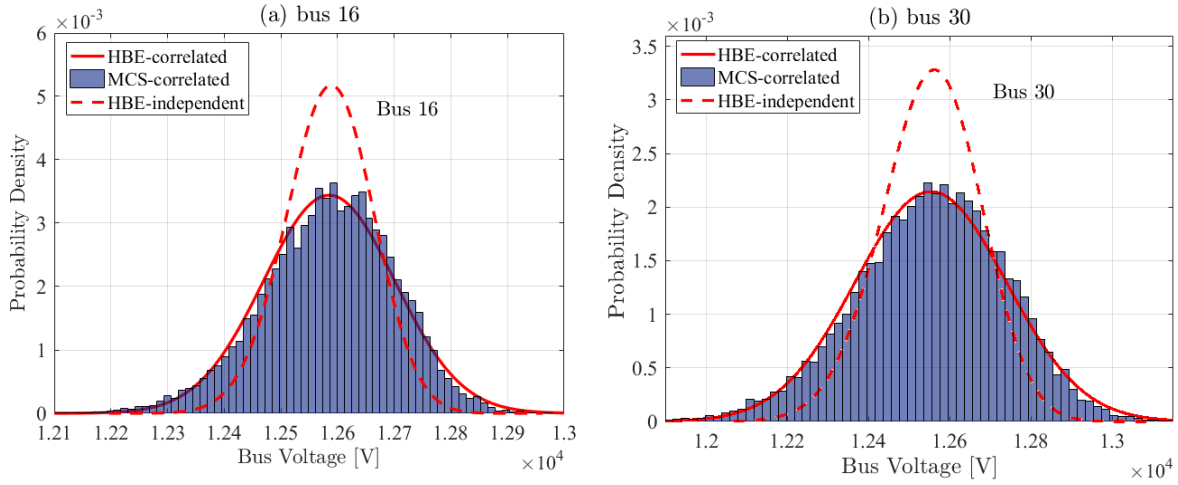

Figure 7.26: The effects of DG-DG correlations on bus voltage distributions

The plot demonstrates the demonstrates that the effects of the combined effects of interdependencies are correctly applied in the HBE.

7.4.3 Case 3: Considering Interdependence between Loads and DGs

The previous two case studies dealt with correlation between inputs of the same type, for loads and DGs. This case study investigates the performance of the HBE approach with load-DG

interdependencies. Correlation coefficients indicated in Table 7.5 are applied to model the likely dependence between the load demand and PV-DG production in areas I and IV.

The plots in Figure 7.27 compares the profiles for standard deviation on the test feeder. The comparison of the SD plots between the HBE account for correlation and the one without demonstrates that the effects of DG-load dependence can be significant

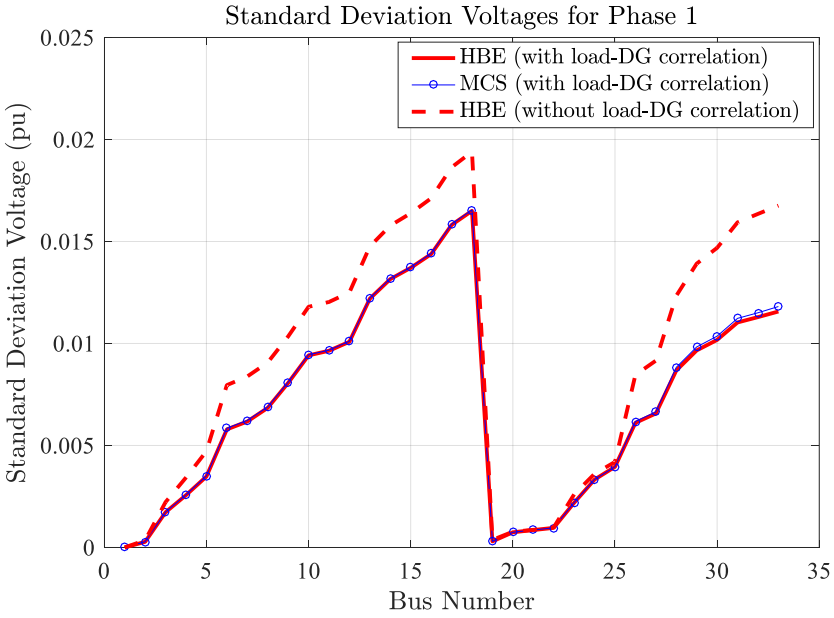

Figure 7.27: Impact of load-DG correlation on the standard deviation of bus voltages

The MCS profile of SDs validate the HBE with DG-load correlation. The PDF plots in Figure 7.28 show that the PDFs of the HBE with correlation are close estimates to the MCS histogram of voltages while the HBE without correlation results in significant variance errors.

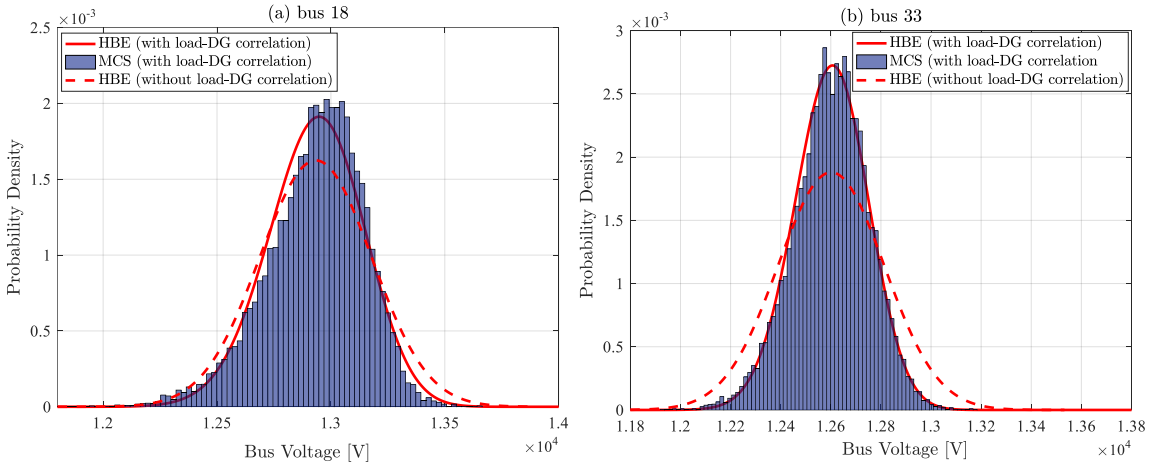

Figure 7.28: The effects of load-DG correlation on bus voltage distributions

The slight mismatch observed in the fit for bus 18 compared to that for bus 33 is as a result of the significant skew in the output distribution. However, the approximation to the data is still acceptable.

7.4.4 Case 4: Correlation between Phases: Balanced Three-phase Loads and DGs

Case studies 1, 2, and 3 have dealt with dependencies between nodal inputs. Building on the test conditions in Case 3 to extend the test to phase-phase correlation, it is assumed all loads and DGs are three-phase systems with fully correlated phase currents. The difference this makes is that the phase currents rise and fall together as opposed to the statistical representation of balanced single-phase connections in which the phase currents may be independent. The correlation matrix in (7.2) is used at each node.

$$\rho_{a,b,c} = \begin{bmatrix} 1 & \rho_{a,b} & \rho_{b,c} \\ \rho_{a,b} & 1 & \rho_{a,c} \\ \rho_{a,c} & \rho_{b,c} & 1 \end{bmatrix}; \quad \rho_{a,b} = \rho_{b,c} = \rho_{c,a} = 1 \qquad (7.2)$$

Figure 7.29 shows the outcomes for standard deviation on the test feeder. The plot demonstrates a dramatic impact of inter-phase correlation on standard deviation. This is expected as the correlation conditions lower the variability in the summated voltage-drops locally at the nodal level.

The close similarity in the SD profiles between the MCS and HBE with correlation demonstrates the validity of the HBE approach to intra-phase dependencies.

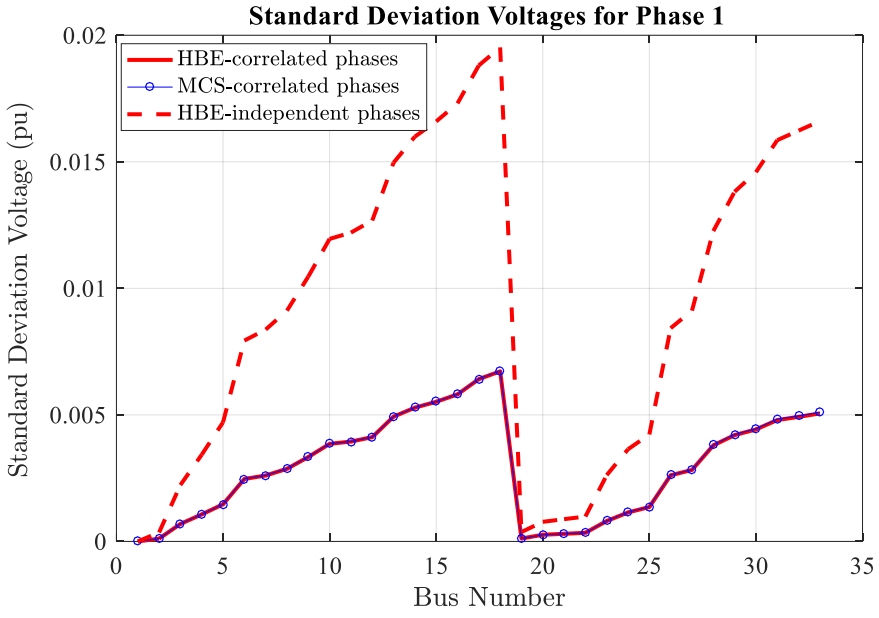

Figure 7.29: Effects of inter-phase correlation on bus standard deviations

The PDFs in Figure 7.30 demonstrate the accuracy of the HBE approach in depicting the mean and variance of the output distributions. The PDFs demonstrate that the effect of correlated phases is accurately applied in the HBE. The plots also demonstrate the impact of perfectly balanced three-phase systems on the voltage conditions of a feeder.

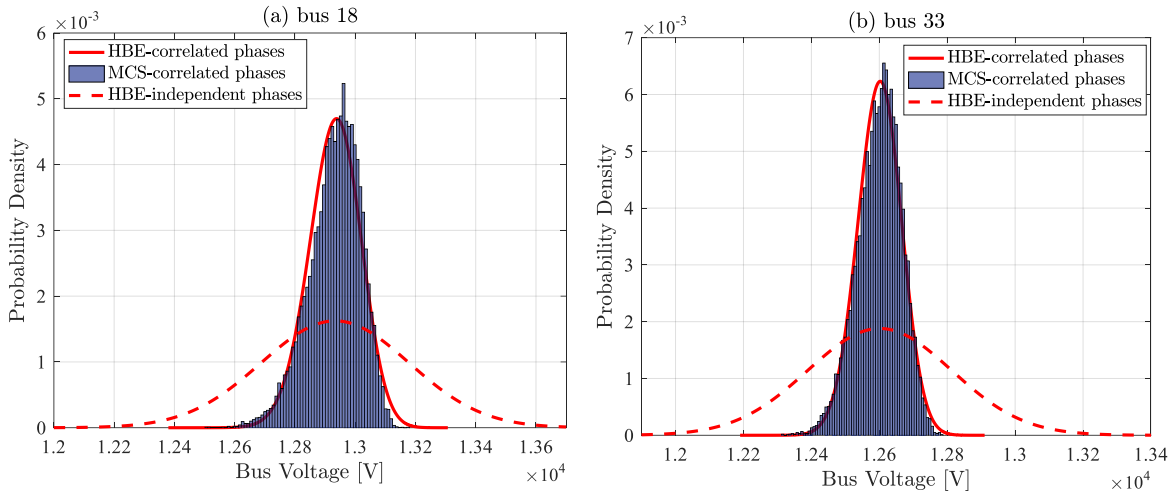

Figure 7.30: Effects of phase-phase correlation on voltage distributions

This finding is consistent with findings from earlier work [132] showing that balanced three-phase PV-DG systems allow for higher penetrations than single-phase systems.

7.5 TESTING THE APPROACH TO BRANCH CURRENTS

The purpose of this case study is to demonstrate the HBE's performance in computing probabilistic branch currents. The IEEE 33-bus in active configuration, as used in Section 7.3.2, is adopted. Due to the unavailability of cable rating data, all cable currents are expressed on a base of a 300 A. The HBE and MCS approaches for line currents are applied to the test network and the outcomes compared.

Figure 7.31 compares the branch currents between the HBE and the MCS.

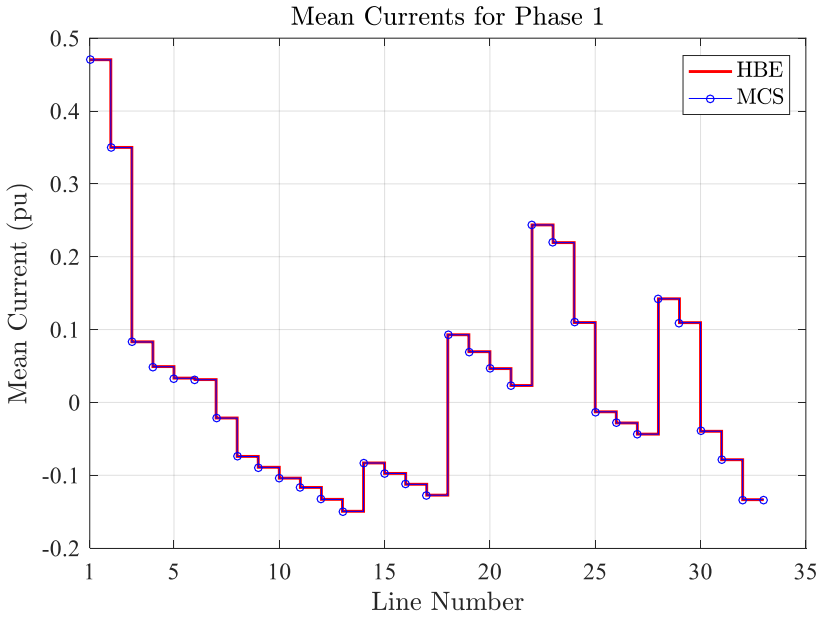

Figure 7.31: Mean branch currents on IEEE 33-bus in active configuration

A staircase plot is used in this case to depict the constant current flow between two successive nodes. Appendix B2 contains the naming of the line numbers as part of the feeder properties.

The plot shows that the mean currents from the HBE accurately match those from the MCS. The current trend also reflects the voltage trend as achieved in Section 7.3.2 and as shown in Figure 7.14, in terms of voltage increase and voltage-drops. Negative mean currents representing reverse current flow are observed where voltage-rise occurs. Similarly, where the voltage falls, the line currents are forward-biased and positive.

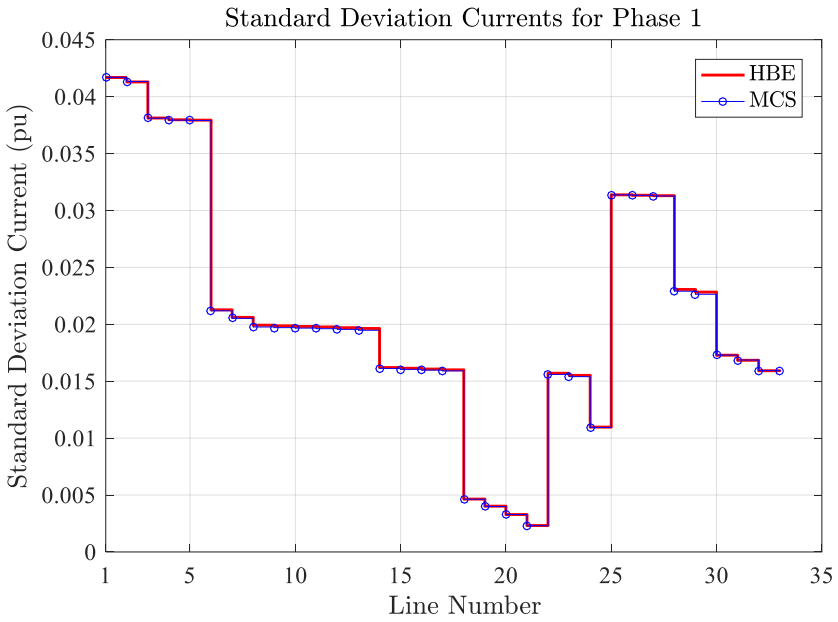

Figure 7.32: Standard deviations of branch currents on IEEE 33-bus in active configuration

Figure 7.33 (a) shows the distribution of currents in a branch with reverse power flow, while Figure 7.33 (b) shows the distribution of currents in a branch with forward flow.

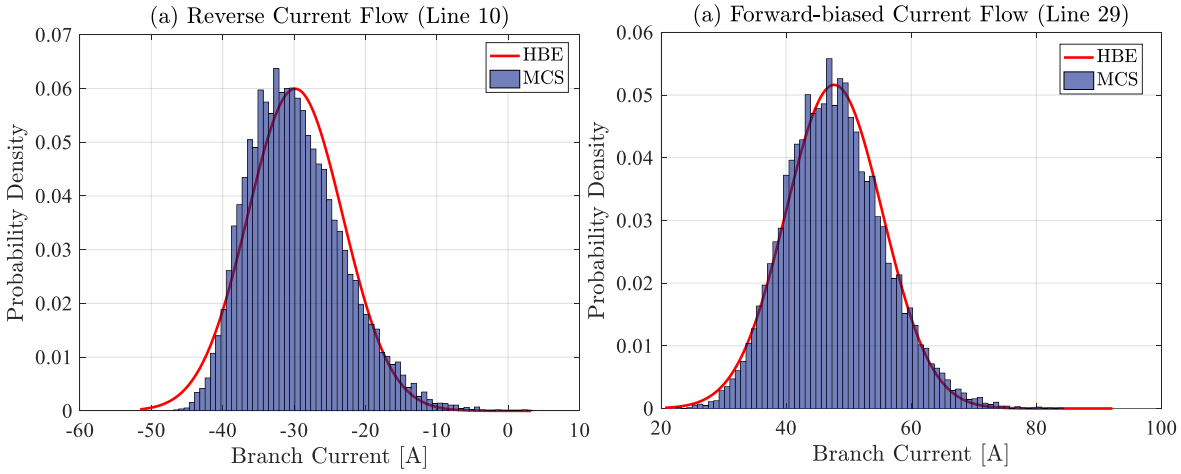

Figure 7.33: Distribution of branch currents on the IEEE 33-bus in active configuration

The fit of the beta PDFs on the MCS histogram of currents demonstrates the validity of the HBE approach for line flows.

7.6 APPLICATION OF THE HBE TO LARGE NETWORKS

This section evaluates the application of HBE to large networks. While doing so, the performance of the developed tool is also compared with the HBA approach for MV feeders (HBA - MV) based on an approximate formulation.

Simulations are conducted on the IEEE 69-bus test systems with published parameters [249]. The feeder is operated at 12.66 kV and consists of 68 deterministic loads characterised as constant-power. The load properties are modified by assigning a standard deviation of 10% to the deterministic load capacities while the feeder is modified by adding DG connections to buses 22, 24, 26, 60, 62, and 64, and shunt capacitors to buses 57, 58, and 61. Figure 7.34 shows the modified networks and the clustering of segmentations of buses into four areas for the analysis of correlations. Buses not included in the four clusters (Area I – IV) are considered independent.

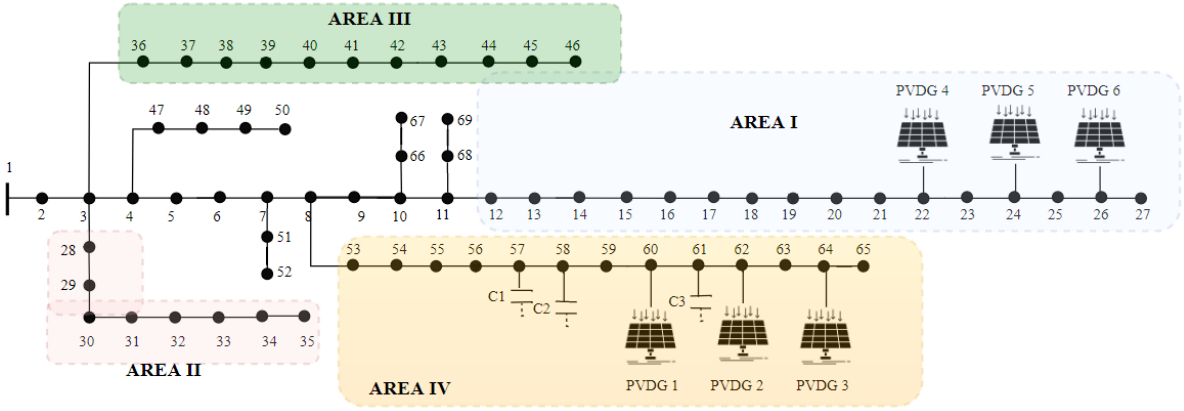

Figure 7.34: Modified IEEE 69-bus network

Table 7.6 provides the DG properties.

Table 7.6: DG properties on the modified IEEE 69-bus feeder

AREA I					AREA IV				
BUS ID.	DG PARAMETERS				BUS ID.	DG PARAMETERS			
	α	β	Power (kVA)	pf		α	β	Power (kVA)	pf
22	12.62	2.21	250	0.90	60	12.62	2.21	500	0.900
24	12.62	2.21	250	0.950	62	12.62	2.21	500	0.950
26	12.61	2.21	250	0.950	64	12.62	2.21	500	0.975

The properties (sizing and placement) of the shunt capacitors provided in Table 7.7 are adopted from a study on the optimal placement and sizing of shunt capacitors on the IEEE 69-bus feeder work [250].

Table 7.7: Shunt capacitor properties on the modified IEEE 69-bus feeder

BUS ID.	SHUNT CAPACITOR PARAMETERS	
	CAPACITOR ID	CAPACITY (kVA)
57	C1	185
58	C2	50
61	C3	1100

The feeder is analysed through four case studies. In Case 1, the feeder is passively configured, and is extended to include capacitors in Case 2, and DG in Case 3. The final case study (Case 4) builds correlation into Case 3. Table 7.3 provides the coefficients of correlation applied according to the service areas represented in Figure 7.34. The loads not included in these areas are considered independent.

Table 7.8: Correlation conditions on the modified IEEE 69-bus

AREA	FEEDER ZONING		DEPENDENCY	
	Load Class	Buses	Load-Load	DG-DG
I	Residential I	12 - 27	0.35	1.00
II	Residential II	28 - 35	0.85	No DG
III	Industrial	36 - 46	0.80	No DG
IV	Commercial	53 - 65	0.25	0.90

In all test cases, the Monte-Carlo simulation provides the validation outcomes.

7.6.1 Case 1: Passive Feeder

The various PLF approaches are conducted on the test feeder in passive configuration. Having assuming the deterministic load capacity as the mean for the statistical loads, the means of the PLFs are comparable to the published deterministic findings [249].

Figure 7.35 shows the profiles of mean voltages from the respective PLFs including the deterministic profile. The comparison of the HBA-MV trace with that of the MCS shows significant differences in the two outputs. This demonstrates the inadequacy of the approximate approach applied in the HBA-MV to compensate for effects of line reactance and reactive power by representing the complex inputs as absolute values. On the other hand, the matching outputs between the HBE (I-model) and the MCS (I-model) show that system size does not affect the HBE's accuracy in computing mean voltages.

As expected, the difference between the HBE (I-model) and the DLF solution increases with the magnitude of voltage-drop. However, as observed in the 33-bus case study, a single iteration of the HBE with updated nodal voltages reduces the differences significantly.

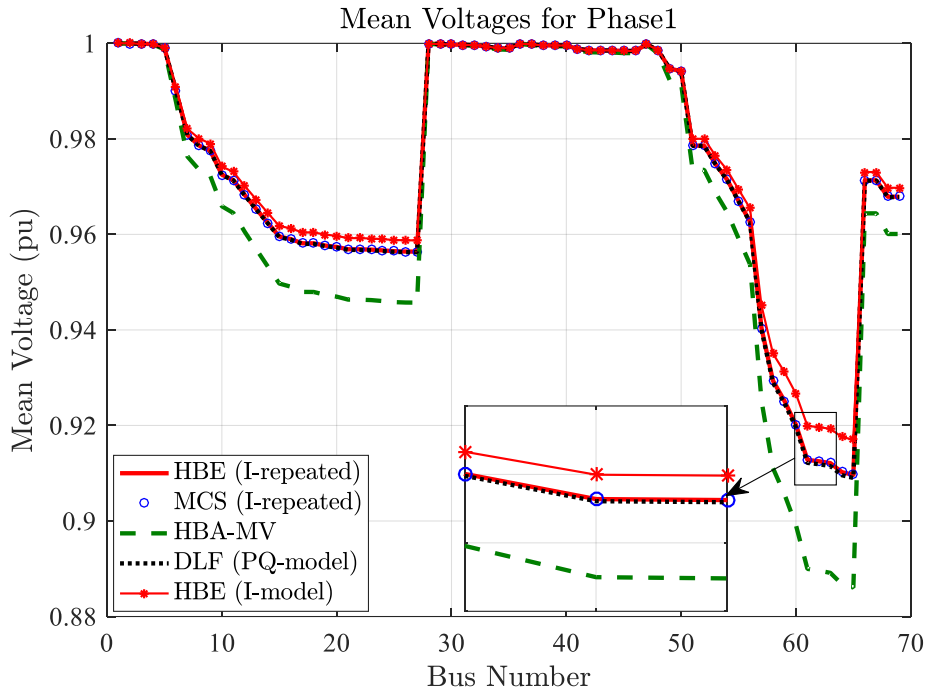

Figure 7.35: Comparison of mean voltages on modified IEEE 69-bus: Passive Configuration

The plots for standard deviations are shown in Figure 7.36. As expected, the HBA-MV results in significant errors while the HBE (I-model) profile and that of the MCS have close similarities. Significant differences in SD profiles between the HBE (I-model) and HBE (I-repeated) demonstrate the impact on standard deviation of an assumed load type.

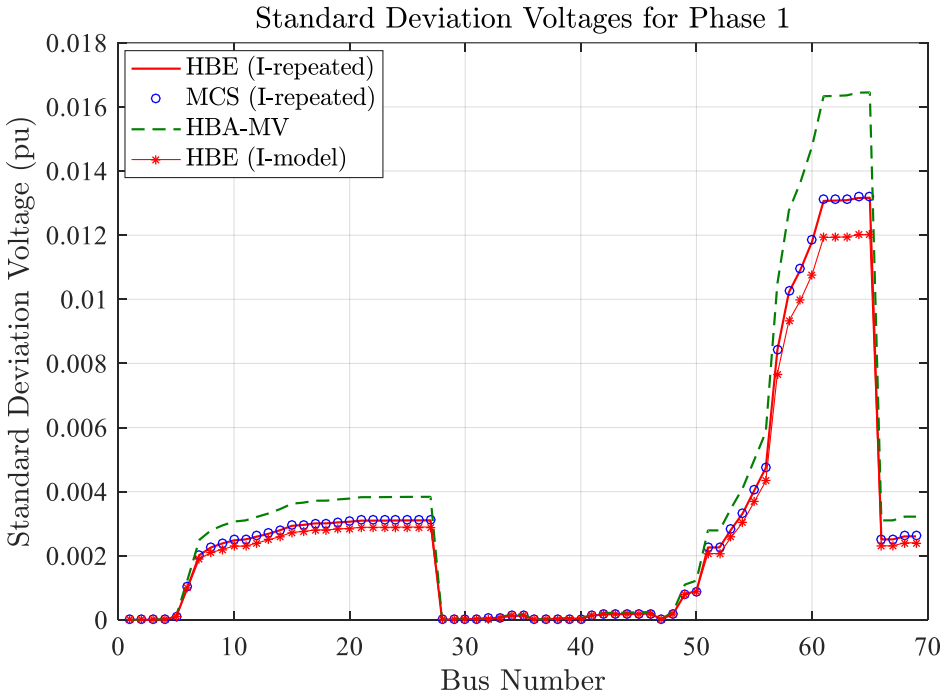

Figure 7.36: Comparison of SD voltages on modified IEEE 69-bus: Passive Configuration

Figure 7.37 compares sample voltage PDFs between the HBE, HBA-MV and the MCS. As anticipated from the results of mean and variance, the HBE provides a significantly better approximation of the output distribution than the HBA-MV.

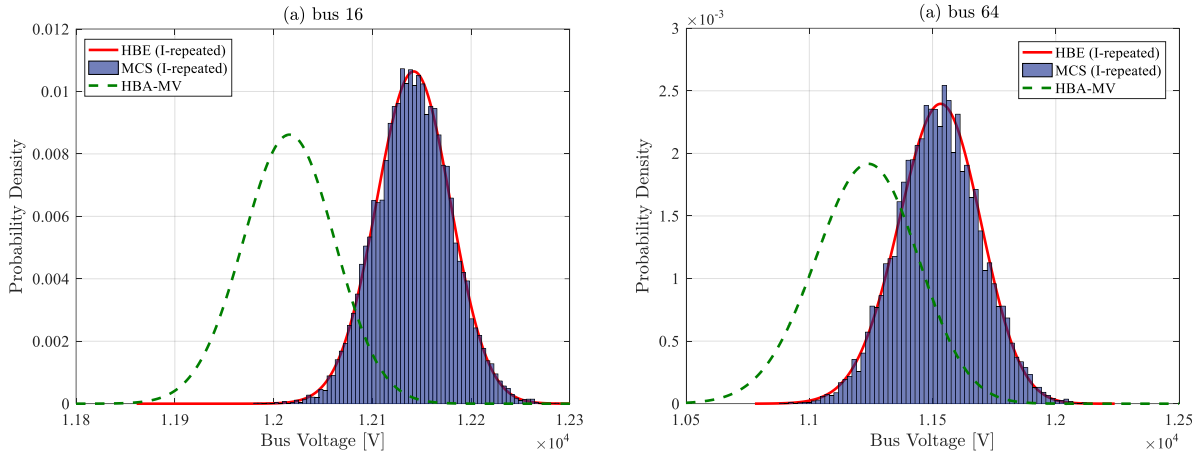

Figure 7.37: Comparison of bus voltage PDFs on the modified IEEE 69-bus: Passive Configuration

The HBA-MV PDF shows errors in the mean and dispersion. The outcomes from this case study demonstrate the suitability of the HBE to larger feeders in passive configuration, with inputs characterized as constant-currents or voltage-dependent models.

7.6.2 Case 2: Compensated Feeder

The PLF is conducted on the feeder with shunt capacitors. Under the conditions of symmetrically distributed spike-like distributions of shunt capacitor currents, the PLF outcomes are comparable with the published deterministic results [250].

As anticipated, the shunt capacitors improve the voltage profile as shown in Figure 7.38.

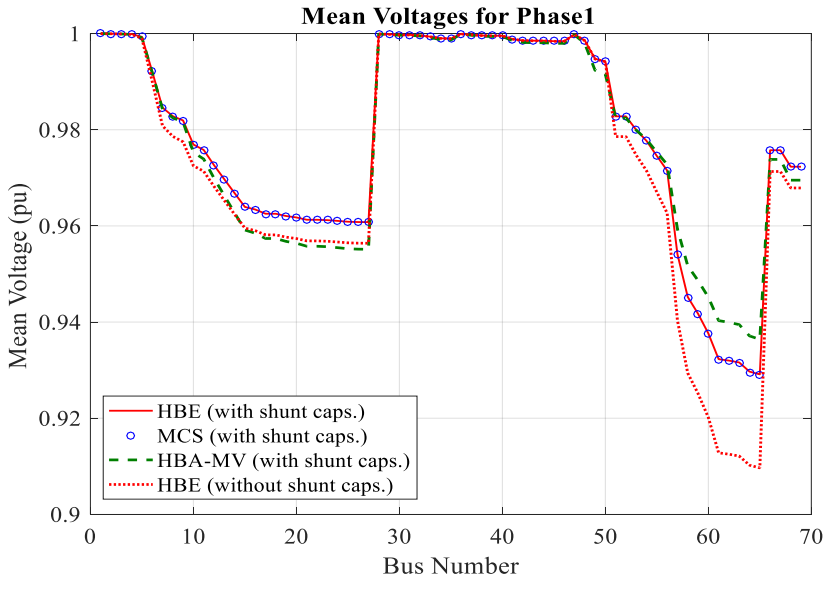

Figure 7.38: Comparison of mean voltages on modified IEEE 69-bus: Compensated

Based on the first HBE calculation, the compensated feeder has a minimum voltage of 0.92923 pu on bus 65, similar to the published result of 0.92975 pu [250] within an error margin of 0.056%. The difference is expected to decrease with repeated calculations, as shown in the passive case.

The plot of SDs in Figure 7.39 confirms the suitability of the HBE for large scale compensated feeders and the inadequacy of the HBA-MV due to errors in the measure of dispersion.

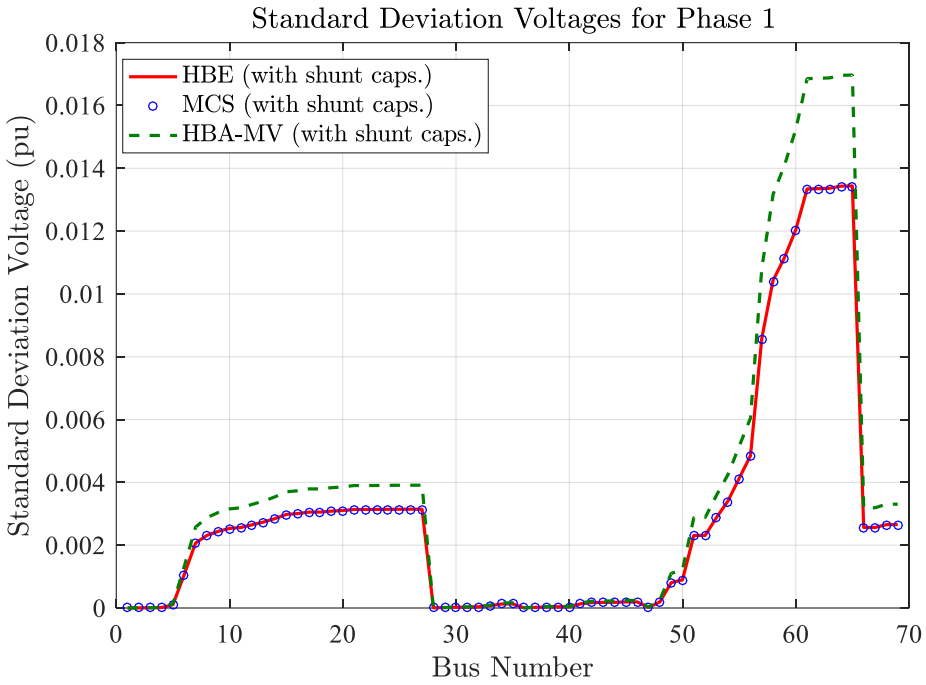

Figure 7.39: Comparison of SD voltages on modified IEEE 69-bus: Compensated

The output PDFs in Figure 7.40 demonstrate the accuracy of the HBE and the inadequacy of the HBA-MV approach.

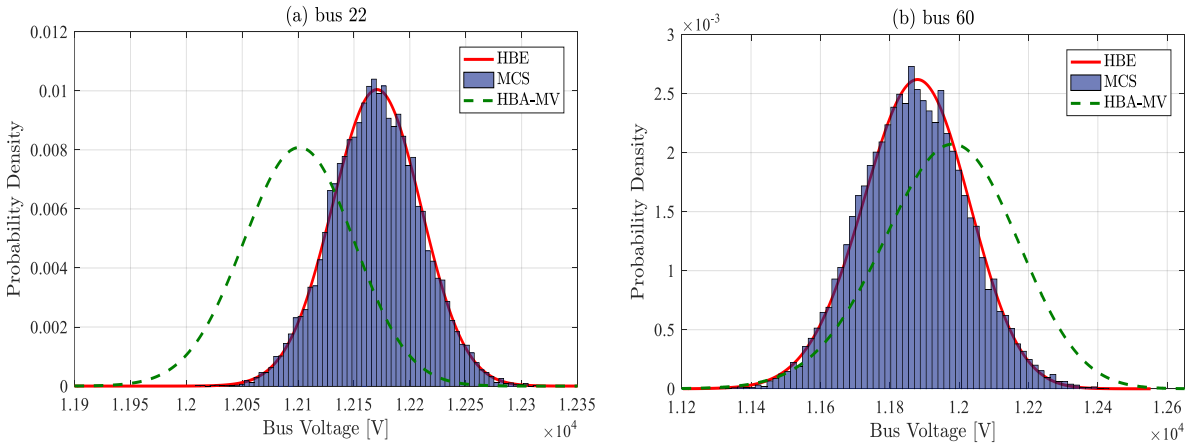

Figure 7.40: Comparison of bus voltage PDFs on the modified IEEE 69-bus: Compensated

7.6.3 Case 3: Active Feeder

The connection of DGs in areas I and IV further enhances the voltage conditions on the feeder as demonstrated by the difference in HBE profiles with and without DG in Figure 7.41. The MCS validates the HBE profile, while the HBA-MV shows considerable errors.

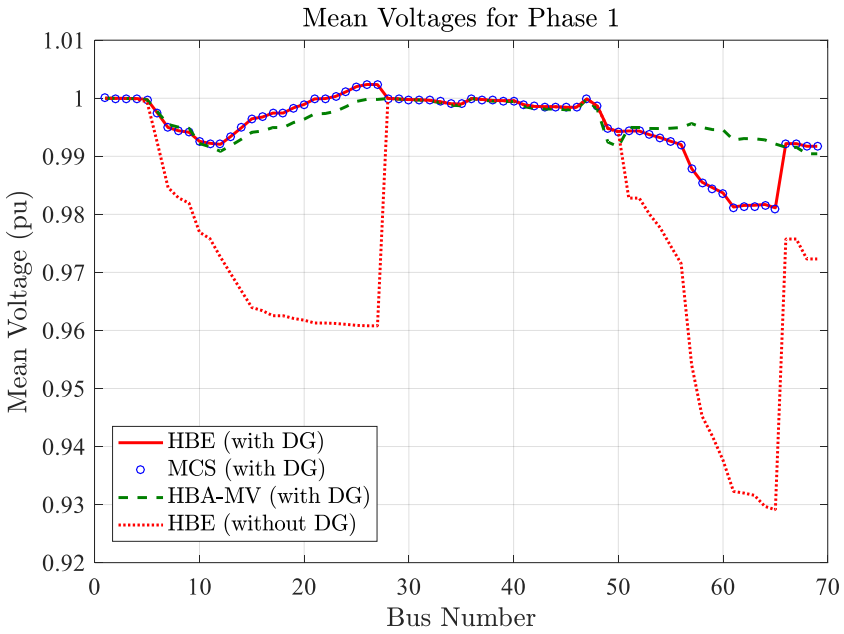

Figure 7.41: Comparison of mean voltages on modified IEEE 69-bus: Active configuration

Figure 7.42 compares the SD plots between the HBE, HBA-MV and the MCS PLF approaches. As expected, the SDs from the HBA-MV result in significant errors while the HBE matches the MCS result.

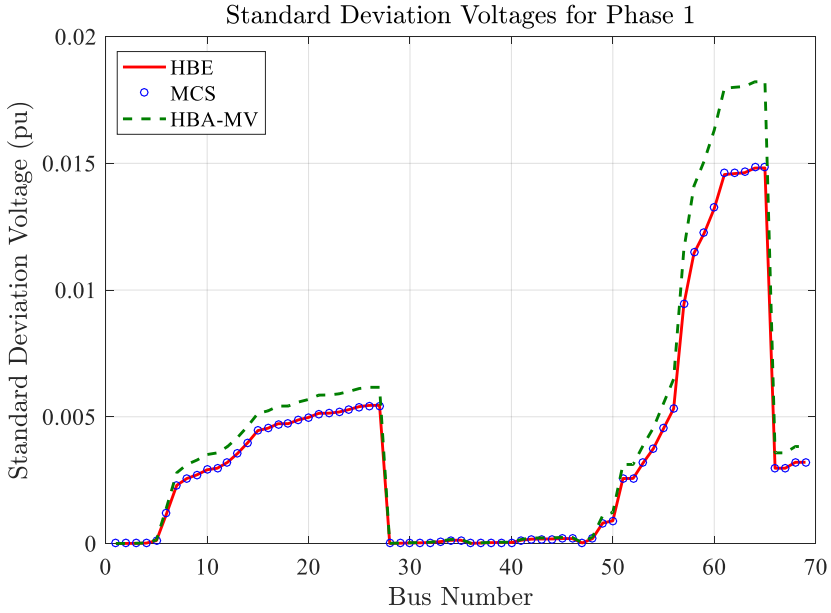

Figure 7.42: Comparison of SD voltages on modified IEEE 69-bus: Active configuration

The voltage PDFs in Figure 7.43 confirm the accuracy of HBE for large active systems and show how errors in the HBA means and SDs affect the voltage distributions.

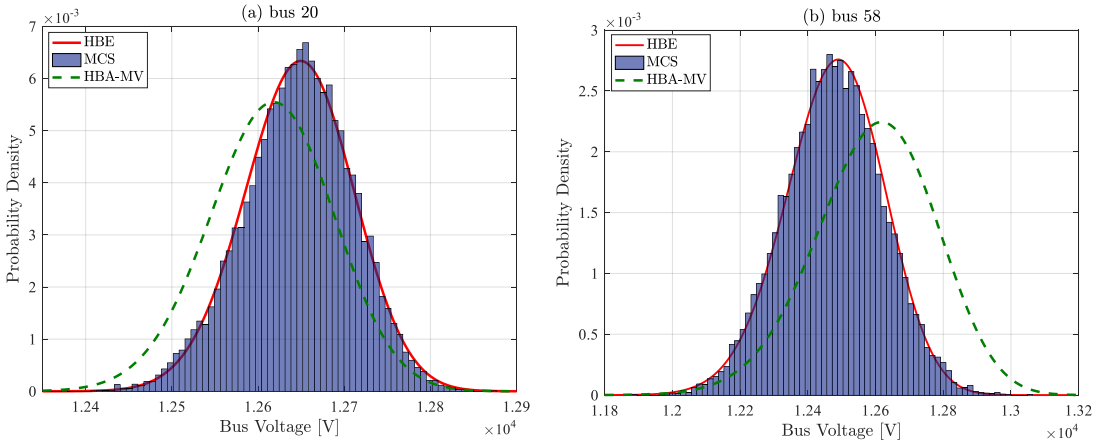

Figure 7.43: Comparison of bus voltage PDFs on the modified IEEE 69-bus: Active Configuration

7.6.4 Case 4: Dependency

Up to this point the RVs in all test cases have been assumed independent. In this case study, correlation between inputs is considered and built into the test conditions in Case 3. The PLFs are applied to the test network and comparisons are made between their outcomes.

Figure 7.44 shows the profiles of SDs resulting from the various PLF approaches. The comparison of the HBE with correlation to the one without demonstrates the effect of correlation on dispersion. The similarity between the MCS and the HBE (with correlation) shows the validity of the HBE approach with large multivariate systems with correlated inputs.

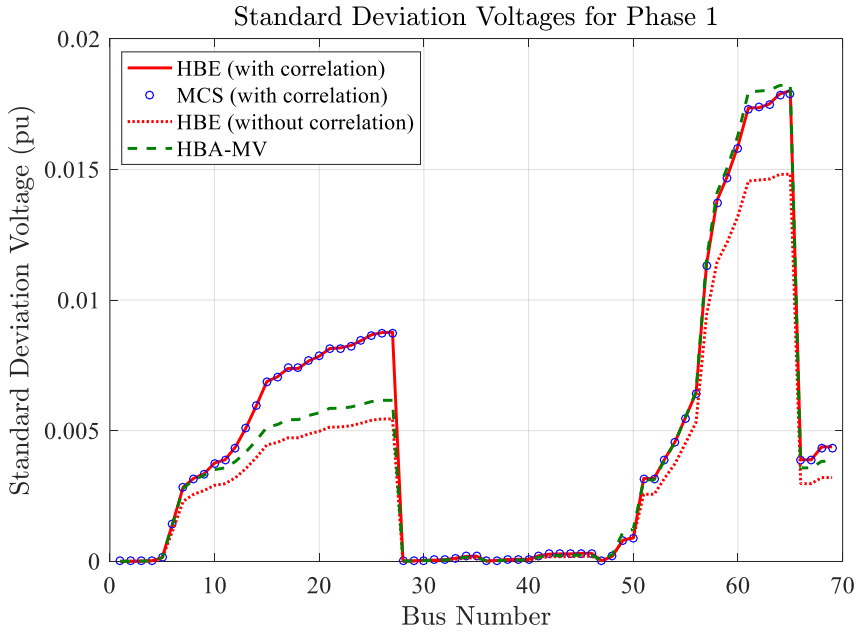

Figure 7.44: Comparison of SD voltages on modified IEEE 69-bus: Dependent inputs

As expected, the HBA-MV approach based on independent RVs results in significant errors. The plot of Figure 7.45 shows PDFs for selected bus voltages. As also reflected from the outcomes of SDs, the PDFs from the HBE (with correlation) fit the MCS histograms with good accuracy, while the HBE (without correlation) fails to capture the dispersion accurately. The HBA-MV has errors in both means and standard deviation as expected.

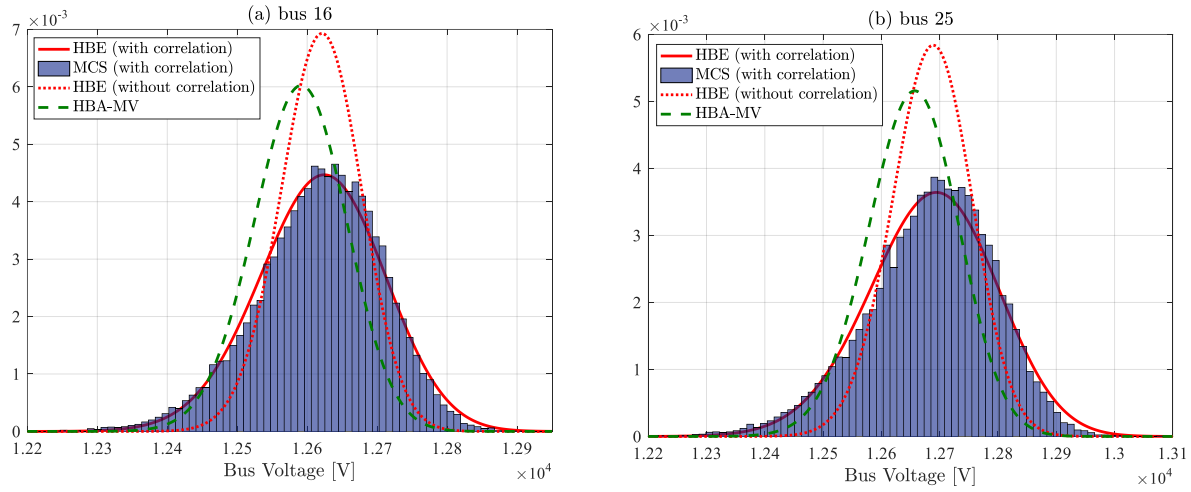

Figure 7.45: Comparison of bus voltage PDFs on the modified IEEE 69-bus: Dependent inputs

The conducted tests demonstrate the suitability of the HBE PLF tool for larger systems with dependent inputs.

7.7 OVERALL PERFORMANCE ANALYSIS

Extensive tests have been carried out in this chapter to evaluate the performance of the developed HBE tool under various network conditions. In this section, the HBE is analyzed for consistency by comparing its performance across the various test feeders (representative 12-bus, IEE 33-bus and 69-bus feeder). The performance is evaluated by looking at the accuracy and computational performance.

7.7.1 Accuracy

Consistency across various feeder conditions

Table 7.9 compares the mean percentage relative errors (based on comparison with the MCS) across different test feeders. Mean errors for the mean, standard deviation, and percentile voltage are compared for three simulation cases; passive, active, and correlation configurations.

Generally, the scale of the ‘errors’ is considerably small. There are various reasons for the differences. Firstly, as discussed in the review of Chapter 2, the MCS has inconsistent outcomes with error margins proportional to the number of simulation trials.

Table 7.9: Comparison of percentage mean errors across different test feeders

TEST SYSTEM	COMPARISON OF MEAN ERRORS								
	Passive			Active			Correlated		
	μ_V	σ_V	$p_{r_{\%}}$	μ_V	σ_V	$p_{r_{\%}}$	μ_V	σ_V	$p_{r_{\%}}$
Representative 12 - bus	0.002	0.465	0.150	0.002	0.382	0.212	--	--	--
IEEE-33 bus	0.002	0.169	0.006	0.003	0.352	0.074	0.005	0.785	0.070
IEEE 69-bus	0.001	0.143	0.005	0.002	0.116	0.025	0.004	0.439	0.050

The factor is common across all test cases. The investigation in Section 7.2.1 demonstrated that the HBE outcomes are consistent with the means of the variable MCS outcomes.

The second source for the difference is linked to modelling discrepancies between the MCS and the HBE. Some extensions, for example the multivariate sampling of non-identically distributed dependent random variables in the MCS, are considerably difficult to model with great precision. This primarily causes variance differences that explain the increased errors in the correlated cases.

Finally, another reason for the errors is related to the approximation of the beta PDF to the output functions. The beta PDF provides the best results when the distributions are symmetrical. Where the functions are skewed, the HBE approximated PDF differ slightly in variance, and hence the percentile values. This is reason for the increased errors in the 12-bus test. The test case is a worst-case with highly skewed inputs (both loads and DGs) at highly differentiated power factors on a short feeder. The results, however, show that the HBE results remain comparable with those from the MCS. The accuracy analysis demonstrates that the performance of the HBE approach is not significantly affected by network configurations, input characteristics and network size. The HBE is very fast in all tests and provides good estimates of the output distributions obtained in a 15,000-run MCS.

The context of 'error'

Throughout the thesis, the outcomes of the HBE have been validated using the MCS. The results confirm that the variation and magnitude of voltages and currents from the HBE are consistent with those from the MCS. Further, the consistency of the HBE under various feeder conditions has been demonstrated by the comparison of mean "errors". It should be noted that this measure determines the level of difference of the HBE results from those of the MCS, and although it is sufficient, given an adequate number of MCS simulations is used, the exact accuracy cannot be determined. The accurate assessment of the 'error' associated with HBE results requires comparison with 'true' values of the outputs. As demonstrated earlier, the MCS results vary creating a Gaussian distribution of variable outputs while the consistent singular HBE result is approximately equal to the mean of the distributions. To further demonstrate the context of accuracy, results from one of the test cases is used.

Figure 7.46 shows the relative errors (in p.u.) between the HBE and MCS (broken red line), HBE and DLF (solid red line) and lastly, the MCS and DLF (solid blue line). The DLF results are based on PQ-models while the HBE and MCS results are estimates achieved from iterative calculations (see Section 7.6.1).

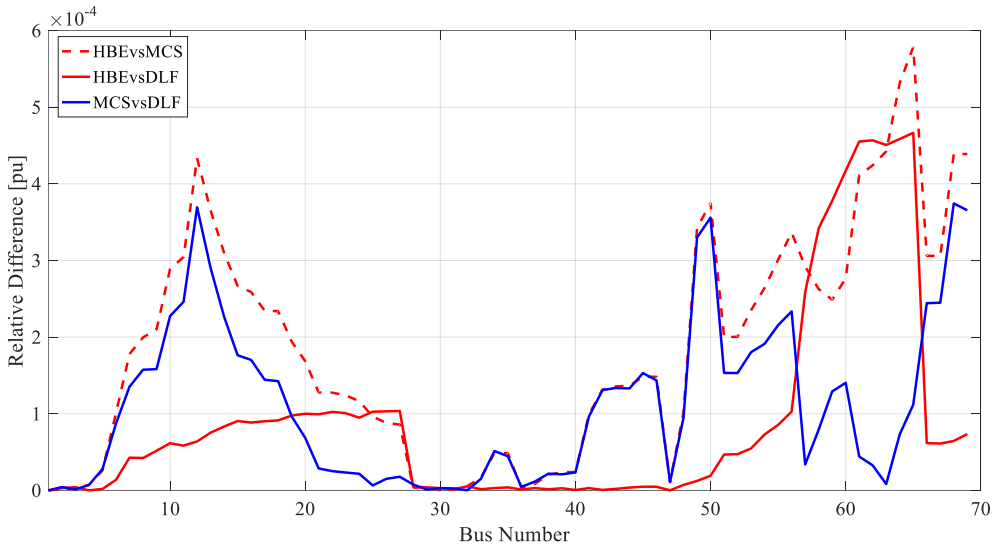

Figure 7.46: Relative errors between the MCS, HBE and DLF results on a passive IEEE 69-bus

The results demonstrate three important aspects;

(1) Using either the MCS or DLF as the true value, the margin of HBE errors is acceptable for all buses; the relative error for all bus voltages lies within a tolerance of 0.0006 p.u.

(2) The MCS also differs from the DLF results. This is due to the variation as a result of random simulations and the limited number of conducted simulations. In some cases, the MCS errors are higher than those of the HBE. However, they are also within a small tolerance. Accordingly, when comparing results with the MCS, care needs to be taken on the interpretation of the outcomes. In this specific case, the error in the HBE based on the MCS results (broken red trace) is in most cases overestimated as shown by the difference between the broken and solid red traces.

(3) The magnitude of the HBE error varies with voltage deviation. The reasons for this related to the approximation of the HBE to inputs of constant-PQ type have already been covered in Sections 7.3.1 and 7.6.1.

7.7.2 Computational Efficiency

This section aims to assess the HBE's computational performance compared to the MCS. This is achieved by comparing the CPU utilization time of the MATLAB-coded PLF programs on three selected networks. This comparison is presented in Table 7.10.

Table 7.10: Computational efficiency evaluation

TEST SYSTEM	COMPUTATIONAL TIME (seconds)	
	MCS	HBE
Representative 12 - bus	112.6543	0.1012
IEEE-33 bus	912.6326	0.7015
IEEE 69-bus	3112.1952	3.1296

As expected, the results show the HBE's is highly computational efficient. On the IEEE 33-bus feeder, the HBE based on constant-current models performs up to 1300 times faster than a 15,000-run MCS. This factor is roughly the same across the representative 12-bus and the IEEE 69-bus feeder.

When considering voltage-dependent input models, the computational time would increase by a factor equal to the number of HBE iterations carried out. Based on the validated adequacy of the approximation to the power model using a single iteration, or even a few iterations more, the HBE remains exceptionally efficient.

While the computational advantage demonstrated through these cases might seem unimportant, the speed characteristic of the HBE becomes more apparent in real-time applications and those involving repeated probabilistic load flow calculations. Examples of such applications include DG penetration studies with random placement of generators, optimization studies, and state estimation. The application of the HBE approach in one of these applications is demonstrated in the next chapter.

7.8 CONCLUSION

This chapter, through several case studies on a range of test networks, has rigorously tested the HBE approach for calculating probabilistic load flow on 3p-3w and 3p-4w systems. The conducted tests demonstrate the reformulated approach provides a fast and accurate PLF solution, consistent with the MCS. The approach has been validated with independent or dependent inputs modelled as constant-current or voltage-dependent for application on passive, active and compensated radial feeders. The next chapter demonstrates the practical application of the developed tool in DG penetration studies.

Chapter 8: The Application of the HBE Transform to DG Penetration Studies

The previous chapter validated the HBE transform for voltage and line current calculations on passive, active and compensated radial feeders with independent or dependent inputs, modelled as constant-current or voltage-dependent. The developed tool offers a variety of practical applications. As an example of the possible practical applications, this chapter demonstrates how the developed probabilistic load flow tool can be applied to DG penetration studies.

8.1 INTRODUCTION

The increased interconnection of DG, mostly from PV and wind plants, has offered many possibilities in power systems. Despite the widespread benefits of the integration, diverse forms of technical challenges encountered with high levels of penetration limit the hosting capacities of distribution networks. The assessment for the limits of penetration, considering the uncertainties associated with both loads and DG, and the unknown installed capacities and placement demands the application of probabilistic approaches. Since simulating the effects of randomly placed DG requires repeated PLF calculation with DG capacity and location scenarios, the advantages of a computationally efficient PLF approach becomes very evident.

In previous work [132], the HBA has been applied to the same application but with its inherent assumptions and without considering the impact on thermal limits of conductors. This chapter offers a detailed probabilistic framework for assessing the hosting capacity of active distribution networks using a hybrid approach based on the HBE and MCS approaches.

8.2 PROBABILISTIC FRAMEWORK FOR DG IMPACT ASSESSMENT

As stated earlier, the simulation of a power system network providing more representative results requires consideration of the following non-deterministic characteristics:

1. The stochasticity of the load demand
2. The variability in the PV-DG output
3. The uncertainty associated with the location of PV-DG (it is not known which customers would install PV systems)
4. The uncertainty of the size of the installed PV-DG system (despite the knowledge of the maximum solar hosting capacity reflected by the available roof-space area).

Items (1) and (2) are dealt with using input statistical models (load and DG) and load flow calculations using the HBE which allows the PLF to be analysed quickly and accurately. However, the HBE transform, in its calculations for voltages and branch currents, requires the 'specific' locations and capacity of the installed PV-DG.

The knowledge of the anticipated PV-DG location and capacity is unknown and depends on a variety of characteristics such as customer electricity consumption, level of income (for affordability), customer preference (whether they see value in PV or not), and installation limitations, such as roof space availability. As a result, only a stochastic analysis framework is appropriate in the investigation of the impacts of these ‘randomly’ located and sized PV-DG systems.

An MCS applying uniform probability sampling for phase and node allocation is sufficient to simulate the effects of random placement. This allocation process is bounded by two factors: the level of penetration analysed and the limit of uptake at a given node and phase, depending on the limitations of uptake of the respective properties connected to that point. The power supply circuit breaker is used for the uptake restrictions. To assess the impacts of different levels of PV-DG penetration on the feeder, PV-DG modules with a capacity 2.9 kW are added to the feeder until each customer reaches the uptake limit. The simulation uses 1,000 randomly selected PV-DG allocation combinations for each penetration scenario. Figure 8.1 provides an overview of the implemented methodology.

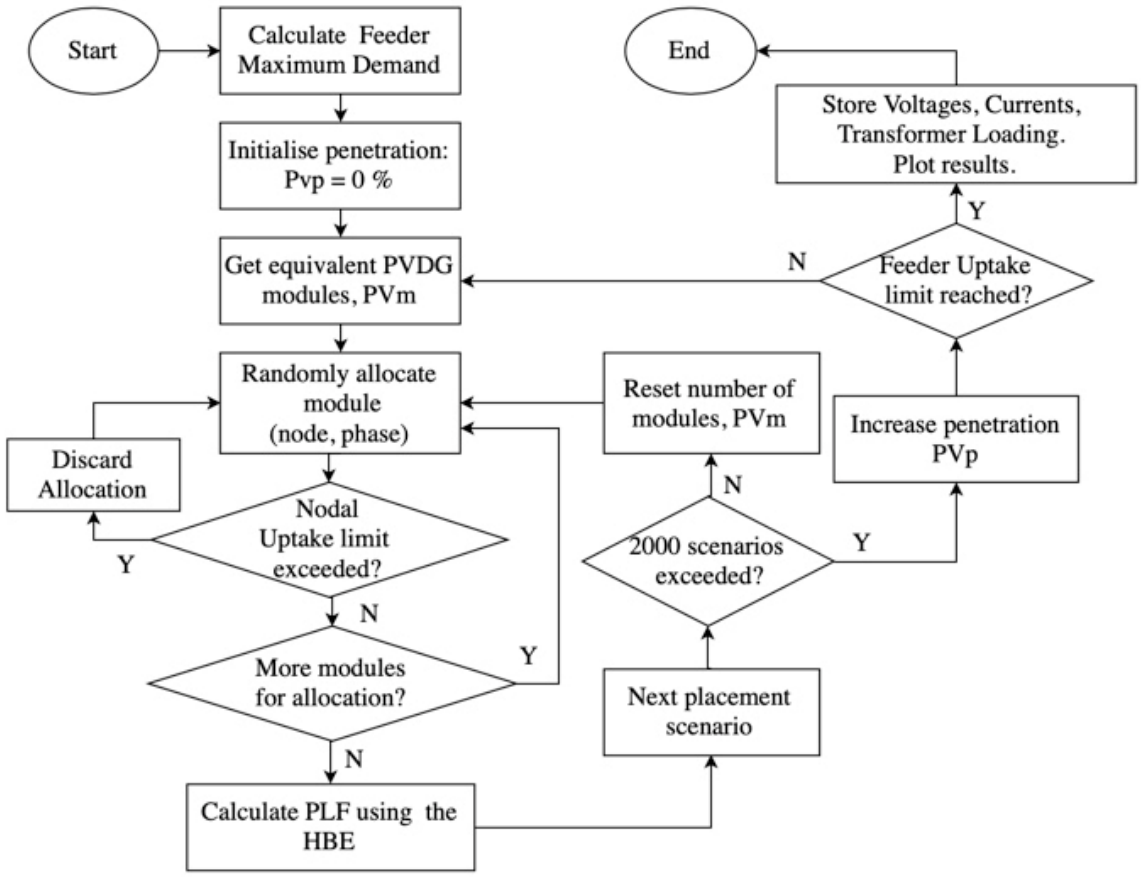

Figure 8.1: Flowchart of the analysis of PV-DG penetration using the combined HBE-MCS approach

While, according to the design guidelines for LV networks in South Africa [94], the permissible voltage limit for both drop and rise is 10%, during the time in which the network under study was designed, voltage regulation on the LV networks allowed for voltage-drop of only up to 8%. Accordingly, the voltage regulation limits of 8% voltage-drop, and 10% voltage-rise are used, a design risk of 5% is used for both voltage and current limits.

8.3 PV PENETRATION ANALYSIS ON DISTRIBUTION FEEDERS

In the work reported in this chapter, the use of the term 'penetration' refers to the total installed PV-DG capacity expressed as a percentage of the feeder's maximum demand (FMD). The FMD is defined as the maximum load (MD) the feeder can supply without exceeding the regulatory voltage limits or the thermal conductor limits. Since the MD is likely to occur on a winter night (characteristic of residential customers in South Africa), the winter beta PDF load models, without any PV-DG, result in the FMD. The penetration analysis is conducted using the load and DG models for summer, during noon-time. The relevant models and simulation methodology used are discussed next.

8.3.1 Network Description

The test network and data are representative of an actual network in South Africa. The utility adopts various limits for the design of feeders, for this feeder the range of voltage must be between 0.92 and 1.1 pu of the rated voltage. The feeder, which will be identified as the 'practical 12-bus feeder', is a four-wire model operated at 230 V and supplying a community of 36 customers in a low-density residential area. The one-line diagram of the feeder is shown in Figure 8.2.

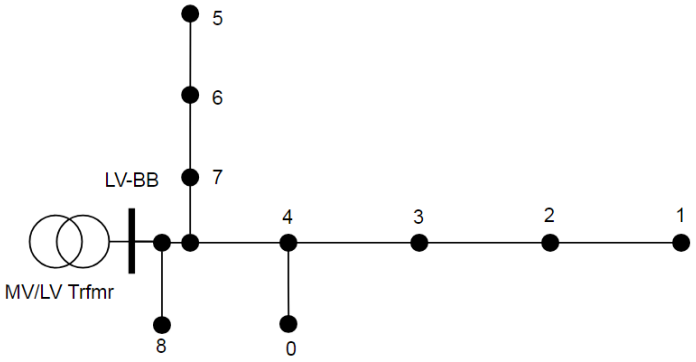

Figure 8.2: A Practical 12-bus feeder

The full network properties are included in Appendix B5.

8.3.2 Statistical Input Parameters

The statistical load models for the group of customers supplied by the test feeder were determined as follows: (1) the after-diversity maximum-demand (ADMD) is extracted from the feeder's yearly consumption data and is taken as the mean demand (μ_D), (2) the standard deviation (σ_D) per phase current is estimated according to the technical guidelines from the utility company [251] given in (8.1).

$$\sigma_D = \frac{1}{\sqrt{3}}\left(2.19 + 0.56\,(3\mu_D)\right) \qquad (8.1)$$

From the mean and variance values, the beta PDF parameters (α_D, β_D, C_D) are calculated as before.

For PV-DG, the statistical models are based on beta PDFs of current at unity power factor and of spike-shape distributions. The low standard deviation in the statistical models of the PV-DG is based on the assumption that the units are optimally installed to produce maximal solar power, ignoring the effects of partial shading, cloud transients, or panel orientation. Accordingly, the models are characterised as means of the efficient output (with inverter output at 80% of the PV panel rating) and a small variance (about 0.05%) to model the variability. The selection of the beta distribution parameters α and β are such that a symmetric spread of currents with little dispersion is achieved.

Table 8.1 specifies the beta load and DG parameters for the test feeder. The parameters are common to all connected customers.

Table 8.1: Load and DG characteristics for practical 12-bus feeder

INPUT TYPE	LOAD CHARACTERISTICS PER HOUSEHOLD						
	Summer		Winter		C [A]	Summer ADD/hh [kVA]	Winter ADMD/hh [kVA]
	α	β	α	β			
Loads	1.083	11.81	1.445	5.118	80.00	1.546	4.051
DG	255.5	255.5	--	--	20.17	2.320	--

With the input parameters determined, save for the capacity and location of PV-DG which will be provided from the MCS, the distribution feeder can be solved.

8.3.3 Simulation Methodology

Three sections describe the simulation methodology followed. The first provides guidelines on how the penetration analysis is conducted while in the next section the characteristics of the MCS stochastic simulator are discussed and in the last section the interpretation of the anticipated simulation results is given.

The PV-DG penetration study in steps

The simulation methodology followed is described in steps below:

1. Load the feeder with the winter load models, and calculate the voltage profile on the feeder, applying a 5% risk (95% confidence interval) to assess the minimum voltages.
2. Increasing the winter load models, model the feeder's highest passive loading it can supply without violating the design voltage. This is the FMD and is not the same as the allocated de-aggregated demand supplied by the feeder.

3. Using the feeder loaded with the summer load model from step 2, randomly (to node and phase, by means of the Monte-Carlo Simulation) allocate PV modules on the feeder and calculate the voltage-rise and conductor thermal loading conditions represented by a risk level (of exceedance) of 2.5% for the respective limits (1.1 pu for voltage conditions and 100% for thermal loading). For each scenario calculate the transformer loading.
4. Repeat this step for an adequate number of placement scenarios; in this study 2000 runs.
5. Successively add further PV modules, repeating steps 3 with each increment, until the feeder is 'full', having reached the limit imposed by the circuit breakers of all households.
6. Plot the results of calculated voltages, line currents and transformer loading for all scenarios against the penetration ratio on a scatter plot. Derive the maximum hosting capacity of the feeder for both voltage-rise and conductor thermal loading, again with a selected confidence risk; in this study a further 2.5%.

Selecting the number of MCS simulation-runs per penetration scenario

The optimum number of simulations for the MCS requires convergence tests which are difficult to carry out especially with variations of the number of random variables with each penetration level. As a general rule, the more the investigated cases the better.

For each level of PV-DG penetration, randomly located, there is a range of the highest 'node and phase' voltages. Plotting this range for all different levels of PV penetration results in a scatter plot of the many different highest voltages that might occur somewhere on the feeder. The blue scatterplots show the highest feeder voltage simulated on any phase at any node as the PV-DG penetration of the feeder increases from none to being full. In these simulations a full feeder is defined as having PV installed at every customer to the maximum allowed by each circuit breaker.

Figure 8.3 depicts the 5% risk curves within two voltage-rise scatter plots for a sample feeder using (a) a 2,000-run randomised MCS and comparing it with those from (b) a 5,000-MCS. There are small differences in the apparent upper limit of the envelope of voltage-rise. The effect of carrying out more simulations is evident in the red line that shows the 95% confidence level, with only 5% of the simulations having a node/phase voltage above the red line. The slightly smoother line with 5,000 simulations, especially compared with 2,000 or fewer simulations, is the result of having more definition in the tail of the PDF with a much larger sample. For the purposes of this study, 'sufficient convergence' is tested by assessing the variations in a sample percentile trend on the scatter plot. Using this method, the optimum number of simulation trials is expected to achieve a smooth percentile trend (or risk trend).

Visually smoothing the trend line, the results do not show significant differences in the risk trends. Accordingly, 2,000 scenarios are used in this study.

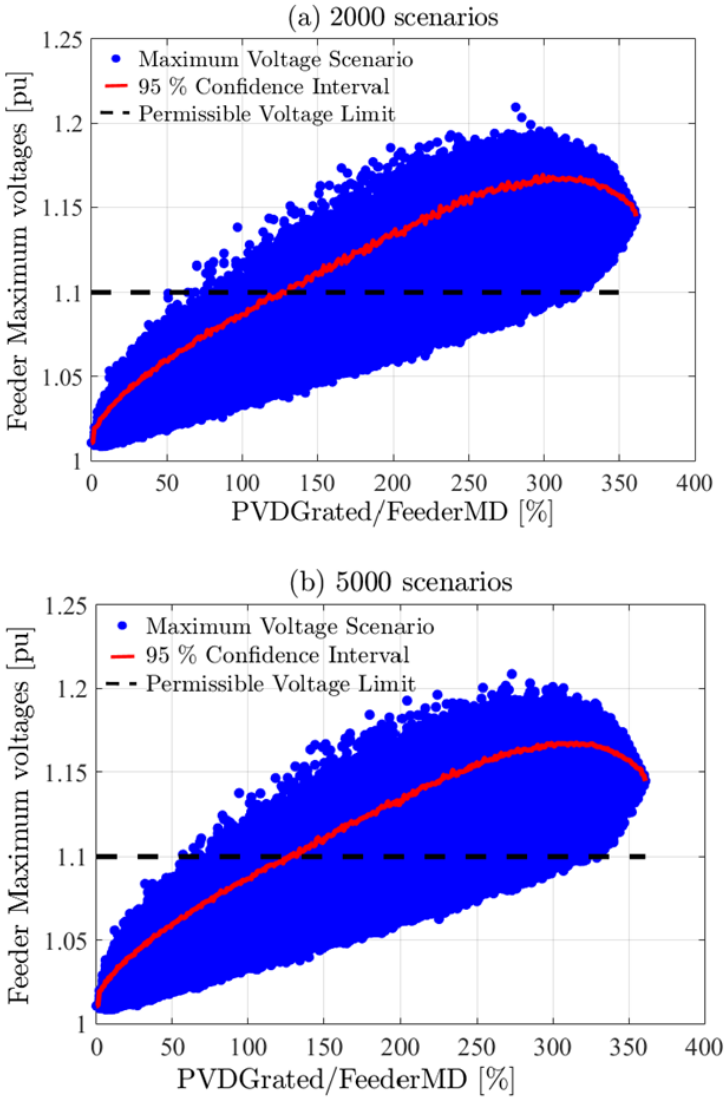

Figure 8.3: Impacts of the number of placement scenarios on voltage scatter plots

Interpretation of scatter plots

When evaluating the technical impacts of DG penetration, it is anticipated that the effects will be minimized and the hosting capacity maximized when the coincidence of load and generation is high, whereas low coincidence can result in surplus generation that can lead to reverse power flow, voltage rise and thermal loading problems. Likewise, high loads with low load coincidence can also lead to voltage-drop problems as well as thermal loading.

The scatter plot of possible voltages and line currents (step 6) replicates the combined effects of the variation of many parameters, including: the different statistical models of the loads on each feeder, the stochastic variation of the loads and the DG, the topology of the feeder, particularly branching, the effect of the correlation between the load location and the DG location, the random placement (node and phase allocation) of the DG, the incremental steps of the DG capacity and the rating of the maximum capacity at each household.

For the voltage scatter plot, since the sending end voltage is taken as 1 p.u. (the nominal feeder voltage), unbalance between the phases causes voltage-rise even without any DG connected, so the scatter plot does not reach the origin at 1 p.u. voltage.

Clearly, there is some risk associated with the upper boundary of the scatter plot represented by a Monte-Carlo simulation and an acceptable risk curve can be plotted with the simulation. In this case, allowing 2.5% simulation risk and 2.5% HBE risk produces a limit curve of about 95% confidence that the penetration of the PV-DG by random uptake by households can be accommodated. In practice, this is a useful limit because up to this level, utility customers might be allowed to connect PV-DG without detailed studies required. More capacity might be possible, but the locational correlation (of nodes and phase allocation) between load and PV-DG would need to be studied in greater detail.

The results from the conducted simulation study are discussed in the next sections.

8.4 RESULTS

8.4.1 Passive Feeder Conditions for Voltage and Current – FMD (Phase A, B and C)

The parameters for each of the feeders in passive operation are presented in Table 8.2, together with the derived feeder capacity. It is noted that the actual winter demand on the feeder is nearly equals the Feeder capacity (FMD) based on supplying the maximum demand without violating the voltage-drop constraints. This is also demonstrated by the voltage-drop profiles in Figure 8.4 because the worst voltages are just beyond the regulated limit of 0.92 pu.

Table 8.2: Feeder parameters in passive configuration

PASSIVE FEEDER PROPERTIES (10% V-DROP RISK)		
Summer After Diversity Demand [kVA]	Winter After Diversity Maximum Demand [kVA]	Feeder Capacity, FMD [kVA]
55.64	145.84	149.49

The results of the simulations have been plotted graphically to assist interpretation.

Figure 8.4 shows the voltage-drop profiles for each phase of the passive feeders. The profiles are color-coded according to the branch currents (calculated at the risk level of 10% adopted for passive feeders) expressed in p.u. of the cable capacity. Blue lines represent branches loaded below 80% of capacity, orange lines for cables approaching their capacity, and red lines for those sections apparently at risk of exceeding the current capacity.

The combined voltage and branch current plots on an axis of electrical distance (measured from the source node) are convenient representations as the profiles are similar to the one-line diagrams. The comparison of Figure 8.2 and Figure 8.4 (a) with all buses numbered demonstrates this. From the plots it is observed that the branch leading to bus 4 presents the bottleneck to the capacity of the feeder. Phases 1 and 2 have thermal violation during the winter ADMD, which demonstrates an under-design

of the system (when using a probabilistic load flow approach). The different (relative) loading condition in Phase 3 is as a result of the unbalance on the network.

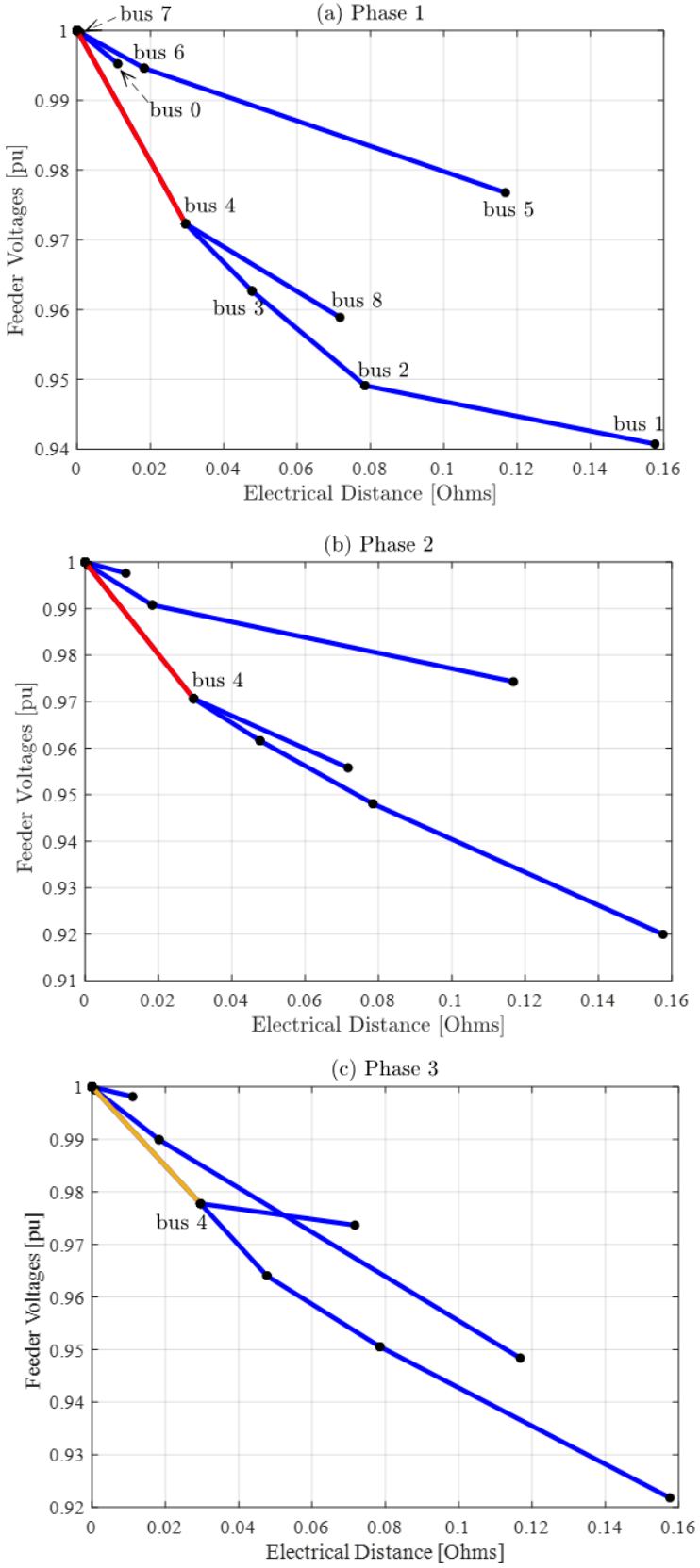

Figure 8.4: Bus voltages and branch currents during the maximum winter demand interval

8.4.2 Active Feeder Conditions under Summer Loads and PV-DG

Figure 8.5 shows the scatter plot for maximum voltages recorded from the repeated placement scenarios and varied PV-DG penetration (2,000 runs per MCS). Based on voltage conditions and applying the 5% design risk, the feeder can accommodate up to 95% PV-DG penetration, provided the voltage can rise to 1.1 pu above the voltage at the source. If the source voltage rises at the same time, then the tolerance for voltage-rise on the feeder is limited to less than 0.1 pu of rated voltage. The plot also demonstrates that the circuit breaker limit is too relaxed as the greater part of the envelope lies above the voltage-rise limit.

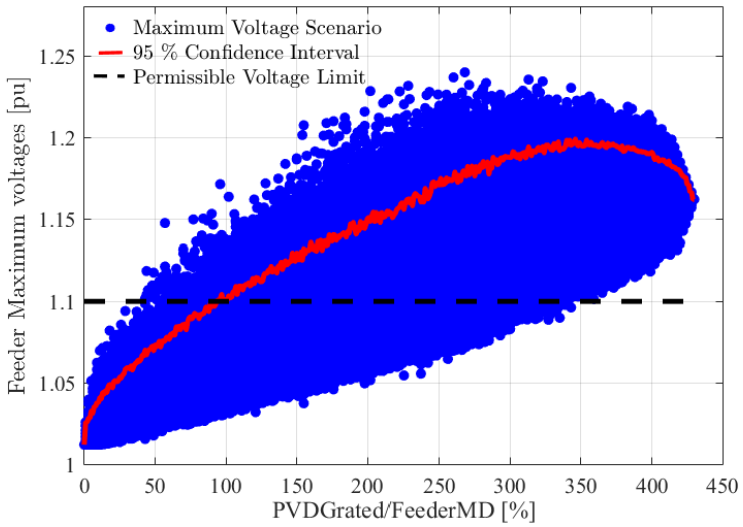

Figure 8.5: Scatter plot for maximum voltages with varying PV penetration and random placement scenarios

The HBE transform also calculates the currents in the branches at the chosen risk level. The maximum ratio of a branch current relative to the current rating of that branch is plotted in Figure 8.6.

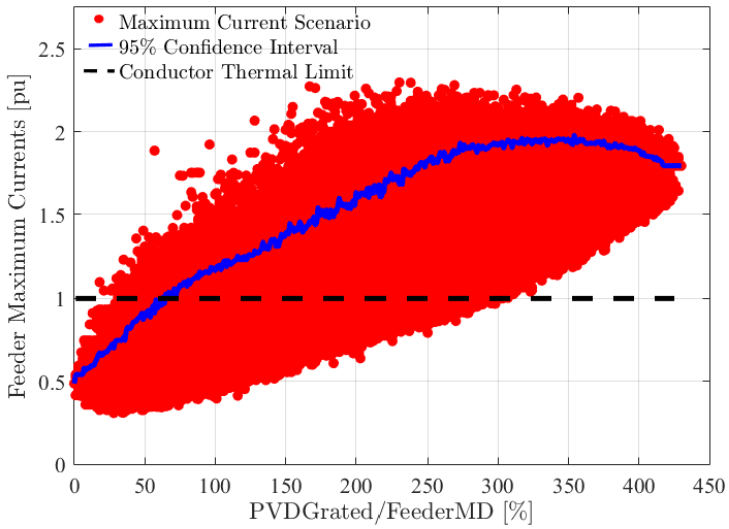

Figure 8.6:Scatter plot for maximum currents with varying PV penetration and random placement scenarios

The plot shows that based on thermal conditions, assuming the current should not be allowed to exceed the conductor rating) a penetration level of only up to 65% is permissible. Compared with the limit imposed by voltage-rise, this demonstrates that the feeder is thermally constrained. The scatter plot in Figure 8.7 show the impacts of randomly allocated PV-DG on the feeder.

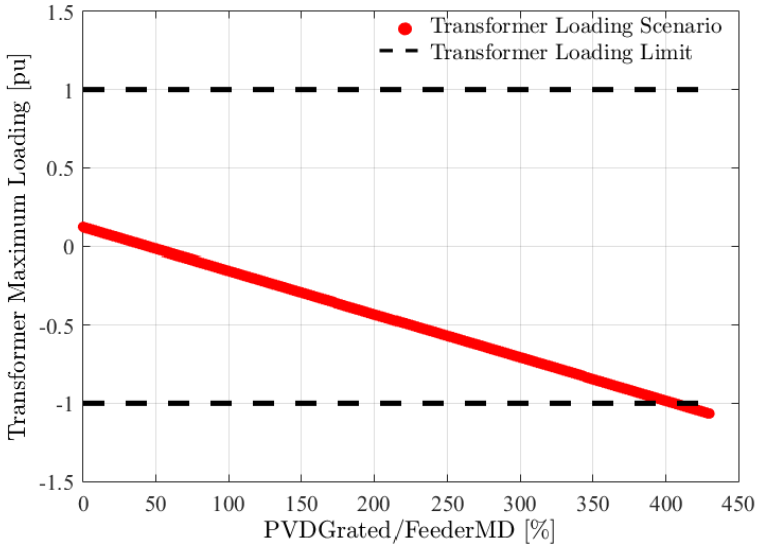

Figure 8.7: Scatter plot for transformer loading with varying PV penetration and random placement scenarios

Unlike the plots for voltage and currents, the scatter plot is not in the form of an envelope. This is because it is not affected by the location of DG on a feeder, the loading impacts of a particular level of DG penetration on the transformer are the same. As a result, the 2,000 scenarios of PV-DG allocation, at a given penetration level, result in the same transformer loading level. Based on the limits for transformer loading (assuming they are at 1 pu or 100%), this lightly loaded transformer under passive feeder conditions can accommodate about 400% penetration under active feeder conditions.

A more detailed interpretation of the plots follows.

8.5 INTERPRETATION

8.5.1 Voltage Magnitude in Active Feeders

Four characteristics are evident from the voltage-rise plots:

- At zero PV penetration, the voltage-rise is above the nominal voltage of the feeder because unbalanced passive loads cause a voltage-rise in the less loaded phases.
- With small levels of PV penetration, the location and phase of the PV-DG might mitigate the unbalance of the loads, so the lower limit of the scatterplot shows that the highest voltage might be lower than occurs without any PV-DG. However, with small numbers of PV-DG (a low level of penetration), if the location and phase of the PV aggravates the unbalance of the loads, then the upper limit of the voltage-rise, with 95% confidence, increases quickly.

- It is possible for a voltage to occur that is much higher than the 95% confidence level. This represents values within the 5% risk. Voltage-rise 40% higher than at the 95% confidence level is not unexpected for PV penetration between 10% and 100% of the feeder MD.
- At high levels of PV penetration, the possibility of unbalance diminishes, until complete balance is reached with a full feeder, bringing the scatterplot to a point.

8.5.2 Thermal (Current) Capacity of Active Feeders

The red scatterplots show the magnitude of the highest current simulated on any phase in any section of the feeder as the PV penetration increases. The currents are plotted as magnitude, as the direction of the current is not relevant to the thermal performance of conductors. The characteristics of the plots are similar to those of the voltage plots, but to comment on a few points:

- At 0% PV-DG penetration, there is only one scenario – the passive feeder with the summer midday loads.
- When PV penetration is low, the highest currents in the feeder branches might be reduced by co-located DG and load. However, at the same level of penetration, the extent of increased currents exceeds the effect of co-location, demonstrating that randomly located DG does not alleviate overloading.
- The 95% confidence level of the highest current (in pu of that conductor's rating) is shown by the blue line. As for the voltage-rise, the risk of even higher currents is evident.

8.5.3 Transformer Capacity

The plot of the transformer maximum loading against PV-DG penetration illustrates the direction of flow of the currents. The onset of reverse power is evident. For this feeder, the limits of voltage-rise or current in the feeder are exceeded before the transformer current reaches the transformer rating.

8.5.4 Simulation Computational Time

The time taken for the penetration study of the selected feeder using the MCS-HBE approach is about 7 minutes. If an MCS approach was used to derive the voltages and currents, 15,000 PLF simulations for each of the 2,000 placement scenarios would be required. Then the same computational efficiency shown in the previous chapter would require 1300 times longer, the MCS-MCS approach would take over 5 days to converge to a solution.

8.6 CONCLUSION

This chapter has demonstrated the application of the HBE transform to DG penetration studies. Through the simulations conducted, it has been shown that an extensive analysis including thermal limit and transformer loading assessments is possible. Due to the need for a stochastic simulator which in this

case is unavoidable, the computational advantage of the HBE is well-appreciated. The next chapter draws conclusions from the findings of the thesis.

Chapter 9: Discussion and Contributions

The new HBE approach has been rigorously tested using numerous case studies on different types of feeders. The previous chapter went on to demonstrate a possible practical application of the HBE approach to distributed generation penetration studies. In the light of these outcomes and the findings made in the other parts of the thesis, this chapter reviews the contributions the thesis makes and assesses the validity of the research hypothesis.

9.1 RECAPTURING THE RESEARCH GOALS AND THE PROGRESS OF THE THESIS

The need for probabilistic load flow analysis of power systems was identified several decades ago. At that time, the uncertainty in the power system was mainly due to the load stochasticity and generation outage rates. With the continual changes to the power system, especially the increased penetration of distributed generation that is highly variable, probabilistic approaches to feeder analysis have become increasingly important.

Chapter 2 identified the existing approaches to the probabilistic load flow, the common challenges in deriving model solutions, limitations of existing solutions and the resulting research opportunities. The review focused on the performance of the existing methods in terms of accuracy (compared to model solutions), computational efficiency, and complexity. The chapter set out the performance and scope requirements of a robust alternative PLF tool in terms of five key features:

1. An efficient probabilistic methodology that minimizes computational burden while preserving the accuracy in its solution.
2. Avoids unnecessary complexity that may limit practical implementation.
3. Suitable for a wide range of conditions common in power systems: unbalance (balanced and unbalanced LF), feeder topologies (3p-3w, and 3p-4w), network size (small and large systems) and feeder configurations (active, passive and compensated).
4. Precise models of the input parameters that characterise adequately the uncertainty associated with the random variables.
5. Representation of interdependency between random input variables.

The accuracies of most PLF approaches discussed in the literature survey are sensitive to factors such as skewness in the input PDFs, conditions of input dependence, and the total number of variables in a system. Further, most iterative approaches have computational issues with large systems. As a result, the reviewed approaches only partially fulfil the established requirements for a robust PLF procedure. The Herman-Beta algorithm (HBA), based on beta-PDF, demonstrates a strong computational

advantage, the flexibility to model various shapes of input distributions, and is relatively uncomplex and suitable for practical applications. However, the approach does not meet all the identified performance requirements due to the assumptions related to the representation of feeder impedance and load currents, which also limits its application to LV feeders, and the independence between inputs. Presuming the HBA could be reformulated without the assumptions, the approach was identified as a candidate solution to meet all key requirements for feeders at all voltage levels.

Chapter 3 surveyed the sources of uncertainty in power systems and their characterisation. It identified that load flow studies mainly address variability only in the load demands and DG. There is a need to include models for all types of loads: constant-current and voltage-dependent. The chapter also identified forms of input interdependence that should be included in PLF analysis.

Chapter 4 introduced the original HBA as a method for probabilistic load flow analysis, studying its theoretical foundations and the progression of its developments. Accordingly, the key features and assumptions underlying the approach were identified. Brief tests confirmed the impacts of the key assumptions on accuracy.

Chapter 5 tested the validity of the beta-PDF as a descriptor of the various inputs identified in Chapter 3in the review of the uncertainty elements. By adapting the Beta PDF to input data derived from established models for loads and DG (wind and PV), and using an MCS-based load flow to determine the effect of input model accuracy, it was shown that the beta PDF is suitable for input modelling.

Chapter 6 and 7 form the core of the thesis.

In Chapter 6, the Herman-Beta extended (HBE) transform was introduced. The initial formulations removed the assumptions related to the representation of feeder impedance and load currents and introduced the approach for 3p-3w feeders. The introduction of a new standardization technique based on the per-unit system allowed for the construction of the output distributions in the reformulated transform. The approach was then extended to include dependence through a direct covariance-based approach. Extensions to voltage-dependent inputs allowed for the analysis of systems with constant-power loads, shunt compensation elements (capacitors and reactors) and line shunt admittances. A transform for the computation of branch currents was introduced.

Chapter 7 rigorously tested the developed algorithms and related extensions against the MCS to test the validity of the approaches. Comparison of the performance of the HBE with the initial HBA demonstrated the impact on accuracy the reformulations in the HBE make.

Chapter 8 demonstrated a practical application of the HBE transform using PV-DG penetration as an example. In the application, the HBE is built into an MCS-operator and calculates the bus voltages and branch currents for randomly selected cases of DG placement at different penetration levels. The

approach showed how the HBE can be used as a component in other applications where probabilistic voltage and current calculations are required.

9.2 ANSWERING THE RESEARCH QUESTIONS

This section discusses the answers to the research questions posed in Chapter 1 to help test the validity of the hypothesis.

The research questions can be grouped into four groups; the modelling of the inputs (RQ1, RQ2, RQ3), the derivation of the statistical solutions (RQ1, RQ6), the scope of applications (RQ5, RQ7) and the performance of the proposed approach (RQ4). The research questions are addressed together in the identified groups.

9.2.1 Input Models

The uncertainty related to power system components can be identified according to the sources of origin; network-based, generation-based and load-based. While the consideration of the full range of uncertainties may enhance the accuracy in the power system analysis, it may not be necessary to model accurately all sources of uncertainty. Network-based uncertainties (e.g. the exact lengths of feeder sections) that have a negligible effect on the outcomes of the PLF (e.g. voltages and currents), can be ignored to avoid increasing unnecessarily the computational burden and data requirements.

For each significant source of uncertainty, two key attributes influence the specific representation; the electrical behaviour and the associated variability. The assumption of an electrical model, whether constant-current, power or impedance without practical justification presents room for misguided PLF outcomes. A full PLF approach must consider all the three forms of electrical behaviour.

Several PDFs have been applied to model the variability in the loads and DG. The Log-Normal and Beta PDF are generally preferred for modelling the load demand as they can characterize skewed distributions. The Weibull is most suitable for characterizing wind speed data for wind generation, but the beta PDF is a good alternative. For PV-DG, the beta PDF is generally preferred for modelling solar irradiance. Although sometimes preferred based on simplicity, Gaussian distributions are mainly limited to modelling aggregate loads and DGs.

Apart from standard distributions, mixture models can be applied to both the load demand and DG and, with an optimal number of components, can offer enhanced input model accuracy. However, the use of many parameters and the complexity in the computation of the associated load flow present significant challenges.

The detailed study of the properties of the beta PDF conducted in Chapter 4 revealed the capacity of the distribution to model various distribution shapes. While it is very flexible, the beta PDF, like any other standard distribution or a mixture model with limited components, does not perfectly fit all forms of

data dispersion. This required investigation of the impact of the potential model inaccuracy on PLF outputs.

The findings reported in Chapter 5, demonstrate that while the beta distribution does not always accurately fit the different forms of input data, the differences in the PLF outputs are not always significant. Where the actual input has a dominant mode (with other low-weighted multi-modes) and is not fat-tailed, the errors in assuming a unimodal, standard distribution are insignificant. And, where the inputs have significant multi-modes and fat-tails, the errors depend on the LF sensitivity factors and the conditions of interdependency between inputs. Where the said factors lead to Gaussian outputs by application of the CLT, the errors are also insignificant. The findings confirmed the validity of the beta PDF as a descriptor of the load demand, PV and wind generation inputs. However, the appropriate application of the beta PDF requires that the load data characterized by a model be well separated. In most cases, well-separated data avoids the complexities of dominant multi-modes [181]

Also related to input models, the review of the modelling of various network topologies for distribution and transmission networks motivated the consideration of complex impedance and complex input currents. The sensitivity tests in Chapter 4 demonstrated the loss of accuracy associated with neglecting the imaginary components of the complex inputs.

Accordingly, the HBE transform takes as inputs deterministic values of resistance and reactance lumped per conductor at rated temperature and with temperature coefficient, so that a conductor temperature can be modelled. For feeders with insignificant susceptance such as LV feeders and short MV feeders, a series impedance model is assumed, while a pi-model with shunt admittances is used for long distribution and transmission feeders.

On the other hand, the complex currents from non-unity power factor loads and DGs are modelled as beta currents of the active power at a deterministic power factor. The adopted model assumes that the load or DG at a selected bus has a constant power factor such that the distributions of the active and reactive components of the load are proportional through the deterministic power factor. While this representation appears limiting, the assumption is widely applied and acceptable in the absence of detailed data and established load models [6], [30], [151], [158], [200], [219], [220]. From the findings of this thesis, it is clear that the statistical characterisation of the load power factor is not clearly understood and could be a topic for future research.

9.2.2 Derivation of the PLF Solutions

In Chapter 6 the equations of the original HBA were reformulated to accommodate the imaginary components of the complex variables.

The initial normalisation approach when applied with the reformulated LF equations resulted in substantial errors. The assumption in the initial approach associating the occurrence of extrema voltages

with unique loading conditions is not always valid; it is sensitive to the variation of line X/R ratios and load power factor. To address the identified limitation, a normalization technique based on the per-unit system was devised to ensure the range of voltages fall in the domain [0, 1] compatible with a two-parameter beta distribution. With this normalization, the statistical moments and distribution parameters of the outputs could be calculated.

The computation of the second moments accounts for linear dependence between RVs directly through covariances. Three forms of interdependence were identified and solved; the correlation between inputs connected to the same phase (intra-phase), across phases (inter-phase), and between nodes (inter-node). This completed the formulation of the HBE solution for radial feeders with dependent inputs.

To extend the statistical solutions to feeders with voltage-dependent inputs, an approach based on repeated HBE calculations was proposed. The approach involves calculating the PLF with nominal-voltage-based current equivalents and repeating the calculations with adjusted nodal currents based on the voltages from the previous iteration. With repetitions sufficient to approximate the convergent result, the true impacts of the voltage-dependent models are determined. On the tested feeders, it was demonstrated that a single repeated calculation of the HBE provides a good estimate of the fully converged result. This is expected to be generally applicable to feeders operating within or close to the regulated voltage limits. Where the excessive voltage-drops or rise exist, more iterations may be required.

In addition to the transforms for voltage, an approach for the computation of line currents was developed. The approach is based on the direct summation of the superposition load currents.

9.2.3 Scope of Applications

With the removal of the assumptions in the HBA and extension to voltage-dependent inputs, several new applications are possible.

The reformulation with complex impedance and currents allows more representative analysis of LV systems than previously possible, and the extension to MV systems with loads and DG characterised as constant currents. Further, the extension to voltage-dependent inputs offers new application to compensated feeders with shunt reactors and shunt capacitors, and the analysis of long distribution and transmission feeders (MV and HV). In addition, reactive power control studies involving DGs under various Volt-Var control schemes are made possible. Lastly, systems with interdependent inputs can also be analysed.

Apart from stand-alone applications, the HBE can be used as a component within other applications requiring fast and accurate probabilistic voltage and current calculations. The application to DG penetration studies as demonstrated in Chapter 8 is an example. State estimation is another, as the HBE

transform resolves the inaccuracy issues identified by Brazilian researchers [229] in trying to use the original HB algorithm for the same application.

9.2.4 Performance of the Proposed Approach

In Chapter 7, the performance of the HBE approach was rigorously tested using modified IEEE feeders (33-bus, 34-bus and 69-bus), a representative 12-bus feeder and a practical distribution network. The conducted investigations assessed the validity of the HBE approach for various network conditions including passive and active systems, balanced and unbalanced loading conditions, compensated feeders, feeders with correlated inputs, and long feeders with shunt admittance.

Compared with the MCS, the achieved results demonstrate that the HBE provides an accurate and considerably faster PLF solution. Regarding computational efficiency, the HBE based on constant-current models performs up to 1300 times faster than a 15,000-run MCS. While the extension to voltage-dependant inputs which requires repeated calculations increases the computation time by a factor proportional to the iterations conducted (typically only a factor of 2), the approach remains exceptionally efficient.

The limited accuracy of the HBE when applied to poorly separated data with multi-modes and fat tails has already been discussed. A recommendation for separation of load data using clustering approaches is made.

9.3 THE VALIDITY OF THE RESEARCH HYPOTHESIS AND CONTRIBUTIONS

The research hypothesis underlying this research was:

> "Reformulating the equations in the existing Herman-Beta transform opens many possibilities for accurate feeder calculations"

Based on the answers to the research questions provided in the previous discussion, the research hypothesis is valid.

The powerful HBE is based on characterising loads and DG by Beta PDFs, including interdependency, allows for voltage responsive loads and DG, shunt capacitors and shunt admittances. The new transform applies to 3p-4w and 3p-3w systems at any operating voltage (LV, MV, HV). Though not tested in this thesis, the application is also applicable bi-phase (phase-phase) and single phase (phase-neutral) which are easily modelled on the tested versions of the 3p-3w and 3p-4w systems. Given these features, the criteria for a robust PLF as identified earlier (in Section 9.1) are met.

The research achieved its intended goals and contributes to the body of knowledge in the field of probabilistic load flow analysis. The developed tool allows engineers in power system design and operation to assess the voltage and branch thermal loading conditions on radially-operated distribution and transmission networks under uncertainty, in an efficient and accurate manner. In addition to being

used independently, the PLF tool can be implemented as a transform in other applications, such as the stochastic assessment of DG penetration, optimal power flow programs, and state estimation. Furthermore, the per-unit normalisation technique opens up the opportunity to model in detail composite feeders with sections operating at different voltages.

The contributions made as a result are summarized as follows:

1. An investigation of the impact of input model accuracy on PLF outcomes.
2. Accurate implementation of a beta-PDF-based PLF without restrictions on feeder impedance and load or DG power factor.
3. Extension to higher voltage classes (MV and HV) with short or long feeders.
4. A transform for the analysis of three-phase three-wire systems with delta-connected loads as well as three phase four wire systems with wye-connected loads.
5. Extension to voltage-dependent load and DG models, including feeders with shunt capacitors and shunt reactors.
6. The incorporation of dependency between loads and generation inputs, including the analysis of systems with balanced three-phase loads and DGs (as a special case in which the phase currents are dependent).
7. An approach to the computation of statistical branch currents.

In addition to the contributions on the developed tool, the work identified gaps in the existing literature that require research; the statistical characterisation of the load power factor and load research leading to practical models of the dependency between random variables input.

9.4 FINAL REMARKS

This thesis has presented a new probabilistic approach, the Herman-Beta Extended method, based on the beta PDF. The work builds on decades of work by others in developing the beta-PDF-based PLF tool introduced in 1993. The approach with its new features fulfils the scope of this thesis, but opportunities for further research to address the needs of evolving active systems with energy storage and electric vehicles will challenge engineers in the years to come.

References

[1] J. Ward and H. Hale, “Digital Computer Solution of Power Flow Problems,” *Trans. AIEEE*, vol. 2, no. 7, p. 961, 1956.

[2] B. Stott, “Review of load-flow calculation methods,” *Proc. IEEE*, vol. 62, no. 7, pp. 916–929, 1974.

[3] B. Borkowska, “Probabilistic load flow,” *IEEE Trans. Power Appar. Syst.*, vol. 93, no. 3, pp. 752–759, 1974.

[4] P. Chen, Z. Chen, and B. Bak-Jensen, “Probabilistic load flow: A review,” in *3rd International Conference on Deregulation and Restructuring and Power Technologies, DRPT 2008*, 2008, pp. 1586–1591.

[5] J. A. Martinez and J. Mahseredjian, “Load flow calculations in distribution systems with distributed resources. A review,” in *Power and Energy Society General Meeting, 2011 IEEE*, 2011, pp. 1–8.

[6] J. Usaola, “Probabilistic load flow with correlated wind power injections,” *Electr. Power Energy Syst.*, vol. 80, no. 1, pp. 528–536, 2010.

[7] R. Herman and J. Kritzinger, “The statistical description of grouped domestic electrical load currents,” *Electr. Power Syst. Res.*, vol. Volume 27, no. 1, pp. 43–48, 1993.

[8] Z. M. Salameh, B. S. Borowy, and A. R. A. Amin, “Photovoltaic module-site matching based on the capacity factors,” *IEEE Trans. Energy Convers.*, vol. 10, no. 2, pp. 326–332, Jun. 1995.

[9] S. Bofinger and H. G. Beyer, “Qualification of wind power forecasts,” in *2002 Global Windpower Conferenc*, 2002.

[10] M. J. Chihota, C. T. Gaunt, and R. Herman, “Applying the Herman-Beta probabilistic method to MV feeders,” in *South African Universities Power Engineering Conference (SAUPEC)*, 2016.

[11] M. J. Chihota, “Applying the Herman-Beta probabilistic method to MV feeders,” University of Cape Town, 2015.

[12] A. M. Leita da Silva, R. N. Allan, S. M. Soares, and V. L. Arienti, “Probabilistic load flow considering network outages,” in *IEE Proceedings C-Generation, Transmission and Distribution*, 1985, pp. 139–145.

[13] M. B. D. Coutto Filho, A. M. Leita Da Silva, V. L. Arienti, and S. M. P. Ribeiro, “Probabilistic load modelling for power system expansion planning,” in *1991 Third International Conference on Probabilistic Methods Applied to Electric Power Systems*, 1991, p. 322.

[14] R. N. Allan and M. R. G. Al-Shakarchi, “Probabilistic a.c. load flow,” *Proc. Inst. Electr. Eng.*, vol. 123, no. 6, p. 531, 1976.

[15] R. N. Allan, A. m. Da Silva, and R. C. Burchett, “Evaluation Methods and Accuracy in Probabilistic Load Flow Solutions,” *IEEE Trans. Power Appar. Syst.*, vol. PAS-100, no. 5, pp. 2539–2546, May 1981.

[16] A. Leita Da Silva, V. L. Arienti, and R. N. Allan, “Probabilistic Load Flow Considering Dependence Between Input Nodal Powers,” *IEEE Trans. Power Appar. Syst.*, vol. PAS-103, no. 6, pp. 1524–1530, Jun. 1984.

[17] M. T. Schilling, A. M. Leite da Silva, R. Billinton, and M. A. El-Kady, “Bibliography on power system probabilistic analysis (1962-88),” *IEEE Trans. Power Syst.*, vol. 5, no. 1, pp. 1–11, Feb. 1990.

[18] M. Aien, A. Hajebrahimi, and M. Fotuhi-Firuzabad, “A comprehensive review on uncertainty modeling techniques in power system studies,” *Renew. Sustain. Energy Rev.*, vol. 57, pp. 1077–1089, May 2016.

[19] B. R. Prusty and D. Jena, “A critical review on probabilistic load flow studies in uncertainty constrained power systems with photovoltaic generation and a new approach,” *Renew. Sustain. Energy Rev.*, vol. 69, pp. 1286–1302, Mar. 2017.

[20] A. M. Leite da Silva and V. L. Arienti, “Probabilistic load flow by a multilinear simulation algorithm,” *IEE Proc. C Gener. Transm. Distrib.*, vol. 137, no. 4, pp. 276–282, 1990.

[21] P. Jirutitijaroen and C. Singh, “Comparison of simulation methods for power system reliability indexes and their distributions,” *IEEE Trans. Power Syst.*, vol. 23, no. 2, pp. 486–492, 2008.

[22] H. Yu, C. Y. Chung, K. P. Wong, H. W. Lee, and J. H. Zhang, “Probabilistic load flow evaluation with hybrid latin hypercube sampling and cholesky decomposition,” *IEEE Trans. Power Syst.*, vol. 24, no. 2, pp. 661–667, 2009.

[23] H. Yu and B. Rosehart, “Probabilistic Power Flow Considering Wind Speed Correlation of Wind Farms,” in *17th Power Systems Computation Conference*, 2011, pp. 1–7.

[24] M. Hajian, W. D. Rosehart, and H. Zareipour, “Probabilistic power flow by Monte Carlo simulation with Latin supercube sampling,” *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1550–1559, 2013.

[25] Tao Cui and F. Franchetti, “A Quasi-Monte Carlo approach for radial distribution system probabilistic load flow,” in *2013 IEEE PES Innovative Smart Grid Technologies Conference (ISGT)*, 2013, pp. 1–6.

[26] M. N. Kabir, Y. Mishra, and R. C. Bansal, “Probabilistic load flow for distribution systems with uncertain PV generation,” *Appl. Energy*, vol. 163, no. 1, pp. 343–351, 2016.

[27] Z. Q. Xie, T. Y. Ji, M. S. Li, and Q. H. Wu, “Quasi-Monte Carlo Based Probabilistic Optimal Power Flow Considering the Correlation of Wind Speeds Using Copula Function,” *IEEE Trans. Power Syst.*, vol. 33, no. 2, pp. 2239–2247, Mar. 2018.

[28] M. M. A. Abdelaziz, “OpenCL-Accelerated Probabilistic Power Flow for Active Distribution Networks,” *IEEE Trans. Sustain. Energy*, vol. 9, no. 3, pp. 1255–1264, Jul. 2018.

[29] M. V. F. Pereira, M. E. P. Maceira, G. C. Oliveira, and L. M. V. G. Pinto, “Combining analytical models and monte-carlo techniques in probabilistic power system analysis,” *IEEE Trans. Power Syst.*, vol. 7, no. 1, pp. 265–272, 1992.

[30] W. Wu, K. Wang, G. Li, X. Jiang, and Z. Wang, “Probabilistic load flow calculation using cumulants and multiple integrals,” *IET Gener. Transm. Distrib.*, vol. 10, no. 7, pp. 1703–1709, 2016.

[31] F. J. Ruiz-Rodriguez, J. C. Hernandez, and F. Jurado, “Probabilistic load flow for photovoltaic distributed generation using the Cornish-Fisher expansion,” *Electr. Power Syst. Res.*, vol. 89, no. 1, pp. 129–138, 2012.

[32] E. Janecek and D. Georgiev, “Probabilistic Extension of the Backward/Forward Load Flow Analysis Method,” *IEEE Trans. Power Syst.*, vol. 27, no. 2, pp. 695–704, May 2012.

[33] Z. Jun, Z. Shuang, G. Feng, L. Xutao, L. Hongqiang, W. Zhiwen, and S. Chen, “An analytical method for probabilistic load flow using Gaussian mixture model,” in *2016 IEEE International Conference on Power System Technology (POWERCON)*, 2016, pp. 1–6.

[34] R. N. Allan, A. M. L. Da Silva, A. A. Abu-Nasser, and R. C. Burchett, “Discrete convolution in power system reliability,” *Reliab. IEEE Trans.*, vol. 30, no. 5, pp. 452–456, 1981.

[35] J. Sakamoto, Y. Mori, and T. Sekioka, “Probability analysis method using Fast Fourier transform and its application,” *Struct. Saf.*, vol. 19, no. 1, pp. 21–36, Jan. 1997.

[36] Y. Wang, N. Zhang, C. Kang, M. Miao, R. Shi, and Q. Xia, “An Efficient Approach to Power System Uncertainty Analysis With High-Dimensional Dependencies,” *IEEE Trans. Power Syst.*, vol. 33, no. 3, pp. 2984–2994, May 2018.

[37] T. Sanabria, LA and Dillon, “Stochastic power flow using cumulants and Von Mises functions,” *Int. J. Electr. Power Energy Syst.*, vol. 8, no. 1, pp. 47–60, Jan. 1986.

[38] L. A. Sanabria and T. S. Dillon, “Stochastic power flow using cumulants and Von Mises functions,” *Int. J. Electr. Power Energy Syst.*, vol. 8, no. 1, pp. 47–60, Jan. 1986.

[39] P. Zhang and S. Lee, “Probabilistic Load Flow using the method of combined cumulants and Gram-Charlier expansion,” *IEEE Trans. Power Syst.*, vol. 19, no. 1, pp. 676–682, 2004.

[40] Z. Hu and X. Wang, “A Probabilistic Load Flow Method Considering Branch Outages,” *IEEE Trans. POWER Syst.*, vol. 21, no. 2, p. 507, 2006.

[41] Julio Usaola, “Probabilistic load flow with wind production uncertainty using cumulants and Cornish-Fisher expansion,” *Int. J. Electr. Power Energy Syst.*, vol. 31, no. 9, pp. 474–481, Oct. 2009.

[42] J. Usaola, “Probabilistic load flow in systems with wind generation,” *IET Gener. Transm. Distrib.*, vol. 3, no. 12, pp. 1031–1041, 2009.

[43] Y. Yuan, J. Zhou, P. Ju, and J. Feuchtwang, “Probabilistic load flow computation of a power system containing wind farms using the method of combined cumulants and Gram–Charlier expansion,” *IET Renew. Power Gener.*, vol. 5, no. 6, pp. 448–454, 2011.

[44] M. Fan, V. Vittal, G. T. Heydt, and R. Ayyanar, “Probabilistic Power Flow Studies for Transmission Systems With Photovoltaic Generation Using Cumulants,” *IEEE Trans. Power Syst.*, vol. 27, no. 4, pp. 2251–2261, Nov. 2012.

[45] D. D. Le, K. V Pham, D. V Ngo, K. V Huynh, N. T. A. Nguyen, and A. Berizzi, “An enhancement to cumulant-based probabilistic power flow methodologies,” in *Innovative Smart Grid Technologies-Asia (ISGT ASIA), 2015 IEEE*, 2015, pp. 1–5.

[46] T. Williams and C. Crawford, “Probabilistic Load Flow Modeling Comparing Maximum Entropy and Gram-Charlier Probability Density Function Reconstructions,” *IEEE Trans. Power Syst.*, vol. 28, no. 1, pp. 272–280, Feb. 2013.

[47] M. Tourandaz Kenari, M. S. Sepasian, M. Setayesh Nazar, and H. A. Mohammadpour, “Combined cumulants and Laplace transform method for probabilistic load flow analysis,” *IET Gener. Transm. Distrib.*, vol. 11, no. 14, pp. 3548–3556, 2017.

[48] M. Madrigal, K. Ponnambalam, and V. H. Quintana, “Probabilistic optimal power flow,” in *Conference Proceedings. IEEE Canadian Conference on Electrical and Computer Engineering (Cat. No.98TH8341)*, 1998, vol. 1, pp. 385–388.

[49] C. Wan, Z. Xu, Z. Y. Dong, and K. P. Wong, “Probabilistic load flow computation using first-order second-moment method,” *IEEE Power Energy Soc. Gen. Meet.*, pp. 1–6, 2012.

[50] R. Herman, “Voltage Regulation Analysis for the Design of Low Voltage Networks Feeding Stochastic Domestic Electrical Loads,” University of Stellenbosch, South Africa, 1993.

[51] G. Celli, S. Mocci, F. Pilo, and R. Cicoria, “Probabilistic optimization of MV distribution network in presence of distributed generation,” in *Proc. PSCC Conf*, 2002, pp. 24–28.

[52] E. Rosenblueth, “Point Estimates for Probability Moments,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 72. National Academy of Sciences, pp. 3812–3814, 1975.

[53] M. E. Harr, “Probabilistic estimates for multivariate analyses,” *Appl. Math. Model.*, vol. 13, no. 5, pp. 313–318, 1989.

[54] C. Chen, W. Wu, B. Zhang, and H. Sun, “Correlated probabilistic load flow using a point estimate method with Nataf transformation,” *Int. J. Electr. Power Energy Syst.*, vol. 65, pp. 325–333, Feb. 2015.

[55] M. Mohammadi, A. Shayegani, and H. Adaminejad, “A new approach of point estimate method for probabilistic load flow,” *Int. J. Electr. Power Energy Syst.*, vol. 51, pp. 54–60, Oct. 2013.

[56] K. S. Li, “Point-Estimate Method for Calculating Statistical Moments,” *J. Eng. Mech.*, vol. 118, no. 7, pp. 1506–1511, Jul. 1992.

[57] H. P. Hong, “An efficient point estimate method for probabilistic analysis,” *Reliab. Eng. Syst. Saf.*, vol. 59, no. 3, pp. 261–267, Mar. 1998.

[58] E. Rosenblueth, “Two-point estimates in probabilities,” *Appl. Math. Model.*, vol. 5, no. 5, pp. 329–335, 1981.

[59] C.-L. Su, “Probabilistic load-flow computation using point estimate method,” *Power Syst. IEEE Trans.*, vol. 20, no. 4, pp. 1843–1851, 2005.

[60] J. M. Morales and J. Perez-Ruiz, “Point estimate schemes to solve the probabilistic power flow,” *Power Syst. IEEE Trans.*, vol. 22, no. 4, pp. 1594–1601, 2007.

[61] P. Caramia, G. Carpinelli, and P. Varilone, “Point estimate schemes for probabilistic three-phase load flow,” *Electr. Power Syst. Res.*, vol. 80, no. 2, pp. 168–175, Feb. 2010.

[62] 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,” *Int. J. Electr. Power Energy Syst.*, vol. 61, pp. 267–278, 2014.

[63] N. Gupta and N. Daratha, “Probabilistic three-phase load flow for unbalanced electrical systems with wind farms,” *Int. J. Electr. Power Energy Syst.*, vol. 87, pp. 154–165, May 2017.

[64] H. Yang and B. Zou, “The Point Estimate Method Using Third-Order Polynomial Normal Transformation Technique to Solve Probabilistic Power Flow with Correlated Wind Source and Load,” in *2012 Asia-Pacific Power and Energy Engineering Conference*, 2012, pp. 1–4.

[65] R. R. Appino, T. Muhlpfordt, T. Faulwasser, and V. Hagenmeyer, “On solving probabilistic load flow for radial grids using polynomial chaos,” in *2017 IEEE Manchester PowerTech*, 2017, pp. 1–6.

[66] Z. Ren, W. Li, R. Billinton, and W. Yan, “Probabilistic Power Flow Analysis Based on the Stochastic Response Surface Method,” *IEEE Trans. Power Syst.*, vol. 31, no. 3, pp. 2307–2315, May 2016.

[67] O. A. Oke, D. W. P. Thomas, G. M. Asher, and L. R. A. X. de Menezes, “Probabilistic load flow for distribution systems with wind production using Unscented Transform method,” in *ISGT 2011*, 2011, pp. 1–7.

[68] M. Aien, M. Fotuhi-Firuzabad, and F. Aminifar, “Probabilistic load flow in correlated uncertain environment using unscented transformation,” *IEEE Trans. Power Syst.*, vol. 27, no. 4, pp. 2233–2241, 2012.

[69] H. R. Baghaee, M. Mirsalim, G. B. Gharehpetian, and H. A. Talebi, “Application of RBF neural networks and unscented transformation in probabilistic power-flow of microgrids including correlated wind/PV units and plug-in hybrid electric vehicles,” *Simul. Model. Pract. Theory*, vol. 72, pp. 51–68, Mar. 2017.

[70] J. Canon and S. Jafarzadeh, “Improved probabilistic load flow using unscented transformations,” *Power Symp. (NAPS), 2017 North Am.*, pp. 1–6, 2017.

[71] Y.-Y. Y. Hong, F.-J. J. Lin, and T.-H. H. Yu, “Taguchi method-based probabilistic load flow studies considering uncertain renewables and loads,” *IET Renew. Power Gener.*, vol. 10, no. 2, pp. 221–227, Feb. 2016.

[72] N. Gupta, “Gauss-Quadrature-Based Probabilistic Load Flow Method with Voltage-Dependent Loads Including WTGS, PV, and EV Charging Uncertainties,” *IEEE Trans. Ind. Appl.*, vol. 54, no. 6, pp. 6485–6497, 2018.

[73] M. Mohammadi, H. Basirat, and A. Kargarian, “Nonparametric Probabilistic Load Flow With Saddle Point Approximation,” *IEEE Trans. Smart Grid*, vol. 9, no. 5, pp. 4796–4804, Sep. 2018.

[74] T. Muhlpfordt, T. Faulwasser, and V. Hagenmeyer, “Solving stochastic AC power flow via polynomial chaos expansion,” in *2016 IEEE Conference on Control Applications (CCA)*, 2016, pp. 70–76.

[75] V. Miranda and M. A. C. C. Matos, “Distribution system planning with fuzzy models and techniques,” in *10th International Conference on Electricity Distribution*, 1989.

[76] J. Saraiva, V. Miranda, and M. Matos, “Fuzzy load flow-new algorithms incorporating uncertain generation and load representation,” 1990.

[77] J. T. Saraiva, V. Miranda, and M. A. C. C. Matos, “Generation and load uncertainties incorporated in load flow studies,” in *[1991 Proceedings] 6th Mediterranean Electrotechnical Conference*, 1991, pp. 1339–1342.

[78] V. Miranda and J. P. Saraiva, “Fuzzy modelling of power system optimal load flow,” *IEEE Trans. Power Syst.*, vol. 7, no. 2, pp. 843–849, May 1992.

[79] M. Pourahmadi-Nakhli, A. R. Seifi, and R. Taghavi, “A nonlinear-hybrid fuzzy/probabilistic load flow for radial distribution systems,” *Int. J. Electr. Power Energy Syst.*, vol. 47, pp. 69–77, May 2013.

[80] Y. Zhang and Z. C. Chen, “The calculation of fuzzy load flow in electric power planning,” in *Automation of Electric Power System*, 1998, pp. 20–22.

[81] M. A. Matos and E. M. Gouveia, “The Fuzzy Power Flow Revisited,” *IEEE Trans. Power Syst.*, vol. 23, no. 1, pp. 213–218, Feb. 2008.

[82] L. Hong, L. Shi, L. Yao, Y. Ni, and M. Bazargan, “Study on fuzzy load flow with consideration of wind generation uncertainties,” in *2009 Transmission & Distribution Conference & Exposition: Asia and Pacific*, 2009, pp. 1–4.

[83] I. S. Arneja and B. Venkatesh, “Probabilistic OPF using linear fuzzy relation,” in *2012 10th International Power & Energy Conference (IPEC)*, 2012, pp. 601–605.

[84] M. Aien, M. Rashidinejad, and M. Fotuhi-Firuzabad, “On possibilistic and probabilistic uncertainty assessment of power flow problem: A review and a new approach,” *Renew. Sustain. Energy Rev.*, vol. 37, pp. 883–895, Sep. 2014.

[85] M. Heleno, J. Sumaili, J. Meirinhos, and M. A. da Rosa, “A linearized approach to the Symmetric Fuzzy Power Flow for the application to real systems,” *Int. J. Electr. Power Energy Syst.*, vol. 54, pp. 610–618, Jan. 2014.

[86] A. Soroudi and M. Ehsan, “A possibilistic–probabilistic tool for evaluating the impact of stochastic renewable and controllable power generation on energy losses in distribution networks—A case study,” *Renew. Sustain. Energy Rev.*, vol. 15, no. 1, pp. 794–800, Jan. 2011.

[87] N. Nikmehr and S. Najafi Ravadanegh, “Heuristic probabilistic power flow algorithm for microgrids operation and planning,” *IET Gener. Transm. Distrib.*, vol. 9, no. 11, pp. 985–995, Aug. 2015.

[88] M. T. Hagh, P. Amiyan, S. Galvani, and N. Valizadeh, “Probabilistic load flow using the particle swarm optimisation clustering method,” *IET Gener. Transm. Distrib.*, vol. 12, no. 3, pp. 780–789, Feb. 2018.

[89] NRS, “NRS 048-2: Electricity Supply - Quality of Supply Part 2 : Voltage characteristics , compatibility levels , limits and assessment methods,” 2003.

[90] IEC, *IEC 600038, IEC standard voltages 1999*. 1999.

[91] H. Kadada, “Designing Low Voltage feeders to meet Quality of Supply specifications for voltage variations,” University of Cape Town, 2012.

[92] R. A. Shayani and M. A. G. de Oliveira, “Photovoltaic Generation Penetration Limits in Radial Distribution Systems,” *IEEE Trans. Power Syst.*, vol. 26, no. 3, pp. 1625–1631, Aug. 2011.

[93] A. Schellenberg, W. Rosehart, and J. Aguado, “Cumulant-Based Probabilistic Optimal Power Flow (P-OPF) With Gaussian and Gamma Distributions,” *IEEE Trans. POWER Syst.*, vol. 20, no. 2, p. 773, 2005.

[94] NRS, “NRS 034-1:2007 Electricity distribution — Guidelines for the provision of electricity distribution networks in residential areas Part 1 : Planning and design of distribution networks,” 1999.

[95] J. D. Glover, M. S. Sarma, and T. J. Overbye, *Power System Analysis and Design*, 5th ed. United States of America: Global Engineering: Christopher M. Shortt, 2010.

[96] L. Hülsmann, “Evaluation of two distribution grids in terms of PV penetration limits and effectiveness of reactive power controls,” KTH Royal Institute of Technology, 2016.

[97] T. Short, *Electric power distribution equipment and systems*. CRC Press, 2018.

[98] A. A. Sallam, M. Desouky, and H. Desouky, “Shunt capacitor effect on electrical distribution system reliability,” *IEEE Trans. Reliab.*, vol. 43, no. 1, pp. 170–176, Mar. 1994.

[99] H. Asano, K. Yajima, and Y. Kaya, “Influence of photovoltaic power generation on required capacity for load frequency control,” *IEEE Trans. Energy Convers.*, 1996.

[100] Y. Gao and R. Billinton, “Adequacy assessment of generating systems containing wind power considering wind speed correlation,” *IET Renew. Power Gener.*, vol. 3, no. 2, p. 217, 2009.

[101] M. D. Galus, M. Zima, and G. Andersson, “On integration of plug-in hybrid electric vehicles into existing power system structures,” *Energy Policy*, vol. 38, no. 11, pp. 6736–6745, Nov. 2010.

[102] K. C. Divya and J. Østergaard, “Battery energy storage technology for power systems—An overview,” *Electr. Power Syst. Res.*, vol. 79, no. 4, pp. 511–520, Apr. 2009.

[103] F. Ahjum, B. Merven, A. Stone, and T. Caetano, “Road transport vehicles in South Africa towards 2050: Factors influencing technology choice and implications for fuel supply,” *J. Energy South. Africa*, vol. 29, no. 3, pp. 33–50, Sep. 2018.

[104] NRS 097-2-3:2017 Edition 1, “Grid interconnection of embedded generation,” 2016.

[105] N. Wight, S. Alahakoon, and P. Pledger, “Voltage drop and unbalance compensation in long distance medium voltage distribution lines a feasibility study,” in *2015 IEEE 10th International Conference on Industrial and Information Systems (ICIIS)*, 2015, pp. 1–6.

[106] Z. Liu and J. V. Milanovic, “Probabilistic Estimation of Voltage Unbalance in MV Distribution Networks With Unbalanced Load,” *IEEE Trans. Power Deliv.*, vol. 30, no. 2, pp. 693–703, Apr. 2015.

[107] P. V. Santos Valois, “Voltage unbalance in low voltage distribution networks,” in *16th International Conference and Exhibition on Electricity Distribution (CIRED 2001)*, 2001, vol. 2001, pp. v2-41-v2-41.

[108] “Bibliography on load models for power flow and dynamic performance simulation,” *IEEE Trans. Power Syst.*, vol. 10, no. 1, pp. 523–538, 1995.

[109] M. H. Haque, “Load flow solution of distribution systems with voltage dependent load models,” *Electr. Power Syst. Res.*, vol. 36, no. 3, pp. 151–156, Mar. 1996.

[110] U. Eminoglu and M. Hakan Hocaoglu, “A new power flow method for radial distribution systems including voltage dependent load models,” *Electr. Power Syst. Res.*, vol. 76, pp. 106–114, 2005.

[111] D. Singh, R. K. Misra, and D. Singh, “Effect of Load Models in Distributed Generation Planning,” *IEEE Trans. Power Syst.*, vol. 22, no. 4, pp. 2204–2212, Nov. 2007.

[112] O. R. Schurig, “A miniature A-C. transmission system: For the practical solution of network and transmission-system problems,” *J. Am. Inst. Electr. Eng.*, vol. 42, no. 10, pp. 1033–1040, Oct. 1923.

[113] G. J. Berg, “Power-system load representation,” *Proc. Inst. Electr. Eng.*, vol. 120, no. 3, p. 344, 1973.

[114] “Load representation for dynamic performance analysis (of power systems),” *IEEE Trans. Power Syst.*, vol. 8, no. 2, pp. 472–482, May 1993.

[115] Y. Li, H.-D. Chiang, B.-K. Choi, Y.-T. Chen, D.-H. Huang, and M. G. Lauby, “Representative static load models for transient stability analysis: development and examination,” *IET Gener. Transm. Distrib.*, vol. 1, no. 3, p. 422, 2007.

[116] A. Bokhari, S. Member, A. Alkan, R. Dogan, M. Diaz-Aguiló, F. de León, S. Member, D. Czarkowski, Z. Zabar, L. Birenbaum, A. Noel, and R. Ebrahem Uosef, “Experimental Determination of the ZIP Coefficients for Modern Residential, Commercial, and Industrial Loads,” *IEEE Trans. POWER Deliv.*, vol. 29, no. 3, 2014.

[117] C. T. Gaunt, R. Herman, M. Dekenah, R. L. Sellick, and S. W. Heunis, “Data collection, load modelling and probabilistic analysis for LV domestic electrification,” in *International Conference on Electricity Distribution (CIRED)*, 1999.

[118] T. . Basu, R. . Misra, and P. Patowary, “A new approach to probabilistic load flow,” in *National Power Systems Conference, NPC 2002*, 2002, pp. 837–843.

[119] C. J. Roy and W. L. Oberkampf, “A comprehensive framework for verification, validation, and uncertainty quantification in scientific computing,” *Comput. Methods Appl. Mech. Eng.*, vol. 200, no. 25–28, pp. 2131–2144, Jun. 2011.

[120] J. V Milanović, “Probabilistic stability analysis: the way forward for stability analysis of sustainable power systems.,” *Philos. Trans. A. Math. Phys. Eng. Sci.*, vol. 375, no. 2100, p. 20160296, Aug. 2017.

[121] L. Min and P. Zhang, “A Probabilistic Load Flow with Consideration of Network Topology Uncertainties,” in *2007 International Conference on Intelligent Systems Applications to Power Systems*, 2007, pp. 1–5.

[122] S. K. E. Awadallah, J. V. Milanovic, P. N. Jarman, and Z. Wang, “Probabilistic Indicators for Assessing Age- and Loading-Based Criticality of Transformers to Cascading Failure Events,” *IEEE Trans. Power Syst.*, vol. 29, no. 5, pp. 2558–2566, Sep. 2014.

[123] Xiu-Ren Lei, Zhen Ren, Wen-Ying Huang, and Bi-Yun Chen, “Fuzzy reliability analysis of distribution systems accounting for parameters uncertainty,” in *2005 International Conference on Machine Learning and Cybernetics*, 2005, p. 4017–4022 Vol. 7.

[124] L. Dong, C.-C. Zhang, Y.-H. Yang, and P. Zhang, “Improvement of Probabilistic Load Flow to Consider Network Configuration Uncertainties,” in *2009 Asia-Pacific Power and Energy Engineering Conference*, 2009, pp. 1–5.

[125] B. Das, “Radial distribution system power flow using interval arithmetic,” *Int. J. Electr. Power Energy Syst.*, vol. 24, no. 10, pp. 827–836, Dec. 2002.

[126] L. Sige, Z. Xiaoxin, F. Mingtian, and Z. Zhuping, “Probabilistic Power Flow Calculation Using Sigma-Point Transform Algorithm,” in *2006 International Conference on Power System Technology*, 2006, pp. 1–5.

[127] A. Seppälä, “Load research and load estimation in electricity distribution,” *VTT Publ.*, no. 289, pp. 3–118, 1996.

[128] M. Meldorf, T. Täht, and J. Kilter, “Stochasticity of the electrical network load,” *Oil Shale*, vol. 24, no. 2 SUPPL., pp. 225–236, 2007.

[129] J. Dickert and P. Schegner, “Residential Load Models for Network Planning Purposes,” 2010.

[130] J. Munkhammar, J. Rydén, and J. Widén, “Characterizing probability density distributions for household electricity load profiles from high-resolution electricity use data,” *Appl. Energy*, vol. 135, pp. 382–390, 2014.

[131] L. M. Korunovic, S. Sterpu, S. Djokic, K. Yamashita, S. M. Villanueva, and J. V. Milanovic, “Processing of load parameters based on Existing Load Models,” in *2012 3rd IEEE PES Innovative Smart Grid Technologies Europe (ISGT Europe)*, 2012, pp. 1–6.

[132] C. T. Gaunt, E. Namanya, and R. Herman, “Voltage modelling of LV feeders with dispersed generation: Limits of penetration of randomly connected photovoltaic generation,” *Electr. Power Syst. Res.*, vol. 143, pp. 1–6, Feb. 2017.

[133] A. Seppãlã, “Statistical distribution of customer load profiles,” in *IEEE Int. Conf. Energy Management and Power Delivery*, 1995, vol. 2, pp. 696–701.

[134] J. Nazarko and Z. A. Styczynski, “Application of statistical and neural approaches to the daily load profiles modelling in power distribution systems,” in *1999 IEEE Transmission and Distribution Conference (Cat. No. 99CH36333)*, 1999, pp. 320–325 vol.1.

[135] W. Labeeuw and G. Deconinck, “Customer sampling in a smart grid pilot,” in *2012 IEEE Power and Energy Society General Meeting*, 2012, pp. 1–7.

[136] B. Stephen, A. J. Mutanen, S. Galloway, G. Burt, and P. Jarventausta, “Enhanced Load Profiling for Residential Network Customers,” *IEEE Trans. Power Deliv.*, vol. 29, no. 1, pp. 88–96, Feb. 2014.

[137] M. Hayn, V. Bertsch, and W. Fichtner, “Electricity load profiles in Europe: The importance of household segmentation,” *Energy Res. Soc. Sci.*, vol. 3, no. C, 2014.

[138] Y. Wang, L. Li, and Q. Yang, “Application of clustering technique to electricity customer classification for load forecasting,” in *Information and Automation, 2015 IEEE International Conference on*, 2015, pp. 1425–1430.

[139] J.-F. J.-F. Toubeau, M. Hupez, V. Klonari, Z. De Greve, F. Vallee, Z. De Grève, and F. Vallée, “Statistical load and generation modelling for long term studies of low voltage networks in presence of sparse smart metering data,” in *Industrial Electronics Society, IECON 2016-42nd Annual Conference of the IEEE*, 2016, pp. 3900–3905.

[140] R. R. Rathod and R. D. Garg, “Regional electricity consumption analysis for consumers using data mining techniques and consumer meter reading data,” *Int. J. Electr. Power Energy Syst.*, vol. 78, 2016.

[141] J. V Paatero and P. D. Lund, “A model for generating household electricity load profiles,” *Int. J. Energy Res.*, vol. 30, no. 5, pp. 273–290, 2006.

[142] R. Herman and C. T. Gaunt, “A Practical Probabilistic Design Procedure for LV Residential Distribution Systems,” *IEEE Trans. Power Deliv.*, vol. 23, pp. 2247–2254, 2008.

[143] C. L. Masters, “Voltage rise the big issue when connecting emmbedded generation to long 11kV overhead lines,” *Power Eng. J.*, vol. 16, no. 1, pp. 5–12, 2002.

[144] S. Repo, H. Laaksonen, and P. Jarventausta, “Statistical Models of Distributed Generation for Distribution Network,” in *CIRED 18th International Conference and Exibition on Electricity Distribution (CIRED)*, 2005, no. June, pp. 6–9.

[145] W. El-Khattam, Y. G. Hegazy, S. Member, and M. M. a Salama, “Investigating Distributed Generation Systems Performance Using Monte Carlo Simulation,” *IEEE Trans. Power Syst.*, vol. 21, no. 2, pp. 524–532, 2006.

[146] C. L. Su, “Stochastic evaluation of voltages in distribution networks with distributed generation using detailed distribution operation models,” *IEEE Trans. Power Syst.*, vol. 25, no. 2, pp. 786–795, 2010.

[147] M. Hasheminamin, V. G. Agelidis, and A. Heidari, “Impact study of high PV penetration in low and medium-voltage networks when considering residential and industrial load profile,” in *Renewable Energy Research and Applications (ICRERA), 2013 International Conference*, 2013, no. 1, pp. 347–352.

[148] K. Zou, A. P. Agalgaonkar, K. M. Muttaqi, and S. Perera, “Distribution System Planning With Incorporating DG Reactive Capability and System Uncertainties,” *IEEE Trans. Sustain. Energy*, vol. 3, no. 1, pp. 112–123, Jan. 2012.

[149] A. Soroudi, M. Aien, and M. Ehsan, “A Probabilistic Modeling of Photo Voltaic Modules and Wind Power Generation Impact on Distribution Networks,” *IEEE Syst. J.*, vol. 6, no. 2, pp. 254–259, Jun. 2012.

[150] S. S. Soman, H. Zareipour, O. Malik, and P. Mandal, “A review of wind power and wind speed forecasting methods with different time horizons,” in *North American Power Symposium 2010*, 2010, pp. 1–8.

[151] G. Carpinelli, P. Caramia, and P. Varilone, “Multi-linear Monte Carlo simulation method for probabilistic load fl ow of distribution systems with wind and photovoltaic generation systems,” *Renew. Energy*, vol. 76, pp. 283–295, 2015.

[152] B. Kroposki, K. Emery, D. Myers, and L. Mrig, “A comparison of photovoltaic module performance evaluation methodologies for energy ratings,” in *Proceedings of 1994 IEEE 1st World Conference on Photovoltaic Energy Conversion - WCPEC (A Joint Conference of PVSC, PVSEC and PSEC)*, vol. 1, pp. 858–862.

[153] G. Tina, S. Gagliano, and S. Raiti, “Hybrid solar/wind power system probabilistic modelling for long-term performance assessment,” 2005.

[154] S. Conti and S. Raiti, “Probabilistic load flow using Monte Carlo techniques for distribution networks with photovoltaic generators,” *Sol. Energy*, vol. 81, no. 12, pp. 1473–1481, Dec. 2007.

[155] R. Preece and J. V. Milanovic, “Assessing the Applicability of Uncertainty Importance Measures for Power System Studies,” *IEEE Trans. Power Syst.*, vol. 31, no. 3, pp. 2076–2084, May 2016.

[156] A. Vaccaro, C. A. Canizares, and D. Villacci, “An Affine Arithmetic-Based Methodology for Reliable Power Flow Analysis in the Presence of Data Uncertainty,” *IEEE Trans. Power Syst.*, vol. 25, no. 2, pp. 624–632, May 2010.

[157] F. Alvarado, Y. Hu, and R. Adapa, “Uncertainty in power system modeling and computation,” in *[EEE International Conference on Systems, Man, and Cybernetics*, 1992, pp. 754–760.

[158] N. Hatziargyriou, “Probabilistic Load Flow in Distribution Systems Containing Dispersed Wind Power Generation,” *Power Syst. …*, vol. 8, no. 1, pp. 159–165, 1993.

[159] B. R. Prusty and D. Jena, “Modeling of correlated photovoltaic generations and load demands in probabilistic load flow,” in *2015 Annual IEEE India Conference (INDICON)*, 2015, pp. 1–6.

[160] A. Augusto, A. de Souza, N. Kagan, K. de Geus, and M. E. M. Udaeta, “Impact of Distributed Generation on the Operational Planning of Medium Voltage Distribution Networks Using Genetic Algorithms,” in *Innovative Smart Grid Technologies Conference Europe (ISGT-Europe), 2014 IEEE PES. IEEE*, 2014, vol. 5, pp. 1–5.

[161] Y. Chen, J. Wen, and S. Cheng, “Probabilistic load flow method based on Nataf transformation and Latin hypercube sampling,” *IEEE Trans. Sustain. Energy*, vol. 4, no. 2, pp. 294–301, 2013.

[162] W. Gu, L. Luo, T. Ding, X. Meng, and W. Sheng, “An affine arithmetic-based algorithm for radial distribution system power flow with uncertainties,” *Int. J. Electr. POWER ENERGY Syst.*, vol. 58, pp. 242–245, 2014.

[163] R. Bernards, J. Morren, and H. Slootweg, “Statistical modelling of load profiles incorporating correlations using Copula,” in *2017 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*, 2017, pp. 1–6.

[164] K. Srinivasan and C. Lafond, “Statistical analysis of load behavior parameters at four major loads,” *IEEE Trans. Power Syst.*, vol. 10, no. 1, pp. 387–392, 1995.

[165] A. K. Ghosh, D. L. Lubkeman, M. J. Downey, and R. H. Jones, “Distribution circuit state estimation using a probabilistic approach,” *IEEE Trans. Power Syst.*, vol. 12, no. 1, pp. 45–51, 1997.

[166] V. Neimane, “Distribution network planning based on statistical load modelingapplying genetic algorithms and Monte-Carlo simulations,” *2001 IEEE Porto Power Tech Proc. (Cat. No.01EX502)*, vol. 3, 2001.

[167] R. Singh, B. C. Pal, and R. Jabr, “Statistical representation of distribution system loads using Gaussian Mixture Model,” *IEEE Trans. Power Syst.*, vol. 25, no. 1, pp. 29–37, 2010.

[168] A. Ahmad, I. Azmira, S. Ahmad, and Nur Ilyana Anwar Apandi, “Statistical distributions of load profiling data,” in *2012 IEEE International Power Engineering and Optimization Conference*, 2012, pp. 199–203.

[169] R. N. Allan and M. R. G. Al-Shakarchi, “Linear dependence between nodal powers in probabilistic a.c. load flow,” *Proc. Inst. Electr. Eng.*, vol. 124, no. 6, p. 529, 1977.

[170] H. R. Sirisena and E. P. M. Brown, “Representation of non-Gaussian probability distributions in stochastic load-flow studies by the method of Gaussian sum approximations,” *IEE Proc. C Gener. Transm. Distrib.*, vol. 130, no. 4, p. 165, 1983.

[171] G. W. W. Irwin, W. Monteith, and W. C. C. Beattie, “Statistical electricity demand modelling from consumer billing data,” in *IEE Proceedings C-Generation, Transmission and Distribution*, 1986, vol. 133, no. 6, pp. 328–335.

[172] A. T. Al-Awami and M. . El-Sharkawi, “Statistical characterization of wind power output for a given wind power forecast,” in *North American Power Symposium (NAPS)*, 2009, pp. 1–4.

[173] H. Louie, “Evaluation of probabilistic models of wind plant power output characteristics,” in *Probabilistic Methods Applied to Power Systems (PMAPS), 2010 IEEE 11th International Conference*, 2010, pp. 442–447.

[174] C. G. Justus, W. R. Hargraves, A. Mikhail, D. Graber, C. G. Justus, W. R. Hargraves, A. Mikhail, and D. Graber, “Methods for Estimating Wind Speed Frequency Distributions,” *J. Appl. Meteorol.*, vol. 17, no. 3, pp. 350–353, Mar. 1978.

[175] V. A. Graham and K. G. T. Hollands, “A method to generate synthetic hourly solar radiation globally,” *Sol. Energy*, vol. 44, no. 6, pp. 333–341, Jan. 1990.

[176] F. Youcef Ettoumi, A. Mefti, A. Adane, and M. . Bouroubi, “Statistical analysis of solar measurements in Algeria using beta distributions,” *Renew. Energy*, vol. 26, no. 1, pp. 47–67, May 2002.

[177] S. S. Parihar and N. Malik, “Interval arithmetic power flow analysis of radial distribution system including uncertainties in input parameters,” in *2017 7th International Conference on Power Systems (ICPS)*, 2017, pp. 69–74.

[178] Y. Huang, Q. Xu, X. Jiang, T. Zhang, and Y. Yang, “Modelling Correlated Forecast Error for Wind Power in Probabilistic Load Flow,” *Elektron. ir Elektrotechnika*, vol. 23, no. 5, pp. 61–66, Oct. 2017.

[179] Q. Xiao, “Modeling Uncertainties in Power System by Generalized Lambda Distribution,” *Int. J. Emerg. Electr. Power Syst.*, vol. 15, no. 3, pp. 195–203, Jan. 2014.

[180] E. Carpaneto and G. Chicco, “Probabilistic characterisation of the aggregated residential load patterns,” in *IET Generation, Transmission & Distribution*, 2007.

[181] G. Chicco, R. Napoli, F. Piglione, P. Postolache, M. Scutariu, and C. Toader, “Application of clustering techniques to load pattern-based electricity customer classification,” in *Electricity Distribution, 2005. CIRED 2005. 18th International Conference and Exhibition on*, 2005, pp. 1–5.

[182] A. Mutanen, M. Ruska, S. Repo, and P. Jarventausta, “Customer classification and load profiling method for distribution systems,” *IEEE Trans. Power Deliv.*, vol. 26, no. 3, pp. 1755–1763, 2011.

[183] M. S. ElNozahy, M. M. A. Salama, and R. Seethapathy, “A probabilistic load modelling approach using clustering algorithms,” in *2013 IEEE Power & Energy Society General Meeting*, 2013, pp. 1–5.

[184] S. Ramos, J. M. Duarte, F. J. Duarte, and Z. Vale, “A data-mining-based methodology to support MV electricity customers’ characterization,” *Energy Build.*, vol. 91, pp. 16–25, 2015.

[185] M. E. Nassar and M. M. A. Salama, “A novel probabilistic load model and probabilistic power flow,” in *Electrical and Computer Engineering (CCECE), 2015 IEEE 28th Canadian Conference on*, 2015, pp. 881–886.

[186] S. Haben, C. Singleton, and P. Grindrod, “Analysis and clustering of residential customers energy behavioral demand using smart meter data,” *IEEE Trans. Smart Grid*, vol. 7, no. 1, pp. 136–144, 2016.

[187] R. N. Allan, C. H. Grigg, D. A. Newey, and R. F. Simmons, “Probabilistic power-flow techniques extended and applied to operational decision making,” *Proc. Inst. Electr. Eng.*, vol. 123, no. 12, p. 1317, 1976.

[188] J. A. Carta, P. Ramírez, and S. Velázquez, “A review of wind speed probability distributions used in wind energy analysis: Case studies in the Canary Islands,” *Renew. Sustain. Energy Rev.*, vol. 13, no. 5, pp. 933–955, Jun. 2009.

[189] A. M. Razali, N. Masseran, K. Ibrahim, A. Zaharim, and K. Sopian, “The Probability Distribution Model of Wind Speed over East Malaysia,” *Res. J. Appl. Sci. Eng. Technol.*, vol. 6, no. 10, pp. 1774–1779, 2013.

[190] M. Anvari, G. Lohmann, M. Wachter, P. Milan, E. Lorenz, D. Heinemann, M. Reza Rahimi Tabar, and Peinke Joachim, “Short term fluctuations of wind and solar power systems,” *New J. Phys.*, vol. 18, no. 6, pp. 1–14, 2016.

[191] H. F. Assuncao, J. F. Escobedo, and A. P. Oliveira, “Modelling frequency distributions of 5 minute-averaged solar radiation indexes using Beta probability functions,” *Theor. Appl. Clim.*, vol. 75, pp. 213–224, 2003.

[192] E. A. Mohamed and Y. G. Hegazy, “A Novel Probablistic Strategy for Modeling Photovoltaic Based Distributed Generators,” *Int. J. Energy Power Eng.*, vol. 9, no. 8, pp. 904–907, 2015.

[193] Y. Lv, L. Guan, Z. Tang, and Q. Zhao, “A Probability Model of PV for the Middle-Term to Long-Term Power system Analysis and Its Application,” *Energy Procedia*, vol. 103, pp. 28–33, 2016.

[194] M. Nijhuis, M. Gibescu, and S. Cobben, “Gaussian Mixture Based Probabilistic Load Flow for LV-Network Planning,” *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 2878–2886, 2017.

[195] G. Valverde, A. T. Saric, and V. Terzija, “Probabilistic load flow with non-Gaussian correlated random variables using Gaussian mixture models,” *IET Gener. Transm. Distrib.*, vol. 6, no. 7, p. 701, 2012.

[196] H. Joe, *Multivariate models and multivariate dependence concepts*. Chapman and Hall/CRC, 1997.

[197] M. Hart and R. de Dear, "Weather sensitivity in household appliance energy end-use," *Energy Build.*, vol. 36, no. 2, pp. 161–174, Feb. 2004.

[198] S. Liu, G. Li, H. Xie, and X. Wang, "Correlation Characteristic Analysis for Wind Speed in Different Geographical Hierarchies," 2017.

[199] G. Papaefthymiou and D. Kurowicka, "Using copulas for modeling stochastic dependence in power system uncertainty analysis," *IEEE Trans. Power Syst.*, vol. 24, no. 1, pp. 40–49, 2009.

[200] C. Carmona-Delgado, E. Romero-Ramos, and J. Riquelme-Santos, "Probabilistic load flow with versatile non-Gaussian power injections," *Electr. Power Syst. Res.*, vol. 119, pp. 266–277, 2015.

[201] Y. Wang, N. Zhang, Q. Chen, J. Yang, C. Kang, and J. Huang, "Dependent Discrete Convolution Based Probabilistic Load Flow for the Active Distribution System," *IEEE Trans. Power Syst.*, 2017.

[202] P. Amid and C. Crawford, "A Cumulant-Tensor-Based Probabilistic Load Flow Method," *IEEE Trans. Power Syst.*, vol. 33, no. 5, pp. 5648–5656, 2018.

[203] X. Vieira Filho, J. W. Marangon Lima, A. M. Leita da Silva, M. T. Schilling, and V. L. Arienti, "Application of probabilistic load flow in operational planning studies: Current status and prospects," in *International Conference on Large High Voltage Electric Systems*, 1991, pp. 1–6.

[204] A. Dimitrovski and R. Ackovski, "Probabilistic load flow in radial distribution networks," in *Proceedings of 1996 Transmission and Distribution Conference IEEE*, 1996, pp. 15–20.

[205] C. Ziser, Z. Dong, T. S.-A. universities power, and undefined 2005, "Investigation of weather dependency and load diversity on queensland electricity demand," in *Australian Universities Power Engineering Conference*, 2005, pp. 457–462.

[206] S. Humeau, T. K. Wijaya, M. Vasirani, and K. Aberer, "Electricity load forecasting for residential customers: Exploiting aggregation and correlation between households," in *2013 Sustainable Internet and ICT for Sustainability, SustainIT 2013*, 2013.

[207] D. Srinivasan and S. Gundam, "Correlation analysis of solar power and electric demand," in *2013 International Conference on Renewable Energy and Sustainable Energy (ICRESE)*, 2013, pp. 221–226.

[208] W. P. Bell, P. Wild, and J. Foster, "Wind speed and electricity demand correlation analysis to determine the ability of wind turbine generators to meet electricity demand in the Australian National Electricity Market," Brisbane, 2015.

[209] P. J. F. Torres, L. Ekonomou, and P. Karampelas, "The Correlation Between Renewable Generation and Electricity Demand: A Case Study of Portugal," 2016, pp. 119–151.

[210] H. E. Thornton, A. A. Scaife, B. J. Hoskins, and D. J. Brayshaw, "The relationship between wind power, electricity demand and winter weather patterns in Great Britain," *Environ. Res. Lett.*, vol. 12, no. 6, p. 064017, Jun. 2017.

[211] J. Ekstrom, M. Koivisto, I. Mellin, R. J. Millar, and M. Lehtonen, "A Statistical Model for Hourly Large-Scale Wind and Photovoltaic Generation in New Locations," *IEEE Trans. Sustain. Energy*, vol. 8, no. 4, pp. 1383–1393, Oct. 2017.

[212] H. Mori and Wenjun Jiang, "A New Probabilistic Load Flow Method Using MCMC in Consideration of Nodal Load Correlation," in *2009 15th International Conference on Intelligent System Applications to Power Systems*, 2009, pp. 1–6.

[213] F. Vallee, J. Lobry, and O. Deblecker, "Impact of the Wind Geographical Correlation Level for Reliability Studies," *IEEE Trans. Power Syst.*, vol. 22, no. 4, pp. 2232–2239, Nov. 2007.

[214] J. M. Morales, L. Baringo, A. J. Conejo, and R. Mínguez, "Probabilistic power flow with correlated wind sources," *IET Gener. Transm. Distrib.*, vol. 4, no. 5, p. 641, 2010.

[215] Z. Qin, W. Li, and X. Xiong, "Generation System Reliability Evaluation Incorporating Correlations of Wind Speeds With Different Distributions," *IEEE Trans. Power Syst.*, vol. 28, no. 1, pp. 551–558, Feb. 2013.

[216] P. Li, X. Guan, J. Wu, and X. Zhou, "Modeling Dynamic Spatial Correlations of Geographically Distributed Wind Farms and Constructing Ellipsoidal Uncertainty Sets for Optimization-Based Generation Scheduling," *IEEE Trans. Sustain. Energy*, vol. 6, no. 4, pp. 1594–1605, Oct. 2015.

[217] C. N. Long, T. P. Ackerman, C. N. Long, and T. P. Ackerman, "Surface Measurements of Solar Irradiance: A Study of the Spatial Correlation between Simultaneous Measurements at Separated Sites," *J. Appl. Meteorol.*, vol. 34, no. 5, pp. 1039–1046, May 1995.

[218] J. Munkhammar and J. Widén, "Correlation modeling of instantaneous solar irradiance with applications to solar engineering," *Sol. Energy*, vol. 133, pp. 14–23, 2016.

[219] G. Stefopoulos, A. P. Meliopoulos, and G. J. Cokkinides, "Advanced Probabilistic Power Flow Methodology," in *15th Power Systems Computational Conference*, 2005, pp. 1–7.

[220] A. Noruzi, T. Banki, O. Abedinia, and N. Ghadimi, "A new method for probabilistic assessments in power systems, combining monte carlo and stochastic-algebraic methods," *Complexity*, vol. 21, no. 2, pp. 100–110, Nov. 2015.

[221] J. Liu, X. Hao, P. Cheng, W. Fang, S. Niu, J. Liu, X. Hao, P. Cheng, W. Fang, and S. Niu, "A Parallel Probabilistic Load Flow Method Considering Nodal Correlations," *Energies*, vol. 9, no. 12, p. 1041, Dec. 2016.

[222] L. Bin, M. Shahzad, Q. Bing, M. Ahsan, M. U. Shoukat, H. M. Khan, and N. A. Fahal, "The probabilistic load flow analysis by considering uncertainty with correlated loads and photovoltaic generation using Copula theory," *AIMS Energy*, vol. 6, no. 3, pp. 414–435, 2018.

[223] B. Su, K. Hou, H. Jia, Y. Mu, and X. Yu, "Maximum entropy based probabilistic load flow calculation for power system integrated with wind power generation," *J. Mod. Power Syst. Clean Energy*, vol. 6, no. 5, pp. 1042–1054, Sep. 2018.

[224] P. Chen, B. Bak-Jensen, and Z. Chen, "Probabilistic load models for simulating the impact of load management," in *2009 IEEE Power & Energy Society General Meeting*, 2009, pp. 1–8.

[225] R. Herman and C. T. Gaunt, "Measurement and representation of maximum demand for individual and grouped domestic consumers including constraints," in *11th International Conference on Electricity Distribution (CIRED) Liege, Belgium*, 1991, vol. 6, pp. 6–14.

[226] C. T. Gaunt, R. Herman, E. Namanya, and M. J. Chihota, "Voltage modelling of LV feeders with dispersed generation: Probabilistic analytical approach using the beta pdf," *Elsevier trans. Electr. Power Syst.*, vol. 143, pp. 25–31, 2017.

[227] R. Herman and S. W. Heunis, "General probabilistic voltage drop calculation method for LV distribution networks based on a beta p.d.f. load model," *Electr. Power Syst. Res.*, vol. 46, no. 1, pp. 45–49, Jul. 1998.

[228] E. Namanya, A. Waligo, A. Ipurale, C. T. Gaunt, and R. Herman, “Voltage calculations program for LV feeders with distributed generation,” in *PowerAfrica, 2016 IEEE PES*, 2016, pp. 174–178.

[229] L. C. Siebert, L. R. Ferreira, A. Aoki, A. Bonelli, A. R. R. Souza, and F. D. O. Toledo, “Deterministic versus Probabilistic approaches to self-healing in Smart Grid,” in *22nd International Conference and Exhibition on Electricity Distribution (CIRED)*, 2013, no. 1413, pp. 10–13.

[230] R. Herman, J. S. Maritz, and J. H. R. Enslin, “The analysis of voltage regulation in residential distribution networks using the beta distribution model,” *Electr. Power Syst. Res.*, vol. 29, no. 3, pp. 213–216, May 1994.

[231] R. Herman and J. S. Maritz, “Voltage regulation algorithm for a bi-phase distribution system feeding residential customers using a Beta p.d.f. load model,” *Electr. Power Syst. Res.*, vol. 43, no. 2, pp. 77–80, Nov. 1997.

[232] KS 1818-1:2010, “Electrical power transmission and distribution - Guidelines for the provision of electrical distribution networks in residential areas Part 1: Planning and design of distribution systems,” 2010.

[233] C. T. Gaunt, R. Herman, G. Celli, and S. Mocci, “MV and LV distribution feeder design using probabilistic approaches to load and DG Introduction,” in *Tutorial at CIRED 21st International Conference on electricity Distribution, Frankfurt Germany*, 2011, no. June, p. Not Published.

[234] E. Namanya, “Voltage calculation on low voltage feeders with distributed generation,” University of Cape Town, 2014.

[235] C. T. Gaunt, R. Herman, and H. Kadada, “Design parameters for LV feeders to meet regulatory limits of voltage magnitude,” in *21st International Conference on Electricity Distribution*, 2011.

[236] H. Cheng, Y. Hou, and F. Wu, “Probabilistic wind power generation model: Derivation and applications,” *Int. J. Energy*, vol. 5, no. 2, pp. 17–26, 2005.

[237] Tai-Her Yeh and Li Wang, “A Study on Generator Capacity for Wind Turbines Under Various Tower Heights and Rated Wind Speeds Using Weibull Distribution,” *IEEE Trans. Energy Convers.*, vol. 23, no. 2, pp. 592–602, Jun. 2008.

[238] F. M. Gonzalez-Longatt, J. L. Rueda, I. Erlich, D. Bogdanov, and W. Villa, “Identification of Gaussian mixture model using Mean Variance Mapping Optimization: Venezuelan case,” in *2012 3rd IEEE PES Innovative Smart Grid Technologies Europe (ISGT Europe)*, 2012, pp. 1–6.

[239] P. G. Jamdade and S. G. Jamdade, “Evaluation of wind energy potential for four sites in Ireland using the Weibull distribution model,” *J. Power Technol.*, vol. 95, no. 1, pp. 48–53, 2015.

[240] F. McLoughlin, A. Duffy, and M. Conlon, “A clustering approach to domestic electricity load profile characterisation using smart metering data,” *Appl. Energy*, vol. 141, 2015.

[241] F. Biscarri, I. Monedero, A. Garcia, J. I. Guerrero, and C. Leon, “Electricity clustering framework for automatic classification of customer loads,” *Expert Syst. Appl.*, vol. 86, pp. 54–63, Nov. 2017.

[242] G. P. H. Styan, “Hadamard products and multivariate statistical analysis,” *Linear Algebra Appl.*, vol. 6, pp. 217–240, Jan. 1973.

[243] R. D. Zimmerman and Hsiao-Dong Chiang, “Fast decoupled power flow for unbalanced radial distribution systems,” *IEEE Trans. Power Syst.*, vol. 10, no. 4, pp. 2045–2052, Nov. 1995.

[244] M. Neaimeh, R. Wardle, A. M. Jenkins, J. Yi, G. Hill, P. F. Lyons, Y. Hübner, P. T. Blythe, and P. C. Taylor, “A probabilistic approach to combining smart meter and electric vehicle charging data to investigate distribution network impacts,” *Appl. Energy*, vol. 157, pp. 688–698, Nov. 2015.

[245] J. E. Anandraj, “Point estimate method of Load Flow for distribution network with photovoltaic generators,” in *2013 International Conference on Energy Efficient Technologies for Sustainability*, 2013, pp. 24–29.

[246] M. E. Baran and F. F. Wu, “Network reconfiguration in distribution systems for loss reduction and load balancing,” *IEEE Trans. Power Deliv.*, vol. 4, no. 2, pp. 1401–1407, Apr. 1989.

[247] R. Srinivasa Rao, S. V. L. Narasimham, M. Ramalinga Raju, and A. Srinivasa Rao, “Optimal Network Reconfiguration of Large-Scale Distribution System Using Harmony Search Algorithm,” *IEEE Trans. Power Syst.*, vol. 26, no. 3, pp. 1080–1088, Aug. 2011.

[248] R. M. Ciric, A. Padilha-Feltrin, and I. F. E. D. Denis, “Observing the performance of distribution systems with embedded generators,” *Eur. Trans. Electr. Power*, vol. 14, no. 6, pp. 347–359, Nov. 2004.

[249] M. E. Baran and F. F. Wu, “Optimal capacitor placement on radial distribution systems,” *IEEE Trans. Power Deliv.*, vol. 4, no. 1, pp. 725–734, 1989.

[250] P. V. Babu and S. P. Singh, “Capacitor allocation in radial distribution system for maximal energy savings,” in *2016 National Power Systems Conference (NPSC)*, 2016, pp. 1–6.

[251] City of Cape Town:CTEF100, “Electrical Reticulation-City of Cape Town Technical Standards: CTEF100,” 2014.

Appendix A: HBE List of Equations

This appendix contains the full description of the calculations in the HBE transform provided in steps. The original HBA algorithm and the modified approach for MV are fully documented in previous work [11].

A1: LIST OF SYMBOLS

SYMBOL	DESCRIPTION
C_b	Normalized load current limit per node
$\alpha_{V_{con}}$	Beta PDF shape parameter, alpha, for scaled voltage
$\beta_{V_{con}}$	Beta PDF shape parameter, beta, for scaled voltage
$\rho_{a,a}$	Correlation coefficient between phases a and a at node i
$\rho_{j,k}$	Correlation coefficient between nodes *j* and k
θi	Power angle for inputs at node i
ai	Beta PDF shape parameter, alpha, for input connected to node i
BETAINV	Beta inverse function
bi	Beta PDF shape parameter, beta, for input connected to node i
ci	Scaling factor (in amperes), usually taken as the supply circuit breaker rating at node i
E(x)	Expected value of x
G	First statistical moment (mean) of a beta PDF
H	Second statistical moment of a beta PDF
Ib	System base current for normalization
kri	Phase to neutral conductor resistance ratio
kxi	Phase to neutral conductor reactance ratio
J,k	Node counters
m_a	Number of customers connected to phase *a*, at node i
m_b	Number of customers connected to phase *b*, at node i
m_c	Number of customers connected to phase *c*, at node i

N	Total number of nodes on the feeder
p	Design risk level applied in the interpretation of probabilistic outputs
Pfi	Input power factor at node i
Rho	Correlation matrix
rn	Temperature-corrected reactance of the neutral conductor per branch
rp	Temperature-corrected resistance of the phase conductor per branch
Sb	System base power for normalization
Vb	System base voltage for normalization
V_{con}	Consumer voltage
V_{con}^{*}	Scaled consumer voltage
ΔV	Voltage drop
Vs	The system's nominal voltage
xn	Temperature-corrected resistance of the neutral conductor per branch
Xn	Normalized temperature-corrected resistance of the neutral conductor per branch
xp	Temperature-corrected reactance of the phase conductor per branch
Xp	Normalized temperature-corrected reactance of the phase conductor per branch

A2: PROCEDURE FOR THE CALCULATION OF BUS VOLTAGES: 3PHASE-4WIRE

STEP 1: Select Network Parameters

1. Nominal system voltage, Vs. Also taken as the slack bus voltage.
2. Specify per-unit base quantities for voltage and power: S_B, V_B
3. Provide shape and scale parameters for the beta-PDF inputs: *ai*, *bi*, *ci*, *pfi*
4. Specify the total number of inputs (customers) at node 'i' according to phase connections: mai, mbi. and mci. A positive number denotes a syncing current for a load (or shunt reactor) and a negative number represents a sourcing current for DG (or shunt capacitance).
5. Specify the phase and neutral electrical properties for each section: rp, rn, xp, xn.
6. Specify the number of nodes on the radial branch section under analysis, *N*.
7. Specify the correlation matrix Rho for all interdependencies; matrix size $N \times N \times$ phases. (covers node-to-node and inter-phase correlations).
8. Specify a design risk value: *p*, in percent.

STEP 2: Express System Inputs in Per-unit: Rp, Rn, Xp, Xn, Ci

$$I_B = \frac{S_B}{V_B}\,; Z_B = \frac{V_B^2}{S_B}$$

$$\mathrm{Rp} = \frac{\mathrm{rp}}{\mathrm{Z_B}}; \mathrm{Xp} = \frac{\mathrm{xp}}{\mathrm{Z_B}};\ \ \mathrm{Rn} = \frac{\mathrm{rn}}{\mathrm{Z_B}}; \mathrm{Xn} = \frac{\mathrm{xn}}{\mathrm{Z_B}};\ \ \mathrm{Ci} = \frac{\mathrm{ci}}{\mathrm{I_B}}$$

STEP 3: Calculate the Superposition Elements: Ri, Rp, Xi, Xp, kxi, kri.

$$R_i = \sum_{j=1}^{i} Rp_{(j)}\,;\ X_i = \sum_{j=1}^{i} Xp_{(j)}\,;\ R_n = \sum_{j=1}^{i} Rn_{(j)}\,;\ X_n = \sum_{j=1}^{i} Xn_{(j)}$$

$$\mathrm{k_r} = \frac{\mathrm{R_n}}{\mathrm{R_i}};\ \mathrm{k_x} = \frac{\mathrm{X_n}}{\mathrm{X_i}}$$

STEP 4: Calculate the Mean and Second Moment for Input Currents, *G*, *H* and $\boldsymbol{\theta_i}$

$$G = \frac{ai}{ai + bi};\ \ H = \frac{ai(ai+1)}{(ai+bi)(ai+bi+1)}; \theta_i = \mathrm{acos}(pfi)$$

STEP 5: Calculate Constants; $\boldsymbol{k_1 - k_4}$; $\boldsymbol{c_1 - c_4}$; $\boldsymbol{d_1 - d_4}$; $\boldsymbol{C_1 - C_4}$

$$\begin{bmatrix} \mathrm{k_1} \\ k_2 \\ k_3 \\ k_4 \end{bmatrix} = \begin{bmatrix} k_r R_i \pm k_x X_i \tan\phi_i \\ k_x X_i \mp k_r R_i \tan\phi_i \\ (1+k_r)\cdot R_i \pm (1+k_x)\cdot X_i \tan\phi_i \\ (1+k_x)\cdot X_i \mp (1+k_r)\cdot R_i \tan\phi_i \end{bmatrix}$$

$$\begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \\ c_6 \end{bmatrix} = \begin{bmatrix} 0.5k_1k_2\sqrt{3} + 0.75k_2^2 + 0.25k_1^2 \\ -0.5k_1k_2\sqrt{3} + 0.75k_2^2 + 0.25k_1^2 \\ k_3^2|m_a|(|m_a|-1) + c_1|m_b|(|m_b|-1) + c_2|m_c|(|m_c|-1) \\ |m_a||m_b|\left(-k_3k_2\sqrt{3} - k_1k_3\right) + |m_a||m_c|\left(k_3k_2\sqrt{3} - k_1k_3\right) + |m_b||m_c|(-1.5k_2^2 + 0.5k_1^2) \\ \rho_{a,a}k_3^2|m_a|(|m_a|-1) + \rho_{b,b}c_1|m_b|(|m_b|-1) + \rho_{c,c}c_2|m_c|(|m_c|-1) \\ \rho_{a,b}|m_a\, m_b|\left(-k_3k_2\sqrt{3} - k_1k_3\right) + \rho_{a,c}|m_a\, m_c|\left(k_3k_2\sqrt{3} - k_1k_3\right) + \rho_{b,c}|m_b m_c|\left(-1.5k_2^2 + 0.5k_1^2\right) \end{bmatrix}$$

$$\begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \\ d_5 \\ d_6 \end{bmatrix} = \begin{bmatrix} -0.5k_1k_2\sqrt{3} + 0.75k_1^2 + 0.25k_2^2 \\ 0.5k_1k_2\sqrt{3} + 0.75k_1^2 + 0.25k_2^2 \\ k_4^2|m_a|(|m_a|-1) + d_1|m_b|(|m_b|-1) + d_2|m_c|(|m_c|-1) \\ |m_am_b|(k_4k_1\sqrt{3} - k_2k_4) + |m_am_c|(-k_4k_1\sqrt{3} - k_2k_4) + |m_bm_c|(-1.5k_1^2 + 0.5k_2^2) \\ \rho_{a,a}k_4^2|m_a|(|m_a|-1) + \rho_{b,b}d_1|m_b|(|m_b|-1) + \rho_{c,c}d_2|m_c|(|m_c|-1) \\ \rho_{a,b}\left(k_4k_1\sqrt{3} - k_2k_4\right)|m_am_b| + \rho_{a,c}\left(-k_4k_1\sqrt{3} - k_2k_4\right)|m_am_c| + \rho_{b,c}\left(-1.5k_1^2 + 0.5k_2^2\right)|m_bm_c| \end{bmatrix}$$

$$\begin{bmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \\ C_5 \\ C_6 \end{bmatrix} = \begin{bmatrix} 0.5(m_b + m_c) \\ 0.5\sqrt{3}(m_b - m_c) \\ k_3^2|m_a| + c_1|m_b| + c_2|m_c| + c_5 + c_6 \\ c_3 + c_4 - c_5 - c_6 \\ k_4^2|m_a| + d_1|m_b| + d_2|m_c| + d_5 + d_6 \\ d_3 + d_4 - d_5 - d_6 \end{bmatrix}$$

STEP 6: Calculate Voltage-drop Moments at Each Node

$$\mathrm{E}(\Delta V_{x_i}) = (k_3 \cdot m_a - k_1 \cdot C_1 - k_2 \cdot C_2) \cdot G \cdot C_b$$

$$E(\Delta \mathrm{V}_{y_i}) = (k_4 \cdot m_a - k_2 \cdot C_1 + k_1 \cdot C_2) \cdot G \cdot C_b$$

$$E(\Delta V_{x}^2{}_i) = C_b^2(C_3 \cdot H + C_4 \cdot G^2)$$

$$E\left(\Delta V_{y}^2{}_i\right) = C_b^2(C_5 \cdot H + C_6 \cdot G^2)$$

STEP 7: Calculate the Summated Voltage-drop Moments at end of Lateral

$$E(\Delta V_{x_t}) = \sum_{i=1}^{N} E(\Delta V_{x_i})$$

$$E(\Delta V_{y_t}) = \sum_{i=1}^{N} E(\Delta V_{y_i})$$

$$E(\Delta V_{x_t}^2) = \sum_{i=1}^{N} E(\Delta V_{x_i}^2) + \sum_{j=1}^{N}\sum_{\substack{k=1 \\ k\neq j}}^{N} \left\{E\left(\Delta V_{x_j}\right)E(\Delta V_{x_k}) + E\left(\Delta V_{x_j}\right)E(\Delta V_{x_k}) + (\sigma_{V_{x_j}} \cdot \sigma_{V_{x_k}} \cdot \rho_{j,k})\right\}$$

$$E(\Delta V_{y_t}^2) = \sum_{i=1}^{N} E(\Delta V_{y_i}^2) + \sum_{j=1}^{N}\sum_{\substack{k=1 \\ k\neq j}}^{N} \left\{E\left(\Delta V_{y_j}\right)E(\Delta V_{y_k}) + E\left(\Delta V_{y_j}\right)E(\Delta V_{y_k}) + \left(\sigma_{V_{y_j}} \cdot \sigma_{V_{y_k}} \cdot \rho_{j,k}\right)\right\}$$

Where: $\sigma_{V_{x_j}} = \sqrt{\left(E\left(V_{x_j}^2\right) - E\left(V_{x_j}\right)^2\right)}$; $\quad \sigma_{V_{y_k}} = \sqrt{\left(E\left(V_{y_k}^2\right) - E\left(V_{y_k}\right)^2\right)}$

$$\sigma_{V_{x_k}} = \sqrt{E\left(V_{x_k}^2\right) - E\left(V_{x_k}\right)^2};\ \sigma_{Vy_k} = \sqrt{E\left(V_{y_k}^2\right) - E\left(V_{y_k}\right)^2}$$

STEP 8: Calculate the Moments of the Consumer Voltage

$$\mathrm{E}(\mathrm{V_{con}}) \approx \mathrm{V_s}\left(1 - \frac{\mathrm{E}(\Delta V_{x_t})}{\mathrm{V_s}} + 0.5\frac{\mathrm{E}(\Delta V_{y_t}^2)}{\mathrm{V_S^2}}\right)$$

$$\mathrm{E}(\mathrm{V}_{\mathrm{con}}^2) \approx \mathrm{V}_{\mathrm{S}}^2 - 2.\mathrm{V}_{\mathrm{s}}.\mathrm{E}(\Delta V_{x_{\mathrm{t}}}) + \mathrm{E}\left(\Delta {V_x}^2{}_t\right) + \mathrm{E}\left(\Delta {V_y}^2{}_t\right)$$

STEP 9: Calculate the Voltage PDF Shape Parameters

$$\alpha_{V_{con}} = \frac{E(V_{con}^2) - E(V_{con})}{E(V_{con}) - \frac{E(V_{con}^2)}{E(V_{con})}}$$

$$\beta_{V_{con}} = \frac{\alpha_{V_{con}}}{E(V_{con}^*)} - \alpha_{V_{con}}$$

STEP 10: Calculate Bus Voltage at Selected Risk Level r%

$$\mathrm{V}_{\mathrm{con}_{\%}}^* = \mathrm{betainv}\left(\alpha_{\mathrm{V_{con}}}, \beta_{\mathrm{V_{con}}}, \mathrm{r}\right)$$

STEP 11: Rescale Outputs

$$\mathrm{V}_{\mathrm{con}_{\%}} = \mathrm{V}_{\mathrm{con}_{\%}}^* \cdot V_B$$

A3: PROCEDURE FOR THE CALCULATION OF BUS VOLTAGES: 3PHASE-3WIRE

STEP 1: Select Network Parameters

1. Nominal system voltage, Vs. Also taken as the slack bus voltage.
2. Specify per-unit base quantities for voltage and power: S_B, V_B
3. Provide shape and scale parameters for the beta-PDF inputs: *ai*, *bi*, *ci*, *pf_i*
4. Specify the total number of inputs (customers) at node 'i' according to phase connections: m_a ,m_b. and m_c. A positive number denotes a syncing current for a load (or shunt reactor) and a negative number represents a sourcing current for DG (or shunt capacitance).
5. Specify the phase electrical properties for each section: rp, xp
6. Specify the number of nodes on the radial branch section under analysis, *N*.
7. Specify the correlation matrix Rho for all interdependencies; matrix size $N \times N \times$ phases. (covers node-to-node and inter-phase correlations).
8. Specify a design risk value: *p*, in percent.

STEP 2: Express System Inputs in Per-unit: Rp, Xp, Ci

$$I_B = \frac{S_B}{V_B}\,; Z_B = \frac{V_B^2}{S_B}$$
$$\mathrm{Rp} = \frac{\mathrm{rp}}{\mathrm{Z_B}}; \mathrm{Xp} = \frac{\mathrm{xp}}{\mathrm{Z_B}};\ \mathrm{C_b} = \frac{\mathrm{ci}}{\mathrm{I_B}}$$

STEP 3: Calculate the Superposition Elements: Ri, Rp, Xi, Xp,$\boldsymbol{\theta_i}$

$$R_i = \sum_{j=1}^{i} Rp_{(j)}\,;\ X_i = \sum_{j=1}^{i} Xp_{(j)}\,; \theta_i = \mathrm{acos}(pfi)$$

STEP 4: Calculate the Mean and Second Moment for Input Currents, *G* and H

$$G = \frac{ai}{ai + bi};\quad H = \frac{ai(ai+1)}{(ai+bi)(ai+bi+1)}$$

STEP 5: Calculate Constants; $\boldsymbol{w_1, w_2, F_1 - F_6}$,

$$\begin{bmatrix} \mathrm{w}_1 \\ w_2 \\ e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \end{bmatrix} = \begin{bmatrix} R_i \pm X_i \tan\phi_i \\ X_i \mp R_i \tan\phi_i \\ 0.5_1 w_2\sqrt{3} + 0.75w_2^2 + 0.25w_1^2 \\ -0.5w_1 w_2\sqrt{3} + 0.75w_2^2 + 0.25w_1^2 \\ 4w_1^2|m_a|(|m_a|-1) + e_1|m_b|(|m_b|-1) + e_2|m_c|(|m_c|-1) \\ (2w_1w_2\sqrt{3} - 2w_1^2)|m_a m_b| + (-2w_1w_2\sqrt{3} + 2w_1^2)|m_a m_c| + (-1.5w_2^2 + 0.5w_1^2)|m_b m_c| \\ \rho_{a,a}4w_1^2|m_a|(|m_a|-1) + \rho_{b,b}e_1|m_b|(|m_b|-1) + \rho_{c,c}e_2|m_c|(|m_c|-1) \\ \rho_{a,b}|m_a\, m_b|\left(2w_1w_2\sqrt{3} - 2w_1^2\right) + \rho_{a,c}|m_a\, m_c|\left(-2w_1w_2\sqrt{3} + 2w_1^2\right) + \rho_{b,c}|m_b\, m_c|(-1.5w_2^2 + 0.5w_1^2) \end{bmatrix}$$

$$\begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \\ f_5 \\ f_6 \end{bmatrix} = \begin{bmatrix} -0.5w_1w_2\sqrt{3} + 0.75w_1^2 + 0.25w_2^2 \\ 0.5w_1w_2\sqrt{3} + 0.75w_1^2 + 0.25w_2^2 \\ 4w_2^2|m_a|(|m_a|-1) + d_1|m_b|(|m_b|-1) + d_2|m_c|(|m_c|-1) \\ \left(-2w_2w_1\sqrt{3} + 2w_2^2\right)|m_a m_b| + \left(2w_2w_1\sqrt{3} + 2w_2^2\right)|m_a m_c| + (-1.5w_1^2 + 0.5w_2^2)|m_b m_c| \\ \rho_{a,a}4w_2^2|m_a|(|m_a|-1) + \rho_{b,b}f_1|m_b|(|m_b|-1) + \rho_{c,c}f_2|m_c|(|m_c|-1) \\ \rho_{a,b}|m_a\, m_b|(-2w_2w_1\sqrt{3} + 2w_2^2) + \rho_{a,c}|m_a\, m_c|(2w_2w_1\sqrt{3} + 2w_2^2) + \rho_{b,c}|m_b m_c|(-1.5w_1^2 + 0.5w_2^2) \end{bmatrix}$$

$$\begin{bmatrix} F_1 \\ F_2 \\ F_3 \\ F_4 \\ F_5 \\ F_6 \end{bmatrix} = \begin{bmatrix} \left(2 \cdot m_a + 0.5 \cdot (m_b + m_c)\right) \\ \frac{\sqrt{3}}{2} \cdot (m_b - m_c) \\ 4k_1^2|m_a| + e_1|m_b| + e_2|m_c| + e_5 + e_6 \\ e_3 + e_4 - e_5 - e_6 \\ 4k_2^2|m_a| + f_1|m_b| + f_2|m_c| + f_5 + f_6 \\ f_3 + f_4 - f_5 - f_6 \end{bmatrix}$$

STEP 6: Calculate Voltage-drop Moments for Each Node

$$\mathrm{E}\left(\Delta V_{x_i}\right) = (F_1 \cdot w_1 + F_2 \cdot w_2) \cdot G \cdot C_b$$

$$E\left(\Delta \mathrm{V}_{y_i}\right) = (F_2 \cdot w_1 - F_1 \cdot w_2) \cdot G \cdot C_b$$

$$E\left(\Delta V_{x}^2{}_i\right) = C_b^2(F_3 \cdot H + F_4 \cdot G^2)$$

$$E\left(\Delta V_{y}^2{}_i\right) = C_b^2(F_5 \cdot H + F_6 \cdot G^2)$$

STEP 7: Calculate the Summated Voltage-drop Moments at End of Lateral

$$E\left(\Delta V_{x_t}\right) = \sum_{i=1}^{N} E(\Delta V_{x_i})$$

$$E\left(\Delta V_{y_t}\right) = \sum_{i=1}^{N} E(\Delta V_{y_i})$$

$$E\left(\Delta V_{x_t}^2\right) = \sum_{i=1}^{N} E\left(\Delta V_{x_i}^2\right) + \sum_{j=1}^{N} \sum_{\substack{k=1 \\ k \neq j}}^{N} \left\{E\left(\Delta V_{x_j}\right) E\left(\Delta V_{x_k}\right) + E\left(\Delta V_{x_j}\right) E\left(\Delta V_{x_k}\right) + \left(\sigma_{V_{x_j}} \cdot \sigma_{V_{x_k}} \cdot \rho_{j,k}\right)\right\}$$

$$E\left(\Delta V_{y_t}^2\right) = \sum_{i=1}^{N} E\left(\Delta V_{y_i}^2\right) + \sum_{j=1}^{N} \sum_{\substack{k=1 \\ k \neq j}}^{N} \left\{E\left(\Delta V_{y_j}\right) E\left(\Delta V_{y_k}\right) + E\left(\Delta V_{y_j}\right) E\left(\Delta V_{y_k}\right) + \left(\sigma_{V_{y_j}} \cdot \sigma_{V_{y_k}} \cdot \rho_{j,k}\right)\right\}$$

Where: $\sigma_{V_{x_j}} = \sqrt{\left(E\left(V_{x_j}^2\right) - E\left(V_{x_j}\right)^2\right)}; \quad \sigma_{V_{y_k}} = \sqrt{\left(E\left(V_{y_k}^2\right) - E\left(V_{y_k}\right)^2\right)}$

$$\sigma_{V_{x_k}} = \sqrt{E\left(V_{x_k}^2\right) - E\left(V_{x_k}\right)^2}; \; \sigma_{V y_k} = \sqrt{E\left(V_{y_k}^2\right) - E\left(V_{y_k}\right)^2}$$

STEP 8: Calculate the Moments for the Consumer Voltage

$$\mathrm{E}(\mathrm{V_{con}}) \approx \mathrm{V_s}\left(1 - \frac{\mathrm{E}\left(\Delta V_{x_t}\right)}{\mathrm{V_s}} + 0.5\frac{\mathrm{E}\left(\Delta V_{y_t}^2\right)}{\mathrm{V_S^2}}\right)$$

$$\mathrm{E}(\mathrm{V_{con}^2}) \approx \mathrm{V_S^2} - 2.\mathrm{V_s}.\mathrm{E}(\Delta V_{x_t}) + \mathrm{E}\left(\Delta V_{x}^2{}_t\right) + \mathrm{E}\left(\Delta V_{y}^2{}_t\right)$$

STEP 9: Calculate the Voltage PDF Shape Parameters

$$\alpha_{V_{con}} = \frac{E(V_{con}^2) - E(V_{con})}{E(V_{con}) - \frac{E(V_{con}^2)}{E(V_{con})}}$$

$$\beta_{V_{con}} = \frac{\alpha_{V_{con}}}{E(V_{con}^*)} - \alpha_{V_{con}}$$

STEP 10: Calculate Bus Voltage at Selected Risk Level r%

$$V_{con_{\%}}^* = betainv\left(\alpha_{V_{con}}, \beta_{V_{con}}, r\right)$$

STEP 11: Rescale Outputs

$$V_{con_{\%}} = V_{con_{\%}}^* \cdot V_B$$

Appendix B: Test Network Data

B1. LINE AND LOAD DATA FOR THE MODIFIED 12-BUS FEEDER, 11 KV

BUS LAYOUT		ELECTRICAL PARAMETERS										
		Line Impedance		Shunt	Load Parameters (Sj)					Customer Phase Allocation		
From (i)	To (j)	R_{ij} (Ω)	X_{ij} (Ω)	Y/2 (S)	α	β	Cb (A)	pf	ADD (kVA)	A	B	C
1	2	0.092	0.047	--	1.5	5.0	35.34	0.857	45.94	1	0	1
2	3	0.493	0.251	--	1.5	5.0	29.85	0.986	35.94	1	1	0
3	4	0.366	0.186	--	1.5	5.0	43.70	0.832	56.81	0	1	1
4	5	0.381	0.194	--	1.5	5.0	20.33	0.894	26.43	1	0	1
5	6	0.819	0.707	--	1.5	5.0	19.17	0.768	30.77	1	1	0
6	7	0.187	0.619	--	1.5	5.0	67.76	0.894	88.09	0	1	1
7	8	0.712	0.235	--	1.5	5.0	67.76	0.555	142.0	1	0	1
8	9	1.030	0.740	--	1.5	5.0	19.17	0.800	29.55	1	1	0
9	10	1.044	0.740	--	1.5	5.0	19.17	0.949	24.91	0	1	1
10	11	0.197	0.065	--	1.5	5.0	16.39	0.747	23.72	1	0	1
11	12	0.374	0.130	--	1.5	5.0	21.05	0.864	27.36	1	1	0

B2. LINE AND LOAD DATA FOR THE IEEE 33-BUS FEEDER, 12.66 KV

BUS LAYOUT			ELECTRICAL PARAMETERS				
			Line Impedance (Zij)		Shunts	Load (Sj)	
Line number	Sending bus no. (i)	Receiving bus no. (j)	Resistance (Ω)	Reactance (Ω)	Y/2 (S)	Active Power (kW)	Reactive Power (kVAr)
1	1	2	0.0922	0.0470	--	100	60
2	2	3	0.4930	0.2512	--	90	40
3	3	4	0.3661	0.1864	--	120	80
4	4	5	0.3811	0.1941	--	60	30
5	5	6	0.8190	0.7070	--	60	20
6	6	7	0.1872	0.6188	--	200	100
7	7	8	0.7115	0.2351	--	200	100
8	8	9	1.0299	0.7400	--	60	20
9	9	10	1.0440	0.7400	--	60	20
10	10	11	0.1967	0.0651	--	45	30
11	11	12	0.3744	0.1298	--	60	35
12	12	13	1.4680	1.1549	--	60	35
13	13	14	0.5416	0.7129	--	120	80
14	14	15	0.5909	0.5260	--	60	10
15	15	16	0.7462	0.5449	--	60	20
16	16	17	1.2889	1.7210	--	60	20
17	17	18	0.7320	0.5739	--	90	40

18	2	19	0.1640	0.1565	--	90	40
19	19	20	1.5042	1.3555	--	90	40
20	20	21	0.4095	0.4784	--	90	40
21	21	22	0.7089	0.9373	--	90	40
22	3	23	0.4512	0.3084	--	90	50
23	23	24	0.8980	0.7091	--	420	200
24	24	25	0.8959	0.7071	--	420	200
25	6	26	0.2031	0.1034	--	60	25
26	26	27	0.2842	0.1447	--	60	25
27	27	28	1.0589	0.9338	--	60	20
28	28	29	0.8043	0.7006	--	120	70
29	29	30	0.5074	0.2585	--	200	600
30	30	31	0.9745	0.9629	--	150	70
31	31	32	0.3105	0.3619	--	210	100
32	32	33	0.3411	0.5302	--	60	40

B3. LINE AND LOAD DATA FOR A 34-BUS FEEDER, 24.9 KV

BUS LAYOUT			**ELECTRICAL PARAMETERS**				
			Line Impedance (Zij)		**Shunts**	**Load (Sj)**	
Line number	**Sending bus no. (i)**	**Receiving bus no. (j)**	**Resistance (Ω)**	**Reactance (Ω)**	**Y (µS)**	**Active Power (kW)**	**Reactive Power (kVAr)**
1	1	2	1.2591	0.5551	0.001452	19.1	9.87
2	2	3	0.8443	0.3722	0.000968	0	0
3	3	4	15.7294	6.9344	0.017742	5.29	2.74
4	4	5	18.3013	8.0683	0.020967	0	0
5	5	6	14.5093	6.3966	0.016129	0	0
6	6	7	0.4880	0.2152	0	0	0
7	7	8	0.1513	0.0667	0	0.13	0.07
8	8	9	4.9828	2.1967	0.004839	14.9	7.71
9	9	10	0.4099	0.1807	0.000323	2.06	1.07
10	10	11	9.9754	4.3978	0.114514	0	0
11	11	12	0.2538	0.1119	0	1.24	0.64
12	12	13	17.9744	7.9242	0.019355	0	0
13	13	14	0.4880	0.2152	0	4.37	2.26
14	14	15	2.3914	1.0543	0.001613	10	5.17
15	15	16	2.8452	1.2544	0.003226	50	0
16	16	17	0.9858	0.4346	0.000161	46.57	29.72
17	17	18	1.3079	0.5766	0.001613	13.1	6.77
18	2	19	0.4197	0.1850	0.000484	8.859999	7.09
19	19	20	2.8501	1.2565	0.003226	0	0
20	20	21	0.8345	0.3679	0.000806	11.3	5.84
21	21	22	23.4989	10.3597	0.025806	11.84	23.36

22	3	23	6.7056	2.9562	0.075805	0	0
23	23	24	1.4787	0.6519	0.000161	0	0
24	24	25	11.3859	5.0196	0.130482	0	0
25	6	26	1.9000	4.0800	0.009677	0	0
26	26	27	5.1537	2.2720	0.004839	27	21.62
27	27	28	0.7906	0.3486	0.000806	0	0
28	28	29	0.1366	0.0602	1.61E-05	3.04	1.57
29	29	30	0.6588	0.2905	161.2884	149.05	14.9
30	30	31	1.7764	0.7832	0.001613	7.54	3.9
31	31	32	0.2587	0.1140	241.9317	19.45	134.43
32	32	33	0.1366	0.0602	1.61E-05	9.2	4.76
33	33	34	1.5862528	1.0446846	0.001613	0	0

B4. LINE AND LOAD DATA FOR THE IEEE 69-BUS FEEDER, 12.66 KV

BUS LAYOUT			**ELECTRICAL PARAMETERS**				
			Impedance (Zij)		**Shunts**	**Load (Sj)**	
Line number	**Sending bus no. (i)**	**Receiving bus no. (j)**	**Resistance (Ω)**	**Reactance (Ω)**	**Y/2 (S)**	**Active Power (kW)**	**Reactive Power (kVAr)**
1	1	2	0.0005	0.0012	--	0	0
2	2	3	0.0005	0.0012	--	0	0
3	3	4	0.0015	0.0036	--	0	0
4	4	5	0.0251	0.0294	--	0	0
5	5	6	0.366	0.1864	--	2.6	2.2
6	6	7	0.3811	0.1941	--	40.4	30
7	7	8	0.0922	0.047	--	75	54
8	8	9	0.0493	0.0251	--	30	22
9	9	10	0.819	0.2707	--	28	19
10	10	11	0.1872	0.0691	--	145	104
11	11	12	0.7114	0.2351	--	145	104
12	12	13	1.03	0.34	--	8	5.5
13	13	14	1.044	0.345	--	8	5.5
14	14	15	1.058	0.3496	--	0	0
15	15	16	0.1966	0.065	--	45.5	30
16	16	17	0.3744	0.1238	--	60	35
17	17	18	0.0047	0.0016	--	60	35
18	18	19	0.3276	0.1083	--	0	0
19	19	20	0.2106	0.069	--	1	0.6
20	20	21	0.3416	0.1129	--	114	81
21	21	22	0.014	0.0046	--	5.3	3.5
22	22	23	0.1591	0.0526	--	0	0
23	23	24	0.3463	0.1145	--	28	20
24	24	25	0.7488	0.2745	--	0	0
25	25	26	0.3089	0.1021	--	14	10
26	26	27	0.1732	0.0572	--	14	10
27	3	28	0.0044	0.0108	--	26	18.6

28	28	29	0.064	0.1565	--	26	18.6
29	29	30	0.3978	0.1315	--	0	0
30	30	31	0.0702	0.0232	--	0	0
31	31	32	0.351	0.116	--	0	0
32	32	33	0.839	0.2816	--	14	10
33	33	34	1.708	0.5646	--	19.5	14
34	34	35	1.474	0.4673	--	6	4
35	3	36	0.0044	0.0108	--	26	18.55
36	36	37	0.064	0.1565	--	26	18.55
37	37	38	0.1053	0.123	--	0	0
38	38	39	0.0304	0.0355	--	24	17
39	39	40	0.0018	0.0021	--	24	17
40	40	41	0.7283	0.8509	--	1.2	1
41	41	42	0.31	0.3623	--	0	0
42	42	43	0.041	0.0478	--	6	4.3
43	43	44	0.0092	0.0116	--	0	0
44	44	45	0.1089	0.1373	--	39.22	26.3
45	45	46	0.0009	0.0012	--	39.22	26.3
46	4	47	0.0034	0.0084	--	0	0
47	47	48	0.0851	0.2083	--	79	56.4
48	48	49	0.2898	0.7091	--	384.7	274.5
49	49	50	0.0822	0.2011	--	384	274.5
50	8	51	0.0928	0.0473	--	40.5	28.3
51	51	52	0.3319	0.1114	--	3.6	2.7
52	9	53	0.174	0.0886	--	4.35	3.5
53	53	54	0.203	0.1034	--	26.4	19
54	54	55	0.2842	0.1447	--	24	17.2
55	55	56	0.2813	0.1433	--	0	0
56	56	57	1.59	0.5337	--	0	0
57	57	58	0.7837	0.263	--	0	0
58	58	59	0.3042	0.1006	--	100	72
59	59	60	0.3861	0.1172	--	0	0
60	60	61	0.5075	0.2585	--	1244	888
61	61	62	0.0974	0.0496	--	32	23
62	62	63	0.145	0.0738	--	0	0
63	63	64	0.7105	0.3619	--	227	162
64	64	65	1.041	0.5302	--	59	42
65	11	66	0.2012	0.0611	--	18	13
66	66	67	0.0047	0.0014	--	18	13
67	12	68	0.7394	0.2444	--	28	20
68	68	69	0.0047	0.0016	--	28	20

B5. LINE AND LOAD DATA FOR THE PRACTICAL 12-BUS FEEDER, 0.23 KV

BUS LAYOUT		ELECTRICAL PARAMETERS										
		Line Impedance		Shunt	Load Parameters (Sj)					Customer Phase Allocation		
From (i)	**To** (j)	R_{ij} (Ω)	X_{ij} (Ω)	**Y/2** (S)	**α**	**β**	**Cb** (A)	**pf**	**ADD** (kVA)	**A**	**B**	**C**
4	0	0.0421	0.0128	--	1.445	5.118	80	1.00	4.051	2	2	1
2	1	0.0791	0.0075	--	1.420	4.194	80	1.00	4.654	1	2	2
3	2	0.0308	0.0056	--	1.418	4.146	80	1.00	4.690	2	1	1
4	3	0.0181	0.0055	--	1.446	5.945	80	1.00	3.600	1	1	2
LV-BB	4	0.0296	0.0133	--	1.444	5.029	80	1.00	4.104	1	1	0
6	5	0.0985	0.0093	--	1.446	5.237	80	1.00	3.982	1	1	2
7	6	0.0175	0.0079	--	1.296	2.658	80	1.00	6.030	1	2	1
LV-BB	7	0.0008	0.0004	--	1.445	5.118	80	1.00	4.051	1	2	1
LV-BB	8	0.0111	0.0050	--	1.420	4.194	80	1.00	4.654	2	1	1
4	0	0.0421	0.0128	--	1.418	4.146	80	1.00	4.690	2	2	1
2	1	0.0791	0.0075	--	1.446	5.945	80	1.00	3.600	1	2	2