KU LEUVEN

ARENBERG DOCTORAL SCHOOL
Faculty of Engineering Science

Stochastic Hosting Capacity of Low Voltage Distribution Systems Using Polynomial Chaos Expansion

Arpan Koirala

Supervisors:
Prof. dr. ir. D. Van Hertem
Dr. ir. R. D'hulst
(VITO, Mol, Belgium)

Dissertation presented in partial fulfillment of the requirements for the degree of Doctor of Engineering Science (PhD): Electrical Engineering

January 2023

Stochastic Hosting Capacity of Low Voltage Distribution Systems Using Polynomial Chaos Expansion

Arpan KOIRALA

Examination committee:
Prof. dr. ir. P. Verbaeten, chair
Prof. dr. ir. D. Van Hertem, supervisor
Dr. ir. R. D'hulst, supervisor
(VITO, Mol, Belgium)
Prof. dr. ir. J. Driesen
Prof. dr. T. Steenberghen
Dr. dipl.-ing. H. Ergun
Prof. dr. L. Roald
(University of Wisconsin-Madison, Madison, USA)
Prof. dr. ir. F. Vallée
(University of Mons, Belgium)

Dissertation presented in partial fulfillment of the requirements for the degree of Doctor of Engineering Science (PhD): Electrical Engineering

January 2023

© 2023 KU Leuven – Faculty of Engineering Science
Uitgegeven in eigen beheer, Arpan Koirala, Kasteelpark Arenberg 10, Bus 2445, B-3001 Leuven (Belgium)

Alle rechten voorbehouden. Niets uit deze uitgave mag worden vermenigvuldigd en/of openbaar gemaakt worden door middel van druk, fotokopie, microfilm, elektronisch of op welke andere wijze ook zonder voorafgaande schriftelijke toestemming van de uitgever.

All rights reserved. No part of the publication may be reproduced in any form by print, photoprint, microfilm, electronic or any other means without written permission from the publisher.

Preface

> *The river has taught me to listen; you will learn from it, too. The river knows everything; one can learn everything from it. You have already learned from the river that it is good to strive downwards, to sink, to seek the depths. -Herman Hesse (Siddhartha)*

The PhD journey I started four years ago was like the *river* in the above quote. I want to thank all my friends, colleagues and family, who were constant support during this journey of seeking the depth.

Firstly, I would like to thank my supervisors, prof. Dirk Van Hertem and Reinhilde D'hulst for shaping my raw ideas into solid work. I would also like to thank members of my PhD committee, François Vallée, Johan Driesen, and Thérèse Steenberghen, for their constant support and creative feedback in every committee meeting. Thank you, Hakan Ergun, not only as a jury of the thesis but also as a mentor. Thank you, prof. Line Roald, for agreeing to be part of my jury. Your works were the inspiration for some parts of this thesis.

This work was supported by the energy transition funds project BREGILAB funded by the FPS economy, S.M.E.s, Self-employed and Energy. Thanks to project manager Marc Meuris for taking the lead and organizing the project. Gratitude to the colleagues from EnergyVille for the collaboration in the project.

I want to thank collaborators who made some tangible contributions to this work in alphabetical order. Thank you, Pablo Arboleya, for being a mentor that started as a master thesis supervisor. It was your effort to make the LVDS data public, which is used throughout the thesis. Thank you, Frederick Geth, for generalizing the impedance model and collaboration in developing Julia package `StochsticPowerModels.jl`. I want to thank Md Umar Hashmi for his active contribution in shaping the last part of this thesis and making me a collaborator in some of his work. Chelsa Sisk, your contributions in making the text readable even for a common audience require hats off. Tom Van Acker, your role was pivotal in shaping this thesis and the direction of my research.

Thanks for being an active support, for fancy plots in tikz, and for taking the lead on developing the package `StochasticPowerModels.jl`. Thank you, Marta Vanin, for numerous help in debugging the Julia code in the initial days and for your collaboration in supervising the master thesis and summer job students.

Thank you, Katja, Katleen, Veerle and Veronica, for your assistance in figuring out the complex bureaucracy of KU Leuven and making them look easy.

Energyville colleagues who always forced me to leave Genk and move to Hasselt were the motivators of this journey. If you are a member of *Fun in Hasselt* group, this thanks is for you. I would also like to thank friends from Nepal, Aakash, Pritam, Riwaj and Nischal. Iros, you will be missed.

Finally, I want to thank my parents for their encouragement. Thanks to my sisters and brothers-in-law for always being there to support me. Shruti, thanks for being with me on this journey as a life partner and a good friend. I want to dedicate this thesis to my nephews, Sworaj and Aaryav. Talking to you two was the real stress buster in the whole PhD.

Abstract

The increased penetration of renewable resources, such as photovoltaics (PV) and new loads, such as electric vehicles (EV) and heat pumps (HP), have increased uncertainty levels in the low voltage distribution system (LVDS). The traditional approach to planning such LVDS is the conservative "fit and forget" approach investing for the worst-case scenario, which means installing a lot of additional infrastructure. However, considering LVDS planning as a stochastic problem is deemed more efficient, reducing grid infrastructure investment cost. Determining the PV hosting capacity (HC) is one such planning problem, yielding the capacity of the system to incorporate new PV resources. Stochastic HC calculations are currently done using computationally expensive, iterative Monte Carlo (MC) based methods, which require solving the power flow equations thousands of times for each iteration of the PV installation scenario. To calculate the PV HC of a large service area consisting of thousands of LV feeders, a computationally tractable and accurate probabilistic power flow (PPF) tool is required. Various approaches exist to make MC-based iterative HC calculation method computationally more efficient, e.g., linearizing the power flow equations, replacing some of the uncertain variables with deterministic values, etc. Alternatively, MC-based methods can be replaced by faster analytic methods, or the iterative HC calculation method can be replaced by stochastic optimization. This thesis aims to build analytical alternatives for MC-based methods and stochastic optimization tools to replace iterative LVDS HC calculation methods.

First, comparing the existing PV HC methods is challenging as they are demonstrated on different test feeders using different assumptions. Therefore, this thesis proposes a benchmark wherein the effects of assumptions made for calculating the HC are evaluated. The comparison of the PV HC obtained from existing methods shows a huge spread, mainly affected by the assumptions on the grid and stochastic limits and the size and number of the PV installations. Standardizing these limits is a major step towards defining a realistic HC of LVDS feeders. As a benchmark, a stochastic limit of 5% and a grid voltage

limit of 0.95-1.05 pu are recommended.

Second, MC simulations are the bottleneck in computing the stochastic PV HC of a large service area. This thesis proposes non-intrusive general Polynomial Chaos (gPC) expansion based PPF as an alternative. The gPC-based PPF outweighs MC and quasi-Monte Carlo (qMC) methods clearly in terms of computational effort with comparable accuracy, as the complex power flow equations are replaced by their polynomial surrogate. The proposed PPF, when used to calculate the congestion probability of a European LVDS, is ten times faster than the MC-based method with the same accuracy.

Third, stochastic HC calculations require an understanding of the uncertainties in LVDS. The operational uncertainties due to load and generation variations and PV scenario uncertainties due to size, type and phase of the PV installations are usually sampled together in MC-based probabilistic HC approaches. This dissertation presents a decoupled approach to calculate the stochastic PV HC, where the probability of violation of the operational limits is computed for possible PV planning scenarios. Then, the PV scenario that results in the highest total PV power installed without violating the stochastic operational limits of the LV feeder is proposed to be the stochastic HC of that feeder. Decoupling the PV (planning) scenarios from the operational uncertainties instead of sampling them together enables to study of the impact of planning policies and operational rules on the overall PV HC.

Fourth, the iterative approach of computing PV HC, where the PV size installed is increased in every iteration, is a computationally demanding process. The decoupled method is also a brute force approach where the probability of congestion is calculated for each possible scenario. In contrast, a chance-constrained stochastic optimal power flow (SOPF) based method reduces the decoupled HC calculation into a single-shot problem. This thesis proposes an intrusive gPC-based SOPF for calculating stochastic PV HC. This method reduces the computation time from days to seconds while giving the upper bound of the HC that can be identified using conventional methods.

Finally, when investigating the needs for a larger region, it is generally too hard to compute the stochastic HC for all individual LV feeders, as a small service area already can have hundreds of feeders. An approximation of the PV HC of the entire service area can be obtained by scaling up the HC of a small number of representative feeders. This thesis presents a clustering scheme that captures the most relevant features of LVDS feeders for PV HC to obtain such representative feeders. A case study shows that the PV HC of a small service area can be approximated by using only 3% of the carefully chosen representative feeders with an error of 20%.

Beknopte samenvatting

De toegenomen installatie van hernieuwbare bronnen, zoals PV-panelen en nieuwe belastingen zoals elektrische voertuigen (EV) en warmtepompen (HP), hebben de onzekerheidsniveaus in het laagspanningsdistributiesysteem (LVDS) vergroot. De traditionele benadering voor het plannen van dergelijke LVDS is de conservatieve 'fit and forget'-benadering waarbij geïnvesteerd wordt voor het worstcasescenario. Dit houdt in dat er vaak veel extra infrastructuur moet worden voorzien. Het beschouwen van LVDS-planning daarentegen als een stochastisch probleem wordt als efficiënter beschouwd, omdat er dan rekening gehouden kan worden met de kans op en het risico horende bij een planningsscenario, dat niet noodzakelijk het worstcasescenario hoeft te zijn, waardoor de investeringskosten voor de netinfrastructuur mogelijk worden verlaagd. Het bepalen van de PV-hostingcapaciteit (HC) is zo'n planningsprobleem: hierbij wordt de capaciteit van het systeem om nieuwe PV-bronnen op te nemen, berekend. Stochastische HC-berekeningen worden momenteel uitgevoerd met behulp van iteratieve, op Monte Carlo (MC) gebaseerde methoden, waarvoor de vermogensstroomvergelijkingen duizenden keren moeten worden opgelost voor elke iteratie van een PV-installatiescenario. Om de PV HC van een groot gebied, bestaande uit duizenden LV-feeders te berekenen, is een probabilistische powerflow (PPF)-tool vereist die nauwkeurig is, maar ook de berekeningen in redelijke tijd kan uitvoeren. Er bestaan verschillende benaderingen om de op MC gebaseerde iteratieve HC-berekeningsmethode rekenkundig efficiënter te maken, bijv. het lineariseren van de vermogensstroomvergelijkingen, het vervangen van enkele van de onzekere variabelen door deterministische waarden, enz. Als alternatief kunnen de op MC gebaseerde methoden worden vervangen door snellere analytische methoden, of de iteratieve HC-berekeningsmethode kan worden vervangen door stochastische optimalisatie. Dit proefschrift heeft tot doel analytische alternatieven te ontwikkelen voor de op MC gebaseerde methoden om iteratieve LVDS HC-berekeningsmethoden te vervangen.

Ten eerste is het een uitdaging om de bestaande PV HC-methoden te vergelijken, omdat ze worden gedemonstreerd op verschillende testfeeders met verschillende

belastingaannames. Daarom wordt in dit proefschrift een benchmark voorgesteld, waarin de effecten van de aannames die zijn gemaakt voor het berekenen van de HC worden geëvalueerd. De vergelijking van de PV HC verkregen uit bestaande methodes laat een enorme spreiding zien, voornamelijk beïnvloed door de aannames over het net en stochastische limieten en de grootte en het aantal van de PV-installaties. Het standaardiseren van deze limieten is een belangrijke stap in de richting van het definiëren van een realistische HC van LVDS-feeders. Als benchmark wordt een stochastische limiet van 5% en een net spanning limiet van 0,95-1,05 pu aanbevolen.

Ten tweede vormen de MC-simulaties het knelpunt bij de klassieke berekening van de stochastische PV HC van een groot gebied. Dit proefschrift stelt een niet-intrusieve algemene Polynomial Chaos (gPC) expansie-gebaseerde PPF voor als alternatief. De op gPC gebaseerde PPF bereikt veel sneller de oplossing in vergelijking met MC en quasi-Monte Carlo (qMC) methoden met vergelijkbare nauwkeurigheid. De reden hiervoor is dat de complexe vermogensstroomvergelijkingen worden vervangen door hun polynoomsurrogaat. Wanneer de voorgestelde PPF wordt gebruikt om het risico op een congestie van een Europees LVDS te berekenen, is deze tien keer sneller dan de op MC gebaseerde methode met dezelfde nauwkeurigheid.

Ten derde vereisen stochastische HC-berekeningen inzicht in de onzekerheden in LVDS. De operationele onzekerheden als gevolg van belastings- en opwekkingsvariaties en onzekerheden in het PV-scenario als gevolg van de grootte, het type en de fase van de PV-installaties worden gewoonlijk samen gesampled in op MC gebaseerde probabilistische HC-benaderingen. Dit proefschrift presenteert een ontkoppelde benadering om de stochastische PV HC te berekenen, waarbij de kans op overschrijding van de operationele limieten wordt berekend voor mogelijke PV-planningsscenario's. Vervolgens wordt het PV-vermogen in het PV-scenario dat resulteert in het hoogste totale geïnstalleerde PV-vermogen zonder de stochastische operationele limieten van de LV-feeder te schenden, de stochastische HC van die feeder. Door de PV-scenario's (planning) los te koppelen van de operationele onzekerheden in plaats van ze samen te nemen, kan de impact van planningsbeleid en operationele regels op de algehele PV HC worden bestudeerd.

Ten vierde is de iteratieve benadering van het berekenen van de PV HC, waarbij de geïnstalleerde PV-grootte bij elke iteratie wordt vergroot, een rekenkundig veeleisend proces. De ontkoppelde methode is ook een brute force-benadering waarbij voor elk mogelijk scenario de kans op congestie wordt berekend. Daarentegen reduceert een chance constrained, op stochastische optimale power flow (SOPF) gebaseerde methode de ontkoppelde HC-berekening tot een enkelvoudig probleem. Dit proefschrift stelt een intrusieve, op gPC gebaseerde SOPF voor als berekeningmethode van de stochastische PV HC.

Het gebruik van deze methode vermindert de rekentijd van dagen tot seconden. Met deze methode wordt de bovengrens van de HC berekend die kan worden geïdentificeerd met behulp van conventionele methoden.

Ten slotte is het over het algemeen moeilijk om de stochastische HC voor alle individuele LV-feeders binnen een bepaalde regio te berekenen, aangezien een kleiner gebied al honderden feeders kan hebben. Een benadering van de PV HC van het gehele gebied kan worden verkregen door de HC van een klein aantal representatieve feeders op te schalen. Dit proefschrift presenteert een clusterschema dat de meest relevante kenmerken van LVDS-feeders voor PV HC vastlegt om dergelijke representatieve feeders te verkrijgen. Een case study toont aan dat de PV HC van een klein gebied kan worden benaderd door slechts 3% van de zorgvuldig gekozen representatieve feeders te gebruiken met een fout van 20%.

List of Abbreviations

ACR	Rectangular Power-Voltage.
BFS	Backward Forward Sweep.
CC	Chance Constraints.
CDF	Cumulative Distribution Function.
CI	Confidence Interval.
CVaR	Conditional Value-at-Risk.
DR	Difference from the Representative feeder.
DSO	Distribution System Operator.
ENWL	Electricity North West.
EV	Electric Vehicles.
GIS	Geographical Information System.
gPC	general Polynomial Chaos.
HC	Hosting Capacity.
HP	Heat Pumps.
IVR	Rectangular Current-Voltage.
kWp	kilowatt-peak.
LV	Low Voltage.
LVDS	Low Voltage Distribution System.
MAAE	Mean Average Absolute Error.
MC	Monte Carlo.
MV	Medium Voltage.
OPF	Optimal Power Flow.
PCE	Polynomial Chaos Expansion.
PDF	Probability Density Function.
PEM	Point Estimate Method.
PF	Power Flow.
PMF	Probability mass function.
PPF	Probabilistic Power Flow.
PV	Photovoltaics.

qMC	quasi Monte Carlo.
QoI	Quantity of Interest.
RES	Renewable Energy Resources.
SM	Smart Meter.
SOPF	Stochastic Optimal Power Flow.
TSO	Transmission System Operator.
TTR	Three-Term Recursive.
UQ	Uncertainty Quantification.

List of Symbols

Operators

$\mathbb{E}[.]$ Expectation operator

$f(.)$ Probability density function

$g(.)$ probabilistic power flow

M Pre-calculated multiplication tensor of orthogonal polynomials

$\mathbb{P}$ Probability operator

$\mathbb{V}[.]$ Variance operator

Indices and sets

$c \in \phi$ Set of phases in a line or a bus

$d \in \mathcal{D}$ Set of LV consumer

$di \in \mathcal{T}^{\mathrm{d}}$ Tuple set of device-d in bus-i

$f \in \mathcal{F}$ Set of all LV feeders in the network

$i \in \mathcal{I}^{\mathrm{ref}}$ Set of reference node

$j, i \in \mathcal{I}$ Set of nodes

$k \in \mathcal{K}$ Set of the coefficient of orthogonal polynomials bases

$l \in \mathcal{L}$ Set of branches in the feeder

$lij \in \mathcal{T}^{\mathrm{l}\rightarrow}$ Tuple set of branch-l from bus-i to bus-j

$lji \in \mathcal{T}^{\mathrm{l}\leftarrow}$ Tuple set of branch-l from bus-j to bus-i

$p \in \mathcal{P}$ Set of installed PV devices

$pi \in \mathcal{T}^{\mathrm{p}}$ Tuple set of PV installation-p connected to bus-i

$s \in \mathcal{S}$ Set of source generator

$si \in \mathcal{T}^{\mathrm{s}}$ Tuple set of source-s and bus-i

$t \in \mathcal{T}$ Set of all transformer in the network

$u \in \mathcal{U}$ Set of units, load, PV or source

$\mathcal{T}^{\mathrm{l}}$ Union Set branch tuples: $\mathcal{T}^{\mathrm{l}\leftarrow} \cup \mathcal{T}^{\mathrm{l}\rightarrow}$

$\mathcal{P}^{\mathrm{rt}}$ Set of available PV ratings for installation

Variables and parameters

b^{sh} Shunt susceptance

δ_j Deterministic parameter j in LVDS

ε *Probabilistic limit* for probability of parameters

E Solar Irradiance

g^{sh} Shunt conductance

HC_f PV hosting capacity of feeder f

$\mathrm{HC}^{\mathrm{pl}}$ PV hosting capacity with respect to Peak load $P^{\max}$

$\mathrm{HC}^{\mathrm{s}}_f$ Stochastic PV hosting capacity of feeder f

HC^{t} PV hosting capacity with respect to $\mathsf{S}^{\mathrm{rated}}_t$

I Current

$\bar{\mathsf{I}}_l = [\mathsf{I}_{l,c}]_{c\in\phi}$ Three phase current vector in line l

I^{im} Imaginary part of current

I^{ov} Ratio of current in a branch to its current rating

I^{re} Real part of Current

I^{rated} Rated current of a line

J Lifted current variable

$\bar{\mathsf{J}}$ vector of lifted current variables

N_p Number of PV installations in feeder f

P	Real Power
P^{kWp}	Deterministic kWp rating of a PV installation
P_u	Real power consumption or generation of a unit $u \in \mathcal{U}$
$P^{\max}$	Peak real power of the network N
$P^{\max}$	Upper limit of real power
$P^{\min}$	Lower limit of real power
P^{nom}	Nominal real power
P^{R}	PV penetration ratio
P^{ref}	Reference Real power
Q	Imaginary Power
$Q^{\max}$	Upper limit of imaginary power
$Q^{\min}$	Lower limit of imaginary power
Q^{ref}	Reference Imaginary power
r^{s}	Series resistance
S	Power
S_t	Load of transformer t in kVA
$\mathsf{S}_t^{\mathrm{rated}}$	Rated kVA of transformer t
U	Voltage
$\mathbf{U}_i = [\mathsf{U}_{i,c}]_{c\in\phi}$	Three phase voltage vector at bus i
$\mathbf{U}^{\mathrm{s}}$	Three-phase source voltage 1x3 vector $\in \mathcal{C}^{1x3}$
U^{im}	Imaginary part of the voltage
$U^{\max}$	Maximum voltage in a feeder
$U^{\min}$	Minimum voltage in a feeder
$\mathsf{U}^{\max,\,\mathrm{th}}$	Threshold voltage for maximum voltage in feeder
$\mathsf{U}^{\min,\,\mathrm{th}}$	Threshold voltage for minimum voltage in feeder
U^{re}	Real part of the voltage

$\mathsf{U}^{\mathrm{ref}}$	Voltage at reference bus
$\mathsf{U}^{\mathrm{ub,\ th}}$	Threshold for voltage unbalanced factor in feeder
W	Lifted voltage
$\overline{\mathsf{W}}$	Vector of Lifted voltage variables
$\mathbf{X}$	Vector of input variables for power flow function
x^{s}	Series reactance
y^{s}	Series admittance
y^{sh}	Shunt admittance
$\boldsymbol{Z}^{\mathrm{n}}$	Network impedance matrix
z^{s}	Series impedance
ξ_i	Stochastic parameter i in LVDS

Contents

Abstract iii

Beknopte samenvatting v

List of Abbreviations x

List of Symbols xiv

Contents xv

List of Figures xix

List of Tables xxiii

1 Introduction 1
- 1.1 Context and Motivation . . . 1
  - 1.1.1 Distribution system Uncertainties . . . 2
  - 1.1.2 Congestion in distribution system . . . 3
  - 1.1.3 Probabilistic Power Flow . . . 4
  - 1.1.4 Stochastic HC calculations using PPF . . . 4
  - 1.1.5 Scaling of LVDS HC . . . 6
- 1.2 Objectives and Scope . . . 6
- 1.3 Project BREGILAB . . . 8
- 1.4 Main contributions . . . 9
- 1.5 Outline . . . 10

2 LVDS modelling and probabilistic power flow analysis 15
- 2.1 Introduction . . . 15
- 2.2 Modelling Three-phase four-wire LVDS . . . 17
  - 2.2.1 Carson's Method and Kron's Reduction . . . 19
  - 2.2.2 Proposed Impedance Transformation . . . 21

2.3 Uncertainty Quantification . . . 23
2.3.1 Stages of Uncertainty Quantification . . . 23
2.3.2 Uncertainty Quantification methods . . . 25
2.3.3 Classical (simulation) methods . . . 25
2.3.4 Analytical methods . . . 26
2.3.5 Other approximate methods . . . 28
2.4 Polynomial chaos expansion . . . 28
2.4.1 Orthogonal Polynomials . . . 29
2.4.2 Multivariate Orthogonal Polynomials . . . 31
2.4.3 Non-Intrusive Polynomial Chaos Expansion . . . 31
2.4.4 Intrusive Polynomial Chaos Expansion . . . 33
2.4.5 Stochastic Galerkin Method . . . 34
2.4.6 General Polynomial Chaos (gPC) . . . 35
2.5 Uncertainties in LVDS . . . 36
2.5.1 Classification of Uncertainties . . . 36
2.5.2 Sources of Uncertainty . . . 38
2.6 Summary and conclusion . . . 45

3 Hosting Capacity of Photovoltaic Systems 47
3.1 Introduction . . . 48
3.2 Hosting capacity definitions . . . 50
3.2.1 Generic HC definition . . . 50
3.2.2 Deterministic HC . . . 52
3.2.3 Stochastic HC . . . 52
3.3 Hosting capacity calculation . . . 53
3.3.1 Deterministic HC calculations . . . 54
3.3.2 Stochastic HC calculations . . . 56
3.3.3 Analytical methods . . . 59
3.4 Reference method for stochastic HC . . . 61
3.4.1 Reference constraints for stochastic HC . . . 63
3.5 Evaluation of HC in Non-synthetic European Network . . . 64
3.5.1 Test Network . . . 64
3.5.2 Sensitivity of the reference method . . . 64
3.5.3 Deterministic HC . . . 66
3.5.4 Stochastic HC . . . 67
3.5.5 Computational advantages . . . 71
3.5.6 Discussion . . . 73
3.6 Conclusion and Future Works . . . 74

4 Probabilistic Power Flow using General Polynomial Chaos 77
4.1 Introduction . . . 78
4.1.1 Background and Motivation . . . 78
4.1.2 Literature Overview . . . 79

4.1.3 Contribution and Organization . . . 82
4.2 Monte Carlo Simulations . . . 83
4.3 General Polynomial Chaos Expansion . . . 85
4.4 Numerical Illustration . . . 87
4.4.1 Test system . . . 89
4.4.2 MC vs gPC: Testing points . . . 90
4.4.3 gPC: Impact of polynomial degree . . . 90
4.4.4 Comparison of gPC, qMC, and MC . . . 93
4.4.5 Network analysis: Impact of uncertainty sources . . . 96
4.5 Conclusion . . . 97

5 Decoupled Stochastic Hosting Capacity 99
5.1 Introduction . . . 100
5.1.1 Motivation and Literature Review . . . 100
5.1.2 Contribution and organization . . . 102
5.2 Decoupling of uncertainties in LVDS . . . 103
5.3 gPC based decoupled PV HC . . . 106
5.4 Numerical results . . . 109
5.4.1 Decoupled gPC-based HC . . . 110
5.4.2 HC calculation for whole network . . . 114
5.5 Conclusion and Future Work . . . 115

6 Stochastic Optimal Power Flow based Hosting Capacity 117
6.1 Introduction . . . 118
6.2 Deterministic HC . . . 121
6.3 Stochastic OPF HC . . . 126
6.4 Numerical Illustration . . . 131
6.4.1 Test Feeder . . . 131
6.4.2 Deterministic OPF HC . . . 132
6.4.3 gPC-CC-OPF PV HC assuming Gaussian output . . . 133
6.4.4 Tuning of moment-based chance-constraints . . . 135
6.4.5 OPF HC vs gPC-CC-OPF HC in actual network . . . 137
6.5 Conclusion . . . 138

7 Scaling Hosting Capacity Using Representative Feeders 139
7.1 Introduction and Motivation . . . 140
7.2 Methodology . . . 141
7.2.1 PV Hosting Capacity . . . 141
7.2.2 Representative Feeders . . . 142
7.2.3 Scaling PV HC . . . 145
7.3 Case Study . . . 146
7.3.1 Role of number of representative feeders in scaling . . . 148
7.3.2 Feature Selection . . . 148

7.3.3 Impact of feature selection 151
7.4 Conclusion 151

8 Conclusion and Future work **155**
8.1 Summary and conclusions 155
8.1.1 LVDS PV HC 156
8.1.2 Calculation tools 157
8.1.3 Scaling of PV HC 158
8.2 Recommendations and Future work 158
8.2.1 LVDS uncertainties 158
8.2.2 LVDS power quality limits 159
8.2.3 Hosting capacity definition and calculations 160
8.2.4 Stochastic optimal power flow 160
8.2.5 Scaling of PV HC 161
8.2.6 Open LVDS data 161
8.2.7 LVDS modelling 161

A LVDS Test network **163**

Curriculum vitae **187**

List of publications **189**

List of Figures

1.1 Position of the author's work (WP2) in the global overview of project BREGILAB 9
1.2 Overview of the chapters of thesis along with the research objective addressed by the chapter 10

2.1 LVDS schemes: North American versus European 17
2.2 Schematics for LVDS with multi-grounded neutral as assumed in Kron's reduction method 18
2.3 Schematics for LVDS with Single-grounded neutral 18
2.4 General stages of Uncertainty Quantification 23
2.5 Types of quantities of interest 24
2.6 UQ (or PPF) methods for power systems 25
2.7 Illustration for MC 33
2.8 Illustration of non-intrusive PCE 33
2.9 Illustration of intrusive PCE 35
2.10 Versatility of Beta distribution to represent positively skewed, bimodal symmetrical, or negatively skewed shapes of consumer load 39
2.11 Different consumer load profile uncertainty models 40
2.12 Different irradiance uncertainty models 42

3.1 Simple representation a three-phase low voltage (LV) feeder . . 50
3.2 Flow chart for HC calculation of a network of LVDS feeders. . . 53
3.3 Uncertainties considered in $\mathrm{M}^{\mathrm{ref}}$ 61
3.4 Single line Diagram of the test network 63
3.5 $\mathbf{M}^{\mathbf{ref}}$ with different *operational limits* (voltage: L1, L2 and L3) and *probabilistic limits* (C00, C01, C05 and C10). 65
3.6 Comparison of HC_{N} obtained from stochastic and deterministic hosting capacity (HC) methods with M^{ref}. 68

3.7 Comparison of different HC methods and their respective reference methods. 69
3.8 Comparison of time for HC calculation, stochastic, deterministic and reference. 71

4.1 Classical and emerging techniques for probabilistic power flow. The methods indicated in blue are the focus of this study . . . 82
4.2 PDF for load and irradiance for typical spring day 88
4.3 CI of mean (dotted) and tail event (dashed) of QoIs obtained from degree 2 gPC based on (a) MC, (b) sobol and (c) halton samples, plotted along with CI from MC (dark gray: mean and light gray: tail QoI) 91
4.4 CI of mean (dotted) and tail event (dashed) of QoIs obtained from degree 1-4 gPC based on sobol samples (a)-(d), plotted along with CI from MC (dark gray: mean and light gray: tail QoI) 92
4.5 Mean and tail values of QoIs obtained from qMC using Sobol and Halton sequences plotted along with CI from MC and gPC 94
4.6 Error percentage of probability of congestion per feeder 95
4.7 Error variation on p^{ov}: qMC vs gPC with number of uncertainties for selected feeders 96

5.1 Decoupled vs reference HC method 104
5.2 HC calculation using decoupled gPC based tool for one feeder based on overvoltage 112
5.3 HC calculation using decoupled gPC based tool for feeder 1 using probability of undervoltage and overcurrent. 112
5.4 Comparison of 6 kWp single phase PV vs 6 kWp three phase PV in the feeder 1. 114
5.5 Position of HC for selected feeder in HC map for the complete network 115

6.1 Simple representation of a three-phase LV feeder 122
6.2 π-model of a branch lij, with series impedance z_l^{s} and from- and to-side shunt admittances y_{lij}^{sh} and y_{lji}^{sh}. 123
6.3 Single phase representation of a LV feeder with all consumer with PV for OPF-HC 124
6.4 Histogram of feasible and infeasible scenarios compared with OPF- HC 133
6.5 PV HC scenarios obtained from OPF and Scenario based method (PV HC for each scenario is given in the legend) 134
6.6 PV HC scenarios obtained from gPC-CC-OPF with different PV size limits 135
6.7 Moment based reformulation of stochastic limits. 136

6.8 Probability of voltage in each node crossing 1.05 pu 136
6.9 Histogram of mismatch between OPF HC and gPC-CC-OPF HC normalized by gPC-CC-OPF HC 137

7.1 Covariance matrix of normalized to standard form used for feature correlationship. 144
7.2 Silhouette coefficient for different sized feeder clusters for 160 Spanish distribution network feeders 146
7.3 MAAE of stochastic PV HC scaling from representative feeders using all features 147
7.4 MAAE of stochastic PV HC from representative feeders after feature selection 152

A.1 Schematic of Non-synthetic European LVDS network 164

List of Tables

2.1 Name and support interval of orthogonal polynomials on *Askey scheme* 30

3.1 Different Deterministic PV HC methods for distribution feeders 55
3.2 Review of stochastic HC calculation literature for different types of uncertainties considered in Low Voltage Distribution System (LVDS) 57
3.3 Different Stochastic HC methods to be evaluated 58
3.4 Mathematical representation of different Stochastic HC methods 60
3.5 Evaluation deterministic HC for European Feeders with respect to $\mathbf{M}^{\mathbf{ref}}$ 67
3.6 Evaluation of SHC in Non-synthetic European Feeder 72

4.1 LV consumer group of test feeder 89
4.2 Computation time and error for gPC with 100 000 final samples compared to MC with the same sample size 93

5.1 Features of selected feeders 110
5.2 Residential PV size scenario for each penetration case 111
5.3 PV hosting capacity: single phase and three phase 113

6.1 Uncertainty PDF for load and irradiance for τ. 131
6.2 Computation time of Scenario based MC, OPF HC, and gPC-CC-OPF HC for test feeder (7 consumer, each consumer can get 0-15 kWp in step of 1 kWh) 132

7.1 Representative feeders selected 147
7.2 Rank of different indices when taking representative feeders, considering all features (C1) and individual features (C2) . . . 149

7.3 Rank of different indices when taking representative feeders considering all features (C1), individual features (C2), and combinations of features (C3) 150
7.4 Comparison after feature selection 150

Chapter 1

Introduction

1.1 Context and Motivation

The energy transition is happening now and accelerating, replacing traditional fossil fuel based generation by electricity generation from renewable energy resources (RES) at an increasingly rapid pace. The speed at which the transition is happening can be seen by comparing the EU directives over time. The initial target in 1997 to have 12% renewables by 2010 has been revised multiple times and is replaced by more ambitious targets in 2021 [1]. By the start of this PhD thesis in 2018, the reference documents for the EU were the 2030 Climate and Energy Policy Framework [2] and the European Commission Renewable Energy Directive [3]. In these directives, Europe had set itself a target to collectively reach a share of at least 27% renewable sources in the final energy consumption mix by 2030, with at least half of the EU's electricity generation coming from renewable sources. The European Council had also endorsed an indicative energy saving target of 27% by 2030. The EU-level target of 'at least 32% cut' in greenhouse gas emissions (from 1990 level) by 2030 was set in those directives. However, the proposal was revised in 2021 [1] by setting new targets for endorsing renewable energy resources. The target of greenhouse gas emission reduction (from 1990 level) by 2030 was revised to 'at least 40% cut'. It also put forward an EU *ambition* to reduce net greenhouse emissions by at least 55% by 2030 and become *climate neutral* by 2050. The target of the overall energy mix supplied by renewable energy sources is revised to 'at least 40%' of the overall energy mix. It means that the steady renewable evolution in recent years and decades must become a revolution.

These future ambitions result in a lot of new renewable resources and new types of emission-reducing loads being introduced in (residential) buildings, such as photovoltaics (PV), heat pumps (HP) and electric vehicles (EV), more than ever. These devices will be directly connected to the Low Voltage Distribution System (LVDS). The 'Fit-for-55' package introduced by the European council [4] sets a binding target of 13% greenhouse gas intensity reduction in transport by 2030. European Environment Agency reports that the electric car registrations for the year 2020 were close to 1 325 000 units, compared to 550 000 units in 2019 [5] which is 11% of total new registrations. A report by Eurelectric mentions that by 2030, 7% to 12% of all cars and trucks in use will be fully electric or hybrid [6]. In a small country such as Belgium, these numbers might be higher, in particular with the Flemish target to phase out the sale of new passenger cars or vans with combustion engine from 2029 in mind, targeting one million EV (roughly 15%) by 2030 [7]. Another estimate believes that the EU alone will have 40-50 million HP and 510 GW of RES of which 70% will be connected to distribution networks [8]. All these revolutions mean that the impact on LVDS will be huge, both in the planning and operational domain.

The 'Fit-for-55' Package envisions this revolution to be supported by an investment of €29bn in power grid enhancement [4]. European Distribution System Operators (E.DSO) argue that the electricity distribution grid is one of the biggest challenges for this EU ambition to realize and forecast that Distribution System Operators (DSOs) in Europe will require a new investment of more than €400bn by 2030 to meet this ambition [8, 9]. Fluvius, the DSO in the Flanders region of Belgium estimates an investment of €4bn on grid infrastructure in 2023-2032 to meet the energy transition goals, with 75% of the total investment anticipated for LVDS [10]. The investment cost is eventually reflected in the consumer's bill and using a risk-based advanced planning tools reduces network enhancement costs which is indeed reflected in the form of reduced bills for the consumer [9].

1.1.1 Distribution system Uncertainties

Most of the abovementioned changes directly affect the distribution system and its operation. This distribution system is fundamentally differently organised and managed than the transmission system. Contrary to transmission systems, they are more distributed, longer, virtually unmonitored and operated in a radial network topology. The energy density of distribution networks is also much smaller. Traditionally, these networks had only one role, supplying the electricity generated by large generators to the end consumers. Most of the loads in those end-consumers were passive household loads, e.g. for lighting and cooking. The philosophy used by the distribution network planner was "*fit and*

forget", where oversized cables were installed in anticipation that future demands would never exceed the cable capacity, even in a worst-case situation. Additional grid investments were done when the system evolved to exceed capacity limits. However, the revolution due to the emergence of HP, EV and RES in the distribution networks introduces not only different loads to the system, but in particular a higher level of stochasticity and a higher risk for synchronicity in consumer load and generation. In this changed scenario, planning for the worst-case snapshot will demand a very high investment cost [11].

Given that these new components can also be controlled to adapt their power exchange pattern and offer flexibility, they can be actively used to manage the grid. The distribution system planners can account for the probability that such a worst-case instant occurs based on the uncertainties on consumer load and generation and account for the available flexibility in the system. Solving the distribution system planning and operational problems based on such probabilities in network will eventually reduce the grid investment cost [12].

The uncertainties in the distribution networks could originate from the consumer behavior (demand uncertainties), RES (generation uncertainties) and/or grid (reference voltage and impedance uncertainties). These uncertainties are time and location dependent. Modelling them and using them to obtain probabilities of worst-case situation in the distribution network is challenging as they increase the complexity of the problem. However, if modelled correctly, the resulting output can be used for stochastic planning and operations.

1.1.2 Congestion in distribution system

Congestion in power grids occurs due to the lack of transmission line capacity to deliver electricity without exceeding voltage, thermal or stability limits. In the case of distribution system, the occurance of congestion is defined by power quality limits proposed in standards such as EN 50160 [13]. Such power quality limits are usually for the voltage at consumer node and current ratings of the branches having fuse or circuit-breakers; for example, the terminal voltage of $\pm 10\%$ around the nominal voltage should be maintained for 95% of the time and current should not cross the overcurrent settings of the circuit-breaker. If the instantaneous power consumption in a network is known along with the network information, a simple deterministic power flow gives the congestion status in the network corresponding with the deterministic input state. However, if the aim is to forecast a future congestion based on the stochastic load and injection variables, the result has to be in terms of the probability of congestion. These congestion risks are referred to as the *stochastic limits*, e.g., the probability of nodal voltage being greater than a prescribed *operational limit* of voltage

should be less than ϵ. The limit ϵ serves as the risk associated with congestion. This congestion risk calculation can be done using Probabilistic Power Flow (PPF) methods.

1.1.3 Probabilistic Power Flow (PPF)

A PPF method considers all the stochastic variables of distribution system and supports the application of *stochastic limits* in the output [14]. The stochastic Quantities of Interest (QoIs), such as the risk of congestion, needs a probability distributions of the output voltage and current. The commonly used deterministic power flow is simple but insufficient to calculate such distributions. From the early days of computing, Monte Carlo (MC) methods were proposed to solve such problems. Using MC methods avoids extensive reformulation of the original problem as the deterministic functions are applied many times, using randomly generated input samples. However, it has a very high computational cost due to the required repetition of complex calculations. Quasi Monte Carlo (qMC) methods are similar to MC methods but uses deterministic samples of input from algorithmic series [15]. Analytical methods like Polynomial Chaos Expansion (PCE) expansion have gained increasing interest in PPF calculations. It has been shown that PCE outweighs other analytical methods used for PPF calculation as this method preserves the non-linearity of the power flow equation [16]. In this thesis, different PPF for LVDS hosting capacity (HC) calculations are compared, and PCE is proposed as a viable solution to reduce the computational cost compared to alternative probabilistic methods such as MC methods.

1.1.4 Stochastic HC calculations using PPF

The most accepted definition of HC is, *the amount of new generation or consumption that can be accommodated on a given feeder without impacting system operation under existing control and infrastructure configuration* [17]. HC calculations that assume a deterministic worst-case scenario result in a deterministic HC. In contrast, consideration of uncertainties in the input and risk based indices in nodal voltages and currents results in what is referred to as stochastic HC [18]. The PV HC an LVDS feeder is the sum of rated PV kWp that can be installed in the feeder:

$$\mathrm{HC}_f = \sum_{p \in \mathcal{P}_f} p. \tag{1.1}$$

The HC of the LVDS is the total additional capacity of PV that the given system can host and is then obtained as sum of HC_f of each feeder in that

LVDS:

$$\mathrm{HC_N} = \sum_{f \in \mathsf{N}} \mathrm{HC}_f. \tag{1.2}$$

PPF calculations are used to propagate the uncertainties in the consumer load and generation to the decision parameters in the power system, such as the risk of congestion. These calculations could be performed for different time-frames: a real-time calculation to mitigate congestion, a day-ahead calculation for load and generation scheduling, or year-ahead planning calculations. One of the important questions raised by distribution network planners is how much more load or generation can be accommodated by an existing network without any grid congestion. This extra available capacity depends on the location, type of load or generation, the possibility of flexible operation of installed devices, and the limits of the grid.

After the review of existing HC methods, stochastic methods like PPF are recommended in [18,19] for photovoltaic (PV) HC calculations of LVDS. However, selecting the best available stochastic method from the literature is difficult as each method is implemented with different feeders, at a different voltage level, with different assumptions on load, PV generation. The inconsistency in the input uncertainties and operational and stochastic limits makes the comparision more complicated. Given the importance of the hosting capacity, both in academic literature and in the development of actual distribution systems, a consistent benchmarking of existing stochastic methods was deemed necessary. In Chapter 3 of this thesis, a consistent comparison of the HC calculation methods in literature is compared with a proposed reference method and applied to large scale LVDS.

The stochastic PV HC calculation methods in the literature are mostly calculated with an iterative approach, using MC based methods. In such approach, the total PV installed in an LVDS feeder is increased at each iteration until the grid operational and stochastic limits are violated. This process is repeated for all the feeders in the LVDS to get $\mathrm{HC_N}$. The MC methods hit the computational limits very fast when the PV HC of a large service area with thousands of LV feeders has to be computed. The common approaches taken reduce computation time are: using deterministic methods [20], using representative feeders only [21], reducing the input source of uncertainties [22], checking fewer operational parameters [23], replacing the power flow equations with analytical transformations [24,25], using simplified power flow equations [26], sampling of both planning and operational uncertainties together [27], etc. A computationally efficient method which relies on very few assumptions and simplifications was observed as a gap in the available literature. In this thesis, computationally efficient stochastic methods to compute the PV HC of an LVDS feeder are proposed. The methods are

based on PCE and the input demands and PV generations are in the form of a univariate stochastic variables, and is presented in Chapters 5 and 6.

Polynomial Chaos Expansion (PCE) is a method for representing a stochastic variable in terms of a polynomial function of stochastic input variables [28]. This method forms the orthogonal polynomial bases from the known input distributions, weighted by respective coefficients to represent stochastic variables in the system. Essentially, the PCE of a stochastic variable is comparable to the Fourier transform of a periodic waveform, where an infinite-dimensional object is represented in terms of finitely many scalar coefficients [29]. The uncertainty propagation using PCE can be *intrusive*, where the physical equations of power flow are intruded with the orthogonal polynomials of the input uncertainties or *non-intrusive*, where the surrogate model of power flow equations is formed using the orthogonal polynomial bases to act as a black-box model. The in-depth discussion on PCE will follow in Chapters 2 and 4. Whereas in Chapter 5, a non-intrusive PCE based PV HC calculation method is presented and in Chapter 6 stochastic optimal HC calculations based on the intrusive PCE is presented.

1.1.5 Scaling of LVDS HC

LVDS feeders represent a large portion of any power system. A small suburban town may have 160 such LV feeders [30], while the total number of LV feeders operated by a single operator can go up by hundreds of thousands. Calculating stochastic HC of LVDS by an operator to get a HC in their entire service area requires repeating this complex calculation for each feeder. However, in literature, HC methods are presented only for a few randomly selected feeders or synthetic feeders to reduce computational effort. The variability in the feeder's features is neglected in such feeders hence, scaling hosting capacity using them is not straightforward. A gap was observed in the literature regarding how to obtain better representative feeders that can be used to obtain the PV HC of an LVDS. A method to obtain the features of LVDS feeders that are relevant for scaling the PV HC is proposed in Chapter 7 of this thesis.

1.2 Objectives and Scope

This thesis aims to tackle the challenges and gaps proposed in the previous section, obtaining a method for calculating stochastic HC of a large scale LVDS, reducing the need for assumptions or simplifications. The thesis uses the capacity of LVDS to absorb additional PV installations as case study, although

the methods are also suitable to investigate other changes in the power system such as the increase of HP and EV. The objectives of this thesis are three-fold. The first is to understand and benchmark existing PV HC calculation methods in a realistic scenario with consistent assumptions and limits, and propose a better way to realize LVDS HC. The second is to explore and propose a better computational method for the LVDS planning problems. For this, PPF and Stochastic Optimal Power Flow (SOPF) are investigated and improved. The third is to explore the possibility of scaling stochastic PV HC calculations for a larger service area with minimum computation efforts and limited data availability.

Objective 1: To identify the appropriate method to calculate PhotoVoltaics Hosting Capacity in low voltage distribution feeders through benchmarking of existing methods

The first goal of this thesis is to investigate the different proposed PV HC methods in literature, benchmark them and to use this exercise to develop a robust definition of PV HC.

The research questions addressed with this objective are:

- What are the different LVDS PV HC calculation methodologies available in the literature and how do they compare? Is there a way to benchmark them in a fair, transparent and efficient manner?
- How sensitive are the calculated stochastic PV HC calculation methods to the grid *operational limits* and *stochastic limits*?
- What are different *planning* and *operational* scenarios in LVDS arising due to the uncertainties from the high amount of PV deployment, and what are the benefits of dealing with these uncertainties separately?
- How can a SOPF be formulated for calculating the LVDS PV HC as one-shot problem to replace the iterative methods?

Objective 2: To reduce computation efforts of calculating stochastic hosting capacity (HC) by using efficient probabilistic power flow (PPF) and stochastic optimal power flow (SOPF)

The second goal of this thesis is to investigate the state-of-art in stochastic methods and propose PPF and SOPF methods, which are faster and more

accurate in calculating stochastic HC of LVDS feeders. The stochastic HC calculations for LVDS feeders in literature are done in an iterative approach using computationally expensive MC method in each iteration. The gain in computational time could be through replacing the MC method with an efficient PPF method or replacing iterative HC calculation with optimization-based methods.

To this end, the following research questions are addressed for this objective:

- What are the essential features of PPF methods for the stochastic PV HC calculations and which methods are efficient for such computation?
- How SOPF based on PCE can be used to determine the stochastic HC of LVDS?

Objective 3: To scale stochastic photovoltaic hosting capacity to a larger region

The third objective of this thesis is to identify the most relevant features which can be used to obtain representative feeders to calculate cumulative PV HC of LVDS of a large service area.

The research questions addressed by this objective are:

- What are the most relevant features of LVDS feeders when scaling the stochastic PV HC spatially for a large service area using representative feeders?
- How many representative feeders are required for scaling the stochastic PV HC?

1.3 Project BREGILAB

This research was done within the broader scope of the Belgian Energy transition fund, project BREGILAB [31] supported by the FPS economy, S.M.E.s, Self-employed and Energy. The study's overall objective is to understand and map the actual PV potential for the Belgian electricity system based on land availability, grid capacity, and market limitations. The aim of the work package (WP2) in which the author was involved was to understand the grid limitations for very high PV penetration in LVDS (see Fig. 1.1).

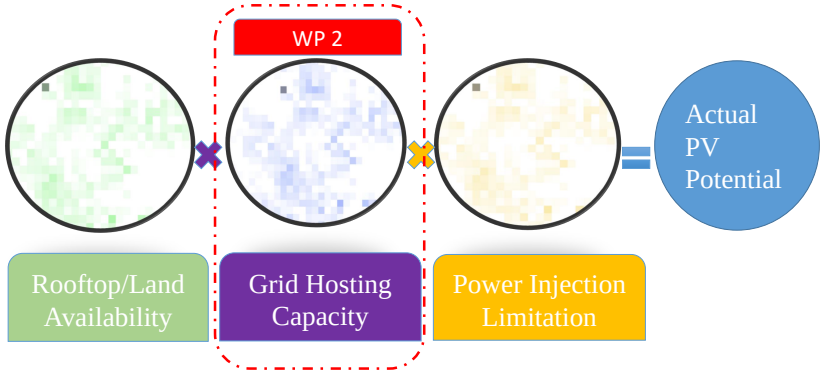


Figure 1.1: Position of the author's work (WP2) in the global overview of project BREGILAB

1.4 Main contributions

The main contributions of this thesis are:

- Benchmark of existing deterministic and stochastic PV HC methods using an actual European LVDS network.
- Recommendation of the best variant of PCE based PPF method for capturing tail-events (risk of congestion) in an LVDS with very high PV penetration.
- Development of a stochastic, decoupled PV HC calculation tool to study the congestion risk of different planning scenarios in LVDS.
- Presentation of a SOPF based PV HC calculation method, considering a stochastic limit, without reliance on sampling, linearization of power flow equations, or relaxations.
- Presentation of a methodology to use representative feeders for scaling the stochastic PV HC of a larger service area.

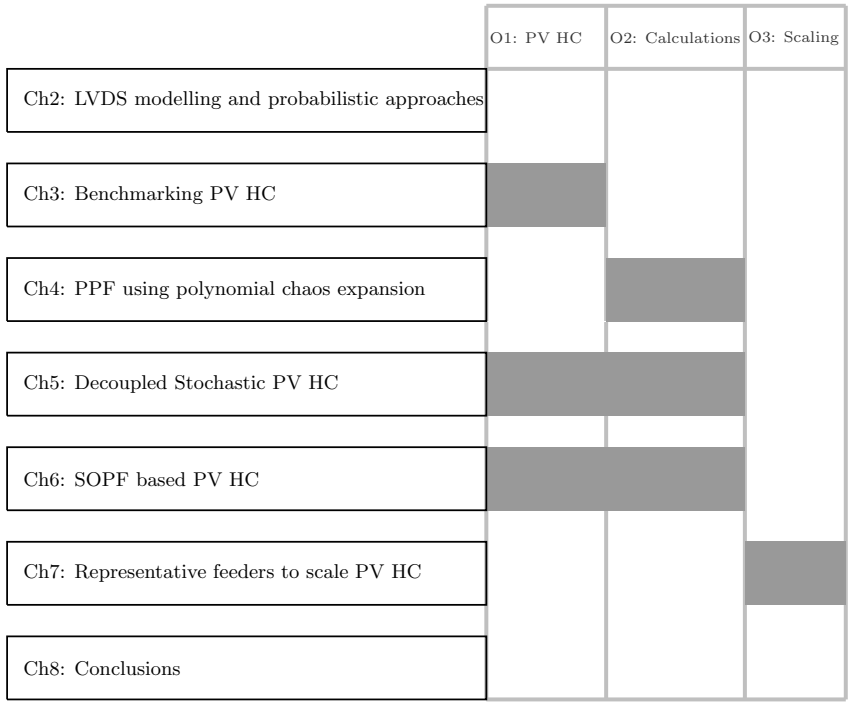


Figure 1.2: Overview of the chapters of thesis along with the research objective addressed by the chapter. The gray shaded region represents respective objectives covered by the particular chapter.

1.5 Outline

The outline of the thesis is visualized in Figure 1.2 with respect to the different outlined objectives. The thesis is split into 8 chapters. After this introductory chapter, the following individual chapters are described below.

Chapter 2: LVDS modelling and probabilistic approaches

This chapter explains the LVDS modelling assumptions taken throughout the paper. It also provides insight into the uncertainties existing in LVDS and different ways of dealing with such uncertainties. This chapter provides the necessary background and aspects regarding data handling, grid modelling and probabilistic tools used in the remainder of this work. This chapter is based

on the following four papers where the author was involved as the main and co-author:

> **Koirala, A.**, D'hulst, R., Van Hertem, D. (2019). *Impedance modelling for European style Distribution Feeder.* Presented at the 2019 International Conference on Smart Energy Systems and Technologies (SEST), Porto, Portugal, Sep 2019. doi:10.1109/SEST.2019.8849015
>
> **Koirala, A.**, Van Acker, T., Van Hertem, D., Camargo, J., D'hulst, R. (2020). *General Polynomial Chaos vs Crude Monte Carlo for Probabilistic Evaluation of Distribution Systems.* Presented at the 16th International Conference on Probabilistic Methods Applied to Power Systems, Liege, Belgium, Aug 2020. doi: 10.1109/PMAPS47429.2020.9183453
>
> Geth, F., Heidari, R., **Koirala, A.** (2022). *Computational Analysis of Impedance Transformations for Four-Wire Power Networks with Sparse Neutral Grounding.* In Proceedings of the Thirteenth ACM International Conference on Future Energy Systems (pp. 105-113). doi: 10.1145/3538637.3538844
>
> Van Acker, T., Geth, F., **Koirala, A.**, Ergun, H. (2022). *General polynomial chaos in the current–voltage formulation of the optimal power flow problem.* Electric Power Systems Research, 211, Art.No. 108472. doi: 10.1016/j.epsr.2022.108472

Chapter 3: Benchmarking PV HC methods of LVDS

This chapter reviews the existing PV HC calculation methods in LVDS. These methods are expressed in a consistent mathematical framework and are used to calculate the PV HC of a typical European LVDS. The calculated PV HC are compared with a reference method which includes the majority of sources of uncertainty existing in the LVDS with very high PV. Finally, a sensitivity analysis of PV HC on the *operational* and *stochastic limits* is performed for the reference method proposed. This chapter is published as:

> **Koirala, A.**, Van Acker, T., D'hulst, R., Van Hertem, D., *Hosting capacity of photovoltaic systems in low voltage distribution systems: A benchmark of deterministic and stochastic approaches.* Renewable and Sustainable Energy Reviews, 155, Art.No. 111899. doi: 10.1016/j.rser.2021.111899

Chapter 4: Probabilistic Power Flow (PPF) using Polynomial Chaos Expansion (PCE)

This chapter discusses a PCE based PPF method, tailored for LVDS applications. It is compared with MC and qMC methods in terms of accuracy and computational efficiency. Different variants and degrees of the polynomial are investigated for non-intrusive PCE. The best variant chosen from the evaluation is used to compute the probability of congestion in a European LVDS with very high PV penetration. This chapter has been published as:

> **Koirala, A.**, Van Acker, T., D'hulst, R., & Van Hertem, D. (2022). *Uncertainty quantification in low voltage distribution grids: Comparing Monte Carlo and general polynomial chaos approaches.* Sustainable Energy, Grids and Networks, Art. number 100763. https://doi.org/10.1016/j.segan.2022.100763

Chapter 5: Decoupled Stochastic PV HC

In this chapter, the sources of uncertainty in LVDS are decoupled into two types: planning uncertainties (due to future PV installation scenarios) and operational uncertainties (due to PV generation and consumer load). The probability of congestion is calculated for each planning scenario, and the HC of LVDS feeders is computed as the scenario with maximum PV installation within the defined stochastic limit of congestion. The PPF introduced in Chapter 4 is used to calculate the probability of congestion. This method allows for observing the impact of planning uncertainties and policy on operational constraints. This chapter has been presented at Power System Computation Conference, PSCC 2022 and published as:

> **Koirala, A.**, Hashmi, Md. U., D'hulst, R., Van Hertem, D., *Decoupled Probabilistic Feeder Hosting Capacity Calculations Using General Polynomial Chaos.* Electric Power Systems Research, 211, Art.No. 108535. doi: 10.1016/j.epsr.2022.108535

Chapter 6: Stochastic Optimal Power Flow (OPF) based PV HC calculation

This chapter proposes a chance-constrained optimization-based hosting capacity calculation technique, which eliminates the necessity of repetitive solving of power flow equations. General Polynomial Chaos expansion is used to translate the input uncertainties defined by the probability density function to the hosting

capacity of the network. Chance constraints are applied for nodal voltages and thermal overloading. A case study for an actual low voltage (LV) feeder shows that the computational time of using the MC based method is reduced from days to the order of seconds without any sampling, relaxation, or linearization. The result is also compared with the PV HC obtained from a deterministic OPF and the decoupled method proposed in Chapter 5. This chapter is submitted for publication as:

> **Koirala, A.**, Van Acker, T., Hashmi, Md. U., D'hulst, R., Van Hertem, D., *Chance-constrained optimization based PV hosting capacity calculation using general Polynomial Chaos.* Prepared to submit in IEEE Transactions of Power System

Chapter 7: Representative feeders to scale PV HC

This chapter aims to capture appropriate clustering schemes and the most relevant features of low voltage distribution feeders so that an accurate estimate of the hosting capacity of a full service area can be calculated by scaling up the hosting capacity of a small number of representative feeders. The stochastic PV HC computed in Chapter 5 is used to scale the PV HC of an European LVDS based on representative feeders. This chapter has been presented at the PES ISGT Europe Conference, Novi Sad, Serbia 2022 as:

> **Koirala, A.**, Hashmi, Md. U., D'hulst, R., Van Hertem, D., *Representative feeders for spatial scaling of stochastic PV hosting capacity*, 2022 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe), 2022, pp. 1-6, doi: 10.1109/ISGT-Europe54678.2022.9960332.

Chapter 8: Conclusions and future work

This chapter contains a summary and the conclusion of the thesis and presents recommendations for future research.

Chapter 2

LVDS modelling and probabilistic power flow analysis

2.1 Introduction

This chapter briefly explains low voltage distribution system (LVDS) modelling challenges and assumptions taken throughout the thesis. It also provides insight into the uncertainties existing in such a system, different ways of dealing with such uncertainties and quantification of those uncertainties to support decision making. Details regarding data handling, grid modelling and probabilistic power flow tools are discussed. This Chapter is based on the publications [32–35] where the author was involved as co-author.

The power system is traditionally split into three parts: generation, transmission and distribution. The generation refers to power stations such as hydro-power plants, coal-powered thermal plants, gas-powered plants, nuclear power plants and centralized renewables such as large, centrally dispatched wind farms. The transmission system is formed by the high voltage lines and nodes that provide mass transportation of the generated electricity towards the load centres and is organized in a predominantly meshed manner. The meshed structure means power could flow over multiple paths and flows depending on the actual generation and load demand pattern. Industrial multi-Megawatt (MW) loads are usually connected to such systems. These systems are operated by a

Transmission System Operator (TSO). The distribution system is the final capillary power system aiming to distribute power from the transmission nodes to the individual consumers in a reduced voltage level. In Europe, these are operated by a Distribution System Operator (DSO).

The distribution system is again split into three parts based on their voltage level:

- **High Voltage (HV) system**: These segments of the distribution systems are responsible for the transportation of power from the transmission system to the regional distribution substations. This system is meshed and operated on a meshed or a weakly meshed structure. The usual voltage level is between 30 kV to 70 kV, while the actual voltage level differs from country to country. Most of the inter-connectors between the distribution and transmission system are at this level.
- **Medium voltage (MV) system**: MV system comprises the distribution networks between 1 kV and 30 kV. These supply power from the sub-transmission substations to the distribution transformers. The networks are meshed but operated radially, e.g., open ring configuration. The structure is meshed to allow a higher security of supply during contingencies. Larger industrial and commercial consumers are connected to the MV level [36].
- **Low voltage (LV) system**: Most residential consumers are connected to the low voltage part of the distribution network. Common Low Voltage Distribution System (LVDS) is a three-phase four-wire system with dual voltage levels, i.e., 400/230 V or 208/120 V. In some countries, a delta-connected three-wire LVDS with a single voltage level of either 230 V or 120 V is common, e.g. the US and in some regions of Belgium. The LVDS could be radial or meshed but is usually operated radially. The three-phase four-wire system has a neutral wire that might be grounded only at the distribution transformer or be sparsely grounded in a few nodes. The focus of this thesis is on the LV part of the distribution system.

The distribution system worldwide can be broadly divided into two major types: North American and European styles (shown in Fig. 2.1). In Europe, each distribution transformer connected to the MV system creates a fully functional LV system serving many consumers and loads. The transformer ratings are generally below 1000 kVA, and each transformer has two to eight radial LV feeders. In most of North America, the distribution system is organised differently. Single-phase transformers with a very short LV side serving a few consumers are common. These single-phase transformers are connected to the four-wire MV system.

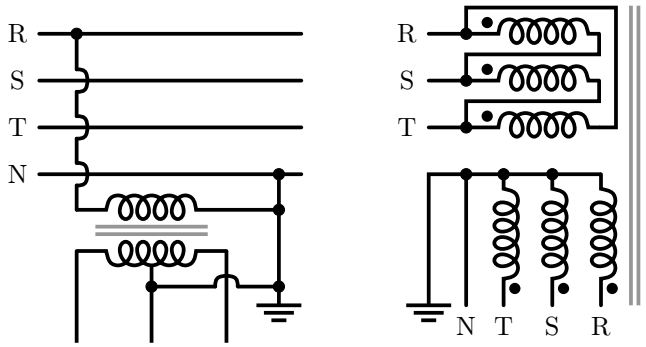


Figure 2.1: LVDS schemes: North American (left) and European (right) [30]

Most textbooks and papers on distribution systems focus mainly on North American style distribution systems. For in-depth details on modelling such systems, *Distribution System Modeling and Analysis* by Kersting [37] is recommended. The focus of this thesis is more on European LVDS. The modelling of European style three-phase four-wire LVDS has some challenges and scope for improvement, which will be discussed in Section 2.2. This section mainly details the newly proposed impedance transformation for sparsely grounded four-wire LVDS in [32] and [33].

Furthermore, the need for Uncertainty Quantification (UQ) in LVDS and different Probabilistic Power Flow (PPF) approaches for UQ in such a system is discussed in Section 2.3. In this thesis, a Polynomial Chaos Expansion (PCE) based analytical methods are proposed for computationally efficient PPF. Section 2.4 discusses the types of PCE and their properties. Identification of uncertainties is very important for any UQ and Section 2.5.2 focuses on the literature on modelling uncertainties in a European LVDS and the assumptions taken in this thesis.

2.2 Modelling Three-phase four-wire LVDS

A three-phase four-wire LVDS has three wires for the three-phase supply and the fourth wire as neutral. Such an LVDS is a three-phase dual-voltage system as two voltage levels are provided (e.g., 400/230 V or 208/120 V). This LVDS can have different schemes for grounding the neutral wire, e.g., *multi-grounded* (see Fig. 2.2), *single-grounded* (see Fig. 2.3), etc. In case the system is operated unbalanced, a neutral current will flow through the neutral wire. The existence of a neutral current means that there will be a non-trivial amount of neutral

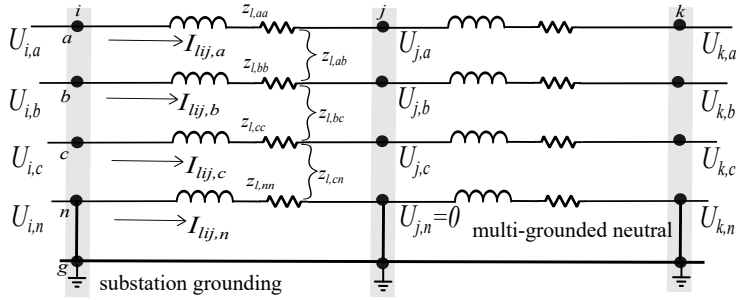


Figure 2.2: Schematics for LVDS with multi-grounded neutral as assumed in Kron's reduction method

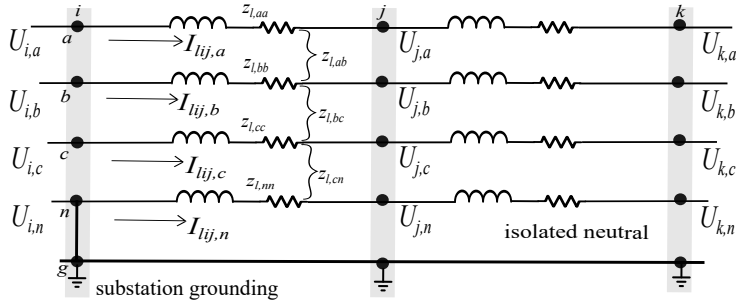


Figure 2.3: Schematics for LVDS with Single-grounded neutral

voltage *shift*, downstream of the transformer [32, 33], impacting the phase-to-neutral voltage. This voltage *shift* will be higher in the *single-grounded* case as the neutral current has to travel throughout the length of the feeder to the substation. The neutral point at the substation is always grounded to avoid *floating-neutral*. Extra grounding in the remaining part of the LVDS feeder usually depends upon the network code, requirements to limit fault currents, or operator practices [38].

Most of the algorithms used to analyse three-phase four-wire systems reduce the four-wire LVDS to a three-wire model [39–42] using Kron's reduction as explained in [43]. This reduces the computational complexity of the power flow or optimal power flow calculations by eliminating the neutral voltage and neutral current variables. Kron's reduction assumes that the neutral voltage at every node is zero. It has been proven that the error introduced by this assumption is minimal for a *multi-grounded* LVDS [43]. However, this reduction cannot accurately capture the neutral voltage rise for a sparsely grounded European

LVDS. The accuracy is worse for a *single-grounded* LVDS with significant load unbalance. Such *single-grounded* LVDS is common in Spain, Belgium, Germany and the Netherlands and has the consumer neutral isolated from the consumer ground. The return current flows back to the substation through the neutral conductor, increasing both neutral voltage drop and neutral current [44]. In such a scenario, Kron's reduction gives a significant error in phase-to-neutral voltage calculated as it neglects the neutral voltage *shift* [45].

To keep the accuracy of the model acceptable, four or five-wire LVDS models are proposed [45–48]. In [45], a 5×5 impedance matrix is proposed where the ground and neutral are treated as two different cables. A reduction of this 5×5 impedance matrix to a 4×4 matrix to reduce computational complexity is proposed in [46]. The reduction is achieved based on the fact that all the unbalanced return current is through the neutral wire, and the ground has no role in this case. The phase to neutral voltage is included in a Backward Forward Sweep (BFS) method when calculating nodal current injection in [47] but uses a 4×4 impedance matrix for the power flow solution. In [48], it is shown that a 4×4 impedance matrix for power flow solutions helps to capture the neutral voltage accurately in the LVDS with *single-grounded* neutral. The four-wire simulations calculate the neutral voltage *shift*. However, it also increases the number of variables to be calculated, increasing the complexity and calculation time.

The European LV test feeder in [49] uses positive, negative, and zero sequence impedance to represent the three-phase four-wire LVDS feeder. However, a method to calculate the sequence impedance of a European style LVDS with sparse grounding is not available in the literature. The sequence impedance can be calculated using Carson's equation as explained in [50] modified for distribution systems as explained in [51]. This method is applicable only for *multi-grounded* LVDS. To this end, we have proposed a new transformation for a *single-grounded* LVDS so that all algorithms designed for three-wire model of *multi-grounded* LVDS could be used without increasing computational complexity [32]. In Section 2.2.1, we briefly describe the traditional Kron's reduction followed by the proposed new transformation in Section 2.2.2.

2.2.1 Carson's Method and Kron's Reduction

Carson used the conductor imaging method in his paper [50] to calculate the impact of the ground return current. Kersting modified it in [43] to give new equations for self-impedance and mutual impedance of each conductor, where

r^{s} represents the self-resistance and $r^{\mathrm{d}} = 0.0953\Omega/mi$ [1], the resistance due to the ground return path of the current. The self impedance of phase a ($z^{\mathrm{s}}_{l,aa}$) and mutual impedance between phase a and b ($z^{\mathrm{s}}_{l,ab}$) is given as:

$$z^{\mathrm{s}}_{l,aa} = r^{\mathrm{s}} + 0.0953 + j0.12134\left[ln\frac{1}{GMR_i} + 7.934\right] \quad \Omega/mi \tag{2.1}$$

$$z^{\mathrm{s}}_{l,ab} = 0.0953 + j0.12134\left[ln\frac{1}{D_{ij}} + 7.934\right] \quad \Omega/mi \tag{2.2}$$

(2.1) and (2.2) are used to calculate the primitive series impedance matrix of a line l i.e., $\mathbf{z}^{\mathrm{s}}_l$ given as:

$$\mathbf{z}^{\mathrm{s}}_l = \begin{bmatrix} z^{\mathrm{s}}_{l,aa} & z^{\mathrm{s}}_{l,ab} & z^{\mathrm{s}}_{l,ac} & z^{\mathrm{s}}_{l,an} \\ z^{\mathrm{s}}_{l,ba} & z^{\mathrm{s}}_{l,bb} & z^{\mathrm{s}}_{l,bc} & z^{\mathrm{s}}_{l,bn} \\ z^{\mathrm{s}}_{l,ca} & z^{\mathrm{s}}_{l,cb} & z^{\mathrm{s}}_{l,cc} & z^{\mathrm{s}}_{l,cn} \\ z^{\mathrm{s}}_{l,na} & z^{\mathrm{s}}_{l,nb} & z^{\mathrm{s}}_{l,nc} & z^{\mathrm{s}}_{l,nn} \end{bmatrix}. \tag{2.3}$$

In the real world, cables and overhead lines induce current into the ground due to capacitive coupling with the earth [37, 46]. On the other hand, the shunt current contribution is limited in normal operation in 50 or 60 HZ due to the low voltage and is neglected for LVDS.

Kirchoff's voltage law applied to four-wire *multi-grounded* LVDS (Fig. 2.2) gives following relationship:

$$\underbrace{\begin{bmatrix} \mathsf{U}_{i,a} \\ \mathsf{U}_{i,b} \\ \mathsf{U}_{i,c} \\ \mathsf{U}_{i,n} \end{bmatrix}}_{\mathbf{U}_i} = \underbrace{\begin{bmatrix} \mathsf{U}_{j,a} \\ \mathsf{U}_{j,b} \\ \mathsf{U}_{j,c} \\ \mathsf{U}_{j,n} \end{bmatrix}}_{\mathbf{U}_j} + \underbrace{\begin{bmatrix} z^{\mathrm{s}}_{l,aa} & z^{\mathrm{s}}_{l,ab} & z^{\mathrm{s}}_{l,ac} & z^{\mathrm{s}}_{l,an} \\ z^{\mathrm{s}}_{l,ba} & z^{\mathrm{s}}_{l,bb} & z^{\mathrm{s}}_{l,bc} & z^{\mathrm{s}}_{l,bn} \\ z^{\mathrm{s}}_{l,ca} & z^{\mathrm{s}}_{l,cb} & z^{\mathrm{s}}_{l,cc} & z^{\mathrm{s}}_{l,cn} \\ z^{\mathrm{s}}_{l,na} & z^{\mathrm{s}}_{l,nb} & z^{\mathrm{s}}_{l,nc} & z^{\mathrm{s}}_{l,nn} \end{bmatrix}}_{\mathbf{z}^{\mathrm{s}}_l} \underbrace{\begin{bmatrix} \mathsf{I}_{lij,a} \\ \mathsf{I}_{lij,b} \\ \mathsf{I}_{lij,c} \\ \mathsf{I}_{lij,n} \end{bmatrix}}_{\mathbf{I}_{lij}} \tag{2.4}$$

The $\mathbf{z}^{\mathrm{s}}_l$ matrix is then reduced by Kron's reduction [43] assuming that the neutral voltage is zero at each node i.e., $\mathsf{U}_{i,n} \approx 0 \approx \mathsf{U}_{j,n}$. (2.4) is partitioned and the neutral voltages are substituted with 0.

$$\begin{bmatrix} \mathbf{U}_{i,abc} \\ \mathbf{U}_{i,n} \end{bmatrix} = \begin{bmatrix} \mathbf{U}_{j,abc} \\ \mathbf{U}_{j,n} \end{bmatrix} + \left[\begin{array}{c|c} \mathbf{z}^{\mathrm{PP}}_{\mathbf{l}} & \mathbf{z}^{\mathrm{PN}}_{\mathbf{l}} \\ \hline \mathbf{z}^{\mathrm{NP}}_{\mathbf{l}} & \mathbf{z}^{\mathrm{s}}_{\mathbf{l,nn}} \end{array}\right] \cdot \begin{bmatrix} \mathbf{I}_{lij,abc} \\ \mathbf{I}_{lij,n} \end{bmatrix} \tag{2.5}$$

$$\mathbf{I}_{lij,n} = -\frac{1}{\boldsymbol{z}^{\mathrm{s}}_{\boldsymbol{l,nn}}}\mathbf{z}^{\mathrm{NP}}_{\mathbf{l}}\mathbf{I}_{lij,abc} \tag{2.6}$$

[1] Note that the imperial unit Ω/mi is used in (2.1) and (2.2), however, it is converted to the SI unit Ω/km when computing the impedance of a European LVDS.

Replacing the value from (2.6) in the top entries of (2.5):

$$\mathbf{U}_{\boldsymbol{i,abc}} = \mathbf{U}_{\boldsymbol{j,abc}} - \left(\mathbf{z}_{\mathbf{l}}^{\mathrm{PP}} - \frac{1}{\boldsymbol{z}_{\boldsymbol{l,nn}}^{\mathrm{s}}} \mathbf{z}_{\mathbf{l}}^{\mathrm{PN}} \mathbf{z}_{\mathbf{l}}^{\mathrm{NP}} \right) \mathbf{I}_{\boldsymbol{lij,abc}} \tag{2.7}$$

Kron's reduction of the neutral is now defined as the transformation $\mathcal{K}$,

$$\mathcal{K}(\mathbf{z}_l^{\mathrm{s}}) = \mathbf{z}_l^{\mathrm{PP}} - \frac{1}{\boldsymbol{z}_{\boldsymbol{l,nn}}^{\mathrm{s}}} \mathbf{z}_{\mathbf{l}}^{\mathrm{PN}} \mathbf{z}_{\mathbf{l}}^{\mathrm{NP}}. \tag{2.8}$$

2.2.2 Proposed Impedance Transformation

For European *single-grounded* LVDS, the assumption of $\mathsf{U}_{i,n} \approx 0 \approx \mathsf{U}_{j,n}$ is not valid, as the multi-grounding condition does not hold. We propose a new impedance transformation in [32] for such an LVDS without any mutual impedance. Furthermore, Geth et al. generalized the transformation proposed where the author was involved as a co-author [33]. The numerical illustration of the validity of this transformation and computational gain is skipped in this thesis, and interested readers are referred to [33].

This new transformation is valid under the following assumptions:

- The neutral grounding is perfect where it occurs, i.e., fixes the voltage at the ground voltage of 0 V (in this case, the transformer neutral).
- Current injections into the *ground* through capacitive effects, i.e. the line shunt admittance $\mathbf{y}_{lij}^{\mathrm{sh}}$ is small and therefore shunt currents are small in LVDS and can be neglected.
- Power consumption or generation devices connected to the network inject a negligible amount of current into the ground. This injection in *single-grounded* LVDS is identified as a fault, protected by ground fault circuit interrupters.

Under these assumptions, the total current through the ground is zero, except for the bus with neutral grounding. Consequently:

$$\mathsf{I}_{lij,g} = 0 = -(\mathsf{I}_{lij,a} + \mathsf{I}_{lij,b} + \mathsf{I}_{lij,c} + \mathsf{I}_{lij,n}), \tag{2.9}$$

and therefore, the neutral current can be expressed as a linear combination of the phase currents:

$$\mathsf{I}_{lij,n} = -(\mathsf{I}_{lij,a} + \mathsf{I}_{lij,b} + \mathsf{I}_{lij,c}). \tag{2.10}$$

Next, the scalar voltage-to-ground variables at bus i in the vector $\mathbf{U_i}$ are used to derive the phase-to-neutral voltages $\mathbf{U_i^{pn}}$:

$$\mathbf{U_i} = \begin{bmatrix} \mathsf{U}_{i,a} \\ \mathsf{U}_{i,b} \\ \mathsf{U}_{i,c} \\ \mathsf{U}_{i,n} \end{bmatrix}, \mathbf{U_i^{pn}} = \begin{bmatrix} \mathsf{U}_{i,an} \\ \mathsf{U}_{i,an} \\ \mathsf{U}_{i,an} \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{bmatrix}}_{\mathsf{T}} \begin{bmatrix} \mathsf{U}_{i,a} \\ \mathsf{U}_{i,b} \\ \mathsf{U}_{i,c} \\ \mathsf{U}_{i,n} \end{bmatrix} = \mathsf{T}\mathbf{U_i}. \tag{2.11}$$

Similarly, the current vector can be written as the transformation of the phase current vector:

$$\mathbf{I_{lij}} = \begin{bmatrix} \mathsf{I}_{lij,a} \\ \mathsf{I}_{lij,b} \\ \mathsf{I}_{lij,c} \\ \mathsf{I}_{lij,n} \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & -1 & -1 \end{bmatrix}}_{\mathsf{T}^{\mathrm{T}}} \underbrace{\begin{bmatrix} \mathsf{I}_{lij,a} \\ \mathsf{I}_{lij,b} \\ \mathsf{I}_{lij,c} \end{bmatrix}}_{\mathbf{I_{lij,abc}}} = \mathsf{T}^{\mathrm{T}}\mathbf{I_{lij,abc}}. \tag{2.12}$$

Multiplying (2.4) with the transformation matrix T:

$$\mathsf{T}\mathbf{U_j} = \mathsf{T}\mathbf{U_i} - \mathsf{T}\mathbf{z}_l^{\mathrm{s}}\mathbf{I_{lij}}, \tag{2.13}$$

and substituting (2.11) and (2.12), the new expression in terms of phase-to-neutral voltage will be:

$$\mathbf{U_j^{pn}} = \mathbf{U_i^{pn}} - \underbrace{\mathsf{T}\mathbf{z}_l^{\mathrm{s}}\mathsf{T}^{\mathrm{T}}}_{\mathcal{T}(\mathbf{z}_l^{\mathrm{s}})}\mathbf{I_{lij,abc}}. \tag{2.14}$$

The transformed impedance matrix $\mathcal{T}(\mathbf{z}_l^{\mathrm{s}})$=$\mathsf{T}\mathbf{z}_l^{\mathrm{s}}\mathsf{T}^{\mathrm{T}}$ is given as:

$$\begin{bmatrix} z^{\mathrm{s}}_{l,aa} - z^{\mathrm{s}}_{l,na} - z^{\mathrm{s}}_{l,an} + z^{\mathrm{s}}_{l,nn} & z^{\mathrm{s}}_{l,ab} - z^{\mathrm{s}}_{l,nb} - z^{\mathrm{s}}_{l,an} + z^{\mathrm{s}}_{l,nn} & z^{\mathrm{s}}_{l,ac} - z^{\mathrm{s}}_{l,nc} - z^{\mathrm{s}}_{l,an} + z^{\mathrm{s}}_{l,nn} \\ z^{\mathrm{s}}_{l,ba} - z^{\mathrm{s}}_{l,na} - z^{\mathrm{s}}_{l,bn} + z^{\mathrm{s}}_{l,nn} & z^{\mathrm{s}}_{l,bb} - z^{\mathrm{s}}_{l,nb} - z^{\mathrm{s}}_{l,bn} + z^{\mathrm{s}}_{l,nn} & z^{\mathrm{s}}_{l,bc} - z^{\mathrm{s}}_{l,nc} - z^{\mathrm{s}}_{l,bn} + z^{\mathrm{s}}_{l,nn} \\ z^{\mathrm{s}}_{l,ca} - z^{\mathrm{s}}_{l,na} - z^{\mathrm{s}}_{l,cn} + z^{\mathrm{s}}_{l,nn} & z^{\mathrm{s}}_{l,cb} - z^{\mathrm{s}}_{l,nb} - z^{\mathrm{s}}_{l,cn} + z^{\mathrm{s}}_{l,nn} & z^{\mathrm{s}}_{l,cc} - z^{\mathrm{s}}_{l,nc} - z^{\mathrm{s}}_{l,cn} + z^{\mathrm{s}}_{l,nn} \end{bmatrix}.$$

With this transformation, the measurement of the neutral voltage is lost, but it can be recovered if required [33]. The advantage of this transformation is that the algorithms and power flow equations developed for *multi-grounded* LVDS with Kron's reduction can be easily reused. Also, it eliminates all the assumptions regarding the neutral voltage to solve power flow equations in European *single-grounded* LVDS without increasing computation time, complexity and number of variables. The remainder of the chapters use this transformation whenever a three-phase unbalanced power flow solution is required.

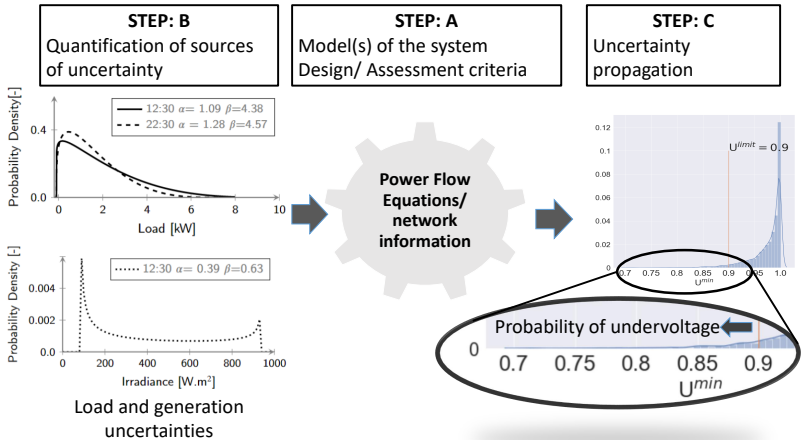


Figure 2.4: General stages of Uncertainty Quantification [52]

.

2.3 Uncertainty Quantification

Uncertainty Quantification (UQ) is the process of propagating uncertainties through a physics based model to obtain the stochastic output variables [52]. In T.J. Sullivan's words, *UQ is the blending of probability theory and statistical practice with 'the real world'* [28]. For most of the power system planning and operation problems the output or *Quantities of Interest (QoIs)* are related to voltages and currents. UQ is an essential tool to obtain the QoIs in the form of Probability Density Function (PDF) of voltages or currents, or the probabilities of the voltages and currents crossing their pre-defined limits. In short, UQ methods enable to propagate the power demand or injection uncertainties through the power flow equations to obtain their influences in the nodal voltages and currents uncertainties. In this section, different stages of a typical UQ problem and commonly used UQ methods are discussed keeping in the mind UQ in LVDS.

2.3.1 Stages of Uncertainty Quantification

The three generalized stages of UQ are visualized in Fig.2.4 with an example for LVDS:

- **Step A:** Define the model and governing physical equations. For power

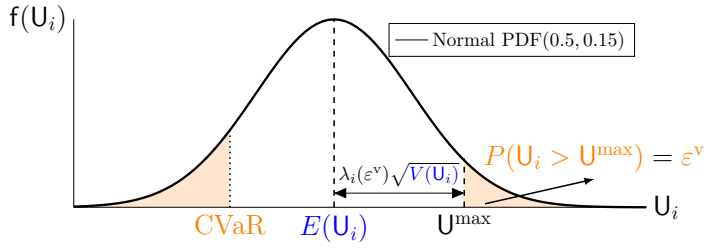


Figure 2.5: Types of QoIs, statistical indices (in blue), risk-based indices (in orange) and graphical representation as PDF (black line).

system applications, the system model for this step is in the form of power flow equations. Booth et el. used the term Probabilistic Power Flow (PPF) for the first time to refer such UQ methods involving power flow equations [53,54]. Specific to LVDS, this step consists of gathering network topology information, unbalanced power flow equations, and consumer information.

- **Step B:** *Uncertainty source quantification.* In simple words, sources of uncertainty are the unknowns in any system. These unknowns could be due to external factors or due to incomplete information about the system. The most important sources of uncertainty in LVDS are discussed in details in Section 2.5.2. The end product will be a vector of random variables taken as the problem's input.

- **Step C:** *Uncertainty propagation* through the physical model equations. This step involves calculating the impact of random input variables in the output, for example, the probability of congestion in LVDS. The different types of QoI expected from any UQ methods or PPF (demonstrated Fig. 2.5) are:

    1. **statistical indices** such as mean and variance of nodal voltage, branch current and distribution transformer load,
    2. **risk-based indices** such as Conditional Value-at-Risk (CVaR) and probability of nodal voltage crossing minimum or maximum limits, and
    3. **graphical output or a function** such as PDF and Cumulative Distribution Function (CDF) of the nodal voltage.

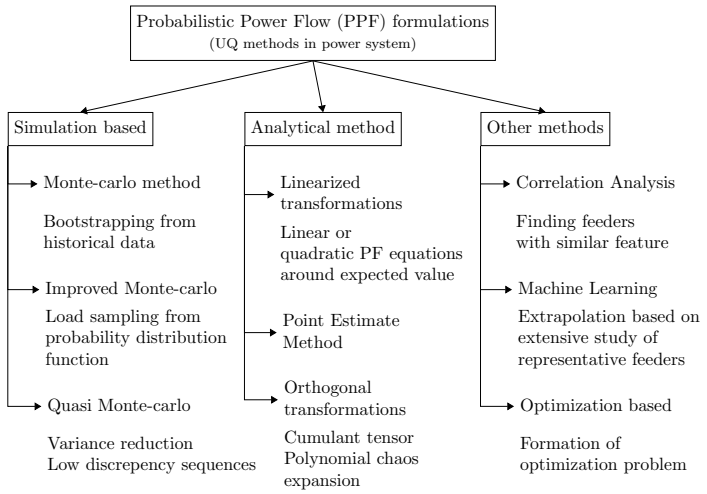


Figure 2.6: UQ methods for power systems (PPF formulations)

2.3.2 Uncertainty Quantification methods

The different methods for UQ can be classified into simulation based methods, analytic methods and other approximate methods (Fig. 2.6). These methods differ in required inputs, accuracy and computational effort. There is not a single best method. The UQ method chosen depends upon the output expected or QoI, the accuracy, or the computational effort [55]. The different UQ methods used in power systems, i.e., PPF (See Fig. 2.6), are briefly described in the following sub-sections.

2.3.3 Classical (simulation) methods

2.3.3.1 Monte Carlo method

Monte Carlo (MC) simulation is the classical simulation based method to solve PPF [53]. In the MC method, a set of random samples are created and passed through the equations representing a deterministic model of the problem to obtain the output samples. This model is ideally based on the physics governing the behavior of the system under investigation. The power flow equations based on first principles represent the system model of the LVDS. The output samples are then processed to get the true QoIs such as probability of congestion. The MC input can be obtained from simple *bootstrapping*, where the input is sampled

from a pool of existing historical data [21, 27]. The data can be fitted to a PDF if historical data are abundant. The input stochastic variables are then sampled using those PDFs and passed through the power-flow equations to obtain the stochastic output variables, which are then used to obtain the QoIs. A MC based method can be used to obtain all three types of QoI.

2.3.3.2 Quasi Monte Carlo

MC methods using variance reduction methods such as importance sampling [56], or using low discrepancy sequences such as Sobol and Halton sequences [57, 58] are referred to as quasi Monte Carlo (qMC) methods. These qMC methods reduce the number of samples required to get the statistical indices QoIs, e.g., mean and variance and risk based indices such as CVaR. However, additional schemes, such as a kernel density estimator, are required if the QoI required is a PDF [59].

The error percentage, also known as the convergence rate of the QoIs, defines the number of samples required for a MC simulation based PPF [60]. For statistical indices such as mean and variance the convergence rate of MC method with N^{mc} samples is $(N^{\mathrm{mc}})^{-0.5}$ [61], while it gets reduced to $(N^{\mathrm{mc}})^{-1}$ on using qMC [62, 63]. In MC methods, the confidence interval (CI) is an alternative measure of accuracy of convergence if the QoIs are statistical indices or risk based indices. The CI of a QoI is obtained by repeating MC simulations with N^{mc} samples for N^{i} instances resulting in N^{i} estimates of the QoI. The 2.5% and 97.5% quartile of the QoI is then calculated to get the lower and the upper bound, and CI is the space between them. The important factor for being satisfied with the estimate of QoIs after MC are that the CI should be sufficiently small compared to the QoI, e.g., it should be less by at least an order. However, the CI cannot be computed for qMC as they use deterministic low discrepancy sequences, resulting in the need for an accuracy estimator [62].

The convergence rate of $(N^{\mathrm{mc}})^{-0.5}$ is usually valid only for the mean value estimation. If the QoIs are risk-based indices, which are usually the events in the tail region of the output PDF, the convergence rate is lower and requires more MC samples [34]. Hence, the computation effort of MC methods for computing risk-based indices is on the higher side compared to computing the mean value.

2.3.4 Analytical methods

Analytical methods rely on transforming the moments of the input uncertainty to the QoIs without the necessity to solve the power flow equations repetitively. For

example, Mulenga et el. in [25] use the impedance transfer function to translate the input uncertainties to the nodal voltage uncertainty. This replacement reduces the necessity of solving power flow equations, but no comments are made about accuracy and gain in computational effort. In [24], a Hilbert transform-based method is presented, which is designed to translate the parameters of Beta PDFs representing the stochastic input variables to the parameters of the Beta PDFs representing the output nodal voltages and branch currents. However, the power flow equations are linearised using Taylor's approximation. The PDF of the output variables are regenerated, assuming they follow the Beta distribution. In [64], the Gaussian Component Combination Method is used where the PDF of wind generation is represented by a combination of Gaussian distributions. However, this approximation is only advantageous when mean values are calculated as part of the original distribution is discarded.

2.3.4.1 Point Estimate Method

The Point Estimate Method (PEM) [65–69] maps selected points from PDFs of stochastic input variables to make an estimate of the QoIs. The power flow equations are solved at the selected three points for each random input, with a weight provided for each point. The final output expectation is obtained as the weighted sum of the expected value at each point. Variations of this method differ regarding the regeneration of the PDFs of the output variables and the number of samples used. In [65], only two points per stochastic input variable are used, and the results are presented in the form of mean and standard deviation. In [66], the three points are used to do analytic power flow, and the resulting PDF is reconstructed using the Gram-Charlier series while requiring only one extra deterministic power flow. Rodriguez et al. in [67] shows that the Cornish-Fischer expansion is more suitable than the Gram-Charlier series for the PDF re-construction with non-Gaussian inputs. In [68], the authors compare different methods of PDF reconstruction after PEM and conclude that all those methods are suitable if the inputs are Gaussian-like distributions. The authors also conclude that such analytic methods would not estimate the nodal voltages efficiently in a system with high penetration of renewable energy sources (RES) and suggest exploring orthogonal transformation-based methods.

2.3.4.2 Orthogonal transformation based methods

An unscented transformation method [70], cumulant tensor method [71], and PCE based method [72–74] are orthogonal transformation based methods used for uncertainty propagation in LVDS. These methods are based on the conversion of the infinite-dimensional space of uncertain variables represented

by a continuous PDF to a finite-dimensional space represented by orthogonal polynomials or vectors weighted by respective coefficients. Of the proposed methods, the PCE based method is more flexible to obtain all three types of QoI similar to MC method without loss of accuracy [75, 76] and hence will be the tool mostly used and investigated in this thesis. Section 2.4 details more on the theory related to PCE based methods.

2.3.5 Other approximate methods

Other approximate methods of handling uncertainty are through indirect approaches such as correlation analysis, and machine learning. In [77], the relation between the LVDS feeder and the hosting capacity (HC) is mapped using a linear regression function based on the features of the feeder. Machine learning tools like neural networks are recently used for minimizing the cost of generation under uncertainties as in [78]. In [79], a deep neural network (DNN) is used to approximate the power flow calculation, making a surrogate model to reduce computational effort of MC based methods. However, the authors mostly focus on building a DNN than exploring the potential of this method in the UQ.

Robust-optimization methods are used to find a solution which satisfies all the possible scenarios in a LVDS [80]. The QoI in such methods are usually the cost of generation or optimal set-point of voltage control devices. The optimization of a power system with stochastic variables and limits is refereed to as Stochastic Optimal Power Flow (SOPF) where the QoI is the risk-averse cost of generators, optimal HC of the system without congestion, etc. One such application is discussed in detail in Chapter 6, where SOPF is used to calculate the optimal hosting capacity of an LVDS.

2.4 Polynomial chaos expansion

Polynomial Chaos Expansion (PCE) is a method for representing a stochastic variable in terms of a polynomial function of stochastic input variables [28]. This method forms the orthogonal polynomial bases from the known input distributions, weighted by respective coefficients to represent stochastic variables in the system. The PCE for a random variable is similar to the Fourier series for a periodic signal: a way to represent infinite-dimensional variables in terms of a finite number of coefficients [29]. The sum of the product of these coefficients, along with their orthogonal polynomial bases, represents an infinite-dimensional random variable. When there is more than one independent uncertainty in the

considered system, multi-variate orthogonal polynomials obtained by tensor product (commonly termed as Hilbert space) of univariate polynomials form the base for uncertain parameters and the variables [28]. PCE can be done in two ways:

a) a **non-intrusive** way, where the power flow equations are solved for specific test points and the scalar QoI is fitted to the orthogonal polynomials to obtain the coefficients, or

b) an **intrusive way**, where the random variable represented by an affine policy is projected into physical equations using Galerkin's projection in the Hilbert space.

In PCE, a stochastic QoI y is expressed as the sum of the basis polynomials of the stochastic inputs weighted by respective coefficients as in (2.15) which can extend up to infinity:

$$y = \sum_{k=1}^{\infty} c_k \Phi_k, \tag{2.15}$$

where c_k and Φ_k denote the coefficient and basis polynomial, respectively.

2.4.1 Orthogonal Polynomials

The elements of an univariate orthogonal vector $\boldsymbol{\phi} = \{\phi_k\}_{k\in\mathcal{K}}$ of an uncertain parameter ξ_i are called an ordered set of monic orthogonal polynomials with respect to ξ_i if it satisfies the following [81]:

$$\phi_0 = 1, \tag{2.16}$$

$$\phi_{k+1}(x) = (x - A_k)\phi_k(x) - B_k\phi_{k-1}(x),\ A_k, B_k \in \mathbb{R}, \forall k \in \mathcal{K}\backslash\{0\} \tag{2.17}$$

$$\langle \phi_l, \phi_k \rangle = \int_{\mathbb{R}} \phi_l(x)\phi_k(x) d\xi_i(x) = \gamma_l \delta_{lk},\ \forall k, l \in \mathcal{K}, \tag{2.18}$$

$$\gamma_l > 0, \tag{2.19}$$

where δ_{lk} is the Kronecker-delta and the numbers γ_k are normalization constants. The index k is a natural ordering of the elements of $\boldsymbol{\phi}$ based on the degree of the polynomial, with the first polynomial always being ϕ_0=1 [82]. The property (2.17), is the classical way of generating orthogonal polynomials, also commonly known as the Three-Term Recursive (TTR) relation among the monic polynomials. The initial conditions used for TTR relation are: ϕ_{-1}=0 and ϕ_0=1, and A_k and B_k are real constants which are uniquely determined by

the type of PDF [83]. Property (2.18) is used to construct the Hilbert space of the orthogonal polynomials as $\gamma_l = ||\phi_k||^2$ represents the norm of the l-th orthogonal polynomials. The Kronecker delta δ_{lk} is a function of two variables: $\delta_{lk} = 1\ \forall l \neq k$, and $\delta_{lk} = 0\ \forall l = k$.

For a stochastic variable ξ_i, represented by a Normal distribution with mean and variance μ and σ, the monic orthogonal polynomials, also known as Hermite polynomials, are given as:

$$\begin{aligned} \phi_0 &= 1, \\ \phi_1 &= x - \mu_1, \\ \phi_2 &= x^2 - \sigma^2 - 2\mu\sigma + \mu^2 \end{aligned} \tag{2.20}$$

The interest in orthogonal polynomial was raised in the 19th century when Chebyshev studied the least-squares problem to fit experimental data [84]. For well-defined PDFs, the orthogonal polynomials are well defined as summarized in Table 2.1. The first four schemes in Table 2.1, Hermite, Jacobi, Legendre, and Laguerre, are also known as classical orthogonal polynomials. Askey and Wilson in their work [85] proposed a group of orthogonal polynomials, whose name are listed in Table 2.1. This group of orthogonal polynomials is collectively known as the *Askey scheme*, and are considered to be a complete set of hyper-geometric orthogonal polynomials [82]. Details on these hyper-geometric sequences are not within the scope of this thesis and interested readers are pointed to the book *Hypergeometric Orthogonal Polynomials and their q-Analagous* by Koekoek, Lesky & Swarttouw [86] for details in the current extension of the *Askey scheme*.

Table 2.1: Name and support interval of orthogonal polynomials on the *Askey scheme* [85, 87]

	Type	Support Interval	Name of polynomial basis
Continuous	Gaussian	$(-\infty,\infty)$	Hermite
	Beta	[a,b]	Jacobi
	Gamma	$[0,\infty)$	Laguerre
	Uniform	[a,b]	Legendre
Discrete	Poisson	{0,1,2,...}	Charlier
	Binomial	{0,1,...,N}	Krawtchouk
	Negative binomial	{0,1,2,...}	Meixner
	Hypergeometric	{0,1,...,N}	Hahn

2.4.2 Multivariate Orthogonal Polynomials

Multivariate polynomials (Φ_α) are obtained by the tensor product of the univariate orthogonal polynomials of two or more independent random variables. The set of multivariate polynomials truncated to degree β is,

$$\mathbf{\Phi}_\alpha = \left\{\Phi_1, \Phi_2, ...\Phi_\alpha\right\}_\beta \tag{2.21}$$

If the number of random variables is N, the total number of orthogonal polynomials truncated at β degree is given as $\alpha = \frac{(\beta+N)!}{\beta!N!}$. The degree of the multivariate orthogonal polynomial Φ_α is given by the sum of the degree of the constituting individual univariate orthogonal polynomials. There is no natural ordering for multivariate polynomials [29]. The multivariate orthogonal polynomials satisfy the inner product relationship in (2.18) with respect to the joint PDF.

Lets assume two Normal distributions ξ_1 and ξ_2, with mean and variance (μ_1, σ_1) and (μ_2, σ_2) respectively. Using TTR, the univariate orthogonal polynomials are:

$$\boldsymbol{\phi}_1[\xi_1] = \{1,\, x_1 - \mu_1,\, x_1^2 - \sigma_1^2 - 2\mu_1\sigma_1 + \mu_1^2,\, ...\} \tag{2.22}$$

$$\boldsymbol{\phi}_2[\xi_2] = \{1,\, x_2 - \mu_2,\, x_2^2 - \sigma_2^2 - 2\mu_2\sigma_2 + \mu_2^2,\, ...\} \tag{2.23}$$

The multivariate orthogonal polynomials is obtained by the tensor product of (2.22) and (2.23). The set of multivariate polynomials truncated at degree two are:

$$\begin{aligned}
&\Phi_1 : 1,\\
&\Phi_2 : x_1 - \mu_1,\\
&\Phi_3 : x_2 - \mu_2,\\
&\Phi_4 : x_1x_2 - \mu_1x_2 - \mu_2x_1 + \mu_1\mu_2,\\
&\Phi_5 : x_1^2 - \sigma_1^2 - 2\mu_1\sigma_1 + \mu_1^2,\\
&\Phi_6 : x_2^2 - \sigma_2^2 - 2\mu_2\sigma_2 + \mu_2^2,
\end{aligned} \tag{2.24}$$

2.4.3 Non-Intrusive Polynomial Chaos Expansion

Non-intrusive PCE is one of the ways to create the surrogate model of a complex physical system, in our case, power flow equations [52, 88]. Assume G as a physical system model, e.g., the power flow equations. A surrogate model $\widetilde{G}$ is the polynomial approximation of G. All the surrogate modelling schemes, including non-intrusive PCE, have three steps:

- *Selection and design of experiment*: In this step, test points are chosen from the information about the input uncertainties ($\boldsymbol{\Xi}$). These test-points T^{gq} could be: a) the Gaussian nodes with weights [34, 72] or b) selected samples obtained by importance sampling or low discrepancy series [88].
- *Evaluation of computational model*: After the T^{gq} are defined and obtained, the next step is to pass that vector through the physical equations, i.e., power flow equations, to evaluate the QoIs. This step is similar to MC simulation but with relatively fewer samples.
- *Formation of surrogate model* $\widetilde{G}$: Finally, based on the test-points T^{gq} and their respective *evaluations*, the surrogate model of the actual physical model is formed. This is done by Gaussian quadrature integration if test-points are Gaussian nodes and weights or by using regression with least-square minimization if the T^{gq} is formed by importance sampling.

The physical model with complex differentials is replaced by a surrogate polynomial representing the model. The difference of using non-intrusive PCE compared to simple regression fitting is the selection of the polynomial bases. In PCE, the coefficient fitting is done for the multi-variate orthogonal polynomials as in (2.24), to form the surrogate PCE for QoI as in (2.15). The surrogate model for the power system can be sufficiently represented by second-degree orthogonal polynomial bases with a truncation error ϵ^{m} [89,90]. Equation (2.15) that represents the model G can be rewritten to represent its surrogate as:

$$y = G(\boldsymbol{\Xi}) = \sum_{k=1}^{\alpha} c_k \Phi_k + \epsilon^{\mathrm{m}}, \tag{2.25}$$

$$y \approx \widetilde{G}(\boldsymbol{\Xi}) = \sum_{k=1}^{\alpha} c_k \Phi_k. \tag{2.26}$$

The test-points T^{gq} and the corresponding evaluations $G(\mathsf{T}^{\mathrm{gq}})$ are now used to find the coefficient-vector $\boldsymbol{y}_\alpha = [c_1, c_2, ..., c_\alpha]$ by solving following linear equations:

$$\boldsymbol{y}_\alpha \cdot \boldsymbol{\Phi}_\alpha(\mathsf{T}^{\mathrm{gq}}) - G(\mathsf{T}^{\mathrm{gq}}) = 0. \tag{2.27}$$

The difference between non-intrusive PCE and MC is illustrated using Figs. 2.7 and 2.8. In MC simulation, the randomly generated samples from the uncertainty sources are passed through the power flow equations to get QoIs. In non-intrusive PCE, the test points are selected and power flow equations are solved at those points to get respective QoIs. A surrogate polynomial based on orthogonal polynomials is formed to represent each QoI. The PDF of the QoI is then obtained by passing the MC samples through the surrogate polynomials. Since

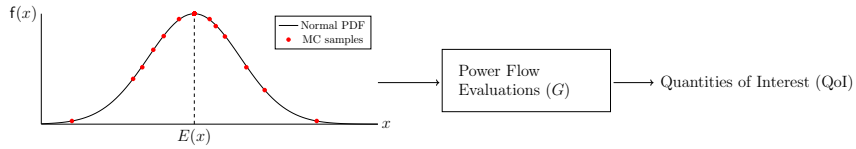


Figure 2.7: Illustration for MC

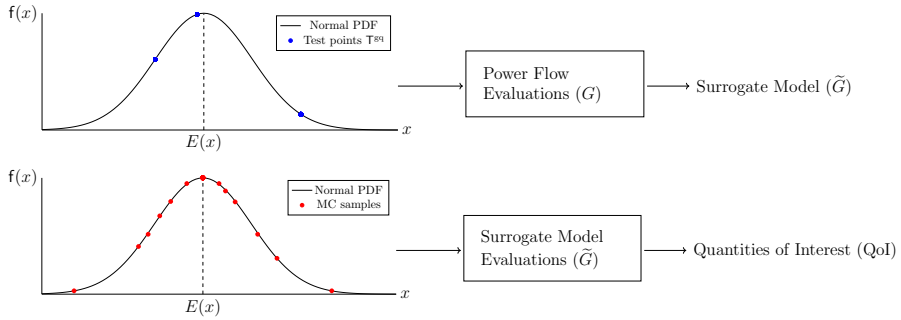


Figure 2.8: Illustration of non-intrusive PCE

the final step is similar to MC method, all three types of QoI can be evaluated using non-intrusive PCE [88].

Non-intrusive PCE is widely used in applications where there is limited knowledge about the system model, or where the system model is computationally expensive. The advantage of this method is that the physics based equations do not need to be re-formulated and can be replaced by a black box model represented by a polynomial. Application of non-intrusive PCE can be seen in a wide array of fields such as structural engineering [91], fluid dynamics [92], and aeronautics [93]. Non-intrusive PCE and their variants are explored in detail in Chapter 4 for LVDS application, where the objective will be to use it to obtain risk based indices such as probability of congestion.

2.4.4 Intrusive Polynomial Chaos Expansion

In non-intrusive PCE, the power flow equations are treated as a black-box as they are not *intruded*. The surrogate model of the system is formed based on test points and their respective evaluations. On the other hand, in intrusive PCE, the polynomial representation of the uncertainty sources is projected on physics based equations requiring reformulation of those equations. It means writing a

new set of equations for the system under investigation, which is unnecessary in non-intrusive PCE. However, with this reformulation, the system equations with stochastic variables can be solved as a single-shot problem without the need of any sampling [89].

In intrusive PCE, the input uncertainties represented by a PDF are first converted into polynomials. If the uncertainties are independent, the known PDF can be represented by orthogonal polynomials weighed by the coefficients. The approximated PCE $\hat{\mathsf{x}}$ of stochastic variable x of a finite variance can be written as a linear combination of orthogonal basis $[\Phi_k]_{k\in\mathcal{K}}$,

$$\hat{\mathsf{x}} = \sum_{k\in\mathcal{K}} x_k \Phi_k = \sum_{k\in\mathcal{K}} \frac{\langle \mathsf{x}, \Phi_k \rangle}{\langle \Phi_k, \Phi_k \rangle} \Phi_k, \tag{2.28}$$

where x_k denotes the PCE coefficients of the approximated random variable $\hat{\mathsf{x}}$. The x_k of a given stochastic variable x is given by the inner product of x and the corresponding polynomial basis function Φ_k divided by the inner product of the Φ_k with itself as given in (2.28). The expectation $\mathbb{E}$ and variance $\mathbb{V}$ of $\hat{\mathsf{x}}$ are then,

$$\mathbb{E}[\hat{\mathsf{x}}] = x_0, \tag{2.29}$$

$$\mathbb{V}[\hat{\mathsf{x}}] = \sum_{k\in\mathcal{K}\setminus\{0\}} \langle \Phi_k, \Phi_k \rangle x_k^2. \tag{2.30}$$

2.4.5 Stochastic Galerkin Method

The Galerkin projection is used to propagate the approximated stochastic variables to the set of non-linear physical equations [28], i.e., power flow equations, to obtain the vector of unknown stochastic variables or QoIs. There are two distinct operations: summation and multiplication. Let $\hat{\mathsf{x}}$, $\hat{\mathsf{y}}$, and $\hat{\mathsf{z}}$ be three stochastic variables defined using the same orthogonal polynomial basis. The Galerkin projections for both operations are,

$$\hat{\mathsf{z}} = \hat{\mathsf{x}} + \hat{\mathsf{y}} \overset{\text{gp}}{\rightarrow} z_k = x_k + y_k, \qquad \forall k \in \mathcal{K}, \tag{2.31}$$

$$\hat{\mathsf{z}} = \hat{\mathsf{x}} \cdot \hat{\mathsf{y}} \overset{\text{gp}}{\rightarrow} z_k = \sum_{k_1,k_2\in\mathcal{K}} M(x_{k_1} \cdot y_{k_2}), \quad \forall k \in \mathcal{K}, \tag{2.32}$$

where M denotes the multiplication tensor of the orthogonal polynomials,

$$M = \frac{\langle \Phi_{k_1}, \Phi_{k_2}, \Phi_k \rangle}{\langle \Phi_k, \Phi_k \rangle}, \tag{2.33}$$

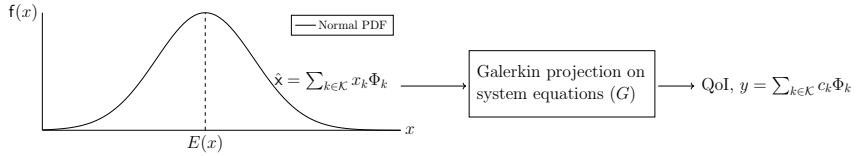


Figure 2.9: Illustration of intrusive PCE

and can be computed ahead of time. The Galerkin projection translates the system equations by employing $|\mathcal{K}|$ deterministic relationships to get a polynomial realization of the QoI as illustrated in Fig. 2.9 [35]. The intrusive PCE can be used to obtain the simple output realization under uncertainty, i.e., UQ, or PPF, or can be extended to obtain optimal operation points under uncertainty, i.e., SOPF.

The intrusive PCE is complex due to the necessity to rewrite all physics based system equations and has a relatively higher CPU time than non-intrusive PCE on UQ problems [94]. However, there are advantages when solving optimization problems, such as shape optimization under uncertainties [95], stochastic optimal AC power flow [89] and topology optimization [96]. In Chapter 6 of this thesis, intrusive PCE based SOPF proposed in [89] is improved and expanded for calculating stochastic optimal photovoltaics (PV) HC analytically without any sampling, linearization, and relaxations.

2.4.6 General Polynomial Chaos (gPC)

PCE was introduced by Wiener in 1938 [97], where Hermite polynomials with Gaussian distributions were used. The general Polynomial Chaos (gPC) expansion was introduced as a generalized form of PCE, valid for *Askey schemes* of orthogonal polynomials. The term gPC expansion was framed by Xiu and Karniadakis [87] when they extended PCE for general hyper-geometric orthogonal polynomials. In this thesis, we deal primarily with non-Gaussian PDFs. Hence, the term gPC is used in the remaining of the chapter as a correct abbreviation. However, in many works in literature, PCE and gPC are being used interchangeably.

2.5 Uncertainties in LVDS

In simple words, uncertainties are the unknowns in any system. These unknowns could be due to external factors or due to incomplete information about the system. As mentioned Section 2.3, identification of those uncertainties and their sources is an important step in any UQ problem. In this thesis, we focus mainly on the uncertainties existing in LVDS.

2.5.1 Classification of Uncertainties

The sources of uncertainties in any physical problem can always be broadly classified into *aleatoric* and *epistemic* uncertainties [52]:

- *aleatoric uncertainty* refers to the natural variability in the phenomenon under consideration. For LVDS, consumer momentary active power consumption, PV active power generation, MV voltage, etc. fall under this category. Such uncertainties are mostly well-defined and represented as known PDFs or in a non-parametric way as a CDF.
- *epistemic uncertainty* refers to the lack of knowledge. In the case of LVDS, all the planning uncertainties (PV/EV/HP size, their type and location in the LVDS) are epistemic. Along with that, the network topology and connected phase of the single-phase loads are the epistemic uncertainties. However, the scarce data due to the lack of measurement devices or limited measurements also causes such uncertainties.

This classification is vague, and most of the work in LVDS dealing with uncertainties does not differentiate between these two except for [25]. In [25], it is pointed out that both of these types are essential in LVDS. It is very reasonable to have a robust solution for the operational problem, such as optimal tap setting, or real-time congestion management, which satisfies all possible scenarios. However, for a planning problem such as PV HC calculation, and optimal storage allocation, scenarios of PV or storage deployment are dependent on the future policies or market. For such a problem, defining an empirical uncertainty for future installations and obtaining a robust solution for such uncertainty does not give detailed information about the best possible solution for future PV placements. The better approach would be to find the best planning scenario from the uncertainty space or a set of planning scenarios which satisfy the existing *grid limits* to a certain *stochastic limit*. Such a solution will be a driver for future policy and disruptive technology.

As such, we propose a new classification of the uncertainties for LVDS: a) *operational*, b) *planning* and c) *feeder* uncertainties as a more reasonable classification. With this division, the three types of stochastic problems usually existing in LVDS, i.e., operational, planning and scaling to a large service area, can be solved decoupled. First, it can be used to obtain real-time operational set points such as optimal tap settings and PV inverter set-points, where the only stochastic variables are the *operational* variables. Second, the impact of uncertain planning scenarios on the operational *grid limits* and *stochastic limits* given by the resulting probability of congestion can be easily observed. In such problems, both *planning* and *operational* uncertain variables are decoupled. Third, it can be easily used for building operational guidelines, strategy for new PV/electric vehicles (EV)/heat pumps (HP) deployments, or investment policy robust to all feeder uncertainties within the *grid limit* and *stochastic limits*. Detailed example of one such advantage of this classification is shown in Chapter 5, where uncertainties in *planning* and *operational* variables are decoupled to obtain maximal PV HC of a given LVDS.

The different uncertainties in this new classification are:

- *operational uncertainties*: PV generation, active power demand of consumers, MV voltage magnitude, HP/EV load, etc.
- *planning uncertainties*: Future PV/EV/HP size, type, phase and location.
- *feeder uncertainties*: The modelling uncertainties of a feeder (impedance, topology, etc.) and characteristics of feeders (length, number of consumers, peak and the annual load of feeder,

Modelling these uncertainties in the presence of abundant data is purely a statistical and computational problem. In the case of LVDS with few smart meters (SMs), the measurement is limited and the uncertain variables need to be modelled. For example, the operational uncertainties have to be framed based on the limited pool of measurement data [25–27, 98]. Similarly, the planning scenarios have to be framed based on practical limitations, e.g., residential PV cannot be more than 15 kilowatt-peak (kWp), each consumer can have at most one PV installation [23], etc. These assumptions reduce the possible scenarios to find a solution within the limited search-space. The one feature of this classification we often exploit in this thesis is that the *operational* uncertainties can be represented by a continuous PDF. In contrast, the *planning* uncertainties are mostly discrete. This separation enables us to use a different analytical technique for *operational* and *planning* variables represented by continuous and discrete uncertainties, respectively.

2.5.2 Sources of Uncertainty

In this section, the different sources of uncertainty existing in a European LVDS system are discussed in brief. These sources of uncertainty might be external to the grid e.g., consumer behavior, weather, or from the unknowns in grid itself, e.g., the topology and impedance of feeder, ratings of connected PV installations, etc. For UQ the sources of uncertainty has to be understood and defined using parametric or non-parametric methods.

2.5.2.1 Consumer behavior and electrical load

The electrical load of a single consumer in LVDS is the measure of the electric energy and power consumed by that consumer over a time-period. This directly depends upon their *electrical behavior* and varies from consumer to consumer, according to time of day, day in a week or season [99, 100]. The electric load when aggregated in the level of transmission level usually have a well-defined seasonal and daily pattern. However, there is not an easy way to disaggregate such load to obtain load profile of individual consumers. Traditionally, the consumer load for planning application is deterministic based on the average load at the system-peak, termed after-diversity-maximum-demand (ADMD) [101]. An empirical factor is usually associated with such system peak to represent the consumers' individual peak, e.g. demand factor or diversity factor. Such factors cannot be used for UQ in LVDS. To perform a UQ, the uncertainty in the load profiles of the consumer in household level is required. These load profiles are difficult to model compared to the aggregated load [102]. Furthermore, with the increase in the energy intensive and controllable devices such as EV and HP, a detailed uncertainty model of consumer load has an increased importance [103].

In literature, different ways to model a single consumer load exist. In [99], the consumers are divided into 46 sub-classes, and the loads for a day, month or season are modelled as log-normal distributions. The high resolution data of consumers was fitted to Log-Normal and Weibull distributions to give hourly, daily and monthly PDF of the household load of Sweden in [100]. The goodness-of-fit was shown to be above the acceptable limit for $\approx$70% of cases. In [104], the consumer load is modelled by assuming the uncertainty of the operation time of the household devices and summing the individual devices' profiles to obtain the consumer load profile. Consumer load uncertainty is modelled in a non-parametric way in [27], where consumer loads are picked from a pool of historical profiles. Authors of [105] use a pool of worst-case loads instead of taking daily load profiles. In [26], the data from SM is converted to a CDF and load of consumer is sampled from that. Residential load from South Africa were fitted to Weibull, Normal and Beta distribution in [106]. The outcomes of fitness

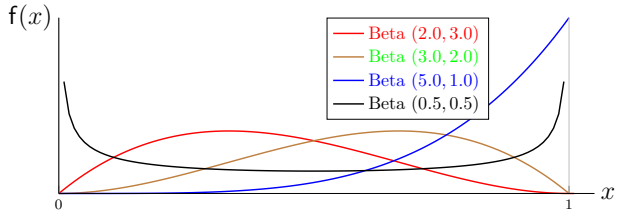


Figure 2.10: Versatility of Beta distribution to represent positively skewed (red), symmetrical (black), or negatively skewed (brown and blue) shapes of consumer load PDF

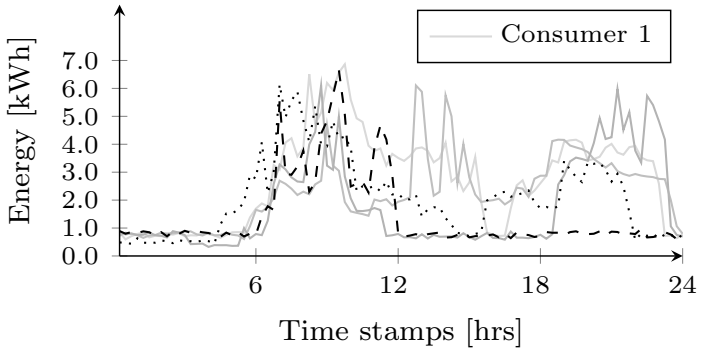
test and flexibility in modelling skewed as well as symmetrical distribution led to the recommendation of a Beta PDF. In [103], the LVDS load is modeled as a Beta distribution:

$$f(x) = \frac{1}{B(a,b)} x^{a-1}(1-x)^{(b-1)} \text{where}, B(a,b) = \int_0^1 x^{(a-1)}(1-x)^{(b-1)} dx, \tag{2.34}$$

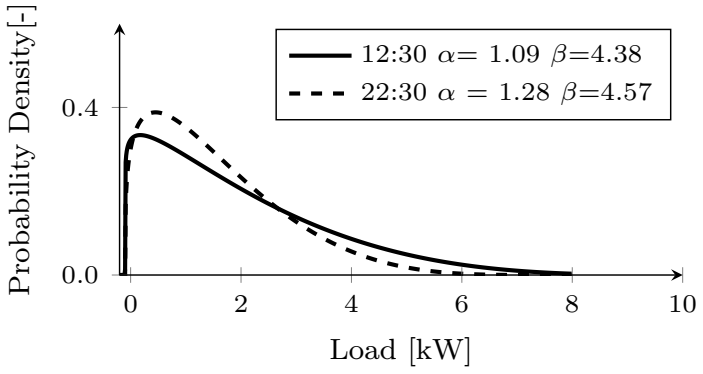
and a and b are alpha and beta parameters. The parameters a and b of Beta distribution can be tuned to model the differently distributed data over the limits 0 and 1. This feature makes it suitable for representing consumer loads PDF with positively skewed, symmetrical, or negatively skewed shapes (see Fig. 2.10). An aggregated consumer load fitted to a Normal distribution in [102], and tested with the Chi-square and Kolmogorov goodness-of-fit. The assumption of Gaussian model was found valid for the aggregated consumers' load.

In this thesis, two models to define uncertainty in consumer load are chosen: a) the non-parametric bootstrapping from the pool of historical load profiles; and b) parametric quarter-period Beta distribution. These are the most common models in the literature for such individual household load profile for planning problems [103].

Chapter 3 of this thesis uses load profiles picked from the pool of historical profiles (see Fig 2.11a). Pools of load profiles are formed for different consumer types, based on the similarity of the consumer annual energy intake and contracted load capacity. Due to the limited number of load profiles in the test network [30] (details of test LVDS in Appendix A), it is supplemented by the load data of Belgian consumers [107]. The information on factors affecting uncertainties, such as consumer behavior, temperature, and humidity, etc. are lost while grouping load profiles based on contracted load and annual energy intake only. This information is retained by splitting the annual load profile to daily load profiles and forming a sub-set of daily load profiles for different



(a) Few selected load profiles from pool of historical profile of same consumer type in one day (spring)



(b) Beta distribution for quarter-hour periods for one consumer-type (spring season)

Figure 2.11: Different consumer load profile uncertainty models

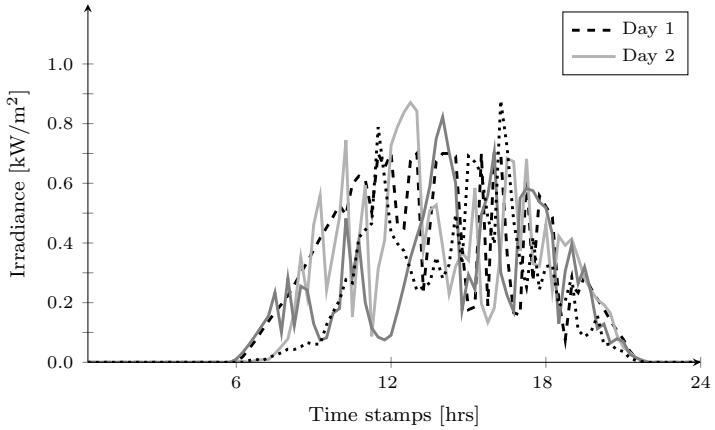
seasons for each consumer type. The uncertainty of consumption during the day is incorporated by taking a 15-minutes averaged load profile.

In the remaining chapters, the load is modelled as a Beta distribution per 15 minutes time-period (see Fig 2.11b). The beta distribution is fitted on the pool of historical daily load profiles for each consumer type and season. The beta distribution fitting was done with `scipy.stats` library of python. In this thesis, we mainly deal with residential consumer loads. Hence, the reactive load is modelled as a constant 10% of active power, as seen in the smart meter data given in [30].

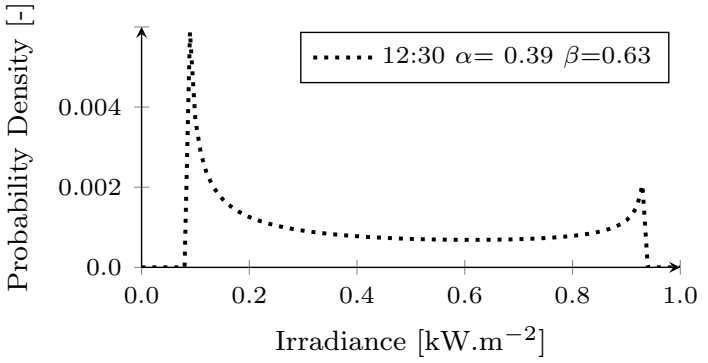
2.5.2.2 Irradiance and Generation from PV

Irradiance is the measure of *the amount of light energy from one thing hitting a square meter of another each second.* The amount of electrical energy generated by a PV installation is directly dependent upon the irradiance from the sun at a given time and location. However, the actual conversion of the light energy to electrical energy from a PV installation depends upon diffusion of solar irradiation by atmospheric components like cloud, and various other local parameters such as ambient temperature, the orientation of installation, shadow by nearby trees, or buildings, dust on the PV surface, ageing, faulty PV cells, etc. [108]. Due to all these factors, generating a predicted power based only on theoretical solar irradiance is not always possible. Irradiance data are easily available from weather services and they are the measure of actual irradiance at the earth surface. Since this measurement excludes the solar radiation that is scattered or reflected by atmospheric components, it can be taken as a stochastic input variable to model generation from a PV installation [109, 110]. The uncertainty of PV generation due to orientation of the house, soiling, etc., can be modelled if historical measurements of PV generation of each installation are known. Modelling these uncertainties for long-term planning problems is a research topic in itself [111] and is not considered in this thesis. These uncertainties are replaced by a deterministic efficiency factor of η_p=0.95. The power output of a PV installation is then represented as the measured irradiance E on the surface of the PV installation scaled by η_p and the kWp rating of PV installation.

Similar to load uncertainty modelling, there are different schemes for the irradiance profile uncertainty modelling described in the literature. The irradiance distribution is highly skewed and cannot be represented by a simple Gaussian model [112]. Hourly transmittance of a day was modelled using a Beta distribution in [113] and scaled to obtain the synthetic hourly irradiation value on a horizontal surface. In [114], long-term hourly irradiance data of 30 years was fitted using three PDF, Beta, Weibull and Log-Normal. Chi-square and Kolmogorov-Smirnov goodness-of-fit tests were applied, and Beta distribution was found suitable for most cases, followed by a Weibull distribution. The hourly irradiance data of the four different locations of the UK was fitted to Beta distribution [115] and observed that the goodness-of-fit test indicates a poor result. However, this method intrinsically considered all phenomena like cloud shading and incident angle. It also added simplicity to the simulation after the shape parameters were obtained for all hours of the year. Alternatively, the irradiance or PV generation profiles can be sampled from a set of available historical profiles as in [27]. Non-parametric CDF is proposed in [26], based on the historical data of SM. Beta distribution to the available profiles per period is recommended in [116] considering the versatility of such distribution



(a) Few selected irradiance profiles from pool of historical profile from rooftop of Energyville, Genk (different days of spring)



(b) Beta distribution for a quarter-hour period irradiance (spring)

Figure 2.12: Different irradiance uncertainty models

in representing both skewed and symmetrical bimodal distributions. Modelling irradiance uncertainty as a Beta distribution or bootstrapping from historical data seems to be most chosen way for modelling PV generation at least for the planning applications that are dealt with in this thesis.

The temporal correlation among the generation of the PV installations can be assumed for a large service area [117]. However, the length of an LVDS feeder is usually less than one kilometre long, where the irradiance can be assumed equal.

In Chapter 3 of this thesis, the daily irradiance profile for the spring season is obtained by sampling from a pool of historical profiles from the measurement at rooftop of the Energyville building in Genk, Belgium for the same season (Fig. 2.12a). In the remaining of the thesis, the PDF of the irradiance per quarter hour fitted as a Beta distribution based on the historical measurement data of days of the same season (Fig.2.12b).

2.5.2.3 PV/EV/HP size, type and location

The size, location and type of the PV and new loads (EV and HP) that is going to be installed in future LVDS is unknown. Most of the policy focuses on the number of new installations or devices but the actual location and nature of installation is not necessarily defined [4, 7]. The uncertainty might be due to future policies, technology innovation, disruptive changes in consumer behavior, market circumstances, or new types of contracts. The uncertainties related to future PV, EV and HP size, location and type directly impact the future operation of the LVDS.

In the literature, the uncertainties in PV/HP/EV size, type and location are considered while solving planning problems such as calculating HC of an LVDS [118], planning active distribution networks [119], or optimal integration of charging stations of electric vehicles [120]. The review of uncertainty modelling techniques for distribution planning presented in [121] points out that there is no single best uncertainty modelling technique for future installations, but it depends upon the type of the planning problem. For PV HC calculations, most of the work on literature consider PV size and location as an uncertain variable [27, 98, 122–125]. In [122], the type of PV (one-phase or three-phase) is considered as a stochastic variable along with the number of PV installations in the LVDS. The HC calculations in [23] and [21] consider the PV size as a deterministic variable with PV location as the only stochastic variable. Chapter 3 of this thesis has details of the different assumptions on uncertainties related to PV installations size, type and number and their impact on HC calculation in LVDS.

2.5.2.4 LVDS topology and impedance uncertainty

LVDS topology uncertainty is created due to the mismatch between the actual installations and the record of the operators. Errors might have occurred in the DSO's effort to have a digital version i.e., Geographical Information System (GIS), of their LVDS. While some other variables like the phase to which a consumer is connected is not recorded as the phases of underground LV cables

are not marked. Similarly, the GIS mapping of most DSOs does not exclusively record the information about the connectivity of the consumer to the feeder. This creates an uncertainty in the actual position of load-point in an LVDS feeder.

The impedance uncertainty in the LVDS model is the error in the calculated series impedance and shunt impedance. The impedance value obtained from DSOs might not be entirely accurate due to a) wrong cable-type or topology, b) wear and tear of the cable with time, c) presence of joints and hot-spot within the cable, d) reconfiguration of the network, e) conductor temperature, etc. [126, 127]. The different sources of impedance uncertainty in an LVDS feeder are described below:

1. **Series impedance**: The value of the resistances given in the datasheet are usually for an ambient temperature of $20°C$ but the operating temperature of the cables can be above the ambient temperature. Similarly, the uncertainty in the geometry of the underground cables and overhead lines can create some fluctuations in self and mutual reactance. There might be systematic errors due to conductor ageing, corrosion, and joints after cable breakage or faults. These uncertainties might lead us to different results while solving the power flow equations compared to reality.

2. **Shunt Admittance**: Shunt admittances exist due to the capacitance between the cable and ground. In [126], it is pointed out that there might be some uncertainties in the value calculated due to an unknown configuration. Neglecting the shunt admittance in LVDS might introduce a systematic error, although it is quite common to do.

3. **Line Type and Length**: The digital databases of DSOs might have a record differing from the actual cable type and length under the ground. The discrepancy might be due to unrecorded maintenance activity, load shifting from one feeder to another, splitting of feeder, etc. This inconsistency will influence the actual impedance matrix of the LVDS and the recorded line limits. This is, so far, the most critical uncertainty affecting impedance accuracy.

The power flow simulations created based on LVDS topology from GIS, impedance matrices, and SM measurement of power, referred as *the digital twin* of the LVDS, is gaining interest among the DSOs and the Distribution Management System (DMS) providers. These uncertainties in the grid database challenge building an accurate *digital twin* of the LVDS. Research effort is spent in defining methods to remove these uncertainties: e.g., with a lot of measurement and data-driven methods as in [128], or using state-estimation as

in [129]. The presence of phase measuring units (PMU) in LVDS could help to identify the phase of the consumer and topology of the network as shown in [130] although this technology is currently only seldomly used. The topology and impedance identification of LVDS using SM data is shown in [131], but requires very high SM penetration. A mixed-integer programming approach is used in [127] for phase identification which requires time synchronized voltage and power data of all consumers. In this work, it is assumed that the LVDS topology and consumer phase is correctly marked in the digital record of the DSO and impedance of the LVDS feeders are known and deterministic.

2.5.2.5 Other variabilities

Variability might also arise from the upstream grids, e.g., voltage unbalance in the MV distribution system and voltage magnitude at the MV side of the transformer. These variabilities propagate to the LVDS. This type of uncertainty is case specific and LVDS operation should have some headroom on voltage limits to manage those upstream variabilities.

2.6 Summary and conclusion

This chapter discusses the challenges of European Low Voltage Distribution System (LVDS) feeder modelling, describes the uncertainty modelling for LVDS and presents details on different Probabilistic Power Flow (PPF) tools for Uncertainty Quantification (UQ) in LVDS.

Firstly, a new transformation method is presented to reduce a three-phase four-wire single-grounded European LVDS to its three-wire equivalent. Secondly, the various stages of UQ for power system are discussed. The features of different PPF approaches are presented keeping in mind the UQ in an LVDS. The Polynomial Chaos Expansion (PCE) based method is then described in detail as it is the mostly used method in this thesis. Finally, a review of uncertainty source quantification of a European LVDS is done. Different types of uncertainties and their source are detailed as well as possible ways to model them for a planning problem, i.e., the computation of the photovoltaics (PV) hosting capacity of an LVDS.

Chapter 3

Hosting Capacity of Photovoltaic Systems

Koirala, A., Van Acker, T., D'hulst, R., Van Hertem, D., *Hosting capacity of photovoltaic systems in low voltage distribution systems: A benchmark of deterministic and stochastic approaches.* Renewable and Sustainable Energy Reviews, 155, Art.No. 111899. doi: 10.1016/j.rser.2021.111899

The first author is the main author of the article. The contributions of the first author include the literature review, methodology, software, and the writing of the article.

Abstract: This paper compares and benchmarks different methods to calculate the hosting capacity of low voltage distribution systems for residential photovoltaic generation. The low voltage distribution system hosting capacity provides insight to the network planner and operator regarding the capability of the distribution system to accommodate new photovoltaic installations. In the literature, many different methods to calculate the hosting capacity of the distribution network exist, but they are demonstrated using different feeders and with different assumptions, making it difficult to compare different approaches. In this paper, a consistent benchmark of existing hosting capacity methods is performed. A publicly available real network of a semi-urban area is used to compare the different approaches in a consistent manner. A reference stochastic hosting capacity method is proposed, which considers all sources of uncertainty

as well as *operational* and *probabilistic limits*. Hosting capacity definitions, based on both deterministic and stochastic methods, are evaluated based on their performance parameters and benchmarked. The performed analysis allows selecting the most appropriate hosting capacity definition and comparing results amongst systems.

Keywords: Low voltage, distribution feeder, photovoltaics, hosting capacity, stochastic

3.1 Introduction

Low voltage distribution systems (LVDS) were commonly planned and built using a "fit and forget" approach. They were designed to distribute the bulk power from centralized power plants to the end-users connected to the low voltage grid. Such were developed to accommodate the worst case scenario, and operated without significant monitoring or intervention. However, the rapid growth of distributed renewable energy resources (RES) like solar photovoltaics (PV) systems or the introduction of new loads such as electric vehicles (EV) and heat pumps (HP) requires a change of manner in which the distribution system is managed. The distributed RES can lead to bidirectional power flows which require Distribution System Operators (DSOs) to be more aware of the network operational state. The new load types can lead to congestion in the network which DSOs should be able to identify or forecast.

DSOs and regulators are interested to understand to what extent existing distribution networks are capable of handling the future power demand and generation without hitting the system limits and to what extent connection requests can be accepted. For example, a high PV penetration in low voltage (LV) distribution networks can create several problems ranging from overvoltage [20, 132–134], deteriorated voltage quality [123, 135], thermal stress, increased feeder loss [136], increased tap operations [137] and deterioration of transformer insulation [138]. Determining whether such system limits are reached under a wide range of circumstances is hindered by large uncertainties and significant complexity, even more so for smart distribution systems with active load and generation management. DSOs use different techniques to evaluate new connection requests for PV,for example, by individually studying the feasibility or applying some empirical limit. The term *hosting capacity (HC)* is often used to define such limitations of the distribution network. The earlier work on the calculation of the HC aimed to find a single number to describe the whole system, using worst-case scenario-based approaches. However, new calculation methods and tools suggest the use of probabilistic calculation methods and to determine the HC on individual feeders level, rather than the system level.

Despite the availability of these new techniques, DSOs still use “rules of thumb”, e.g., *“the HC corresponds to 15 % of the transformer’s capacity”* [18].

The most accepted definition of HC is “the amount of new generation or consumption that can be accommodated on a given feeder without impacting system operation under existing control and infrastructure configuration” [17, 139]. However, there are many other definitions and many calculation methods to obtain a numerical value for HC. It is unclear to what extent they are compatible or even comparable. The proposed definitions or methodologies are generally illustrated in the literature by using different types of networks and different assumptions, making the comparison between the methods cumbersome. Existing review papers on Low Voltage Distribution System (LVDS) HC calculation techniques [18,132,139,140] provide a very broad conclusion regarding different HC methods. After reviewing previous works, [140] recommends the adoption of stochastic methods, but the comparison between stochastic methods and the development of new methods are presented as open research questions. The authors of [139] conclude that the different HC methods have different implications, but this is not quantified. Article [132] reviews only the impact of high PV penetration. Authors propose a new method in [17], but the significance of this new method is not compared against the results from existing methods. Similarly, [18] enlists different HC studies, but a detailed comparison is missing. The existing review papers reach similar conclusions: a) hosting capacity studies are very relevant to DSOs; b) stochastic methods are more suitable than the deterministic methods; c) the method selected depends on the data available; and d) further research is required to understand the HC limitation of LVDS. This shows the need for evaluating all available deterministic and stochastic HC methods to understand their strength as well as usefulness. This paper focuses on the comparison of the existing HC calculation methods in a realistic grid with load and PV generation data. The focus will be on calculating a one-dimensional HC capacity of LVDS, i.e., for increased PV installation only. Nevertheless, the conclusions can be extended to include increased electric vehicles and heat pumps. The benchmarking is done by implementing existing HC methodologies on the "Non-synthetic European low voltage test system" described in [30]. All methods are compared with a reference method with similar stochastic and voltage limits.

The paper is organized into 6 sections. After this introduction, the hosting capacity definitions are discussed in Section 3.2. The hosting capacity calculations methods are discussed in Section 3.3. Section 3.4 discusses the reference method for benchmarking both deterministic and stochastic HC. In Section 3.3.1, the HC methods discussed in Section 3.3 and 3.4 are evaluated using an actual European LV network. Finally, Section 3.6 concludes the work.

3.2 Hosting capacity definitions

In this section, a unified mathematical representation for calculating the hosting capacity of a low voltage feeder is defined. This representation is applied to an extensive list of HC methods found in literature in a simplified and uniform way. This representation is also used to define the reference HC method in Section 3.3.2.

3.2.1 Generic HC definition

The mathematical formulation of hosting capacity is discussed using a simple representation of European LV feeders (Fig. 3.1). In the figure, the light gray part represents the LV substation with a transformer t. In European style LV systems, the transformers usually have high ratings and each transformer has one to eight LV feeders connected to them. Given the ease associated with upgrading transformers compared to underground cables, transformers are not considered a limiting factor for hosting PVs. The LV network N is

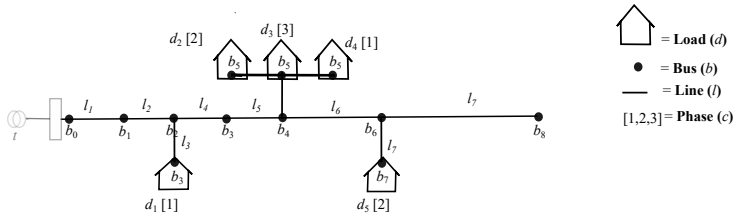


Figure 3.1: Simple representation a three-phase low voltage (LV) feeder

the set of all LV feeders ($\mathcal{F}$). These feeders are radial and can be studied independently. The network comprises of multiple LV transformers to which these feeders are connected. The PV hosting capacity HC_f of a feeder f is defined as the total PV that can be added to the feeder without violating its *operational limits* during any realistic operational scenario. These *operational limits* could be minimum voltage threshold $\mathsf{U}^{\mathrm{min,\,th}}$, maximum voltage threshold $\mathsf{U}^{\mathrm{max,\,th}}$, maximum voltage unbalance limit $\mathsf{U}^{\mathrm{ub,\,th}}$, cable current rating I_l^{rated} or transformer rating $\mathsf{S}_t^{\mathrm{rated}}$. The realistic operation scenarios are achieved by assigning the consumer load from historical pool of profiles and PV generation from the irradiance profiles. Even though $\mathsf{S}_t^{\mathrm{rated}}$ is considered to be a weak constraint, it is accounted to prevent unrealistic scenarios, e.g., high rated PV at the beginning of the feeder. A mathematical expression for a feeder's directional hosting capacity is the sum of additional PV capacity of a subset of consumers that can be accommodated by a feeder f without violating its *operational limits*

as given in (3.1) where, $\mathcal{P}_f$ is the set of new PV capacity in the feeder. It is assumed that $\mathcal{P}_f$ is mapped to the set of LV consumers $\mathcal{D}_f$ one-to-one and the cardinality of $\mathcal{P}_f$ is less than or equal to that of $\mathcal{D}_f$. This is a reasonable assumption in case of LV network where each LV consumer can have at most one PV installation.

$$\mathrm{HC}_f = \sum_{p \in \mathcal{P}_f} p \qquad \text{s.t.,} \quad |\mathcal{P}_f| \leq |\mathcal{D}_f|. \tag{3.1}$$

The HC of the network N is the total additional capacity of PV that the given network can host and is then obtained as sum of HC_f:

$$\mathrm{HC}_{\mathsf{N}} = \sum_{f \in \mathsf{N}} \mathrm{HC}_f. \tag{3.2}$$

This work considers the inverter size equal to the PV installation's rating and generates only active power. Hence, the kiloWatt-peak (kWp) of the PV installation is taken as a preferred unit of the HC. Alternatively, suppose the size of the inverters is smaller than the PV peak capacity and generates only active power. In that case, the HC can be expressed in kiloWatt (kW) of inverters. The PV HC can also be expressed in kiloVolt-Ampere (kVA) of the inverters when the inverters used are used for generating both active and reactive power.

In many works, the HC is not given as an absolute value in kWp but is expressed relatively to quantities like transformer capacity $\mathsf{S}_t^{\mathrm{rated}}$, peak active load in the network $P^{\max}$, etc. For instance, the transformer hosting capacity (HC^{t}) is the additionally installable PV capacity normalized to the LVDS transformer ratings. This is expressed in (3.3) as the sum of feeder HC defined in (3.1) divided by the sum of transformer ratings in the network. Taking the sum of transformer ratings as denominator in (3.3) is applicable in case of European style LVDS where transformers are of high ratings, mostly of the same size and are weak constraints for LV PV installations. However, if the LVDS has wide range of transformer ratings, it is more appropriate to define the HC per transformer.

$$\mathrm{HC}^{\mathrm{t}} = \frac{\sum_{f \in \mathsf{N}} \mathrm{HC}_f}{\sum_{t \in \mathsf{N}} \mathsf{S}_t^{\mathrm{rated}}} \tag{3.3}$$

Similarly, the peak load hosting capacity ($\mathrm{HC}^{\mathrm{pl}}$) is defined as the additionally installable PV capacity in a network normalized to the peak load ($P^{\max}$) in the network (3.4). If $\mathrm{HC}^{\mathrm{pl}}$ is calculated for a particular feeder, then the peak load should be of the same feeder.

$$\mathrm{HC}^{\mathrm{pl}} = \frac{\sum_{f \in \mathsf{N}} \mathrm{HC}_f}{P^{\max}} \tag{3.4}$$

The concepts of HC formulated in (3.1), (3.3) and (3.4) are the generic representation of LVDS HC and can be used for both deterministic and stochastic HC.

3.2.2 Deterministic HC

In deterministic HC formulations, the commonly considered variable is the number of consumers having PV ($N_p = |\mathcal{P}_f|$) while keeping the PV size and type of PV deterministic. The other parameters like consumer load, PV generation, PV location, etc are also deterministic for such formulations. The common assumptions in deterministic HC calculations are: a) each consumer's PV size is the same, b) the consumers installing PV are fixed and c) the *operational limits* are deterministic. The penetration ratio P^{R} of the PV in a feeder is the ratio of the number of PV installations to the total number of consumers as expressed in (3.5). The deterministic HC of a feeder with a penetration ratio P^{R} and PV rating of P^{kWp} is expressed in (3.6) which is deduced from (3.1) where every $p \in \mathcal{P}_f$ has the same size, i.e., P^{kWp}.

$$P^{\mathrm{R}} = \frac{N_p}{|\mathcal{D}_f|} \tag{3.5}$$

$$\mathrm{HC}_f = P^{\mathrm{kWp}} \cdot P^{\mathrm{R}} \cdot |\mathcal{D}_f| \tag{3.6}$$

These deterministic expressions are valid for worst-case scenarios, which undermine the actual HC of real networks. These assumptions make calculations easy and fast, but the realistic and practical aspects of the problem are ignored.

3.2.3 Stochastic HC

Unlike deterministic methods, in stochastic methods stochasticity can be added to the inputs such as load, PV generation, PV and consumer location, or network topology. The stochastic HC expression can also be represented using (3.1), but with following differences: a) the new PV capacity for installation is randomly selected from a predefined set of PV ratings $\mathcal{P}^{\mathrm{rt}}$, b) the technical grid limits are enforced up to a *probabilistic limit*, and c) for each PV deployment scenario, different combinations of PV location, PV size, PV generation profile, etc. are considered.

$$\mathrm{HC}_f^{\mathrm{s}} = \sum_{p \in \mathcal{P}_f} p \quad \text{s.t.,} \quad p \in \mathcal{P}^{\mathrm{rt}} \tag{3.7}$$

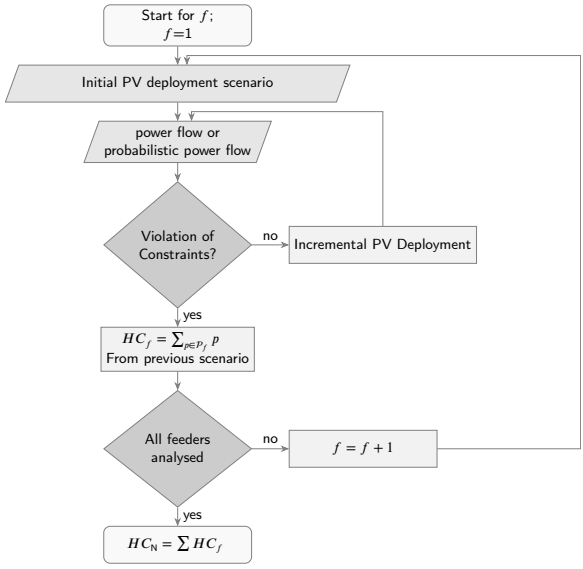


Figure 3.2: Flow chart for HC calculation of a network of LVDS feeders.

The HC calculated by deterministic method represents one snapshot of many scenarios considered in the stochastic method.

3.3 Hosting capacity calculation

The hosting capacity problem for LVDS with unbalanced power flow is generally solved in an iterative way. Optimization-based methods exist for medium voltage (MV) networks assuming balanced systems [141] or unbalanced systems [142]. However, these methods simplify the network equations to a linearized power flow.

Fig. 3.2 represents a generic block diagram of an iteration based method for an LVDS. The first step is to select one of the feeders and assign a initial PV deployment scenario, i.e., assign the number of PV installations that the feeder can have, represented by penetration ratio (P^{R}) for the deterministic HC. For stochastic HC, the initial PV deployment comprises of number of scenarios for a fixed P^{R}, depending upon the Monte Carlo (MC) samples considered. Then, full ac power flow simulations are used to check if the grid limits are violated by those scenarios of PV deployment. For LV systems, unbalanced power flow is preferred considering their unbalanced nature to check the voltage and current

in the each phase of the feeder bus and branches. The Backward Forward Sweep (BFS) algorithm proposed by [143] with the 3-wire reduction method explained in [32] is used in this paper for all power flow calculations. PV is added using some algorithms until the *operational limit* is reached. Different ways of adding PV in a feeder is one of the factors which makes PV HC methods different and will be discussed in details later. In deterministic HC, the *operational limits* are deterministic as in (3.8) and are checked after running a power flow for each PV scenario.

$$\mathsf{U}^{\min,\,\mathrm{th}} \leq |\mathsf{U}_{i,c}| \leq \mathsf{U}^{\max,\,\mathrm{th}} \quad \forall \quad i \in \mathcal{I}, c \in \phi \tag{3.8a}$$

$$|\mathsf{I}_{l,c}| \leq I_l^{\mathrm{rated}} \quad \forall \quad l \in \mathcal{L}, c \in \phi \tag{3.8b}$$

$$\mathsf{U}_i^{\mathrm{ub}} \leq \mathsf{U}^{\mathrm{ub},\,\mathrm{th}} \quad \forall \quad i \in \mathcal{I} \tag{3.8c}$$

$$\mathsf{S}_t \leq \mathsf{S}_t^{\mathrm{rated}} \quad \forall \quad t \in \mathcal{T}. \tag{3.8d}$$

The increase in PV deployment for a feeder is stopped once any of these limits are violated. The resultant HC_f is the total PV in the scenario previous to the violation. The process is continued for all LV feeders (Fig. 3.2).

In iteration based stochastic HC calculation methods, scenarios with additional PV deployment are evaluated until a probabilistic *operational limit* is reached. In (3.9), these probabilistic *operational limits* are listed where $(1-\epsilon)$ represents the *probabilistic limit of acceptable scenarios.*

$$\mathbb{P}(|\mathsf{U}_{i,c}| \leq \mathsf{U}^{\min,\,\mathrm{th}}) \leq 1-\varepsilon_u \quad \forall \quad i \in \mathcal{I}, c \in \phi \tag{3.9a}$$

$$\mathbb{P}(|\mathsf{U}_{i,c}| \geq \mathsf{U}^{\max,\,\mathrm{th}}) \leq 1-\varepsilon_u \quad \forall \quad i \in \mathcal{I}, c \in \phi \tag{3.9b}$$

$$\mathbb{P}(|\mathsf{I}_{l,c}| \geq I_l^{\mathrm{rated}}) \leq 1-\varepsilon_i \quad \forall \quad l \in \mathcal{L}, c \in \phi \tag{3.9c}$$

$$\mathbb{P}\big(\mathsf{U}_i^{\mathrm{ub}} \geq \mathsf{U}^{\mathrm{ub},\,\mathrm{th}}\big) \leq 1-\varepsilon_{ub} \quad \forall \quad i \in \mathcal{I} \tag{3.9d}$$

$$\mathbb{P}(\mathsf{S}_t \geq \mathsf{S}_t^{\mathrm{rated}}) \leq 1-\varepsilon_t \quad \forall \quad t \in \mathcal{T} \tag{3.9e}$$

3.3.1 Deterministic HC calculations

The available literature on deterministic hosting capacity is listed in Table 3.1. The common features of these methods are: a) HC is represented by a single number, b) HC is calculated for the worst case scenario, and c) *operational limits* are conservative. From Table 3.1 it is quite clear that there is not a single way to come up with a single number which defines hosting capacity. However, all the existing methods can be classified into four methodologies which can be

represented by mathematical formulation developed in Section 3.2, named **D1**, **D2**, **D3** and **D4** in this paper.

Method **D1** uses the formulation (3.1), where the PV kWp is incremented keeping penetration ratio P^{R} constant till the grid constraints are violated. In method **D2**, (3.5) and (3.6) are used, where the percentage of consumers that can have a fixed kWp (P^{R}) is incremented till the grid constraints are violated or 100% consumers have PV installations. In method **D3**, the P^{R} is incremented in each step like in **D2**. The difference is that the PV installed are not of the same size but selected from the set of available PV ratings ($\mathcal{P}^{\mathrm{rt}}$). Method **D4** is different from other methods in terms of how the final HC result is expressed, e.g., the ratio of transformer rating as in (3.3), the ratio of peak load as in (3.4), etc.

The works [144–148, 161] use method **D1**. Although all of these studies use similar methods, different HC results are found due to differences in the following features: a) test feeders used, i.e., real, synthetic or simplified, b) location of

Table 3.1: Different Deterministic PV HC methods for distribution feeders

Ref	Network Type	Voltage	Network/feeder	Reference	Control	$\mathrm{HC_N}$ or P^{R}	With control	HC method
[144]	real rural grid	MV	feeder	total PV	yes	1.2 MW	2.3 MW- 4 MW	**D1**
[145]	rural and urban representative	LV	feeder	total PV	no	20 kW - 140 kW rural; 200 kW urban	-	**D1**
[146]	10 node test feeder	LV	feeder	additional PV	no	7.5 kWp/phase	-	**D1**
[147]	synthetic Brick network	MV	network	Bricks nodal injection limit	no	9-13 MW	-	**D1**
[148]	EPRI Ckt7	MV	feeder	total PV	no	average 6.55-6.86 MW	-	**D1**
[20]	European real grid	MV and LV	network	consumers with PV	no	30% for 2 kWp	-	**D2**, with P^{kWp} = 2 kWp
[149]	European rural	LV	feeder	consumers with PV	yes	100% for 3.5 kWp	5.5 kWp	**D2**, with P^{kWp} = 3.5 kWp
[150]	US real grid	MV	network	consumers with PV	no	100% for2 kWp	-	**D2**, with P^{kWp} = 2 kWp
[151]	European new grid	MV and LV	network	consumers with PV	no	100% (0.8 to 4.4 kWp)	-	**D3**, with P^{kWp} = {0.8 to 0.44 kWp}
[134]	North American	MV	network	consumers with PV	no	100% for 2.5 kWp	-	**D2**, with P^{kWp} = 2.5 kWp
[152]	IEEE European	LV	Feeder	consumer with PV	no	100% for 4 kWp	-	**D2**, with P^{kWp}= 4 kWp
[153]	European grid with real PV installations	LV	feeder	transformer size	no	70% for 400kVA	-	**D4**, with $\mathrm{S}_t^{\mathrm{rated}}$ = 400 kVA
[154]	real grid of a town in Europe	MV	network	peak load	no	63.2%	-	**D4**
[155]	US real grid	MV	network	annual peak load	no	47%	-	**D4**
[156]	US real grid	MV	network	annual peak load	no	50%	-	**D4**
[157]	Australian grid with real PV installation	LV	Network	annual peak load	no	22%	-	**D4**
[158]	EPRI Feeder	MV	feeder	peak load	no	50%	-	**D4**
[159]	European real grid	LV	feeder	transformer size	yes	45% of tx size	upto 100%	**D4**, with $\mathrm{S}_t^{\mathrm{rated}}$ = 630 kVA
[160]	synthetic radial feeder	LV	feeder	peak load	no	132% of 75 kW/phase load	-	**D4**, with P^{max}=75 kW/ phase
[161]	simplified equivalent network	MV	feeder	peak load	no	urban 8.3 MW (40%); rural 6.3 MW (30%)	-	**D1**, **D4**
[162]	IEEE-30 bus system	HV & MV	network	transformer size	no	60% for 150 MVA	-	**D4**, with $\mathrm{S}_t^{\mathrm{rated}}$ = 150 kVA

the feeder, i.e., rural or urban, c) control schemes, i.e., active control system or no control and d) final representation of HC. Similar conclusions are applicable for methods **D2** as seen studies [20, 134, 149, 150, 152]. Additionally, method **D2** has difference in the fixed P^{kWp}. The HC calculations in [20] and [150] use P^{kWp}= 2 kWp. The other works consider fixed PV capacity of 2.5 kWp [134], 3.5 kWp [149] and 4 kWp [152].

Method **D4** is widely used compared to other methods because of its applicability even with limited network information [153–162]. However, the method **D4** is only the way of representation of HC calculated from **D1**, **D2** and **D3** and hence not completely treated as a different method. Similar to previous cases, the results obtained in these studies are unique but in all cases the transformer HC is less than 100%; showing that transformer rating plays the least role in LVDS PV hosting capacity.

These variations in deterministic HC results are due to the different assumptions, simplifications and unique properties of the selected feeder. Despite being an easier method, deterministic HC methods have their shortcomings. These methods: (a) do not account for the uncertainties in PV power production and LV consumer demand, and (b) neglect the variations and uncertainties in size, location and properties of PV installations [19]. Every feeder has some uncertainties linked with them which are neglected while using a deterministic "rule of thumb". Further, LVDS operation always have some margin for violation of voltage and current limits, and are not as rigid as the constraints in (3.8). Stochastic hosting capacity methods proposed to overcome these shortcomings and give more realistic HC results are discussed in Section 3.3.2.

3.3.2 Stochastic HC calculations

Stochastic HC method considers the uncertainties in feeder characteristics, the load profile and the PV distribution [18, 19].

The set of uncertain input variables can be broadly be classified into following sets:

$$\text{Feeder uncertainty} : \{\mathcal{D}_f, \mathcal{P}_f, ...\} \tag{3.10a}$$

$$\text{Location uncertainty} : \{\phi_d : d \in \mathcal{D}_f, \quad \phi_p : p \in \mathcal{P}_f, \quad \mathcal{P}_f \mapsto \mathcal{D}_f, ...\} \tag{3.10b}$$

$$\text{Power injection uncertainty} : \{\mathsf{P}_d, \mathsf{P}_p\}. \tag{3.10c}$$

The feeder uncertainty refers to: (a) the number of LV consumers in the feeder, and (b) the number of LV consumers with PV. The location uncertainty deals with: (a) the phase that the load and PV are connected to, and (b) the location

of PV along the feeder. The power injection uncertainty includes the power consumed by the connected LV consumers and generated by PV installations .

The common approach is then to convert these uncertainties to voltage and current uncertainty in each phase of buses and branches of the feeder represented as function of uncertainties ξ in the system:

$$\underline{\mathsf{U}}_{i,c}, \underline{\mathsf{I}}_{l,c} : g(\xi_1, \xi_2, ...,). \tag{3.11}$$

Table 3.2 shows which uncertainties are considered in different studies in the calculation of the stochastic HC. The point to be noted is that not all uncertainties are considered in all studies. The input uncertainties along with other deterministic inputs are propagated to calculate the uncertainties in bus voltage and branch current.

$$\underline{\mathsf{U}}_{i,c}, \underline{\mathsf{I}}_{l,c} : g(\xi_1, \xi_2, ..., , \delta_1, \delta_2, ...) \tag{3.12}$$

A number of studies [17, 21–23, 25–27, 98, 105, 118, 122–125, 163] present probabilistic approaches to calculate the hosting capacity of a distribution grid, considering different sources of uncertainty (Table 3.2). All these approaches sample all the uncertainties at the same time without splitting the feeder uncertainties, location uncertainties, and power injection uncertainties except [25]. In [25], the so-called "mixed aleatory-epistemic" is proposed to quantify the effect of different input uncertainties. Ideally, stochastic HC methods should give a PV HC of a feeder in the form of a Probability Density Function (PDF). But as shown in Table 3.2, only eight out of twelve methods reviewed consider PV size as an uncertain parameter in stochastic HC calculations. That means the remaining four methods have PV HC represented by a single number rather than a PDF. These methods are counted as hybrid

Table 3.2: Review of stochastic HC calculation literature for different types of uncertainties considered in LVDS

	PV uncertainty						Consumer and feeder uncertainty			
	PV size	location	phase	type(1-ϕ or 3-ϕ)	number	generation	Load	consumer location	transformer voltage	Total
[122]	x	x		x	x					4
[17]		x	x				x		x	4
[123]	x	x			x					3
[124]	x	x					x			3
[27]	x	x					x			3
[163]	x						x	x		3
[23]		x			x		x			3
[98]	x	x			x					3
[125]	x	x								2
[26]						x	x			2
[22]	x									1
[21]		x								1
	8	9	1	1	4	1	6	1	1	-

methods as these methods consider the other uncertainties or use stochastic constraints, but the final HC is a single number instead of a distribution.

From Table 3.2, it is seen that uncertainties in PV location, PV size and load are considered in most studies while the uncertainties in PV-phase, type of PV, PV generation, consumer location and transformer voltage gets less attention in PV HC studies for LVDS. None of twelve case studies considers

Table 3.3: Different Stochastic HC methods to be evaluated

Method	ref	Input uncertainties		Other Input	Output	Remarks
M1	[122]	PV size PV location PV type No of PV	Normal Distribution μ =5 kW range =(1–10) kW randomly assigned to existing consumers $\mathbb{P}(3-\ -\phi)$ = 25%, 50% *or* 75% 25% (0.3), 50% (0.3), 75% (0.3) and 100% (0.1)	Load = 1.5 kW Load uniformly distributed among phases	$\%U^{ub}$ at peak PV	Feeder fails if $\%U^{ub}$>2 at peak PV
M2	[17]	worst case load No–load V PV PV phase	uniform distribution (0–250 W) Uniform (238–242 V) randomly allocated (6 kW) randomly distributed	PV all 1–ϕ PV rating=6 kW	Maximum PV in feeder without Overvoltage	consumers with PV is increased till 110% voltage limit(90th pc)
M3	[123]	PV rating PV location No of PV	Uniform distribution (1–5)kWp 1–ϕ Random distribution 25%, 50% or 75%	Load 360 kW ph–1 60 kW, ph–2 120 kW, ph–3 180 kW	Failure index based on %VUF	$\mathbb{P}(U^{ub}>2\%)$ is calculated for each feeder
M4	[27, 124]	Load profiles PV profiles PV location	2000 load profiles (basis: UK household statistics) 2000 profiles with PV size 1–4 kWp randomly selected	100 PV profiles based on irradiance of 30 sunny days considering 15% efficiency	%voltage problem and transformer loading level	%PV increased in step of 10% till problem
M5	[163]	Consumer location Load Profiles PV size	randomly allocated to different nodes from pool of historical worst-case load increased uniformly	probability mass function (PMF) voltage 0 and ±5% PV type: 1–phase and 3–phase Rural, intermediate and urban feeder	ratio of PV to maximum network load	simulation stops when 5% violation
M6	[23]	load samples no. of PV PV node	Beta distribution randomly add 1kW PV random (<7 kW/consumer)	1–ϕ PV at assigned phase PV generation 800 W for 1 kW inst	installed PV/ Load size	HC is the ratio where $\lvert U_{i,c}\rvert$<1.1
M7	[98]	PV size No. PV PV location	selected from pdf for residential and commercial PV 2% increment in each 100 MC-scenario 50 different placement for each no. PV	min load 6 MW was selected	1st PV HC, 50% PV HC and 100% PV HC	over-voltage, V deviations and VUF
M8	[125]	PV location size of PV	randomly allocated randomly allocated	PV gen from irradiance and temperature constant hourly load	Scenarios with V (0.95–1.05) and thermal limit	Best scenario is selected from all combinations
M9	[26]	PV gen Load	sampled from CDF of SM with PV sampled from CDF of SM data	Other parameter are fixed Different scenarios with new PV for consumers w/o PV	$P_{\text{overvoltage}}$, $P_{\text{undervoltage}}$ and $P_{\text{unbalance}}$	Output from diferent scenarios to choose best HC
M10	[22]	PV size	Fixed total PV allocated randomly	All other input parameter are fixed	risk = $\frac{N(violations)}{N(scenarios)}$	Max total DG with risk=0 as HC
M11	[21]	PV location	20% Consumers are randomly selected to have PV	PV gen is fixed Only 2 hours peak gen is taken	Total DG without violation of following basis:	U (over, under and unbalance), cable and tx cap

all the LVDS uncertainties. It is also seen that a maximum of four uncertain variables is considered per study. Despite the differences, all methods: a) use MC simulations for probabilistic power flow; b) use one or more probabilistic *operational limits* in (3.9), and c) consider one or more uncertainties listed in Table 3.2. There is only one agreement among the methods, i.e., on the usage of MC approach. In MC based methods, the power flow equations are solved for multiple combinations of the samples of uncertain variables to get the stochastic output or quantity of Interest (QoI) [34]. But there are differences between the methods in terms of: a) uncertainties considered and their PDF or PMF, b) sampling style of the uncertainties, c) selection of stochastic output or QoIs and their threshold, and d) selection of the *probabilistic limit* (1-ε). To evaluate the effect of these differences in calculating stochastic HC, eleven different methods in the literature are evaluated in this paper; named M1, M2, ..., M11 (Table 3.3/3.4).

Table 3.3 lists the methodologies used with input uncertainties and sampling methods, other specific deterministic input, output and rules based on which HC is determined in these methods.

In Table 3.4, the mathematical formulation for the methods enumerated in Table 3.3 is presented in terms of the mathematical formulation proposed in Section 3.2. The columns of this table are: a) the HC expression used: (3.3), (3.4), (3.6) or (3.7), b) the selected constraints from (3.9), the respective voltage and current thresholds and the *probabilistic limit* (1-ε), c) the uncertainties considered and their sampling methods, d) the assumptions regarding deterministic variables, e) the way results are expressed in the literature, and f) type: fully stochastic or hybrid. The iterative HC method shown in Fig. 3.2 is used for all the methods, where deterministic power flow is replaced by probabilistic power flow.

The direct comparison of methods just by reading the papers is impossible as they use different grids and feeders with different simplified assumptions. The comparison of the methods is hence done by using them to evaluate PV HC of a real LVDS. Additionally, a reference method has to be framed which acts as the pivot for comparison.

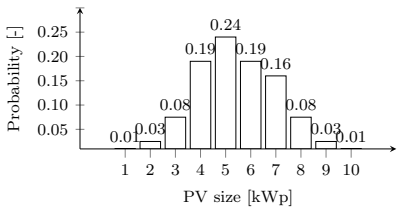
3.3.3 Analytical methods

The HC calculation methods in 3.3.1 and 3.3.2 are power flow based methods. There are few methods in the literature that use analytical or optimization-based methods to reduce the high computational time required for the MC-based methods. In [141, 142], an optimization based method for MV lines with a fixed location of PV installations is proposed. In [164] a robust optimization framework to calculate the HC of a radial MV network is proposed considering

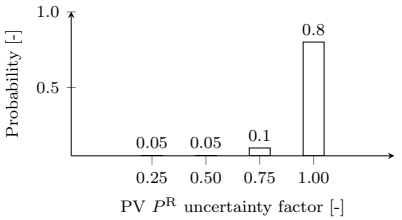
Table 3.4: Mathematical representation of different Stochastic HC methods

Method	HC	Limits	Uncertainties	Deterministic	Result	Type
M1	(3.7)	$\mathbb{P}(\mathsf{U}_t^{\text{ub}} > 2) \leq 0.00$ rigid constraint as $\mathbb{P} \not< 0$	$p \in \mathcal{P}^{\text{rt}}$= {1,2,...,10 kWp} $\forall \quad p \in \mathcal{P}_f$ $\mathcal{P}_f \mapsto \mathcal{D}_f$, one to one, stochastic map PV-type: $\mathbb{P}(3-\phi)$=25%, 50% or 75% {N_p: $\mathbb{P}(P^{\text{R}} == 25\%) = 0.3$, $\mathbb{P}(P^{\text{R}} == 50\%) = 0.3$, $\mathbb{P}(P^{\text{R}} == 75\%)$=0.3, $\mathbb{P}(P^{\text{R}} == 100\%)$=0.1}	P_d=1.5 kW	{Set of feeders having problem}	stochastic
M2	(3.6)	$\mathbb{P}(\lvert\mathsf{U}_{i,c}\rvert > 1.1\,pu) \leq 0.10 \quad \forall \quad i \in \mathcal{I}, c \in \phi$	P_d= 0–250 KW U^{s}= 238–242 V $\mathcal{P}_f \mapsto \mathcal{D}_f$, one to one, stochastic map PV phase={$c : c$=A, B or C}	P^{kWp}=6kW PV type=1–ph	N_p in f	hybrid
M3	(3.7)	-	$p \in \mathcal{P}^{\text{rt}}$= {1,2,...,5 kWp} $\forall \quad p \in \mathcal{P}_f$ $\mathcal{P}_f \mapsto \mathcal{D}_f$, one to one, stochastic map N_p:P^{R} = 25%, 50% or 75%	P_d=360 KW	Failure index based on $\mathbb{P}(\mathsf{U}^{\text{ub}})$	stochastic
M4	(3.7)	$\mathbb{P}(0.9 < \lvert\mathsf{U}_{i,c}\rvert < 1.1) \geq 0.95$ $\mathbb{P}(\lvert\mathsf{I}_{l,c}\rvert > I_l^{\text{rated}}) \leq 0$ for $l = 1$ (main line)	$\mathsf{P}_d \in$ Set of 2000 profiles $\mathsf{P}_p \in$ Set of 100 profiles $p \in \mathcal{P}^{\text{rt}}$= {1,1.5,2,...,4 kWp} $\forall \quad p \in \mathcal{P}_f$ $\mathcal{P}_f \mapsto \mathcal{D}_f$, one to one, stochastic map	-	max(P^{R}) obtained by increment of P^{R} by 10%	stochastic
M5	(3.7), (3.4)	$\mathbb{P}$(violating any of the grid constraints)<5%	P_d= From pool of worst-case load d location: $\{i : i \in \mathcal{I}\}$ P^{kWp}: increased uniformly	U^{s} = 1, 0.05 and 1.05 3 types of simplified feeder	HC^{pl} for simplified feeder	stochastic
M6	(3.4)	$\mathbb{P}(\lvert\mathsf{U}_{i,c}\rvert \geq 1.1) \leq 0$	P_d= Beta dist Additional N_p= 1 (1 kWp PV added in each iteration) max($p \in \mathcal{P}_f$) ≤7 kWp	PV type = 1-ph	HC^{pl} when constraint is violated	hybrid
M7	(3.7)	$(\text{HC}_f)_1 : \mathbb{P}(\lvert\mathsf{U}_{i,c}\rvert > 1.05) \leq 0.01$ for 100 MC $(\text{HC}_f)_{100}$: $\mathbb{P}(\lvert\mathsf{U}_{i,c}\rvert > 1.05) = 1$ for 100 MC	p sampled from pdf $\forall p \in \mathcal{P}_f$ N_p: plus 2% per 100 MC $\mathcal{P}_f \mapsto \mathcal{D}_f$:50 random scenarios	$\sum_{d \in \mathcal{D}_f} \mathsf{P}_d = 6$ MW	$(\text{HC}_f)_1$:0.01 violation; $(\text{HC}_f)_{100}$: all violation	hybrid
M8	(3.7)	$\mathbb{P}(0.95 < \lvert\mathsf{U}_{i,c}\rvert < 1.05\,pu) \geq 1$	$\mathcal{P}_f \mapsto \mathcal{D}_f$, one to one, stochastic p= randomly generated $\forall p \in \mathcal{P}_f$	P_d= constant hourly load N_p = 1, 2 and 4	max{HC} without violations	stochastic
M9	(3.7)	$\mathbb{P}(\lvert\mathsf{U}_{i,c}\rvert > 1.1\,pu) \leq 0.05$ $\mathbb{P}(\lvert\mathsf{U}_{i,c}\rvert < 0.90) \leq 0.05$ $\mathbb{P}(\mathsf{U}_t^{\text{ub}} > 2) \leq 0.05$	P_p: sampled from cdf of SM data P_d: sampled from cdf of SM data	$\mathcal{P}_f \mapsto \mathcal{D}_f$, one to one and fixed Fixed real grid with $\lvert\mathcal{P}_f\rvert = \lvert\mathcal{D}_f\rvert$	Best HC from different scenarios	stochastic
M10	(3.7)	$(0.9 < \lvert\mathsf{U}_{i,c}\rvert < 1.1)$	p randomly allocated to d $\forall p \in \mathcal{P}_f, d \in \mathcal{D}_f$	$\sum p$ = constant for each iteration	HC_{min} and HC_{max}	hybrid
M11	(3.3)	All constraints in (3.9)	P^{R}= 20% $\sum p$ increased by 1kWp till violation	p: based on capacity of d	distribution of HC^{t}	hybrid
M^{ref}	(3.7)	$\mathbb{P}(\lvert\mathsf{U}_{i,c}\rvert < 0.95) \leq 0.05$ $\mathbb{P}(\lvert\mathsf{U}_{i,c}\rvert > 1.05) \leq 0.05$ $\mathbb{P}(\lvert\mathsf{I}_{l,c}\rvert > I_l^{\text{rated}}) \leq 0.05$ $\mathbb{P}(\mathsf{U}_t^{\text{ub}} > 2) \leq 0.05$ $\mathbb{P}(\mathsf{S}_t > 1.5 \cdot \mathsf{S}_t^{\text{rated}}) \leq 0.05$	PV size : M7 PV location, no and type: M1 PV gen: M4 Load: M4 and M11 Phase(Load and PV): M2	max(N_p) increased in step of 10% P^{R}	mean PV capacity before violations	-

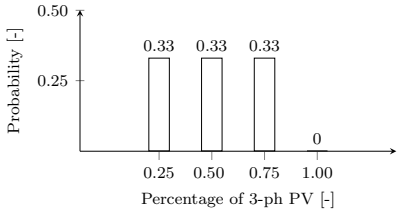
online tap changers and static VAR compensators. A multi-objective HC method for renewable energy optimization is proposed in [165]. These methods usually work for MV networks, with simplified linear power flow, assuming a single-phase equivalent of the network. The resulting HC is for a specific location of a single large size PV. The method proposed in [164] and [142] considers some of the uncertainty in the input variables but the *operational limits* are not stochastic. A data-driven method proposed in [166] calculates HC based on stochastic load and PV generation model without considering the grid constraints. A spatiotemporal and stochastic voltage sensitivity analysis proposed in [167] evaluates the HC of LVDS without relying on Monte-carlo simulations. In [25], a transfer impedance method is proposed to reduce the computational effort of the flow-based method. Compared to the power flow-based methods, the computation time is lower in analytical methods, but at the expense of simplified assumptions and design for the worst case. Although interesting, these studies are not included in the scope of the present work as they need simplification of the power-flow equations and dropping stochastic



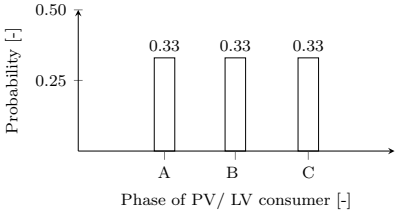
(a) PMF of PV size in M^{ref}



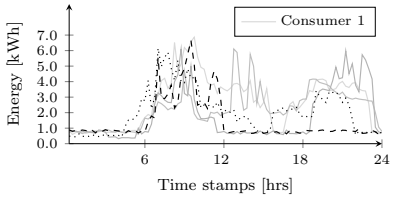
(b) PMF of PV penetration



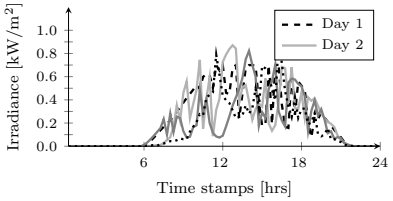
(c) PMF of percentage of 3-ph PV in M^{ref}



(d) PMF of phase of 1-ph PV/ consumer in M^{ref}



(e) Selected load profiles for a consumer-type



(f) Selected profiles from Irradiance profile pool

Figure 3.3: Uncertainties considered in M^{ref}

chance constraints for robust design.

3.4 Reference method for stochastic HC

A reference HC method is proposed in this paper as $\mathbf{M^{ref}}$ which takes all uncertainty and limits in consideration and acts like a base case for comparison. All the limits in (3.9) were considered, with $U^{max,\,th}$=1.05 pu, $U^{min,\,th}$=0.95 pu and $(1-\varepsilon) = 0.05$. The different uncertainties considered are:

i) PV size: The PV sizes are the samples from the PMF shown in Fig. 3.3a; which is similar to method M7.

ii) PV number uncertainty: The actual PV numbers in different MC samples are obtained by multiplying the maximum number of PV installations N_p multiplied by the factor obtained from PMF in Fig. 3.3b (similar to M1).

iii) PV type: The percentage of three phase PV devices is sampled from PMF in Fig 3.3c (similar to method M1).

iv) PV location: The PV location is randomly selected with equal probability for all LV consumers to get a PV installation (similar to M4).

iv) PV/LV phase: The phase in which single phase PV and LV are connected are randomly sampled with each consumer having equal probability of being connected to any of the phases as shown in Fig 3.3d (similar to M2).

v) Load injection uncertainty: The daily load profiles from SM data as well from LINEAR project [107] and Fluvius open data[1] are pooled according to the consumer annual kWh and connection kW into 25 categories. Each category has the daily load profiles obtained from the annual load profile. Then, each LV consumer is assigned with a one-day load profile randomly selected from the category matching their features per MC run. This method is a mixture of methods M4 (sampling from historical data) and M11 (based on consumer features). The 25 categories were formed based on annual energy consumption and connection capacity, taking in mind that each category has enough load profiles to sample from. Fig. 3.3e shows five randomly chosen profiles for one of the categories.

vi) PV irradiance uncertainty: The irradiance profile is selected randomly from the pool of daily irradiance profiles of Genk, Belgium in the year 2016 (similar to method M4). The 5 randomly selected profiles are shown in Fig. 3.3f. While selecting the PV and load profiles, metadata of similar days was created. The load profile for each scenario was then selected from the days similar to the PV profile chosen.

The transformer voltage is not considered as an uncertainty as no literature was obtained regarding the MV voltage uncertainty representation in form of PDF or PMF. The maximum number of PV installations is increased until the feeder is congested. Once constraints are violated, the loop is terminated and the previous P^{R} is used to calculate stochastic PV HC.

[1] https://www.fluvius.be/nl/thema/open-data

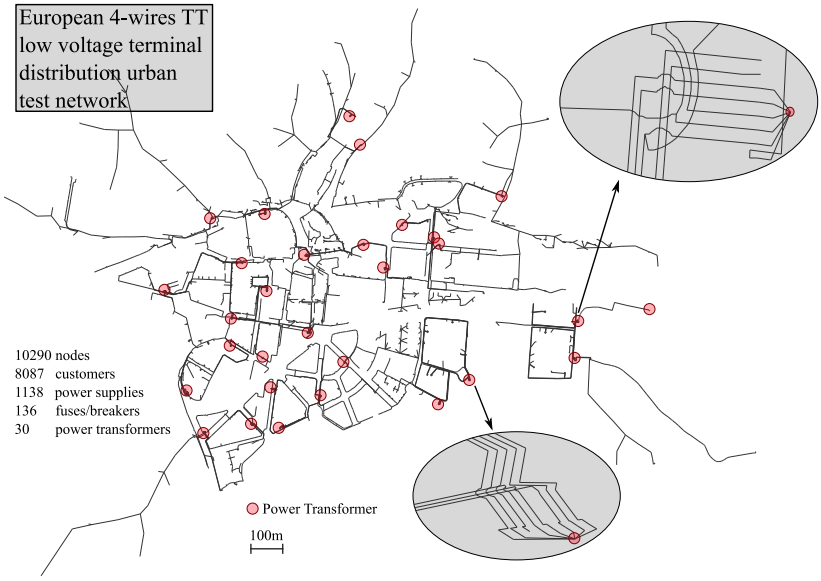


Figure 3.4: Single line Diagram of the test network [30]

3.4.1 Reference constraints for stochastic HC

The methods in comparison not only differs in term of uncertainties but also in terms of the constraints used. The comparison between the methods is not fair if the different limits and constraints used by different methods is not consistent. To understand the effect of different inputs, sampling methods and PV deployment scenarios, the voltage limit and stochastic constraints of the reference method is kept the same as the method under evaluation. For this reason HC from $\mathbf{M^{ref}}$ is evaluated for the combination of three different voltage limits and four different *probabilistic limits*. The three voltage limits are: L1: $0.95 < |\mathsf{U}_{i,c}| < 1.05$, L2: $0.92 < |\mathsf{U}_{i,c}| < 1.08$ and L3: $0.90 < |\mathsf{U}_{i,c}| < 1.10$ and the four *probabilistic limits* are : C00: $(1-\varepsilon) = 0.00$, C01: $(1-\varepsilon) = 0.01$, C05: $(1-\varepsilon) = 0.05$ and C10: $(1-\varepsilon) = 0.10$. In following sections, $\mathbf{M^{ref}}$ refers to voltage limit L1 and *probabilistic limit* C05. All other cases are written with their limit and constraints, e.g., $\mathbf{M^{ref}}$(L1, C00).

3.5 Evaluation of HC in Non-synthetic European Network

In this section, both deterministic and stochastic HC discussed in Section 3.3 are critically evaluated in terms of accuracy, ease of use, data requirement, relevance to stakeholders, computational burden, and ease of interpretation. The different HC methods were evaluated using an actual European LV network to showcase their relevance on a large scale network. Furthermore, the sensitivity of the reference method $\mathbf{M^{ref}}$ proposed in Section 3.4 is evaluated for voltage limit and *probabilistic limit*. Note that voltage limit here refers to the limit $\mathsf{U}^{\max,\,\mathrm{th}}$ and $\mathsf{U}^{\min,\,\mathrm{th}}$ and *probabilistic* limit refers to ε_u in (3.9a) and (3.9b).

3.5.1 Test Network

The test network used is the network of a sub-urban city in Spain presented in [30]. The test network is an existing system consisting of 10290 nodes, 8087 consumers, 160 feeders and 30 power transformers (Fig 3.4). Real load of the consumers for 20 days obtained from smart-meter data are publicly available[2]. The load profiles or the worst-load are then assigned from the historical load profiles as mentioned in Section 3.4. All the studies were done for the maritime climate typical in west-coast of Europe featuring mild summer and not very cold winter (common typically in Northern Spain, Northern France, UK and Belgium). Due to the limited number of load profiles, the historical pool of load data was supplemented by the annual load data of Belgian consumers [107].

3.5.2 Sensitivity of the reference method

The methodologies in the literature use different limits for voltage and different stochastic limits. The reference HC is computed for all combinations of those limits to observe the sensitivity of HC methods towards the choice of *operational grid limits* and *probabilistic limits*. In Fig. 3.5, the sensitivity of the $\mathbf{M^{ref}}$ is plotted against the voltage limit (L1, L2 and L3) and the *probabilistic limits* (C00, C01, C05 and C10). The stochastic HC is represented by a fitted Gaussian distribution (dotted curve), the range of HC (light gray) and 25-75% quartile (dark gray box). These values are used in Tables 3.5 and 3.6 for obtaining the relative accuracy and relative computational time.

[2] https://data.mendeley.com/datasets/685vgp64sm/1

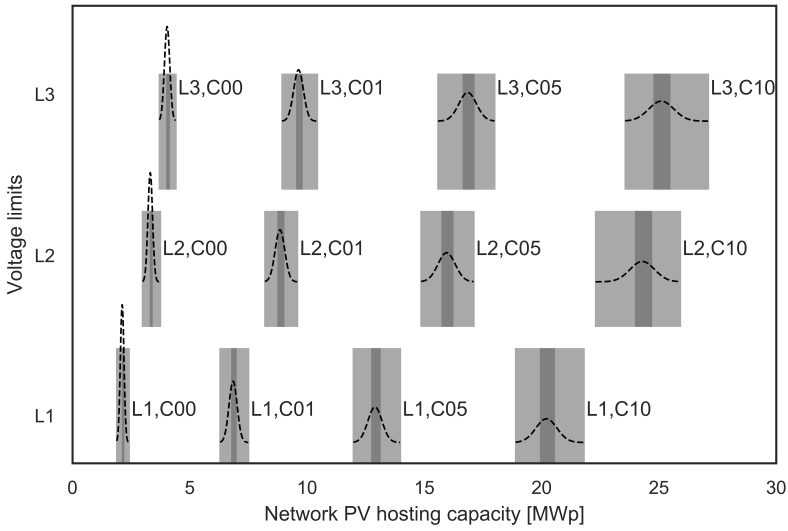


Figure 3.5: $\mathbf{M^{ref}}$ with different *operational limits* (voltage: L1, L2 and L3) and *probabilistic limits* (C00, C01, C05 and C10). (L1: $0.95 < |\mathsf{U}_{i,c}| < 1.05$, L2: $0.92 < |\mathsf{U}_{i,c}| < 1.08$ and L3: $0.90 < |\mathsf{U}_{i,c}| < 1.10$); C00: $(1-\varepsilon)$=0.00, C01: $(1-\varepsilon)$=0.01, C05: $(1-\varepsilon)$=0.05, C10: $(1-\varepsilon)$=0.10.

The reference C00 is a strict *probabilistic limit* where the probability of violation is 0. It is seen that the HC is highly underestimated in this case, even when voltage limits are relaxed. HC increased three-fold when the probability of violation was increased to 0.01. Unlike deterministic methods, where a single scenario is taken, the reference stochastic method considers 100 such scenarios. Keeping a *probabilistic limit* of 0.00 means there should be no congestion at all in any of these 100 scenarios. However in a distribution system, the voltage and current can cross their limits for some instances without hampering the network operations and assuming no violations highly underestimates the real HC of LVDS.

From Fig. 3.5, it is seen that the difference in HC between C05 and C10 is smaller than that between C00 and C05. An interesting observation for C05 and C10 is that relaxing the voltage limit from L2 to L3 does not affect the network HC significantly. At this high PV penetration, the thermal rating of the cables comes to play along with voltage limits to restricts any further PV installations. Based on these observations, C05 seems to be a good choice for *probabilistic limit*

for HC capacity calculation to remain in a slight conservative side. Similarly, voltage limit L2 seems to be more accurate choice for stochastic HC as relaxing voltage constraints from this point does not significantly increase HC. Since, the effect of the medium voltage (transformer voltage) is not considered in the reference method, limit L1 is considered to be standard limit to compensate the effect of voltage fluctuations in medium voltage grid. The key takeaway from this sensitivity analysis is that the voltage and stochastic limits definition plays an important role in defining the actual hosting capacity of the network. Setting the best stochastic limits is not straightforward due to limited measurement in low voltage network. Further research and more measurements in LVDS are required to come up with a practical stochastic limit calculation method.

3.5.3 Deterministic HC

Before moving towards the evaluation of stochastic methods, the deterministic methods of type **D1**, **D2** and **D3** are evaluated for the test network. Instead of implementing all deterministic methods in literature, specific cases for **D1**, **D2** and **D3** are chosen. In **D1**, penetration ratio of 50 % is maintained and in each step 0.5 kWp PV is added to a feeder. The maximum PV a consumer can have is limited to 15 kWp. In **D2**, P^{kWp}=4 kWp is fixed and the penetration ratio is incremented by 5% in each step. In **D3**, the difference from **D2** is that P^{kWp} is selected from $\mathcal{P}^{\mathrm{rt}} = \{1,2,3,..., 10\}$ kWp.

The three deterministic methods, **D1**, **D2** and **D3** calculated using voltage limit L3 have nearly equal HC, 24.69, 22.13 and 22.46 MWp (Table 3.5). The difference between the methods is less than 2 MWp (<10%). For **D3**, three different HC were calculated for all three voltage limits: L1, L2 and L3. The HC are 17.8, 21.3 and 22.46 MWp respectively for L1, L2 and L3. The difference is almost 5 MWp between the strict voltage limit (L1) and a relaxed voltage limit (L3). This shows that the correct choice of voltage limits influences the deterministic HC results more than the choice of the method itself.

The accuracy of these methods was compared with stochastic method $\mathbf{M^{ref}}$(C00, Lx). The *probabilistic limit* C00 converts the stochastic constraints into a deterministic one. From this comparison, it is seen that neglecting the input uncertainties leads towards at least six times over-estimation of the HC in LVDS. The methods **D1**, **D2** and **D3** themselves have little difference in terms of the final output of HC obtained. On the positive side, the computation time of deterministic methods is at most 15% of that of $\mathbf{M^{ref}}$.

Finally, **D4** was evaluated based on HC in **D1**, **D2** and **D3**. The HC of the network as a ratio of the transformer capacity is 118% and as a ratio of peak

load in network is 42.36%. This supports our previous hypothesis that the transformer does not play limiting role in the PV HC of LVDS.

3.5.4 Stochastic HC

The stochastic HC obtained from methods in Table 3.4 and the deterministic PV HC obtained from methods in Table 3.5 are presented along with the M^{ref} in Fig. 3.6. In Table 3.6, the HC results of these methods are presented and compared with both $M^{\mathbf{ref}}$(Lx, Cx) and $M^{\mathbf{ref}}$. As previously expressed, the HC calculated from different methods are quite different for same network. The method varies both on the stochastic inputs used as well as the *operational limits* and *probabilistic limits*. In following section, a simplified interpretation of the impact of the different assumptions while taking the stochastic inputs are discussed. The effect of difference in *operational limit* and *probabilistic limit* is reduced by comparing the results with the reference method having similar limits, i.e.,$M^{\mathbf{ref}}$(Lx, Cx).

3.5.4.1 Effect of granularity of PV deployment

One of the prominent differences among the stochastic HC methods in the literature is how PV is increased in the feeder. The effect of granularity of PV is shown by a comparison between methods M2 and M6. M2 has a more rigid way of PV allocation where the only available PV capacity is 6 kWp, but uses a *probabilistic limit* C10. In M6, a PV capacity of up to 7 kWp is allocated to each consumer but in steps of 1 kWp, with limits (L3, C00). HC calculated from M2 was 1.005 times the M^{ref}(L3, C10) and that of M6 was 6.7 times of M^{ref}(L3, C00) and 2.8 times of M^{ref}(L3, C01) (Fig. 3.7a). Moreover, the similar HC of

Table 3.5: Evaluation deterministic HC for European Feeders with respect to $\mathbf{M^{ref}}$

	HC def	Voltage Limit	HC equation	PV scenarios	accuracy [ratio of M^{ref}(C00, Lx)]	Computation time [ratio of M^{ref}]	HC_N
D1	(3.8)	L3	(3.1)	$\sum p$ is increased in step of 0.5 kWp (<15 kWp) till violation; P^{R}=0.5	5.915	0.135	24.69 MWp
D2	(3.8)	L3	(3.5),(3.6)	P^{R} is increased in step of 5% for fixed P^{kWp} =4	6.59	0.148	22.13 MWp
D3	(3.8)	L3	(3.1),(3.5)	P^{R} is increased by 5%; PV selected from $\mathcal{P}^{\text{rt}}$ ={1,2, ..., 10}	6.002	0.135	22.46 MWp
		L2			6.31	0.11	21.3 MWp
		L1			9.96	0.152	17.872 MWp
D4	(3.8)	L3	(3.3), (3.4)	The HC obtained from **D1**, **D2** or **D3** is expressed in reference to the $\mathsf{S}_t^{\text{rated}}$ or $P^{\max}$	-	-	HC^{t} = 118%; HC^{pl}=42.36 %

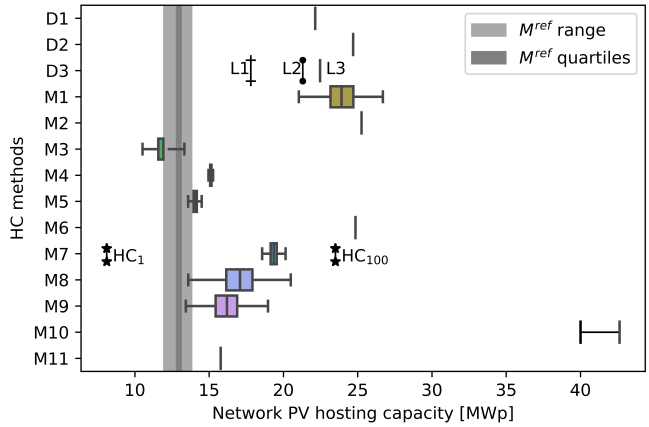


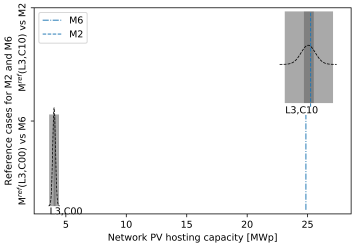
Figure 3.6: Comparison of HC_N obtained from stochastic and deterministic HC methods with M^{ref}.

M2 and M6 in absolute terms shows that finer granularity of assigning the PV size reduces the probability of *operational limits* violations. On the other hand, both assumptions are not realistic for LVDS, where consumers have preferences for random PV sizes.

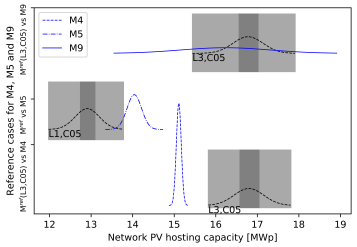
3.5.4.2 Impact of using realistic load and generation

The methods M4, M5 and M9 have HC results similar to the reference method as they consider more realistic scenarios of load and PV compared to the other methods (Fig. 3.7b). In M4, one-day PV and load profiles are picked randomly from the pool of historical load and PV profiles. The worst-case load of each consumer is randomly picked from the pool of load-profiles in M5. The load and generation of the consumers are sampled from the non-parametric PDF obtained from the smart-meter data in M9.

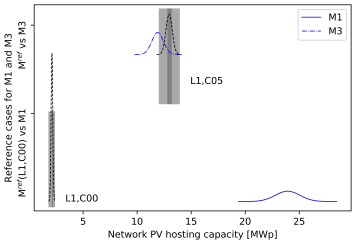
The HC calculated by M4 is less and with lower variance than the reference case because PV size is randomly chosen from a narrower set $\mathcal{P}^{\text{rt}}_{M4} = \{1, 1.5, ..., 4\}$ with lower ratings compared to reference case with $\mathcal{P}^{\text{rt}}_{M^{\text{ref}}} = \{1, 2, ..., 10\}$. The HC calculated from M9 has a higher variance because a) the PV locations are fixed allowing more space for variation of PV capacity, and b) the time-correlation is lost between PV generation and load when sampling them from a PDF. The reason for HC_N calculated in M5 to be slightly higher than reference method is



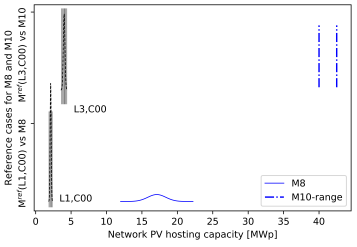
(a) Effect of PV granularity: M2 vs M6



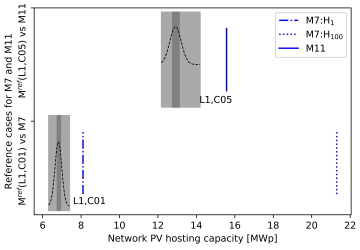
(b) Effect of load profile: M4, M5 and M9



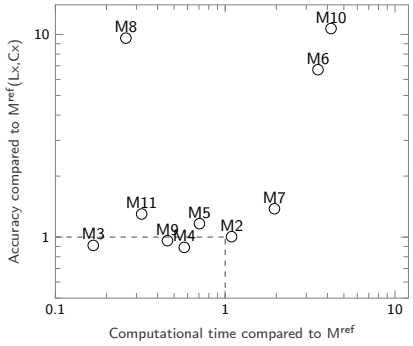
(c) Effect of PV: M1 and M3



(d) Commercial PV HC: M8 and M10



(e) Methods M7 and M11



(f) Relative computational time vs accuracy

Figure 3.7: Comparison of different HC methods and their respective reference methods.

attributed to fact that it considers constant peak load. For residential consumers, normally peak load are in the morning or evening, and considering the peak load at the time of peak generation increases the PV hosting capacity.

Out of the three methods, M4 is nearest to the reference method as it considers the time-correlation between load and PV generation by taking the real one-day profiles and considers PV location as a stochastic variable.

3.5.4.3 Effect of PV size uncertainty

M1 and M3 are compared in terms of PV HC in Fig. 3.7c. M1 samples PV size from $\mathcal{P}^{\text{rt}} = \{1, 2, ..., 10\,kWp\}$ and M3 from $\mathcal{P}^{\text{rt}} = \{1, 2, ..., 5\,kWp\}$ making the total PV HC in M1 1.6 times that in M3. However, the comparison of these methods with $\mathbf{M^{ref}}$ is not possible as both methods: a) check feeder strength for PV deployment uncertainty marking each feeder as strong or weak, b) check only the voltage unbalanced factor, and c) use fixed load of 1.5 kW per consumer (M1) or unbalance 350 kW spread over the feeder (M3). Even-though the HC from M3 broadly corresponds with reference method (Fig. 3.7c), it is not an actual HC as the PV installations in the feeders with problem are also counted.

3.5.4.4 Commercial PV HC

Unlike residential PV,commercial PVs have higher ratings, with three phase inverter and usually at a single point in the feeder. The methods M8 and M10 check the HC of LV feeder for a single three-phase PV installation. The HC calculated is 9.59 (M8) and 10 (M10) times higher than the residential PV HC capacity of the network calculated by the reference methods (Fig. 3.7d). These methods show that a HC of a feeder is higher for a single PV device compared to multiple PV installations. These methods can be used to estimate the size of a large community PV a feeder can host but not for the distributed residential PV hosting capacity.

3.5.4.5 Methods M7 and M11

Method M7 has a unique way of defining HC: the HC_1 is the first scenario with a violation and the HC_{100} is the full violation scenario. 50 scenarios are simulated with each step having 2% increments in the number of PV installations. HC_1 approximates the HC calculated using $\mathbf{M^{ref}}$ with the stochastic limit C01 ($1 - \varepsilon = 0.01$) and HC_{100} is close to the HC derived from deterministic methods (Fig. 3.7e).

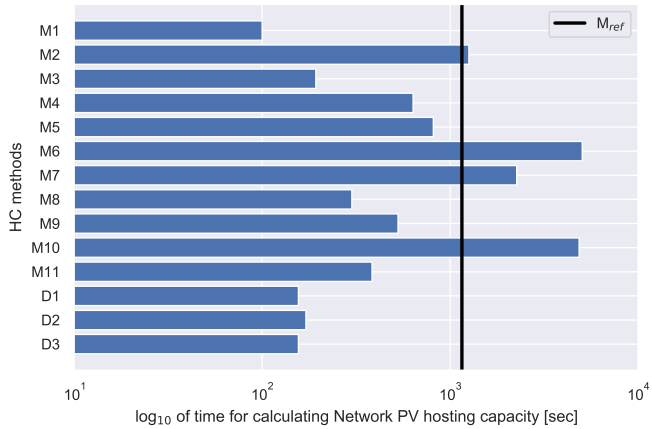


Figure 3.8: Comparison of time for HC calculation, stochastic, deterministic and reference.

M11 proposes HC in terms of PDF of HC^t. The method uses the historical load profiles and PV generation profiles but limited to the two hours of the day when PV is at its peak. The PV penetration is fixed to 20% but PV size is increased by 1 kWp in each step. The HC estimates are 1.3 times higher than the reference method, with the computational time reduced by half (Fig. 3.8). Overestimation of PV HC in M11 is due to consideration of only one uncertainty, i.e., location of PV. It also shows that a feeder can have more HC if a limited number of consumers can have larger size PVs.

3.5.5 Computational advantages

In Fig. 3.7f, the relative computational time of each method is plotted with its relative accuracy ratio (from Table 3.6) when solved using same hardware (16 GB RAM, Inter Core-i7-8650U CPU), `concurrent.futures.ProcessPoolExecutor` of python for parallel computation, and in-house probabilistic power flow tool, for finding the HC_N. The methods M2, M3, M4, M5, M9 and M11 have a relative accuracy ratio of 1. As discussed above, M2 has a quite unrealistic PV allocation style, and HC from M3 is not accurate as the PV of feeder in congestions are also included. The common features among the remaining methods (M4, M5, M9 and M11) is the use of realistic load and generation profiles. Of these four methods, M11 is computationally faster as it takes only two hours during peak generation for checking violations.

Table 3.6: Evaluation of SHC in Non-synthetic European Feeder

	Data used		constraints used	limits		ref: (Lx, Cy)	accuracy [ratio of] (Lx, Cy)	relevance to LVDS	output HC	computational burden [ratio of] (L1, C05)	(Lx, Cy)	specific purpose	HC
	PV generation	Load		Voltage	Stochastic								
M1	maximum; P_p^{kWp} sampled from pdf	minimum	Voltage unbalance	-	C00	(L1, C00)	13.39	only for unbalanced factor	stochastic	0.087	0.193	For VUF related issues	23.96 MWp (mean)
M2	maximum for P_p^{kWp}=6	minimum	Overvoltage	L3	C10	(L3, C10)	1.005	only for high PV	deterministic	1.090	0.881	For high PV	25.25 MWp
M3	maximum; P_p^{kWp} sampled from pdf	randomly generated	-	-	-	(L1, C05)	0.91	only for unbalance factor	stochastic	0.167	0.167	Index of VUF is created for each feeder	11.89 MWp (mean)
M4	pool of historical profiles P_p^{kWp} from pdf	pool of historical profiles	over-voltage; undervoltage & current in branch 1	L3	C05	(L3, C05)	0.89	PV hc of with realistic scenario	stochastic	0.574	0.554	get PV HC (in terms of P^{R}) of real feeders	15.1 MWp (mean)
M5	maximum; P_p^{kWp} incresed unformly	pool of worst-case load	over-voltage; undervoltage; voltage unbalanced & over-current	L1	C05	(L1, C05)	1.165	HC with source voltage uncertainty	stochastic	0.705	0.705	HC^{pl} for simplified feeders	14.054 MWp (mean)
M6	maximum; P_p^{kWp} incremented by kWp	sampled from beta distribution	overvoltage	L3	C00	(L3, C00)	6.7	HC with small granularity of PV increment	deterministic	3.52	6.36	LVDS max PV capacity with small granularity	24.86 MWp
M7	historical profile; P_p^{kWp} pdf	fixed total load distributed randomly	overvoltage	HC_1: L1	C01	(L1, C01)	1.38	multiple scenarios for 2% increment of P^{R}	deterministic	1.95	4.35	HC_1 and HC_{100} are calculated	8.1 MWp
				HC_{100}: L1	-	L1, **D3**	1.19			1.95	15.04		21.3 MWp
M8	historical profile; 1 or 2 PV in each feeder	fixed hourly load	overvoltage and undervoltage	L1	C00	(L1, C00)	9.59	1 or 2 PV per feeder (commercial)	stochastic	0.260	0.579	suitable for commercial type PV in MV	17.13 MWp (mean)
M9	sampled from pool of smart meter data	smart meter data	overvoltage; undervoltage and voltage unbalanced	L3	C05	(L3, C05)	0.96	HC based on real data of LV consumers	stochastic	0.457	0.477	for known scenarios and smart meter dat	16.29 MWp (mean)
M10	maximum; sum of P_p^{kWp} increased uniformly	minimum	overvoltage and undervoltage	L3	C00	(L3, C00)	10.69	maximum possible HC in Medium Voltage	range	4.202	7.613	more suitable for MV level studies	40-42.5 MWp
M11	2 hrs (11:00-1300 hrs); P_p^{kWp} sampled from pdf	real profiles for different consumer type	overvoltage, undervoltage, overcurrent, voltage unbalanced and transformer overload	L1	C05	(L1, C05)	1.30	HC for LVDS with 20% P^{R}	stochastic	0.332	0.332	HC based on two hours PV profile only	15.57 kWp

The logarithmic plot of absolute time for all deterministic and stochastic methods is presented in Fig. 3.8. This plot illustrates several interesting observations:

1. The computational time of deterministic methods **D1**, **D2** and **D3** are the fastest but almost equal to M1 and M3. The single iteration for PV deployment in M1 and M3 makes their speed comparable to **D1**, **D2** and **D3** where iterative methods are used. (In M1 and M3, all scenarios are simulated at the same time and the scenarios with violations are filtered out.)

2. M2, M5 and M6 are computationally expensive because of the very low granularity of PV deployment scenarios (1 kWp incremented for the whole feeder in each iteration). The observed time for M2 and M5 was the result of using higher steps when the feeder was very far from violations.

3. M8 and M10 both have similar PV scenarios. M8 is faster as it uses a single timestamp, runs all the scenarios at the same instance, and filters out the scenarios causing violations.

4. The method HC_{100} results given by M7 can easily be calculated by using deterministic methods in a fraction of time required for M7. It saves the time consumed in iteration from the scenario having HC_1 to that having HC_{100} with an increment of P^{R} by 2%.

3.5.6 Discussion

In the last 20 years, there have been many studies defining PV hosting capacity. The review papers [18, 132, 139, 140] point out the difference between methods for PV hosting calculation in terms of tools used for calculating violations, consideration of uncertainty, or the performance index used. Finding a suitable method for stakeholders such as DSOs, regulators, or other service providers is not an easy task with such studies. This paper aims to bridge this gap in the literature by implementing the proposed methods and applying them to a real-life network. This study shows that for the same set of feeders, different formulations and assumptions lead to an entirely different PV HC (Fig. 3.5 and 3.6). The disadvantage of using a deterministic calculation based on the PV peak is that the resulting HC is based on only one snapshot of many possible situations. Similar is the case when HC capacity is calculated using very low load and peak generation. The best practice for avoiding such pitfalls is by using historical load and PV profiles creating multiple scenarios (similar to **M4** or **M11**). The other critical point is on choosing how PV is allocated to the customers. The lower granularity of incremental PV results in a higher HC

value. Still, methods in which PV size is randomly sampled from a PDF seem to be better considering computational tractability. In methods M8 and M10, a commercial PV,i.e., a single PV of larger size, gives an escalated HC number in the residential network. However, it comes with other practical issues like land availability and the community's eagerness to have shared PV in residential feeders. In most of the studies, time-series analysis is neglected. This works fine if we are interested only in PV HC without considering other factors such as coincident loads. However, it will not be the case if there are new types of loads in the system like EV and heat pumps where the HC of the network should be evaluated in both dimensions (overvoltage and undervoltage) over time.

This study can be of use for both DSOs and regulators. DSOs can use the paper outcomes to select appropriate uncertainties to incorporate in their analysis and provide more precise limits to new installations or operational limits to market players. For regulators, the study can be used to promote future large scale PV deployment by defining *grid operational limits* and *stochastic limits* for the DSOs and encouraging the use of techniques to meet those limits. The critical aspect of these calculations is the selection of the proper *operational limits* and *stochastic limits*. The sensitivity study shows a possibility to DSOs and regulators on manipulating these limits by using the tools and techniques like voltage regulating devices or PV curtailment allowance to increase the PV HC. This study of hosting capacity with different *stochastic* and *voltage limits* can be used as a reference by LVDS planners for high penetration of renewable energy resources.

3.6 Conclusion and Future Works

In this paper, existing deterministic and stochastic hosting capacity methods were evaluated and benchmarked using the "Non-synthetic European LV test feeders" with actual network and smart meter data. The methods were benchmarked using the same network data and a similar computation framework to make the comparison fair. A stochastic reference method was proposed with twelve different combinations of voltage (*operational limits*) and *probabilistic limits* to remove the influence of these limits while studying the role of the input uncertainties in PV HC calculation. With the benchmark scenario, it was shown that the deterministic methods over-estimated the PV HC by six times but with a computation time of 15% of the stochastic reference case ($\mathbf{M^{ref}}$). Deterministic HC is a single snapshot out of a stochastic HC space depending on various scenarios and uncertainties. Hence, deterministic calculations cannot depict the true picture of feeder congestion. For stochastic HC methods, the granularity of PV deployment influences both computation time and accuracy.

The methods where finer granularity (increment of 1 kWp or 1 PV per feeder) in each iteration was adopted had a 10-15% higher estimation of PV HC compared to reference case but had up to 5 times more computation time. Another factor affecting the PV HC result is the number of PV installed. Methods where fewer consumers can have higher size PV installations had relatively high PV HC compared to many consumers with smaller PV. Methods used for medium voltage HC or commercial PV HC overestimates the residential LVDS because the residential consumers have an equal probability of having PV installations of smaller size. PV HC calculations using only PV peak generation save time but over-estimate the PV HC for a realistic system. Finally, the method $\mathbf{M^{ref}}$, with voltage limit 0.95-1.05 pu and *probabilistic limit* of 5%, was proposed as a better approach finding residential PV HC of LVDS . This method samples the PV generation and load from real historical data and considers uncertainty in PV size, location, phase, type, number, and transformer voltage.

Chapter 4

Probabilistic Power Flow using General Polynomial Chaos

Koirala, A., Van Acker, T., D'hulst, R., & Van Hertem, D. (2022). *Uncertainty quantification in low voltage distribution grids: Comparing Monte Carlo and general polynomial chaos approaches.* Sustainable Energy, Grids and Networks, Art. number 100763. https://doi.org/10.1016/j.segan.2022.100763

The first author is the main author of the article. The contributions of the first author include the literature review, methodology, case-study, software implementation, and the writing of the article.

Abstract: Changes in load and distributed generation in low voltage distribution systems (LVDS) have made the individual consumer offtake highly uncertain for the system operator. In order to accurately determine hosting capacity of such systems, congestion related stochastic indices, e.g. probability of undervoltage and overvoltage, must be calculated. These require either too many deterministic calculations (simulation-based methods) or the use of analytical probabilistic power flow. Numerical simulation-based methods often use a Monte Carlo (MC) based approach due to its simplicity. Recently, analytical methods such as general polynomial chaos (gPC) expansion have gained increasing interest. This paper develops a non-intrusive gPC formulation for congestion determination in LVDS and illustrates its effectiveness compared to the MC methods. Both MC and gPC methods are compared for computational time, and

accuracy, for a set of realistic feeders with high photovoltaic (PV) penetration. The PV power injection uncertainty is characterized by a univariate continuous distribution. The paper illustrates the merits of using the general polynomial chaos expansion with degree 2 and Sobol sequences based testing points for the probabilistic assessment of LVDS with many uncertainty sources.

Keywords: probabilistic power flow, Monte Carlo simulation, general polynomial chaos, distribution systems, congestion uncertainty

4.1 Introduction

4.1.1 Background and Motivation

The traditional Low Voltage Distribution System (LVDS) is designed radially, supplying end consumers with electricity delivered by a big centralized generation connected to the upstream high voltage transmission network. This simple role meant the network was developed passively, using the "fit and forget" approach. Such systems are made with over-investment so that distribution system operators (DSOs) need low intervention and rarely any detailed calculations. Traditional deterministic power flow tools are scarcely used. But with high penetration of distributed energy resources and the introduction of new types of load, e.g., electric vehicles and heat pumps, a higher level of randomness is introduced. This change results in a system subjected to more variable power flows, leading to more uncertainty with respect to system congestion. Addressing these changes is new to system operators as these uncertainties are location dependent, substantial in size and time, and have a more significant impact than traditional distribution loads. To adequately operate and develop the system, the current DSO needs to be able to calculate and forecast the grid state under uncertainty [168]. As such, a probabilistic power flow (PPF) tool is needed to take the stochastic parameters introduced in LVDS into account [14]. Specific to congestion type analysis is the interest in the tail region of probability distributions, where the bounds are exceeded. These probabilistic tools are more relevant in congestion assessment of LVDS with very high PV penetration due to a higher tail-region uncertainty.

The commonly used deterministic power flow analysis is simple but is insufficient if the input is stochastic and output is required as a stochastic indication. The operator needs to account for stochastic input parameters to obtain statistical properties of resulting voltages and currents, such as mean, variances, medians or quantiles. From the early days of computing, Monte Carlo (MC) methods were proposed to carry out this type of analysis as it avoids an extensive

reformulation of the original problem [22, 105, 123, 125]. The basic Monte Carlo (MC) method, also known as crude MC, applies the deterministic function to many randomly generated input samples and collects its corresponding outputs. Quasi-Monte Carlo (qMC) methods are similar to the MC methods but use deterministic samples of input represented by algorithmic series [15, 169]. The outcome of the MC and quasi Monte Carlo (qMC) simulations are estimates of the quantity of Interest (QoI), such as nodal voltages or branch currents.

Recently, analytical methods such as the general polynomial chaos (gPC) expansion are gathering interest in power system applications due to their high accuracy with low computational time [16, 71–74, 76, 170, 171]. It has been shown in the literature that the general Polynomial Chaos (gPC) outweighs other analytical methods used for probabilistic power flow as this method preserves the non-linearity of the power flow. However, limited work has been done regarding its usage in LVDS with load and distributed generations represented by non-Gaussian distribution obtained from historical profiles. Furthermore, the comparative analysis of various combinations of gPC with MC or qMC methods to obtain an estimation of mean and tail-event probabilities remains unexplored. Nevertheless, such analysis is quite interesting for DSOs to quickly assess congestion related parameters such as the probability of overvoltage on a large LVDS with large variability or uncertainty, e.g., with high PV, heat pump or electric vehicles penetration.

4.1.2 Literature Overview

MC methods are easy to formulate but require thousands of deterministic simulations yielding in an overall high computation toll [60, 63]. MC simulations in LVDS are considered for reliability assessment [172], voltage unbalance analysis [123], optimum PV hosting capacity calculations [125], calculation of network congestion risk [22], and the economic viability of PV installations [105]. The common interest in most of these studies is to efficiently estimate QoIs related to the violation of *operation limits* in LVDS. Large scale application of MC simulations in [21] used selective time-stamps to reduce the computational burden. In [173], the use of MC simulations for a distribution system with high renewable resources is explored in detail by sampling the load data from a non-parametric cumulative distribution function (CDF). The method simplifies the distribution system to a balanced, single-phase representation, but requires 10 000 deterministic power flows per feeder for the QoI to achieve a convergence of 0.1 %. As a consequence, the MC simulation requires a huge computational effort for large scale LVDS. To be of practical use, the analysis of such a system requires either a less computationally demanding model without losing accuracy [169] or more computing power [125]. The computation time can be reduced using

methods that simplify the power flow equations, e.g., using a balanced, single-phase representation, or using an approximation of power flow equations, e.g. distflow, transfer impedance [25]. However, these directly impact the accuracy. Other probabilistic power flow (PPF) techniques are available in the literature to reduce the computation time while maintaining acceptable accuracy. Those techniques can be widely grouped into three groups a) simulation-based methods: MC and qMC schemes; b) analytical methods: linearized power flow, point estimate method, and gPC; and c) other schemes: correlation based, machine learning based or optimization based methods (Fig. 4.1).

The qMC scheme used in [15] uses importance sampling with Sobol sequences to assess the impact of maximum PV penetration in LVDS. In [169], the effectiveness of qMC with Sobol sequences is shown for generating the probability density functions (PDF) of bus voltages.

Analytical methods such as the Point Estimate Method (PEM) [65–69], unscented transformation method [70], cumulant tensor method [71], and orthogonal transformation based method [72–74] are gaining interest for LVDS applications. They are complicated compared to MC methods, although they are superior in computation time. The PEM needs additional techniques to reconstruct the cumulative distribution function (CDF) after the analytical power flow [68, 69], which has shown outstanding results but is only valid with Gaussian input. Furthermore, this technique is not suitable when the QoI is in the tail-region [68]. In [64], the Gaussian Component Combination Method is used where the Probability Density Function (PDF) of wind generation is represented by a combination of Gaussian distributions. This approximation makes analysis faster for the test case with a few sources of uncertainty with the cost of losing the information in the tail-region. Orthogonal polynomial based methods are recommended in [16, 72–74, 76, 170, 171, 174] for power systems related problems as they directly give the PDF of the output, sustain the non-linearity of power flow equations and have an easier formulation. However, few works are there for radial LVDS [72, 170, 171, 174]. In [72], the authors focus on using gPC based on a stochastic collocation method to reduce the computation time, but the method used is limited to mean value estimation for time-series load data, using simplified Gaussian loads. In [170], the authors reduce the number of uncertainties by leveraging the correlation between the distributed generators in LVDS. In [171], gPC is used to propagate uncertainty in radial LVDS to show the computational gain compared to the MC method by using simplified Gaussian uncertainties. In [174], a basis-adaptive sparse polynomial chaos espansion (sPCE) is introduced to perform PPF analysis in distribution feeders, where the nodal voltage is evaluated by their proposed gPC and improved MC. The authors concludes that sparse polynomials take advantage of interactions among the low-degree terms reducing the computational burden

on the higher degree terms.

Although all these studies explore the possibilities of reducing the computational time of gPC in LVDS, the effectiveness, especially in the tail region events compared to MC methods, is not investigated. Most of these studies use a simplified test system where a Gaussian distribution replaces the load by its uncertainty. The output considered in these studies is generally nodal voltage in one node of the feeders. Furthermore, the necessity of solving gPC of all degree for to find the truncation error on every run makes the process computationally expensive, even after using advanced methods like sPCE. The other works using gPC [16, 89] deal with transmission systems, where 2^{nd} degree basis polynomials are found to be sufficient for such systems but, no such study is available in the literature for LVDS. A new analysis of variance (ANOVA) based polynomial construction technique is validated using arbritarty polynomial chaos expansion for transmission networks in [90] and a second order polynomial is found sufficient for the considered case study. In [175], non-linearly correlated variables in transmission network are considered and degree 3 polynomials were found to be suitable for such applications. The authors of [176] propose a data-driven sPCE to access transfer capacity of the transmission network. However, the comparison is with MCs only and not with gPC.

The general conclusion from the literature review is that probabilistic methods are gaining interest among researchers and DSOs for solving multi-dimensional stochastic problems in power systems. Such probabilistic methods need to be a) computationally efficient, b) compatible with non-Gaussian inputs, c) capable of keeping the statistical properties in both the tail and mean regions, and d) well-suited to the non-linear nature of the power flow equations. The gPC-based method suits all these objectives very well [34]. A detailed analysis of gPC with varying degrees of basis polynomials and a comparison of accuracy and computational efficiency with more common alternative methods such as MC and qMC methods is required to choose the better approach for uncertainty quantification. To the best of the authors' knowledge, such a study of LVDS is not available in the literature. A typical LVDS has more nodes and hence more uncertainty sources, and requires a full unbalanced power flow. Furthermore, the literature is missing an evaluation of gPC in an actual LVDS with non-Gaussian load models. Finally, no past studies use gPC to evaluate QoIs on tail regions, for example, where congestion bounds are exceeded in LVDS. As such, a dedicated study of this problem is warranted.

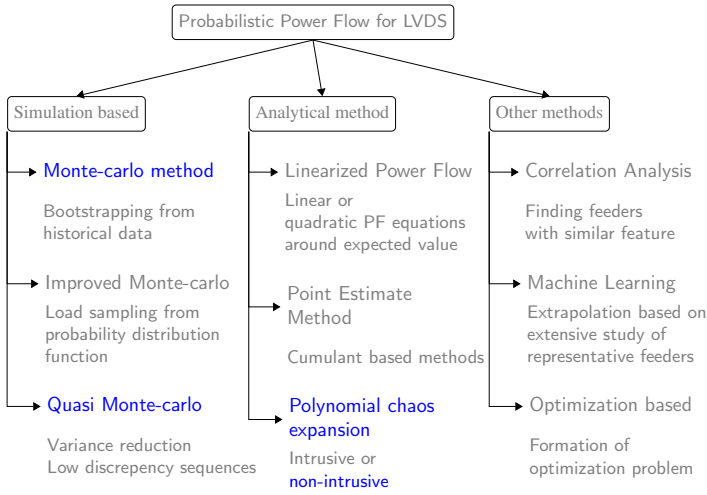


Figure 4.1: Classical and emerging techniques for probabilistic power flow. The methods indicated in blue are the focus of this study [34]

4.1.3 Contribution and Organization

This paper presents a numerical comparison between three PPF methods: MC, qMC and gPC, for LVDS feeders of an existing network. The sources of uncertainties are the load and PV generation, represented by a non-Gaussian PDF obtained from actual measurement data. A detailed study is conducted in a medium-sized feeder for both mean and tail values of a) minimum feeder voltage, b) maximum feeder voltage, and c) maximum current per branch current rating. The chosen QoIs can be directly utilised to obtain the LVDS feeders' congestion probability. The study captures both day and night scenarios in a system with high PV penetration for a typical spring day, and the most appropriate probabilistic analysis methods are identified. Finally, a study is shown where the probability of overvoltage (p^{ov}), undervoltage (p^{uv}) and overcurrent (p^{oc}) are obtained for all LVDS feeders with high PV penetration using all three PPF. This study shows the impact of the number of uncertainty sources in accuracy.

The contribution of this paper can be summarized as follows:

a) the gPC-based approach was compared with MC and qMC schemes to explore the advantages of each method in an actual LVDS for the first time,

b) different degrees and sampling techniques of non-intrusive gPC expansion are compared with MC to get the best variant of gPC for LVDS,

c) a case-study on using gPC for computing probability of congestion is presented and validated with MC and qMC to show the practical applicability in an existing LVDS with high PV penetration. This study is done for 160 actual feeders with 2-12 sources of uncertainty. The proposed gPC is advantageous over MC and qMC in feeders with more uncertainty sources.

The remainder of the paper is organized as follows: Section 4.2 introduces the MC and qMC schemes. Section 4.3 presents the non-intrusive gPC approach for LVDS application. Section 4.4 presents a numerical illustration comparing MC, qMC, and gPC schemes for non-Gaussian variables with different polynomial degrees, sampling schemes and post-processing samples. Furthermore, the congestion probabilities are computed using MC, qMC, and gPC and compared for a set of LV feeders. Section 4.5 concludes the paper.

4.2 Monte Carlo Simulations

This section describes the MC and qMC methods for stochastic power flow analysis in LVDS. The case of observing a QoI y in an LVDS after MC simulations with one sample is represented as:

$$0 = g(\mathbf{U}, \bar{\mathrm{I}}, \mathbf{\Delta}, \mathbf{\Xi}) \tag{4.1}$$

$$y = h(\mathbf{U}, \bar{\mathrm{I}}) \tag{4.2}$$

where $\mathbf{U}$ and $\bar{\mathrm{I}}$ are the vectors of nodal voltages and branch currents obtained from three-phase unbalanced power flow g. The power flow g can have both stochastic inputs ($\mathbf{\Xi} = [\xi_1, \xi_2, ...,]$, i.e., load and power injection by PV) and deterministic inputs ($\mathbf{\Delta} = [\delta_1, \delta_2, ...]$, i.e., network impedance, transformer voltage and installed PV kWp). The QoI is the nodal voltage at a particular node, branch current of a selected branch, or maximum voltage along the feeder. Such QoI is obtained after applying operator h on the output state vectors of power flow equation. Each stochastic input ξ_i is defined by a continuous PDF: $f(\xi)$.

In PPF with MC, the stochastic input variables are randomly sampled from their PDF [177], and the power flow results are obtained for each set of samples. The stochastic input variables can be assumed to be independent so that the joint sampling of those variables is reduced to a series of univariate samples.

The QoI y is then calculated based on the obtained power flow results. In MC with N^{mc} samples, after N^{mc} power flow solutions, the result is a stochastic output vector $\boldsymbol{y}$ (4.3).

$$\boldsymbol{y} = [y_1, y_2, ..., y_{N^{mc}}] \tag{4.3}$$

An increased number of random samples is needed to improve the estimation of the QoI, which means an increase in the number of power flow equations to be solved. As the number of stochastic variables increase, the required number of random samples increases exponentially. The computational burden of MC is directly linked to the sample space corresponding to the amount of expensive power flow solutions required.

The computational burden of MC is usually reduced by improving the sampling techniques. The sample size can be decreased using deterministic sampling methods, often referred to as low-discrepancy sequences (LDS). Commonly used LDS are Halton sequences [58] and Sobol sequences [57]. The *qMC* is hence the MC method which uses these LDS as samples [177]. qMC has proven computational gain in some studies involving distribution grids [169,177]. Unlike MC, where different values of $\boldsymbol{y}$ can be expected in each instance of their run, the output of qMC is deterministic.

In this study, three types of QoIs are obtained:

(a) mean QoIs: mean of minimum feeder voltage ($U^{\min}$), mean of maximum feeder voltage ($U^{\max}$), and mean of maximum branch current per branch rating (I^{ov});

(b) tail-end QoIs: 5% percentile of $U^{\min}$, 95% percentile of $U^{\max}$ & I^{ov}; and

c) congestion related probabilities: p^{ov}, p^{uv}, and p^{oc}.

The key factors for satisfactory estimate of QoIs after MC are: (a) the upper and lower limits of the confidence interval (CI); and (b) the order of magnitude of the confidence interval (CI) [34] . The CI of a QoI is obtained by repeating the MC with N^{mc} samples for N^{i} instances resulting in N^{i} estimates of QoIs. The confidence interval of the QoI is then calculated as the range between the 2.5 and 97.5 percentiles of the N^{i} QoI estimates. For the results of MC to be trusted, the CI should be within the defined acceptable bound of the QoI. The convergence rate of the QoI is proportional to $\sqrt{N^{\mathrm{mc}}}$ for MC, for the estimate of the mean value when input uncertainties are Gaussian [61,63]. This might differ if the stochastic inputs are non-Gaussian or QoIs are in the tail-region. In general much more MC samples are required to increase the confidence in the estimate of tail-region results [63].

In qMC, the output after N^{mc} power flow solutions is a deterministic vector. The QoIs in consideration (mean and percentiles) have deterministic values. Increasing the qMC samples will result in the QoIs to converge towards the result of MC with a large number of samples [62].

4.3 General Polynomial Chaos Expansion

General polynomial chaos (gPC) is a widely used uncertainty quantification (UQ) method based on orthogonal polynomials. The book "Introduction to uncertainty quantification" by Sullivan is recommended for a more detailed study of this method [28]. The gPC expansion is a way to discretize the infinite dimensional problem, similar to the the fourier series for a periodic signal [178]. The stochastic QoI y is expressed as the sum of the basis polynomials of the stochastic inputs weighted by respective coefficients as in (4.4) which can extend up to infinity:

$$y = \sum_{j=1}^{\infty} c_j \Phi_j, \tag{4.4}$$

where c_j and Φ_j denote the gPC coefficient and basis polynomial, respectively. The series in (4.4) is truncated to length L to reduce computational burden. The cardinality of the reduced basis vector $\mathbf{\Phi}_\alpha = [\Phi_1, \Phi_2, \ldots, \Phi_L]$ depends upon the degree β of the basis polynomials and the amount of stochastic input variables N.

$$L = \frac{(\beta + N)!}{\beta! N!}. \tag{4.5}$$

The coefficient vector of the reduced basis vector is $\boldsymbol{y}_\alpha = [c_1, c_2, ..., c_L]$. $\mathbf{\Phi}_\alpha$ is the basis polynomials vector formed by the tensor product of one-dimensional polynomials of stochastic input variables in $\mathbf{\Xi} = [\xi_1, \xi_2, ..., \xi_N]$.

$$\mathbf{\Phi}_\alpha = \left\{ \boldsymbol{\phi}(\xi_1) \otimes \boldsymbol{\phi}(\xi_2) \otimes ... \otimes \boldsymbol{\phi}(\xi_N) \right\}_\beta \tag{4.6}$$

In (4.6), $\boldsymbol{\phi}(.)$ is the vector of univariate orthogonal polynomials for the stochastic input parameter in $\mathbf{\Xi}$. The orthogonal polynomials for most of the continuous univariate distributions are well-defined, e.g., Hermite polynomials for Gaussian-distributed variables, and Jacobi polynomials for Beta distributions [97]. After the tensor product in (4.6), the basis polynomials with order greater than β are excluded [97]. The basis polynomials in $\mathbf{\Phi}_\alpha$ are mutually orthogonal with respect to the PDF ($f(\xi)$), i.e., satisfies the expectation operation ($\mathbb{E}[.]$) in (4.7).

$$\mathbb{E}[\Phi_y, \Phi_z] = 0 \quad \forall \quad y \neq z \quad \& \quad y, z \in 1, 2, ..., L. \tag{4.7}$$

The gPC expansion is done in an intrusive way by using Galerkin projection in power flow equations or by non-intrusive way using regression fitting [88]. The gPC using Galerkin projection is overloaded with large number of integrals to be computed in LVDS due to more sources of uncertainty and nodes [29]. Hence, non-intrusive gPC where the underlining power-flow equations are not "intruded" is favourable for LVDS. The philosophy behind the non-intrusive method is that QoI y is fitted to the orthogonal basis polynomials based on the deterministic power flow results at certain selected samples, also known as testing points T^{gq} of the stochastic input vector $\boldsymbol{\Xi}$. There are different ways of forming the T^{gq}.

In [34] and [72], $P+1$ quadrature nodes with Gaussian weight are taken as the testing points. For the case with N statistically independent random inputs, the dimension of the testing point vector is $(P+1)^N$ and is obtained by the tensor product of individual nodes. This dimension is equal to the number of deterministic power flows required to estimate the coefficients vector $\boldsymbol{y}_\alpha$. The number of deterministic power flows required becomes very high in this method with an increasing number of stochastic input variables. In [72], collocation techniques are used to reduce the number of testing points. The authors of [73] propose to reduce the number of basis polynomials itself by exploiting the correlation between the stochastic input variables. These methods give a similar estimate of the mean value reducing the computation time but none of the studies considers the effect in tail-events.

In this paper, the three sampling techniques discussed in Section 4.2 are used to create the testing point T^{gq}, (a) pseudo-random generator; (b) Sobol sequences; and (c) Halton sequences[1]. With these sampling techniques, the number of deterministic power flows required is equal to the dimension of $\boldsymbol{y}_\alpha$, i.e., L. The coefficients in $\boldsymbol{y}_\alpha$ are calculated by solving L linear equations based on QoI vector $\boldsymbol{y}$ obtained from operation (4.1) and (4.2) and the numerical value of orthogonal basis polynomials at testing points T^{gq}.

The order β of the basis polynomials is one of the key factors affecting the accuracy of gPC. The number of basis polynomials increases with β as given in (4.5). This leads to an increase in the number of deterministic power flows and computation time.

The final step in uncertainty quantification using gPC is the recreation of the PDF of the QoI, which can be done by a) a kernel density estimator; or b) sampling of stochastic input $\boldsymbol{\Xi}$ and passing them through the polynomial

[1]There are many other deterministic sampling techniques available in the literature. A few other sampling techniques were tried and had no extra gain for the considered problem. Halton and Sobol's sequences are shown in this work as they are the most popular sequences among all available, and they resulted in satisfactory outcomes.

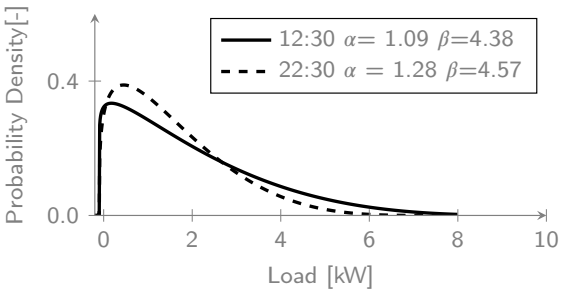
representing y, depending upon the nature of the final output required. In this paper, method b) is used to see the effect of the final sampling size on the accuracy as well as the computational time of the gPC. This process is similar to the MC method discussed in Section 4.2 but the difference is now that the samples are passed through polynomials rather than an iterative power flow solver. It means that the exponential increase in the number of power flow equations that need to be solved in MC and qMC methods with a large sample space can be avoided by using gPC. The input and output of the power flow equations in both MC and qMC methods are complex numbers. Consequently, an increase in MC samples will result in an increased number of iterative calculations involving complex numbers. This is avoided in gPC as the polynomials are formed by solving only a limited number of power flows, and the input array for final sampling and the final output is composed of real numbers.

The two main factors to consider when using gPC are the convergence of the QoI towards the MC solution obtained with infinite sample size and its accuracy. The convergence of the expected value of the output is discussed in detail in [179] when the input variables are Gaussian, and the QoI is in the mean region. However, parameters like the load of the consumers, PV generation, etc., are not necessarily Gaussian variables. Additionally, congestion related QoIs in LVDS require estimation of tail-regions of the output probability density functions. The accuracy of gPC at the tail region is expected to be better than the techniques where the power flow equations are linearized around the expected value. Moreover, the required QoI is expressed as a weighted series of the orthogonal polynomials preserving the non-linearity of power flow equations [76].

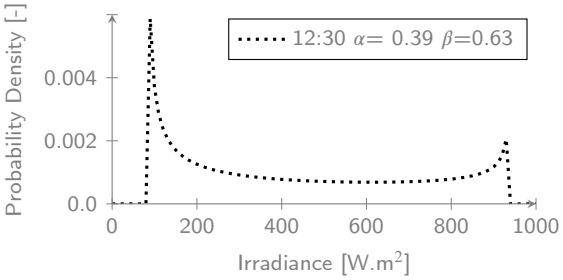
In this article, the testing point selection method, the degree of the polynomial and the final sampling size for QoI are varied to find out the best gPC configuration for an actual LVDS. This type of detailed comparision for LVDS is not yet available in the literature. Furthermore, the gPC based method is designed to calculate the probability of congestion in feeders with very high PV penetration. The output of gPC is then validated with the MC and qMC simulations results.

4.4 Numerical Illustration

In this section, an illustration is presented to show the differences between MC simulation, qMC simulation and gPC when accounting for uncertainty in unbalanced low-voltage power systems. Concretely, their ability to capture



(a) Beta distribution for load (day and night)



(b) Beta distribution for irradiance (day)

Figure 4.2: PDF for load and irradiance for typical spring day

the mean and tail values of feeder minimum voltage, maximum voltage, and maximum current per branch is investigated. The evaluation is done on real LV feeders, with detailed load and PV generation uncertainty definitions obtained from historical data. In summary this section illustrates the following:

a) Comparison of *sobol and halton sequences based gPC* and *MC based gPC* with *MC* in tail and mean regions for three different QoIs in terms of accuracy and computational time. The purpose is to come up with the best variant of gPC for LVDS.

b) Evaluation of the required *polynomial degree* in the best variant of gPC selected from a) to account for uncertainty in LVDS.

c) A comparison of *qMC* results with *gPC* and *MC*. These are shown using a study calculating *probability of congestion* using gPC, qMC and MC for an actual LVDS and illustration of the impact of number of uncertainty sources.

Table 4.1: LV consumer group of test feeder

Load category	1	2	3	4	5	6	7	Total
No. of consumers	14	2	1	2	1	1	1	22

The overall objective is to find the best variant of PPF to deal with the uncertainty quantification in LVDS related to the congestion.

4.4.1 Test system

The non-synthetic European LV feeders presented in [30] are used in this paper. The network has 160 LV feeders of a sub-urban area in the north of Spain. In this study, a historical consumer load data is used to create the probability density function (PDF) of their load. The load data is based on yearly profiles of 15-min energy consumption readings of Belgian consumers [107]. 25 categories of load profiles are defined, based on the contracted connection, annual kWh consumption, and consumer type, i.e., residential consumers or small and medium enterprises (SMEs). A typical spring day is chosen and 96 PDFs, each representing a 15-min load were formed by fitting the historical load profiles in similar days per consumer category. A beta distribution is used for the PDF fitting as it: a) represents the load distribution better than the Gaussian PDF, b) has dispersion, shape, and upper and lower bound parameters, c) fits to the measured data of LVDS consumers [106].

Similarly, the PDF of the irradiance during a typical spring day is created based on historical irradiance measurement data collected from EnergyVille premises in Genk, Belgium. The beta distribution is chosen to model the irradiance [180]. The reasons for choosing Beta distributions for load and irradiance are detailed in Section 2.5.2.

The detailed analysis of the different methods is shown for a single medium-sized LV feeder with 22 consumers, 40 buses, and a feeder length of 1190.18 m. The consumers of this feeder are characterized by one of seven consumer categories (Table 4.1). Consumers in the same category are assumed to have correlated load profiles represented by a common beta PDF. The maximum voltage in the feeder and the maximum current per branch rating is calculated for mid-day with solar-peak (12:30 hrs), and the minimum voltage is calculated for the night (22:30 hrs). Fig. 4.2a shows the beta distributions of one of the load categories for day (12:30) and night (22:30) and Fig. 4.2b shows the beta distribution of irradiance at 12:30 hrs. The irradiance uncertainty and seven uncertainties for seven load types make a total of eight stochastic input variables for the feeder for each quarter-hour. All the consumers are assumed to have 10 kWp PV installed

at their premises to represent a high PV penetration scenario with perfectly correlated PV generations which is directly proportional to the irradiance. In a parallel research [117], the assumption of correlated PV generation was found to be a better assumption than independent sampling of correlated variables. The spatial difference of the irradiance is neglected as the calculation here is done for the low voltage feeder level which is usually very short for geographical irradiance variation.

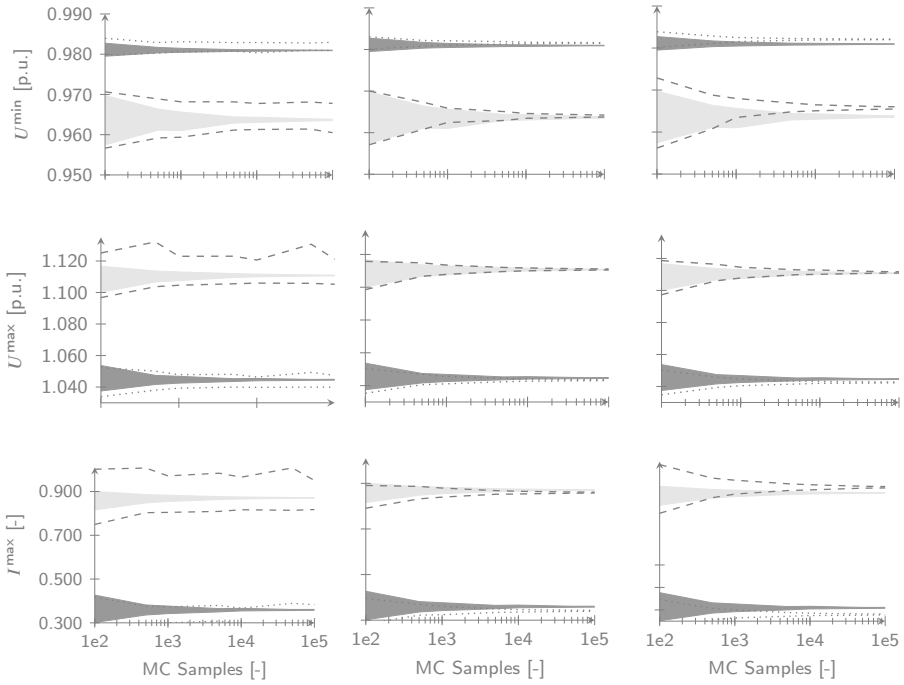
4.4.2 MC vs gPC: Testing points

In this subsection, the CIs of mean and tail values of: a) maximum feeder voltage at daytime; b) minimum feeder voltage at evening load; and c) maximum current per branch rating, are calculated using each of the following four approaches: (i) MC; (ii) gPC with MC based testing point; (iii) gPC with Sobol sequences based testing point; and (iv) gPC with Halton sequences based testing point.

The gPC approach uses the basis polynomials of degree 2 and the PDF of QoI is done by sampling. The polynomials used were of the Askey scheme [85], wherein this particular case Jacobi polynomials were used as all the uncertainties followed a Beta distribution. The polynomials were generated using the "Chaospy" package in Python [88]. The power flows were solved using an inhouse unbalanced power flow solver which uses Backward-Forward sweep [47] but for a three-wire equivalent model, considering the effect of the neutral voltage [32]. All the simulations in this paper were done on a PC with a 2.11 GHz Intel i7-8650U CPU processor with 16 GB memory. The confidence interval is calculated from 100 instances of MC with 100, 500, 1 000, 5 000, 10 000, 50 000 and 100 000 samples. The CI is calculated for both MC and gPC. Fig. 4.3 shows that the gPC based on the Sobol and Halton sequences performs well compared to MC based gPC for both mean and tail values. In all three cases, the computation time for gPC with 100 000 final samples is only 12% of the time taken by the MC simulations with sample size 100 000.

4.4.3 gPC: Impact of polynomial degree

In this subsection, the impact of basis polynomials' degree is evaluated in terms of computation time and accuracy. Fig. 4.4 shows the CI for the tail and mean QoIs obtained from degree one till degree four gPC. The gPC selected is based on Sobol sequences as it has a better performance as shown in previous illustration. In comparison to MC, the accuracy of mean QoIs do not significantly increase beyond degree 2. Higher degree polynomials require more coefficients to be fitted and more computation time. It is interesting to see that a higher degree



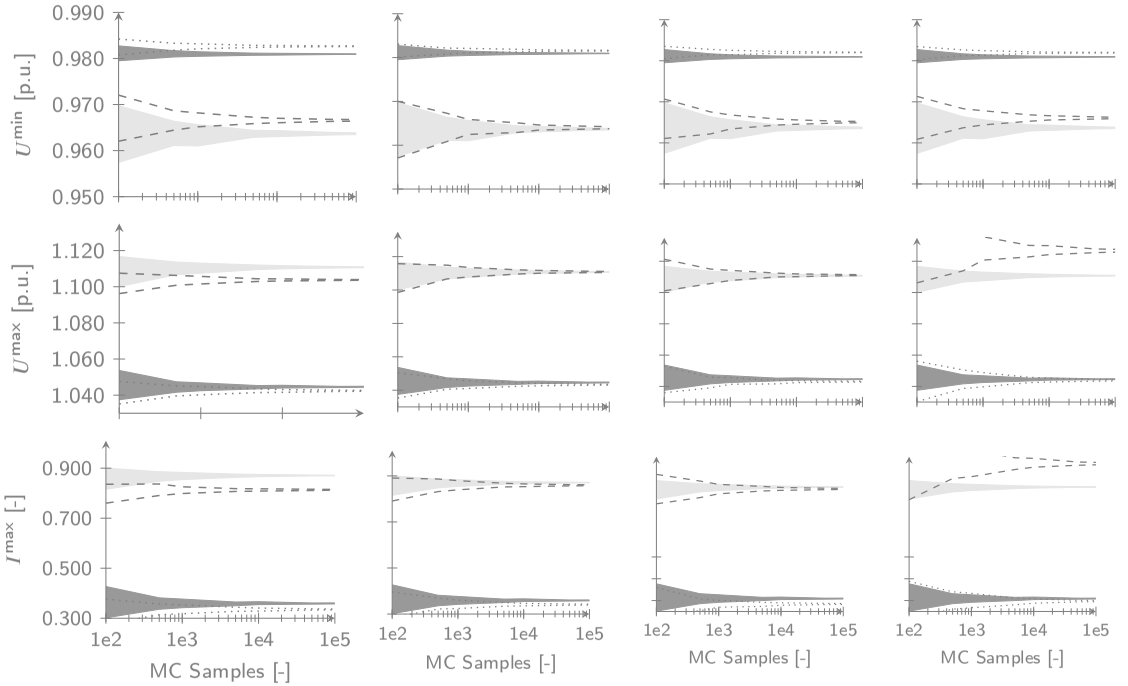
(a) MC / gPC degree-2 (b) sobol / gPC degree-2 (c) halton / gPC degree-2

Figure 4.3: CI of mean (dotted) and tail event (dashed) of QoIs obtained from degree 2 gPC based on (a) MC, (b) sobol and (c) halton samples, plotted along with CI from MC (dark gray: mean and light gray: tail QoI)

polynomial does not necessarily increases the accuracy in the tail-region due to an over-fitting problem which matches with the conclusion made in [181].

As expected, the computation time for gPC increases with increasing polynomial degree. Degree 3 gPC has increased computation time compared to degree 2 without much gain in accuracy in result (Table 4.2 and Fig. 4.4), which seems to match with the result for transmission network studies [89].

Hence, out of the four polynomial degrees compared, degree 2 gPC seems to be a better choice than MC while balancing accuracy and computation time for LVDS applications. The evaluation of model construction time shows that after degree two, increase in degree means exponential growth in the model construction time while the sampling time increases linearly. If the final output is just in terms of mean or variance of QoIs, the computation time of degree 2 gPC is only 0.84% of the MC with the same sample size.



(a) sobol / gPC degree = 1

(b) sobol / gPC degree = 2

(c) sobol / gPC degree = 3

(d) sobol / gPC degree = 4

Figure 4.4: CI of mean (dotted) and tail event (dashed) of QoIs obtained from degree 1-4 gPC based on sobol samples (a)-(d), plotted along with CI from MC (dark gray: mean and light gray: tail QoI)

4.4.4 Comparison of gPC, qMC, and MC

In this subsection, the result of qMC schemes with sample size 100, 500, 1 000, 5 000, 10 000, 50 000 and 100 000 are compared with the CI obtained from MC, and gPC, for all QoIs under consideration. The gPC used is of degree 2 and based on sobol sequences.

The qMC simulations uses Sobol and Halton sequences. The result shown in Fig. 4.5 shows that both Sobol and Halton sequences based qMC have the QoIs within the CI given by MC or gPC. The QoIs are stabilized after 5 000 samples. However, with this method, the results are deterministic and the statistical property of QoIs disappears.

The computation time for qMC with 5 000 samples is equal to MC for the same sample size. The qMC is a good alternative for uncertainty quantification as it results in estimates with a higher accuracy when using a lower sample size compared to MC. However, gPC has some extra advantages over MC and qMC. The gPC expansion represents the system with a surrogate model in the form of a polynomial. The output QoIs for a particular set of samples are obtained by passing the samples through the polynomial. This reduces the need for relying on computationally demanding power flow algorithms while increasing the sample size of the input. As a result, the computation time for MC and qMC simulations increases linearly with increased sample size, while in the case of degree 2 gPC the computation time increases in logarithmic scale. The gPC scheme is the best choice as it reduces the computational time without losing the statistical property and accuracy of the output.

Table 4.2: Computation time and error for gPC with 100 000 final samples compared to MC with the same sample size

gPC-degree			1	2	3	4	MC
Model construction time [sec]			0.20	0.23	1.135	5.31	-
total time [sec]			2.39	3.24	6.58	21.73	27.13
Relative comp. time			0.09	0.12	0.24	0.78	-
error	mean $U^{\min}$	ub	0.0017	0.0006	0.0011	0.0010	
		lb	0.0017	0.0006	0.0011	0.0010	
	tail $U^{\min}$	ub	0.0029	0.0004	0.0014	0.0024	
		lb	0.0029	0.0003	0.0014	0.0024	

ub=upper bound and lb=lower bound of CI

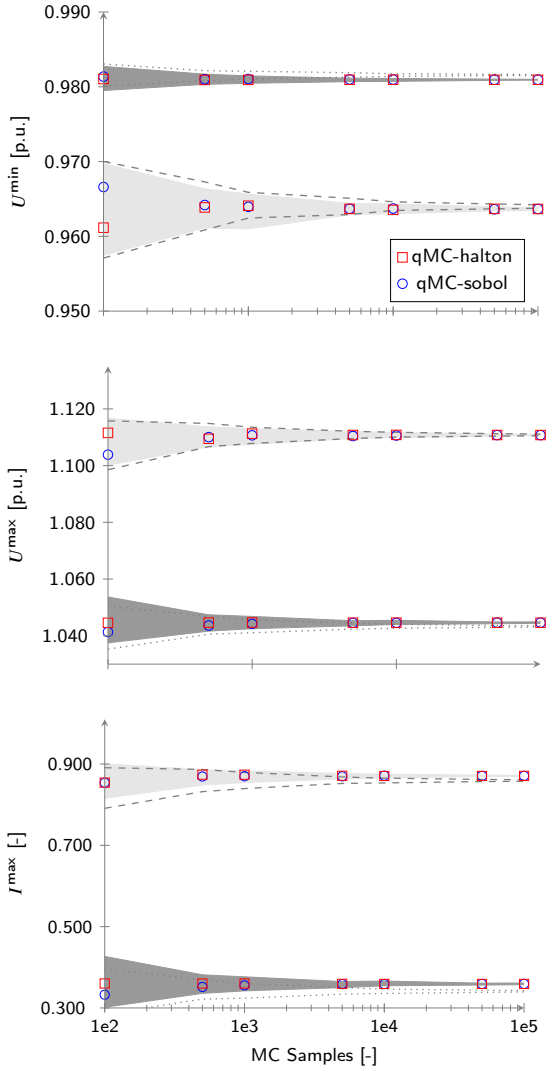


Figure 4.5: Mean and tail values of QoIs obtained from qMC using sobol (blue-circle) and halton (red-square) sequences plotted along with CI obtained from: a) MC (dark gray: mean and light gray: tail QoI) and b) order-2 gPC(dotted: mean and dashed: tail QoI) .

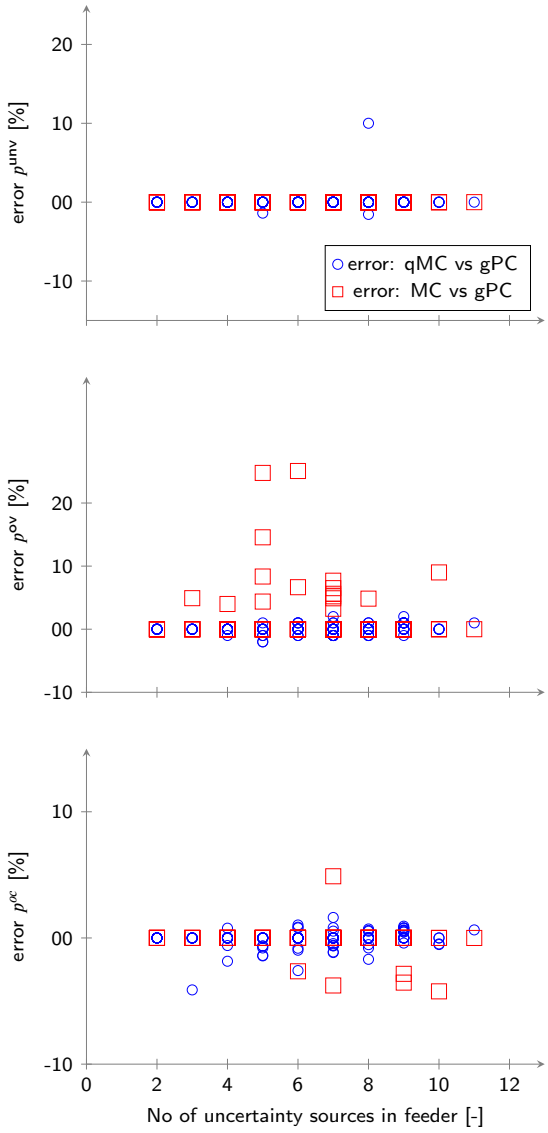


Figure 4.6: Error percentage on $p^{\text{unv}}, p^{\text{ov}}$ and p^{oc} in 160 real LVDS feeders plotted with respect to the number of uncertainty sources. Each blue circle represents error $(p^{\text{x}}_{\text{qMC}} - p^{\text{x}}_{\text{gPC}})/p^{\text{x}}_{\text{gPC}}$ of one feeder and similarly the red square is of $(p^{\text{x}}_{\text{MC}} - p^{\text{x}}_{\text{gPC}})/p^{\text{x}}_{\text{gPC}}$ per feeder.

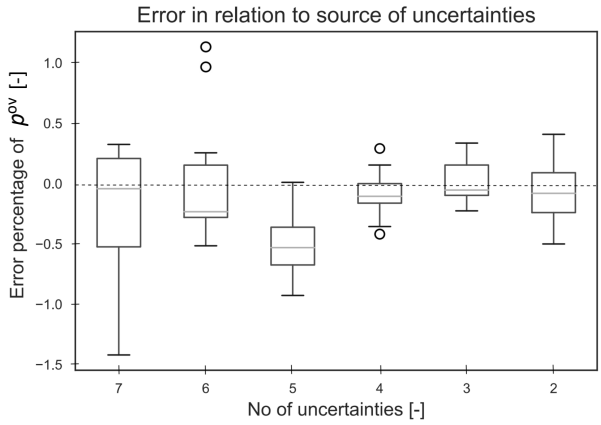


Figure 4.7: Error variation on p^{ov}: qMC vs gPC with number of uncertainties for selected feeders. The circles represent the errors which were between 1.5 interquartile range (IQR) to 3 IQR.

4.4.5 Network analysis: Impact of uncertainty sources

In this subsection, the three methods (MC, qMC and gPC) are used to calculate three probabilities related to congestion in a feeder,

a) probability of undervoltage (p^{unv}): the probability of minimum feeder voltage being less than 0.95 pu,

b) probability of overvoltage (p^{ov}): the probability of maximum feeder voltage being more than 1.05 pu, and

c) probability of overcurrent (p^{oc}): the probability of maximum current normalized by branch rating greater than 1,

for the full system with 160 feeders [30]. The previous assumption of a typical spring day and 10 kWp PV per consumer is carried for this case too. However, the analysis now is for a whole day with 96 quarterly time-stamps. The sample size for MC and qMC is chosen to be 1 000 for all feeders because a higher sample size leads to memory and computational tractability issues for solving the power flow equations using a 16 GB system. The gPC chosen is of degree 2 and based on Sobol sequences with 10 000 final samples. All three methods have the same computation time, even though gPC has ten times more samples.

The error is calculated for qMC vs gPC and MC vs gPC (Fig. 4.6) taking gPC as reference as it is based on more samples. Most of the feeders have a p^{unv} of nearly zero and hence its estimation error using MC and qMC is nearly equal to zero. p^{ov} calculated by MC is seen to be overestimated by 30% in a few feeders compared to gPC. On the other hand qMC has less than 1% estimation error of p^{ov}. The error on the p^{oc} by MC and qMC is less than 5% compared to gPC. From this analysis, it can be concluded that qMC is much better than MC in estimating the probability of congestion than MC and comparable to gPC. However, with gPC, in the same computation time more accurate results are obtained along with possibility to recreate the statistical properties.

Furthermore, to see the impact of the number of uncertainty sources, 10 feeders with 7 sources of uncertainty and $p^{\text{ov}} > 0.01$ are selected. The uncertainty sources related to load are reduced one by one till one source of uncertainty remains. The irradiance uncertainty is kept the same for all cases. The errors in p^{uv} calculated by qMC with 1000 samples are compared with gPC and plotted as a boxplot (Fig.4.7). It is observed that feeders with fewer uncertainty sources tend to have a lower estimation error compared to using qMC with comparable computation time. Increasing the uncertainty sources subsequently increases the estimation error leading to the requirement of a higher sample size. As uncertainties increase, the advantage of using gPC over MC or qMC increases.

4.5 Conclusion

This paper evaluates three probabilistic power flow tools for an LVDS with demand and PV generation uncertainties represented by continuous univariate distributions. From the analysis shown in this paper, it can be concluded that the gPC outweighs MC and qMC clearly in terms of computational effort and accuracy in an LVDS case-study. These advantages are more clearly seen in feeders with more sources of uncertainty. The choice of the testing point vector and the polynomial degree is essential to the accuracy and computational effort, with degree 2 gPC based on Sobol sequences being the best choice. The qMC improves the estimate of QoIs compared to MC but at the cost of losing the statistical properties. Finally, the QoIs considered in this study were linked with congestion issues in LVDS. The calculation of the probability of undervoltage, overvoltage and overcurrent in 160 feeders showed that the gPC can be expanded for other congestion-related studies and probabilistic optimization.

In future research, these congestion probabilities can be applied to LVDS applications such as hosting capacity calculations and congestion-based day-ahead flexibility needs assessment. This study shows that for uncertainties

represented by continuous PDF, gPC outweighs the MC methods. However, if a mixture of continuous and discrete uncertain variables is present, possible ways to decouple these variables must be explored. In Chapter 5 one such method is presented where MC is used for discrete *planning* variables and gPC is used for continuous *operation* variables.

Chapter 5

Decoupled Stochastic Hosting Capacity

Koirala, A., Hashmi, Md. U., D'hulst, R., Van Hertem, D., *Decoupled Probabilistic Feeder Hosting Capacity Calculations Using General Polynomial Chaos.* Electric Power Systems Research, 211, Art.No. 108535. doi: 10.1016/j.epsr.2022.108535

The first author is the main author of the article. The contributions of the first author include the literature review, conceptualization, methodology, data curation, and the writing of the article. The article was also presented in XXII Power systems Computation Conference 2022, Porto, Portugal.

Abstract: This paper introduces a decoupled method of calculating the PV hosting capacity of low voltage distribution system (LVDS) feeders for new solar photovoltaics (PV) installations. The hosting capacity calculation of LVDS is a multidimensional stochastic problem. A general polynomial chaos-based probabilistic power flow is used to solve this problem, as it allows for fast computation times without any compromise in accuracy. Two types of uncertainties exist in the hosting capacity calculation problem: planning level uncertainties such as size, location, type, and number of PV installations and operational uncertainties such as consumer load and PV generation. These two types of uncertainties are usually sampled together in probabilistic hosting capacity approaches. In this paper, a decoupled approach is presented where the

impact of planning level scenarios on the probability of violation of *operational limits* is studied. The highest total PV among the planning scenarios inside the probabilistic bound of the *operational limit* is termed as the hosting capacity of the feeder. The results show that hosting capacity depends upon the planning uncertainties and operational variables, and decoupling them gives a more concise picture of the impact of different uncertainties.

Keywords: Low voltage, distribution feeder, photovoltaics, hosting capacity, stochastic

5.1 Introduction

5.1.1 Motivation and Literature Review

Calculating the hosting capacity (HC) for photovoltaics (PV) installations for an Low Voltage Distribution System (LVDS) is a multidimensional stochastic problem involving different uncertainty sources. It is an important planning problem for distribution system operators and regulators to understand the capability of LVDS to host new loads and generation units. The most accepted definition of HC is, "the amount of new generation or consumption that can be accommodated on a given feeder without impacting system operation under existing control and infrastructure configuration" [17]. HC calculations that assume a deterministic worst-case scenario result in a deterministic HC. In contrast, consideration of uncertainties results in what is referred to as probabilistic HC [18].

In the last 20 years, there have been many studies for calculating the PV hosting capacity of medium and low voltage distribution systems. The review papers [18, 132,139,140] point out the difference between methods for PV hosting calculation in terms of tools used for calculating *grid limitation* violations, consideration of uncertainty, or the criteria to decide hosting capacity. Finding a suitable method for stakeholders is always a challenge as different stakeholders are interested in different dimensions of the problem. Nevertheless, there is growing consensus that PV hosting capacity calculated using deterministic methods represents one snapshot of many possible situations. Hence, a more comprehensive hosting capacity calculation method must consider uncertainties [182].

In the literature, the probabilistic HC is most often calculated using Monte Carlo (MC) based methods [17,21–23,26,27,98,105,118,122–125,163], where the selected uncertain variables are taken to make thousands of scenarios, and power flow equations are solved for each scenario. This repetitive solving of complex power flow equations makes these MC based methods computationally heavy.

These methods were implemented and benchmarked in [182]. It was observed that in these methods, some of the uncertain variables were simplified or even neglected to reduce the dimension of the problem. For example, the seasonal variation of load and generation, the consumer characteristics, or the feeder characteristics were neglected. Some other ways of reducing the computational burden used in literature are: use of peak PV only [122], use of selective time-stamps for simulation [21], use of representative feeders [21], or simplification of the distribution system as a single phase balanced system [26]. In [25], it is pointed out that the probabilistic HC methods in literature usually do not differentiate between the uncertainties, but samples all of them simultaneously. The necessity of differentiating so-called *aleatory uncertainties* (the load and PV generation uncertainties) and *epistemic uncertainties* (feeder and location uncertainties for future PV installation) is then highlighted. However, the calculations are done on peak PV days and time only, assuming the load is minimum at those instants. Furthermore, solving computationally expensive but accurate power flow equations is replaced with a transfer impedance matrix based method to calculate voltage rise due to installed PV.

Few other methods use optimization-based methods to reduce the high computational time required for the MC-based methods. An optimization based method for HC calculations of medium voltage feeders for a fixed location of PV installations is proposed in [141, 142]. In [164], HC calculation using a robust optimization approach for a radial medium voltage (MV) network is proposed, considering online tap changers and static VAR compensators. These methods usually work for MV networks with a simplified linear power flow and assume a single phase equivalent of the network. The resulting HC is for a specific location of a single large size PV. These methods consider uncertainties of PV size in input, but the stochastic limit in the output excluded.

From this analysis, it is evident that there is a necessity for a probabilistic HC method which can: a) differentiate between the input uncertainties and possibility to quantify the effect of those uncertainties, b) take care of *stochastic limit* in *grid limitation* violations, c) computationally tractable to compute the HC of possibly thousands of feeders within LVDS, d) identify the impact of daily time-series load as well as seasonal load. To this end, the decoupling of planning level uncertainties such as PV size, PV number, and PV type from operational uncertainties created due to consumer load and PV generations provides the stakeholder a better way to evaluate the impact of different uncertainties in LVDS PV HC. The usage of the probabilistic power flow tool using general Polynomial Chaos (gPC) expansion offers a tractable solution for such PV HC calculation.

gPC expansion is a method to represent stochastic variables in terms of their orthogonal polynomial basis, similar to how Fourier series represent a periodic

signal [16, 28]. This tool has recently been used for probabilistic analysis of power systems. In [72], the authors focus on using gPC based on a stochastic collocation method for probabilistic analysis of LVDS to reduce the computation time. An arbitrary polynomial chaos expansion is proposed in [73] for stochastic input with a Probability Density Function (PDF) constructed from measurement data for the transmission network. The article [74] proposes a non-intrusive low-rank approximation method for probabilistic analysis of power system with renewable generators to reduce the total number of unknowns to be linear rather than exponential with the increase in random input. In [16], the chance-constrained optimization problem for a transmission network is solved by using gPC. In [34], the computational strength of gPC over crude MC in LVDS is shown. In [183], the gPC of various degrees and types are evaluated in a real LVDS with load and PV generation represented by Beta distribution to come up with the best variant of gPC for LVDS. Furthermore, the gPC-based method is compared with crude-MC and quasi-MC schemes to calculate congestion related parameters. Where most of these studies aim towards establishing gPC based probabilistic power flow as a computationally and accuracy-wise superior method for probabilistic analysis of power systems, this paper introduces an application for a gPC based tool for PV HC calculation.

5.1.2 Contribution and organization

The key contributions of this paper are highlighted in this section:

a) a decoupled stochastic hosting capacity calculation method is introduced to separate the impact of planning level scenarios from operational ones. The planning level uncertainties such as PV size, PV number and PV location are used to form planning level scenarios. The impact of operational uncertainties, i.e., load and PV generation, on each such scenario is computed.

b) a gPC based probabilistic power flow tool is used to determine the PV hosting capacity of LVDS feeders.

c) a demonstration of HC calculation is done for different planning scenarios using various performance indices, i.e., probability of overvoltage, overcurrent and undervoltage. The probability was calculated for each 15-minutes of a day, and maximum, 95% percentile, 75% percentile and mean value were observed as performance indices in three different types of feeder.

The remaining sections are divided as follows: Section 5.2 presents a method to decouple planning level scenarios from the operational scenarios. Section 5.3 introduces the gPC based probabilistic PV HC tool. Section 5.4 demonstrates the numerical illustration of decoupled HC method for an actual LVDS and its spread along temporal and spatial horizon. Finally, Section 5.5 concludes the work in this paper.

5.2 Decoupling of uncertainties in LVDS

The multidimensional nature of the input uncertainties in LVDS with high PV can be broadly categorized into three types:

a) feeder uncertainty: number of consumers in a feeder, number of PV installations in a feeder,

b) location uncertainty: the phase the load and PV are connected to, the location of PV along the feeder,

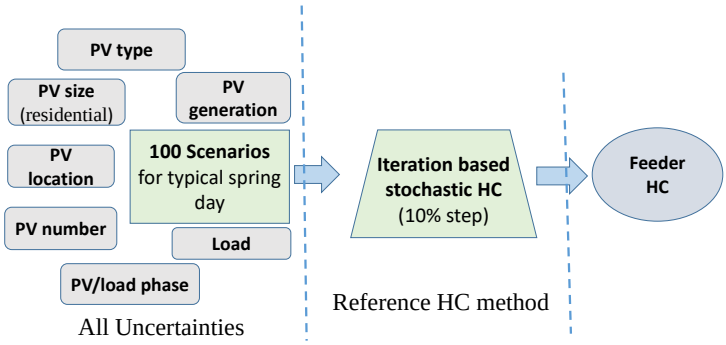
c) power injection uncertainty: power consumed by connected loads and generated by PV installations.

The first two uncertainties arise from the undefined features of present and future PV installations, such as size and location in the network, while the third one is related to the weather uncertainty or consumer behavior. In [25], it is pointed out that the feeder and location uncertainties for future PV installations are epistemic uncertainties, and the load and PV generation uncertainties that determine the *operational limits* of a particular scenario are the aleatory uncertainties.

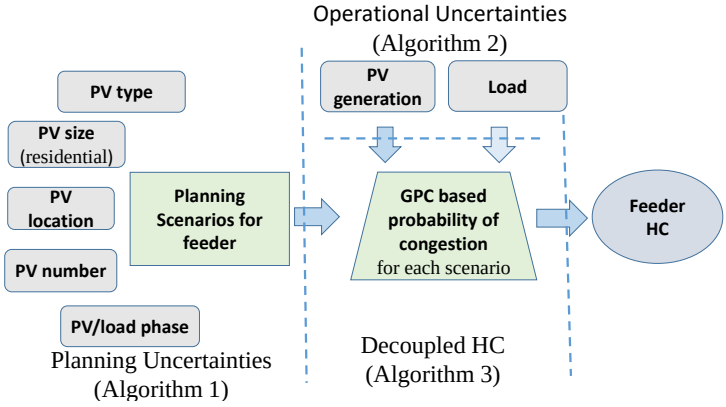
The existing deterministic and stochastic PV HC methods for LVDS in literature were benchmarked to a reference method proposed in [182]. This reference method (schematics in Fig. 5.1a), nevertheless, has the following limitations:

a) There is no separation between the PV size, number and location uncertainties, which are the features of LVDS planning, and the power injection uncertainties arising from the consumer load and power generation. The *probability of congestion* obtained includes the effect of planning level and the power injection uncertainties at the same time, and their impact cannot be separated.

b) Consideration of peak PV generation day presents the scenarios in problematic days omitting the overall annual performance.



(a) Reference method for PV HC of LVDS



(b) Decoupled gPC based HC for LVDS

Figure 5.1: Decoupled vs reference HC method

c) PV size of the consumer is sampled from a probability mass function (PMF).

d) The analysis is done by using a limited number of MC samples, as increasing the number of samples increases the computational time and complexity.

The modelling structure of the decoupled PV HC method is shown in Fig. 5.1b, where the scenarios from location and feeder uncertainties are separated from the power injection uncertainties.

The issues observed for the reference method are addressed as follows:

a) The scenario generation algorithm formulates the possible planning level scenarios of PV location, size, type, and number.

b) For each scenario obtained from (a), the probability of *operational limits* violation due to PV injection and consumer demand is calculated using gPC-based probabilistic power flow.

c) The probability of *grid limit* violations are calculated for an average summer/spring day of a year.

d) The connected PV kilowatt-peak (kWp) is increased uniformly to all consumers.

e) The PV penetration rate is increased in predefined steps while forming the scenario.

The scenario generation part for HC calculation is now split into two parts. The first one involves planning level scenarios with size, location, number, type and phase of PV installations (Algorithm 1). Then, the second part involves generating operational scenarios due to load profile and PV generation for representative days of a year (Algorithm 2).

The next step in the decoupled HC calculation approach is to obtain probabilities of congestion due to operational uncertainties for each of the scenarios generated in Algorithm 1. In Section 5.3, gPC based decoupled PV HC approach is discussed.

Algorithm 1 `Planning level uncertainties`

Inputs: Feeder details

1: Define set of possible size of PV installations,
2: Define type of PV installation, e.g., residential, commercial, single phase, or three-phase,
3: Define possible location of PV, e.g., rooftop of each consumer, PV installed in same phase as consumer,
4: Define the steps for PV penetration ratio increment,
5: Make a set of possible PV scenarios for a given feeder.

Outputs: Set of PV scenarios ($s \in \mathcal{S}$)

Algorithm 2 `Operational uncertainties`

Inputs: representative day, historical load and irradiance profile

1: Collect all the historical load profiles for large set of consumers for the days similar to representative day based on weather and weekends/weekdays,
2: Separate the load profiles to the different consumer groups based on annual kWh consumption and contracted kW,
3: Fit beta distribution for load consumption each time-period (can be 15-minutes or hourly load based on data availability) of the day for each consumer groups,
4: Collect all the historical irradiance profile for days similar to representative day based on weather,
5: Fit beta distribution for each time-period of the day as in step 3.

Outputs: Probability density function (PDF) of all consumer groups and irradiance per time-period of the day

5.3 gPC based decoupled PV HC

The gPC expansion maps the stochastic input vector $\boldsymbol{\Xi} = [\xi_1, \xi_2, ..., \xi_N]$ to the *Quantities of interest* (QoIs). This is achieved by fitting equivalent polynomials for *QoIs* on orthogonal polynomials derived from the stochastic input. The stochastic *QoI* y is expressed as the sum of polynomial bases of stochastic inputs weighted by respective coefficients, as in (5.1).

$$y = \sum_{j=1}^{\infty} c_j \Phi_j \tag{5.1}$$

The series in (5.1) is truncated to length L to reduce computational burden. The cardinality of the reduced basis vector $\boldsymbol{\Phi}_\alpha = [\Phi_1, \Phi_2, ..., \Phi_L]$ and coefficient vector $\boldsymbol{y}_\alpha = [c_1, c_2, ..., c_L]$ depends upon the degree β of the polynomial bases and number of stochastic input variables N:

$$L = \frac{(\beta + N)!}{\beta! N!} \tag{5.2}$$

Let $\boldsymbol{\Phi}_\alpha$ represent the vector of the multi-variate basis polynomials that are formed by the tensor product of univariate polynomials of stochastic input variables in $\boldsymbol{\Xi}$,

$$\boldsymbol{\Phi}_\alpha = \left\{ \boldsymbol{\phi}(\xi_1) \otimes \boldsymbol{\phi}(\xi_2) \otimes ... \otimes \boldsymbol{\phi}(\xi_N) \right\}_\beta, \tag{5.3}$$

where $\boldsymbol{\phi}(.)$ is the vector of univariate orthogonal polynomials for the stochastic input parameters in $\boldsymbol{\Xi}$. These orthogonal polynomials are well-defined for most of the continuous univariate distributions, e.g., Hermite polynomials for Gaussian-distributed variables, and Jacobi polynomials for Beta distributions [97]. After

the tensor product in (5.3), the basis polynomials are truncated to the order of β [97]. The basis polynomials in $\mathbf{\Phi}_\alpha$ are mutually orthogonal with respect to the PDF ($f(\xi)$) fulfilling the following property with expectation operator $\mathbb{E}$:

$$\mathbb{E}[\Phi_y, \Phi_z] = 0 \quad \forall \quad y \neq z; \quad y, z \in 1, 2, ..., L. \tag{5.4}$$

The gPC expansion is done in an intrusive way by using Galerkin's projection in power flow equations or in a non-intrusive way using regression fitting [88]. The gPC expansion used in this paper is a non-intrusive one where the power flow equations are not "intruded". The large number of uncertainties and nodes in LVDS overloads the intrusive gPC using Galerkin projection due to the large number of integrals to be computed [29]. The philosophy behind the non-intrusive method is that *QoI* y is calculated for certain selected samples, also known as testing points T^{gq}, of the stochastic input vector $\mathbf{\Xi}$. There are different ways of forming the T^{gq}. For example, [34] and [72] use quadrature nodes with Gaussian weights, [72] uses collocation techniques to reduce the number of testing points, [170] uses correlation among the DERs to reduce the testing points, [73] proposes to reduce the number of basis polynomial itself based on the correlation between the stochastic input variables. In [183], the sampling technique based on crude-MC, Sobol sequence [57] and Halton sequence [58] were evaluated for different degrees of polynomial. It was observed that degree 2 polynomials with T^{gq} based on Sobol and Halton sequence outperforms the gPC based on crude MC. Hence, a Sobol-sequence based T^{gq} with degree 2 (β=2) polynomial is used in this paper. In this method, the number of deterministic power flow calculations required is equal to the dimension of the $\boldsymbol{y}_\alpha$, i.e., L. The $\boldsymbol{y}_\alpha$ is calculated by solving L linear equations based on *QoI* vector $\boldsymbol{y}$ obtained from (5.5) and (5.6) and the numerical value of orthogonal polynomial bases at testing points T^{gq} .

$$0 \ = g(\mathbf{U}, \bar{\mathsf{I}}, \mathsf{T}_i^{\mathrm{gq}}) \qquad \forall \quad \mathsf{T}_i^{\mathrm{gq}} \in \mathsf{T}^{\mathrm{gq}} \tag{5.5}$$

$$y_i = h(\mathbf{U}, \bar{\mathsf{I}}) \qquad \forall \quad \mathsf{T}_i^{\mathrm{gq}} \in \mathsf{T}^{\mathrm{gq}} \tag{5.6}$$

$$y_i - \boldsymbol{y}_\alpha \{\mathbf{\Phi}_\alpha\}_i = 0 \quad \forall \quad \mathsf{T}_i^{\mathrm{gq}} \in \mathsf{T}^{\mathrm{gq}} \tag{5.7}$$

The expression (5.5) gives the vector of nodal voltages $\mathbf{U}$ and branch currents $\bar{\mathsf{I}}$ obtained from power flow g for each of the testing points. The expression (5.6) has the operator h, which generates the output value of *QoI* y_i at each of those testing points. The expression (5.7) represents the linear equations with $\boldsymbol{y}_\alpha$ as unknown for all testing points. The $\{\mathbf{\Phi}_\alpha\}_i$ is the numerical value of orthogonal basis at the testing point $\mathsf{T}_i^{\mathrm{gq}}$.

The final step in uncertainty quantification using gPC is the estimation of the QoI, which is done by a) a kernel density estimator, or b) sampling of stochastic

Algorithm 3 `gPC-based decoupled HC method`

Inputs: Feeder details, output of Algorithms 1 and 2

1: Get consumer details and assign them one of the pre-defined group (groups from step 2, Algorithm 2),
2: Assign respective PDF of load for each consumer,
3: Collect uncertainties for the feeder due to load, it is equal to number of consumer groups in the feeder,
4: Collect the irradiance uncertainties from the Algorithm 2 to form a full $\boldsymbol{\Xi}$ for all time-periods,
5: Form $\boldsymbol{\Phi}_\alpha$ for each time-period based on $\mathbf{X}$,
6: Form T^{gq} using Sobol sequence for each time-period,
7: Set *grid limits* for congestion, e.g., under voltage limit as 0.95 pu,
8: **for** $s \in \mathcal{S}$ **do**
9: **for** $\mathsf{T}_i^{\mathrm{gq}} \in \mathsf{T}^{\mathrm{gq}}$ **do**
10: Assign the $\mathsf{T}_i^{\mathrm{gq}}$ loads and PV installation,
11: For PV in s, convert irradiance to the PV generation,
12: Run power flow equation (5.5),
13: Obtain QoI y_i based on power flow output,
14: Form the linear equation (5.7),
15: **end for**
16: Compute coefficients $\boldsymbol{y}_\alpha$ from linear equations on step 14
17: Form polynomial for a time-period using $\boldsymbol{y}_\alpha$ and $\boldsymbol{\Phi}_\alpha$
18: Sample N^{mc} values from uncertainties vector $\mathbf{X}$,
19: Pass the samples through the polynomial from step 16,
20: Compute congestion probabilities per time-period for s,
21: Repeat step 9-20 for all time-period
22: **end for**
23: Set percentage of times when *grid limits* can be violated
24: Set a *stochastic limit* for congestion probabilities
25: Sort out the s with maximum PV below the *stochastic grid limit*
26: Total PV installed in scenario obtained in 25 is feeder PV HC

Outputs: PV HC of the feeder

input $\boldsymbol{\Xi}$ and passing them through polynomial representing y, depending upon the nature of the final output required. In this paper, sampling with final sample size of 1000 is used as a compromise between accuracy and computational time [183]. The samples were used to compute three probabilities of *grid limits* violations, namely,

a) probability of under voltage (p^{unv}): the probability of minimum feeder voltage being less than 0.95 pu,

b) probability of overvoltage (p^{ov}): the probability of maximum feeder voltage being more than 1.05 pu, and

c) probability of overcurrent (p^{oi}): the probability of maximum current normalized by branch rating greater than 1.

Note that these limits are more strict than the power quality standards used in distribution systems (EN50160, [13]), which require the voltage to be within ±10% of the nominal voltage for 95% of the measurements. Tighter limits of the grid voltage could be seen as a means to compensate for the inadequacies of the current smart metering infrastructure to measure such power quality limits.

These probabilities of *grid limit* violations are jointly referred to as congestion probabilities and are computed for each scenario obtained from Algorithm 1 based on the load and PV uncertainties from Algorithm 2. The probabilities are recorded for each time-period of the day. The final probability for a day could be:

a) the median: meaning the probability of congestion is lower than this value for 50% of the time,

b) 75% quartile; meaning the probability of congestion is lower than this value for 75% of the time,

c) 95% quartile; meaning the probability of congestion is lower than this value for 95% of the time, or

d) maximum value of the probability over the day.

Once these congestion probabilities are calculated, the cases having probabilities higher than the *stochastic grid limit* are filtered out. The final step in decoupled PV HC calculation is to scan through all the scenarios inside the assigned *stochastic grid limits* and find out the scenario with the highest total connected PV. This total connected PV is referred to as the PV HC of the feeder under consideration. The whole process of this gPC-based decoupled PV HC tool is presented in compact form in Algorithm 3.

5.4 Numerical results

The numerical illustration is done using the actual European feeders presented in [30]. The historical irradiance measurement data from the EnergyVille site in Belgium, along with load data from Belgian consumers, are taken to fit the PDF of each consumer group [107]. The solar generation of different panels is assumed to be correlated. The consumers are categorized into 25 groups, and consumers of the same type are considered co-related. Each consumer group's irradiance

Table 5.1: Features of selected feeders

Feature/Feeder ID	Feeder 1 144	Feeder 2 52	Feeder 3 27
No. of consumer	44	20	2
Avg active energy/year [kWh]	2068.79	4256.73	5424.98
Avg reactive energy/year [kVArh]	202.84	415.14	484.25
Total line length [km]	0.38	0.15	0.31
Main path [km]	0.21	0.13	0.29
Avg length to costumer [km]	0.15	0.13	0.16
Total line impedance [ohm]	4.60	2.52	0.30
Average path impedance [km]	0.10	0.13	0.15

profile and load (kWh) every 15 minutes is represented by a beta distribution obtained by fitting the historical profiles of similar days. This assumption of clustering the historical load was required because the LV consumer data is not easily available for public use. The reason for selecting beta distributions for consumer load and irradiance is discussed in detail in Chapter 3 of PhD thesis by Chihota [103] for the interested readers. The solar generations of all consumers were assumed to be perfectly correlated. Sobol-sequence based non-intrusive gPC of degree two ($\beta = 2$) is used for the gPC-based HC calculations.

In the following section, for detailed numerical illustration, three feeders were chosen as shown in Table 5.1. The selection was done to represent typical types of European LV feeders:

- Feeder 1 is a typical residential feeder with high load power factor, large number of consumers, and high cumulative line impedance with small average path impedance.
- Feeder 2 is a mix of residential and small and medium enterprise (SME) in a single apartment-complex.
- Feeder 3 is a typical commercial/industrial/SME consumer feeder with high consumption, small number of consumers, and low cumulative impedance.

5.4.1 Decoupled gPC-based HC

For the decoupled PV HC calculation in this paper, PV scenarios are considered only for residential roof-top PV with the maximum number of PV installations

equal to the number of consumers in the feeder. For each 5% rise in the PV penetration step, the consumers having PV will have ten possible sizes and types as given in Table 5.2, represented as R1, R2, ..., R10. PV location is assumed to be spread evenly along the feeder length in this work. PV phase is assumed to be the same phase as that of the consumer. The boundary of HC due to location uncertainty can be obtained by placing all PVs towards the end of feeder (lower bound) or near the sub-station (upper bound). The impact of uncertainty location of PV in a feeder is not considered in this paper but will be studied separately in a future publication.

The 15 minutes interval probability of congestion for an average summer/spring day (average summer PV production day) to each planning scenario is computed using the decoupled gPC based tool. Recording probability in each timestamp gives a higher resolution of the time series impact and provides insight into the effects of performance indices in HC. Instead of analyzing probabilities in each time instance, the median, 75% percentile, 95% percentile, and a maximum of the probabilities are taken under consideration for PV HC calculations. In the following section, the maximum congestion probability is taken as a reference for computing decoupled PV HC for whole network and hereafter referred by "probability of congestion".

5.4.1.1 Hosting Capacity from the probability of congestion

In the proposed decoupled HC calculation method, the planning level scenarios have 20 different PV penetration ratios and the PV size is varied from 1 to 15 kWp as shown in Table 5.2. Fig. 5.2 shows the probability of overvoltage for all combinations of single phase PV size and penetration. The hosting capacity for

Table 5.2: Residential PV size scenario for each penetration case

Residential PV scenario	PV size [kWp]	PV type
R1	1	1 phase
R2	2	1 phase
R3	3	1 phase
R4	4	1 phase
R5	5	1 phase
R6	6	1 phase
R7	6	3 phase
R8	9	3 phase
R9	12	3 phase
R10	15	3 phase

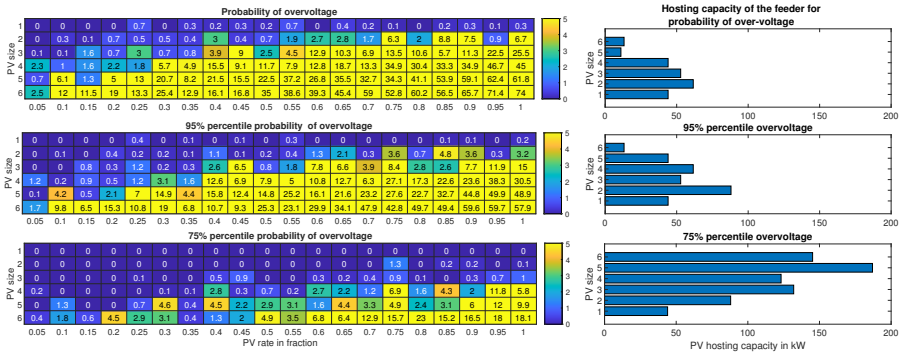


Figure 5.2: HC calculation using decoupled gPC based tool for feeder 1 based on overvoltage: a) maximum probability of overvoltage, b) 95% percentile of probability of overvoltage and c) 75% percentile of probability of overvoltage.

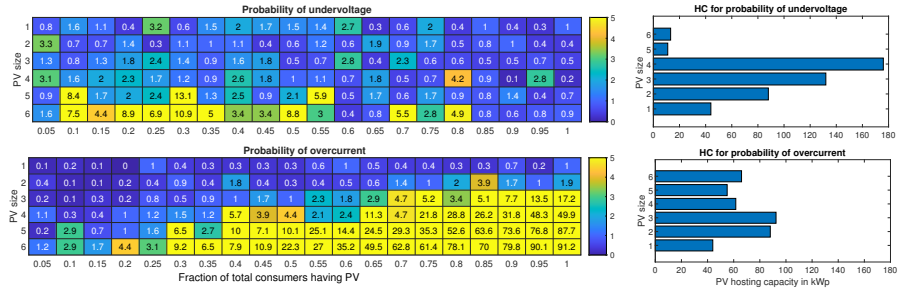


Figure 5.3: HC calculation using decoupled gPC based tool for feeder 1 using probability of undervoltage and overcurrent.

each size of single phase PV was decided based on the maximum penetration percentage with the probability of congestion less than 5%. The HC for each PV size is then plotted in the histogram in the right (Fig. 5.2). For feeder 1, taking 75% percentile of the probability of overvoltage over a day results in PV HC of 190 kWp, which gets reduced to 60 kWp if we consider maximum probability for the same day. This shows that if maximum voltage in a feeder is allowed to be higher than the *grid limit* for 25% of the time, PV HC increases by a factor of 3. This gain in HC capacity can be obtained by intervention in the LVDS with techniques such as PV curtailment. However, the role of intervention in HC is not in the scope of this paper. So, the PV HC for LVDS is calculated on the basis of maximum probability of congestion over the time.

In Fig 5.3, the HC based on other congestion types for the same feeder is shown. The HC calculated based on the probability of overvoltage is the lowest for this feeder. The same conclusion holds for Feeder 2 and 3. It shows that the probability of overvoltage is dominant on calculating PV HC, which is in agreement with comments made in [184] and [185].

5.4.1.2 Single phase vs three phase PV

In Fig. 5.4, the impact of probability of congestions is illustrated for three phase and single phase PV of 6 kWp. It was observed that having larger size single phase PV not only causes an overvoltage in the phase where there are more PV, it also contributes to undervoltage in the remaining phases. This phenomena of undervoltage being induced due to high capacity PV in other phase is neglected if we use the simplified one-phase equivalent network or assume Kron's reduction in a system which is sparsely grounded or grounded only in the substation. The probability of undervoltage diminishes once the penetration level is above 60% with single phase PV. This finding complements the study of voltage sensitivity in remaining phases due to single phase PV in [186]. Nevertheless, the single phase PV HC is constrained by the probability of overvoltage at 5% penetration ratio. On the other hand, for three phase 6 kWp, HC is limited by overcurrent at a penetration ratio of 45%.

In Table 5.3, the single phase and three phase PV HC of the selected feeders are listed. Note, that use of three phase PV devices increases the PV HC in all three feeders. The HC for three phase PV is on average two times more than that of single phase PV.

Table 5.3: PV hosting capacity: single phase and three phase

Feeder	HC: Single-phase PV			HC: three-phases PV		
	HC	PV size	PV rate	HC	PV size	PV rate
1	83.6	2	0.95	186	6	0.7
2	114	6	0.95	300	15	1
3	12	6	1	30	15	1

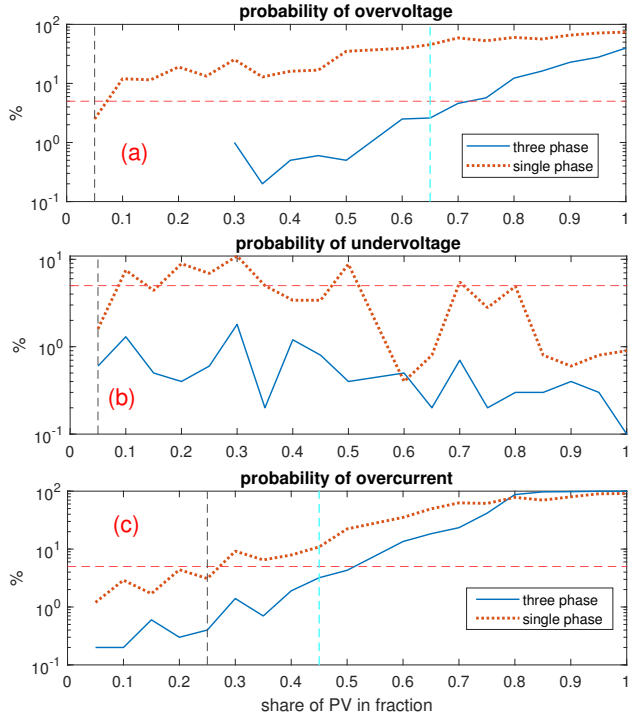


Figure 5.4: Comparison of 6 kWp single phase PV vs 6 kWp three phase PV in the feeder 1. The dashed horizontal line denotes 5% limit; The parameters are plotted on a log scale to capture the low magnitude variability.

5.4.2 HC calculation for whole network

In this section, the PV HC is calculated for the whole network using the decoupled gPC-based PV HC algorithm. In Fig. 5.5, PV HC calculated with a decoupled gPC based tool for the selected three feeders is mapped along with the PV HC of all feeders in the network. The relation between features of feeders, namely: a) average annual consumer load in kWh, b) the number of consumers, c) average feeder resistance, and d) total conductor length, and PV HC was illustrated. The variability of the PV HC with network feature shows that representative or synthetic feeders can be used to estimate the ballpark number for hosting capacity of any given network. However, the detailed HC of each feeder requires a complete run of decoupled HC tool. The key takeaways from Fig. 5.5 are: a) number of consumers and average annual consumption are not strongly linked to HC, b) fewer consumers with more annual consumption means

higher HC (compare representative feeder 1 (red square) and representative feeder 2 (green star)), c) feeders with a lower average line resistance tend to have higher HC, and d) feeders with a lower total length tend to have the highest PV HC. The feeder with SME (black diamond) has relatively low HC because of the limited number of consumers in that feeder.

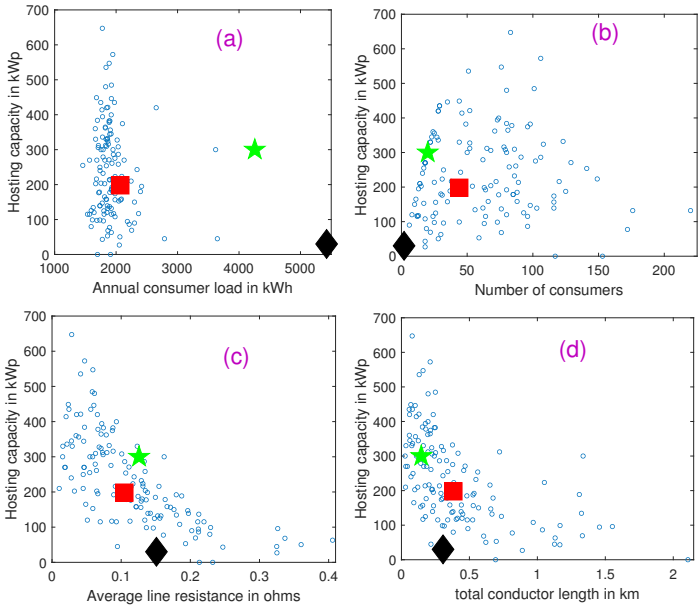


Figure 5.5: Position of HC for selected feeder in HC map for the complete network in relation to: a) annual consumer load in kWp, b) number of consumers, c) average feeder resistance, and d) total conductor length. Selected feeders: red square: feeder 1, green star: feeder 2 and black diamond: feeder 3.

5.5 Conclusion and Future Work

This paper introduced a decoupled general polynomial chaos-based PV hosting capacity tool for low voltage distribution networks. This tool is used to illustrate the detailed probability of congestions for different planning level scenarios, considering uncertainties in load and PV profiles for three selected feeders. The tool is then used to calculate the PV HC of all feeders in an existing European LV network. The decoupled HC tool separates the planning level uncertainties such as size, number, and location of PV from the operational uncertainties

arising from load and PV, making it possible to differentiate the impact of these two types of uncertainty separately. This decoupled PV HC enabled a detailed study of the *grid limit violations* through the probabilities of congestion and their sensitivity towards increased size and penetration of PV installations. For example, a case study shows the influence of size and type of PV on the HC, where three-phase PV installations are more congestion-proof than single phase PV. Similarly, overvoltage is seen as dominant in the case of single phase PV HC. Another case study illustrates how the PV HC of an LVDS feeder increases if the maximum voltage violation is allowed for 25% of the time, concretely by 3 times in the test feeder considered.

The tool presented in this paper opens an avenue to decouple planning level decision problems from operational decision problems. In future work, planning level problems such as the impact of location of PV, the impact of shared PV in an apartment type building, and the impact of community PV will be studied using the decoupled HC tool. Similarly, considering operational uncertainties, the impact of PV curtailment on feeder HC will be studied. In the present work, each planning level scenario is dealt with separately. The extension of this tool is to formulate a single chance constrained optimal power flow (CC-OPF) problem using *Galerkin projection based general polynomial chaos* to calculate an optimal planning scenario. It will solve the decoupled probabilistic PV HC calculation as a one-shot problem.

Chapter 6

Stochastic Optimal Power Flow based Hosting Capacity

Koirala, A., Van Acker, T., Hashmi, Md. U., D'hulst, R., Van Hertem, D., *Chance-constrained optimization based PV hosting capacity calculation using General Polynomial Chaos.* Prepared to submit in IEEE trans. of Power System

The first author is the main author of the article. The contributions of the first author include the conceptualization, software implementation, data curation, visualization and the writing of the article.

Abstract: Increased penetration of renewable resources and new loads have increased the uncertainty levels in low voltage distribution systems (LVDS). This requires considering LVDS planning as a stochastic problem. Low voltage photovoltaics (PV) hosting capacity (HC) calculation is such a planning problem. Traditionally, this is solved by using the iterative Monte Carlo method, which requires solving the power flow equations thousands of times. This paper proposes a chance-constrained optimization-based hosting capacity calculation technique, which eliminates the necessity of repetitive solving of power flow equations. General polynomial chaos expansion is used to translate the input uncertainties defined by probability density function to the hosting capacity of the network. Chance constraints are applied for the nodal voltages and thermal overloading. A case study for an actual LV feeder shows that the computational time of using the Monte Carlo based method is reduced from days to the order

of seconds without any sampling, relaxation, or linearization. Furthermore, the stochastic HC calculated by the proposed Stochastic Optimal Power Flow (SOPF) method gives the upper bound of the distribution network HC which can also be identified using conventional methods.

Keywords: AC optimal power flow, hosting capacity, low voltage distribution system, general polynomial chaos, chance-constrained.

6.1 Introduction

Determining the hosting capacity (HC) of a distribution system is a multidimensional stochastic problem. The level of uncertainty in distribution system has increased due to the increased connection of weather-dependent photovoltaics (PV) generation and new electric vehicles and heat pumps loads. Finding the HC of a distribution system is a planning problem where the additional generation resources or load that can be added to the system without grid congestion is calculated. However, the operation criteria of Low Voltage Distribution System (LVDS) is not very strict and the grid limits for power quality can be violated momentarily [13]. For LVDS planning problem such as HC, risk-based stochastic HC calculations enables more informed decision-making [182].

In [18, 19], existing PV HC methods for LVDS were reviewed and considering uncertainties in calculating HC was deemed important. The available deterministic and stochastic PV HC calculation methods in literature were evaluated and benchmarked in [182]. Stochastic methods consider uncertainties in operational and planning variables while the deterministic methods take only one of the possible operational scenarios, for example, peak PV generation with low load. The HC obtained from deterministic methods takes one snapshot of thousands of possible combinations of stochastic variables, and provides with a rough estimate of the HC. The stochastic HC calculation provides a more comprehensive estimation of the HC than the deterministic one. Another observation from [182] is that every feeder in the network had a different HC and therefore there is no single number to define HC of all feeders. The underlying assumptions of every HC calculation method, such as the possible size of PV, the way of allocating them, and the increment rate of PV size and number, influence the HC of a feeder. Apart from the way of allocating PV, the assumption on the *stochastic* and *operational* constraints applied to the voltage and current directly influences the calculated HC. An important takeaway was the use of a *stochastic* constraints, also known as chance-constraints (CC), in HC calculation which is deemed necessary to prevent calculating the HC for the worst-case [182, 187].

The uncertainties in LVDS can be of different types. In [25], authors differentiated two types of uncertainty that exist in LVDS, i.e., aleatory (background voltage, solar PV production, local consumption) and epistemic (installed capacity per customer, number of customers with solar PV, phase to which single-phase units are connected), stressing the need of treating these two separately. In [188], three types of uncertainties was proposed, operational (load and PV generation), planning (location, size and type of PV installations) and feeder (feeder length, consumer phase) uncertainties. A decoupled approach for PV HC was proposed where the impact of *planning* uncertainties and *operational* uncertainties were decoupled to see the impact of different scenarios in the actual operation of LVDS. The HC of the LVDS was then the scenario with the highest PV within the operational boundary. This method is a quasi brute-force approach to calculate PV HC, which is computationally heavy and time-consuming. To get a true feeder HC, the number of scenarios should be very high, which means a linear increase in computation time.

The stochastic HC methods in literature usually rely on Monte Carlo (MC) based methods [182, 189]. Due to the high computation time associated with MC methods, these methods are usually demonstrated using representative feeders [21], reducing the input source of uncertainties [22], and checking only a few operational parameters [23]. The calculated PV HC is a sub-optimal solution, i.e., the one with the highest total PV amongst the considered PV deployment scenarios. The Optimal Power Flow (OPF) and Stochastic Optimal Power Flow (SOPF) based methods to calculate PV HC in an optimization framework are scarce and in most approaches, many assumptions are made to linearize or relax the problem. In [141, 142], optimization based methods were proposed for medium voltage (MV) feeders with a fixed location of PV. The authors proposed a worst-case analysis using robust optimization in [141], while linearizing the power flow (PF) equations and considering the load of the consumers as the only source of uncertainty. The method improved the computational speed from 84 hrs in the traditional iterative method to 21 s. In [142], a two-stage optimization method was proposed that considers uncertainties in loads and distributed generation, where the operating point is approximated first before calculating the optimal PV HC. The method also used MC to formulate the scenarios of PV location and applied a worst-case analysis. In [190], an orthogonal-polynomial based system was proposed to calculate the PV HC when the uncertainties are represented by continuous probability density functions (PDF). The resultant figure was the risk associated with a particular scenario of PV penetration without any optimization involved. In [24], a analytic-probabilistic approach to distributed generation HC evaluation for radial feeders using an analytical method called Herman Beta Extend transform was demonstrated. This is a moment-based transformation of the power flow equations, where the power flow equations are linearized using Taylor's expansion. The method only works

when the uncertainty of the inputs are represented by a Beta distribution. The proposed method also assumes that the output voltage and current follow a Beta distribution, which is not necessarily true due to the non-linear nature of the power flow equations. However, the method reduced the computational time compared to MC simulations, from days to a few minutes. A scenario-based CC-OPF was presented in [191], where the authors applied chance constraints in the curtailment risk, but an increasing number of scenarios had a huge toll on computation time. Authors of [187] used an actual LV feeder from the UK to demonstrate a stochastic OPF based HC of PV with storage, which calculated the minimum and the maximum HC of each feeder. However, the method relies on linearizing the power flow equations, uses MC samples of load and PV generation, uses worst-case voltage and current limits instead of chance-constraints, and assumes PV size as a constant rather than a variable. The only variable optimized was the number and the location of PV. In [192], a method for calculating stochastic HC for LV networks of the UK is proposed as an optimization problem. A bisection algorithm is used for obtaining optimal PV size and location where the maximum voltage in the feeder is equal to the voltage limit. This method uses linearized power flow and does not consider uncertainties in PV generation and consumer load. However, these methods clearly showed the benefits of moving from scenario-based method or iterative MC based methods to SOPF based calculations. The benefits were: a) reduction on computational time from days to few seconds, b) similar or better accuracy, c) easy scale-up for large service area, and d) minimum input and output data processing requirement. The SOPF based method also makes comparing different policy scenarios faster and easier compared to scenario based MC methods. So far, the methods in literature did not have full ac formulation of CC-OPF HC calculation for low voltage (LV) system that considers the operational uncertainties.

The general Polynomial Chaos (gPC) expansion based CC-OPF introduced in [89] used a full Rectangular Power-Voltage (ACR) formulation of power flow. In [35], authors use the Rectangular Current-Voltage (IVR) formulation of gPC-CC-OPF with auxiliary variables to make problem scalable compared to ACR formulation. This paper proposes a SOPF method to calculate PV HC of LVDS as a unified formulation of gPC and CC-OPF while considering the uncertainties of the operational variables. The CC-OPF HC calculation is presented as a single-shot problem without any sampling, relaxation, or linearization. The method is demonstrated using a single-phase equivalent of a radial network, where load and irradiance are the sources of uncertainty. To the best of our knowledge, this is the first work which proposes a SOPF with chance-constraints to calculate LVDS PV HC without any linearization or sampling. To this end, the main contributions of the paper are:

a) A deterministic OPF based method for calculating PV HC of an LVDS is presented and compared with scenario-based HC calculation methods,

b) A gPC-CC-OPF method for calculating the PV HC of an LVDS feeder is proposed using a single-phase equivalent of the feeder and considering uncertainties in PV generation, PV size, PV number and consumer load,

c) An empirical method to tune the moment-based reformulation of the probabilistic chance constraints when the inputs are non-Gaussian is proposed,

d) The results from gPC-CC-OPF HC method are compared with the OPF HC in a real LVDS setting using European LVDS feeders.

The author believes that the inherent unbalanced nature of the distribution network cannot be correctly represented by a single-phase equivalent. Nevertheless, a single-phase equivalent is used in this paper, assuming all load and PV are balanced. In this way, it proves the application of the gPC-CC-OPF based method for HC calculation. The HC calculated in this chapter is the best planning scenario, which is always three-phase, justifying the balanced assumption. However, if a DSO requires to obtain single-phase PV HC, a three-phase unbalanced gPC-CC-OPF based HC is required. Single-phase PV HC is envisioned as future work and is a natural expansion of this work.

The remaining sections are divided as follows: Section 6.2 presents a deterministic OPF-based PV HC formulation. In Section 6.3, a stochastic formulation (gPC-CC-OPF) for PV HC calculation of low voltage distribution feeder is presented, to solve it as a single-shot problem without any sampling, linearization, or relaxation. Section 6.4 is the numerical illustration of the LVDS PV HC calculations, and Section 6.5 concludes the work in this paper.

6.2 Deterministic HC

In this section, the mathematical formulation of PV HC calculation in LVDS is framed using a simple feeder (Fig. 6.1) and the π-representation (Fig. 6.2) of a segment l of such feeder. The LV substation with a transformer is taken as the source (or slack) bus, with the active and reactive generation of the source limited by the transformer capacity. Although each transformer supplies two to eight LV feeders, this assumption is justified by the ease of upgrading transformers compared to underground cables. Consideration of transformer rating $\mathsf{S}_t^{\text{rated}}$ as the limit of the active power of the grid prevents unrealistic scenarios such as high rated PV at the beginning of the feeder.

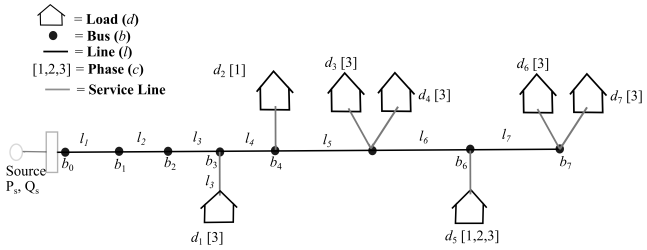


Figure 6.1: Simple representation of a three-phase LV feeder

LV distribution systems in Europe consist of a set of radial LV feeders. The LVDS HC is computed as the sum of the HC of the individual feeders. The PV HC of a feeder is the total PV that can be added to that feeder without violating its *operational limits* during any realistic operational scenario [182, 189]. The usual *operational limits* are minimum voltage $U^{\min}$, maximum voltage $U^{\max}$, maximum voltage unbalance limit, cable current rating I^{rated} or transformer rating $\mathsf{S}_t^{\text{rated}}$. The operational scenarios are the result of the uncertainty in irradiance and load of the consumer while the planning scenarios are due to the uncertainty in the size of the PV, location of PV and type of PV [188]. The feeder's HC is computed as the maximization of the sum of the additional PV capacity a subset of consumers on the feeder can accommodate, i.e., the planning scenario with the highest PV, without violating its *operational limits*:

$$\text{HC} = \max \sum_{p \in \mathcal{P}} P_p^{\text{kWp}} \tag{6.1}$$

where, $\mathcal{P}$ is the set of new PV installations in the feeder, assuming each LV consumer d has the possibility to have a PV p of rating P_p^{kWp}.

Fig. 6.2 shows the simple π-model of the single phase equivalent of a radial LV feeder segment to describe the OPF formulation. Buses $i \in \mathcal{I}$ are the vertices of the graph representing the feeder and branches $l \in \mathcal{L}$ are the edges of the graph. The bus voltage in rectangular coordinates is:

$$\mathsf{U}_i = \mathsf{U}_i^{\text{re}} + j\, \mathsf{U}_i^{\text{im}}. \tag{6.2}$$

In the deterministic formulation, the voltage magnitude $|\mathsf{U}_i|$ is bounded between $U_i^{\min}$ and $U_i^{\max}$. The parameters y_i^{sh}, g_i^{sh} and b_i^{sh} denote the corresponding admittance, conductance and susceptance at a bus i. The current

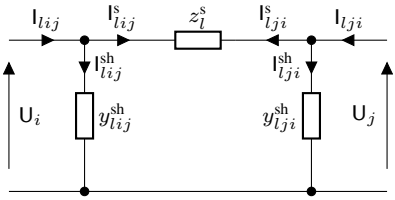


Figure 6.2: π-model of a branch lij, with series impedance z^{s}_l and from- and to-side shunt admittances y^{sh}_{lij} and y^{sh}_{lji}.

flowing through a branch l from bus i to j is:

$$\mathsf{I}_{lij} = \mathsf{I}^{\mathrm{re}}_{lij} + j\,\mathsf{I}^{\mathrm{im}}_{lij} \tag{6.3}$$

$$= \mathsf{I}^{\mathrm{sh}}_{lij} + \mathsf{I}^{\mathrm{s}}_{lij} \tag{6.4}$$

$$= (\mathsf{I}^{\mathrm{sh,re}}_{lij} + \mathsf{I}^{\mathrm{s,re}}_{lij}) + j\,(\mathsf{I}^{\mathrm{sh,im}}_{lij} + \mathsf{I}^{\mathrm{s,im}}_{lij}) \tag{6.5}$$

where, $lij \in \mathcal{T}^{\mathrm{l}\rightarrow}$ links a branch l to its from- i and to-bus j. The union set $\mathcal{T}^{\mathrm{l}} = \mathcal{T}^{\mathrm{l}\rightarrow} \cup \mathcal{T}^{\mathrm{l}\leftarrow}$ has the the reversed branch set $\mathcal{T}^{\mathrm{l}\leftarrow}$ which links a branch l from its to-bus j and from-bus i. The current magnitude $|\mathsf{I}_{lij}|$ is bounded by I^{rated}_l. z^{s}_l denotes the series impedance of a branch l, where r^{s}_l and x^{s}_l are the corresponding series resistance and reactance. y^{sh}_{lij} denotes the shunt admittance of a branch l at bus i, and g^{sh}_{lij} and b^{sh}_{lij} are the corresponding shunt conductance and susceptance.

Units $u \in \mathcal{U}$ generalize source, loads, and distributed generators in the feeder. A tuple $ui \in \mathcal{T}^{\mathrm{u}}$ links a unit u to a specific bus i. The current flowing from the bus i to the unit u and complex power are,

$$\mathsf{I}_u = \mathsf{I}^{\mathrm{re}}_u + j\,\mathsf{I}^{\mathrm{im}}_u, \tag{6.6}$$

$$\mathsf{S}_u = \mathsf{U}_i \mathsf{I}^{*}_u \tag{6.7}$$

$$= (\mathsf{U}^{\mathrm{re}}_i \mathsf{I}^{\mathrm{re}}_u + \mathsf{U}^{\mathrm{im}}_i \mathsf{I}^{\mathrm{im}}_u) + j\,(\mathsf{U}^{\mathrm{im}}_i \mathsf{I}^{\mathrm{re}}_u - \mathsf{U}^{\mathrm{re}}_i \mathsf{I}^{\mathrm{im}}_u) \tag{6.8}$$

$$= \mathsf{P}_u + j\,\mathsf{Q}_u. \tag{6.9}$$

A unit u is either a load $d \in \mathcal{D}_f$ or a PV device $p \in \mathcal{P}$ or a source generator $s \in \mathcal{S}$ in slack-bus. The set-points of the load are P^{ref}_d and Q^{ref}_d and the limits for PV size are P^{min}_p, P^{max}_p, Q^{min}_p and Q^{max}_p, respectively. The size of the PV installations P^{kWp}_p is the only variable in PV HC calculations. The source generator is connected to the feeder head and has maximum real power P^{max}_s and maximum imaginary power Q^{max}_s of the order of the transformer rating.

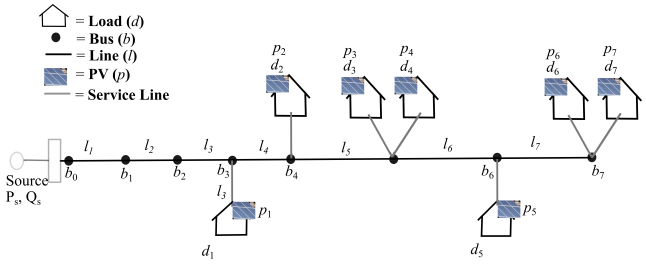


Figure 6.3: Single phase representation of a LV feeder with all consumer with PV for OPF-HC

As previously mentioned, these limits are to avoid unrealistic PV scenarios and should not be confused with transformer loading (Fig. 6.3). To consider reverse power flow, the minimum real and imaginary powers of the source, $P_s^{\min}$ and $Q_s^{\min}$, are considered to be the negative of the maximum limits.

In a deterministic HC problem, the assumption is made that each consumer can have up to $P_p^{\max}$ PV capacity. It is assumed that the PV operates at a unity power factor, such that $Q_p^{\min} = 0$ and $Q_p^{\max} = 0$. The peak irradiance is denoted by E in kilowatt/m^2, so that the actual PV generation is $\mathsf{P}_p = E \cdot P^{\mathrm{kWp}}$, neglecting inverter efficiency, soiling loss, etc. PV systems smaller than 5 kWp are usually connected single-phase, while larger systems are connected three-phase. However, when solving as a single phase-equivalent, all PVs are considered balanced, and so is the load. The variable P_p^{kWp} is assumed to be a continuous variable and is a relaxed version of the discrete PV size. It is assumed that all consumers can have only one PV with a rating between $P_p^{\min}$ and $P_p^{\max}$. Once the maximum PV size is computed, the individual P_p^{kWp} can be rounded to the lower value of PV size in PV ratings set ($\mathcal{P}^{\mathrm{rt}}$) if such set is available.

Consequently, the feasible set of the HC formulation as IVR formulation of the standard deterministic OPF problem is presented below for more clarity:

OPF HC

Reference Bus Constraint:

$$\mathsf{U}_i^{\mathrm{im}} = 0, \qquad \forall i \in \mathcal{I}^{\mathrm{ref}}, \tag{6.9a}$$

Source generator Constraint:

$$-Q_s^{\max} \leq \mathsf{Q}_s \leq Q_s^{\max}, \qquad \forall s \in \mathcal{S}, \tag{6.9b}$$

$$-P_s^{\max} \leq \mathsf{P}_s \leq P_s^{\max}, \qquad \forall s \in \mathcal{S}, \tag{6.9c}$$

$$\mathsf{P}_s = \mathsf{U}_i^{\mathrm{re}}\mathsf{I}_s^{\mathrm{re}} + \mathsf{U}_i^{\mathrm{im}}\mathsf{I}_s^{\mathrm{im}}, \qquad \forall si \in \mathcal{T}^{\mathrm{s}}, \tag{6.9d}$$

$$\mathsf{Q}_s = \mathsf{U}_i^{\mathrm{im}}\mathsf{I}_s^{\mathrm{re}} - \mathsf{U}_i^{\mathrm{re}}\mathsf{I}_s^{\mathrm{im}}, \qquad \forall si \in \mathcal{T}^{\mathrm{s}}, \tag{6.9e}$$

Bus Constraints:

$$\sum_{lij \in \mathcal{T}^{\mathrm{l}}} \mathsf{I}_{lij}^{\mathrm{re}} + \sum_{ui \in \mathcal{T}^{\mathrm{u}}} \mathsf{I}_u^{\mathrm{re}} + g_i^{\mathrm{sh}}\mathsf{U}_i^{\mathrm{re}} - b_i^{\mathrm{sh}}\mathsf{U}_i^{\mathrm{im}} = 0, \quad \forall i \in \mathcal{I}, \tag{6.9f}$$

$$\sum_{lij \in \mathcal{T}^{\mathrm{l}}} \mathsf{I}_{lij}^{\mathrm{im}} + \sum_{ui \in \mathcal{T}^{\mathrm{u}}} \mathsf{I}_u^{\mathrm{im}} + g_i^{\mathrm{sh}}\mathsf{U}_i^{\mathrm{im}} + b_i^{\mathrm{sh}}\mathsf{U}_i^{\mathrm{re}} = 0, \quad \forall i \in \mathcal{I}, \tag{6.9g}$$

$$(U_i^{\min})^2 \leq (\mathsf{U}_i^{\mathrm{re}})^2 + (\mathsf{U}_i^{\mathrm{im}})^2 \leq (U_i^{\max})^2, \qquad \forall i \in \mathcal{I}, \tag{6.9h}$$

Branch Constraints:

$$\mathsf{U}_j^{\mathrm{re}} = \mathsf{U}_i^{\mathrm{re}} - r_l^{\mathrm{s}}\mathsf{I}_{lij}^{\mathrm{s,re}} + x_l^{\mathrm{s}}\mathsf{I}_{lij}^{\mathrm{s,im}}, \qquad \forall lij \in \mathcal{T}^{\mathrm{l}\rightarrow}, \tag{6.9i}$$

$$\mathsf{U}_j^{\mathrm{im}} = \mathsf{U}_i^{\mathrm{im}} - r_l^{\mathrm{s}}\mathsf{I}_{lij}^{\mathrm{s,im}} - x_l^{\mathrm{s}}\mathsf{I}_{lij}^{\mathrm{s,re}}, \qquad \forall lij \in \mathcal{T}^{\mathrm{l}\rightarrow}, \tag{6.9j}$$

$$\mathsf{I}_{lij}^{\mathrm{re}} = g_{lij}^{\mathrm{sh}}\mathsf{U}_i^{\mathrm{re}} - b_{lij}^{\mathrm{sh}}\mathsf{U}_i^{\mathrm{im}} + \mathsf{I}_{lij}^{\mathrm{s,re}}, \qquad \forall lij \in \mathcal{T}^{\mathrm{l}}, \tag{6.9k}$$

$$\mathsf{I}_{lij}^{\mathrm{im}} = g_{lij}^{\mathrm{sh}}\mathsf{U}_i^{\mathrm{im}} + b_{lij}^{\mathrm{sh}}\mathsf{U}_i^{\mathrm{re}} + \mathsf{I}_{lij}^{\mathrm{s,im}}, \qquad \forall lij \in \mathcal{T}^{\mathrm{l}}, \tag{6.9l}$$

$$(\mathsf{I}_{lij}^{\mathrm{re}})^2 + (\mathsf{I}_{lij}^{\mathrm{im}})^2 \leq (I_l^{\mathrm{rated}})^2, \qquad \forall lij \in \mathcal{T}^{\mathrm{l}}, \tag{6.9m}$$

Demand Constraints:

$$\mathsf{U}_i^{\mathrm{re}}\mathsf{I}_d^{\mathrm{re}} + \mathsf{U}_i^{\mathrm{im}}\mathsf{I}_d^{\mathrm{im}} = P_d^{\mathrm{ref}}, \qquad \forall di \in \mathcal{T}^{\mathrm{d}}, \tag{6.9n}$$

$$\mathsf{U}_i^{\mathrm{im}}\mathsf{I}_d^{\mathrm{re}} - \mathsf{U}_i^{\mathrm{re}}\mathsf{I}_d^{\mathrm{im}} = Q_d^{\mathrm{ref}}, \qquad \forall di \in \mathcal{T}^{\mathrm{d}}, \tag{6.9o}$$

PV Constraints:

$$\mathsf{P}_p = \mathsf{U}_i^{\mathrm{re}}\mathsf{I}_p^{\mathrm{re}} + \mathsf{U}_i^{\mathrm{im}}\mathsf{I}_p^{\mathrm{im}}, \qquad \forall pi \in \mathcal{T}^{\mathrm{p}}, \tag{6.9p}$$

$$\mathsf{Q}_p = \mathsf{U}_i^{\mathrm{im}}\mathsf{I}_p^{\mathrm{re}} - \mathsf{U}_i^{\mathrm{re}}\mathsf{I}_p^{\mathrm{im}}, \qquad \forall pi \in \mathcal{T}^{\mathrm{p}}, \tag{6.9q}$$

$$\mathsf{P}_p = EP_p^{\mathrm{kWp}}, \qquad \forall p \in \mathcal{P}, \tag{6.9r}$$

$$0 \leq P_p^{\mathrm{kWp}} \leq P_p^{\max}, \qquad \forall p \in \mathcal{P}, \tag{6.9s}$$

$$\mathsf{Q}_p = 0, \qquad \forall p \in \mathcal{P}. \tag{6.9t}$$

The objective of the optimization is to maximize the total PV capacity as given in (6.1). The uncertainties in both load and irradiance are neglected in this formulation. The final hosting capacity is sum of the individual PV P_p^{kWp}. The output is the highest possible PV that the feeder can accommodate under network constraint for a particular snapshot of the operational scenarios. For PV HC with equal PV, the PV constraints in (6.9p) – (6.9t) can be modified to keep P^{kWp} equal among the $p \in \mathcal{P}$. That means the $|\mathcal{P}|$ number of constraints for (6.9r)-(6.9t) is replaced by single constraints in (6.9w)-(6.9y)

OPF HC with equal PV

Constraints (6.9a) – (6.9o)

$$\mathsf{P} = \mathsf{U}_i^{\mathrm{re}}\mathsf{I}_p^{\mathrm{re}} + \mathsf{U}_i^{\mathrm{im}}\mathsf{I}_p^{\mathrm{im}} \qquad \forall pi \in \mathcal{T}^{\mathrm{p}}, \tag{6.9u}$$

$$\mathsf{Q} = \mathsf{U}_i^{\mathrm{im}}\mathsf{I}_p^{\mathrm{re}} - \mathsf{U}_i^{\mathrm{re}}\mathsf{I}_p^{\mathrm{im}} \qquad \forall pi \in \mathcal{T}^{\mathrm{p}}, \tag{6.9v}$$

$$\mathsf{P} = EP^{\mathrm{kWp}}, \tag{6.9w}$$

$$0 \leq P^{\mathrm{kWp}} \leq P^{\max}, \tag{6.9x}$$

$$\mathsf{Q} = 0. \tag{6.9y}$$

6.3 Stochastic OPF HC

The formulation presented in Section 6.2 assumes deterministic load and irradiance set-points. However, in reality, these set-points are random and time-dependent. All variables in (6.10) become random variables for each time instance t under consideration. In [188], it was shown that overvoltage is the most limiting factor for PV HC. Furthermore, after a Probabilistic Power Flow (PPF) based evaluation, the maximum probability of overvoltage was deemed as a suitable index to calculate PV HC in an LVDS. Instead of running the OPF for HC calculation for all time period $t \in \mathcal{T}$, the search space is reduced by only investigating the time-point τ where the probability that the maximum voltage in the feeder is higher than $U^{\max}$ is highest, i.e., $p_t^{\mathrm{ov}} = \mathbb{P}(\max(\mathsf{U}_{i,t}) > U^{\max})$ is the highest. This instance is found a priori by using PPF from [183] for all timestamps, assuming high PV penetration. The remainder of the CC-OPF formulation is described for single-period optimization for τ time where $p_\tau^{\mathrm{ov}} = \max(p_t^{\mathrm{ov}})$.

The stochastic variables are captured by a real-valued finite-variance stochastic germ $\omega = [\omega_1, ..., \omega_N]^{\mathrm{T}}$ with $N \in \mathbb{N}$, and corresponding set of possible realizations $\Omega \subset \mathbb{R}^N$. The first (N-1) stochastic variables for the selected time τ are due to the uncertainties in load. The last variable ω_N describes the solar irradiance E.

The gPC-CC-OPF used in this paper for stochastic HC is based on the Galerkin projection of the uncertain variables in the power flow equations. The projection is done in form of the inner products of orthogonal polynomials for each uncertain variable considered. The gPC converts the infinite dimensional space of the continuous uncertainty into a finite dimensional vector $\mathcal{K}$ based on the orthogonal polynomials [28, 193]. The value of dimension $\mathcal{K}$ depends on the number of uncertainty variables considered and polynomial degree considered, i.e., for second degree polynomials with N sources of uncertainty, the dimension of $\mathcal{K}$ is given by $\frac{(2+N)!}{2!N!}$. The gPC based reformulation of stochastic problem are mainly based on the properties of orthogonal polynomials summarized below and interested readers are referred to [28, 193] for details:

a) The continuous uncertain variable with known probability density function can be accurately represented by degree one orthogonal polynomials, i.e., with two coefficients.

b) The tensor product of univariate orthogonal polynomials can be used to create the multi-variate orthogonal polynomials $[\Phi_k]_{k\in\mathcal{K}}$ which act as the base for all variables under consideration.

c) The inner product of the orthogonal polynomial bases (also known as expectation operation) can be pre-computed and is given as:

$$\begin{aligned} \langle \Phi_l, \Phi_k \rangle = \mathbb{E}[\Phi_l, \Phi_k] &= \gamma_l \delta_{lk}, \ \forall l, k \in \mathcal{K} \subseteq \mathbb{N}, \\ \delta_{lk} &= 1 \qquad \forall l \neq k, \\ \delta_{lk} &= 0 \qquad \forall l = k \end{aligned} \tag{6.10}$$

where, γ_l is a positive scalar which can be precomputed and δ_{lk} is the Kronecker delta.

d) The first coefficient of the polynomial representation of the variable is the mean of that variable, and the square of the remaining coefficients is the variance of that variable. The approximate polynomial representation of a random variable x is:

$$\hat{\mathsf{x}} = \sum_{k\in\mathcal{K}} x_k \Phi_k, \tag{6.11}$$

where, x_k is the coefficient of Φ_k, the orthogonal polynomial basis. The approximated mean and variance of x are as follows:

$$\mathbb{E}[\hat{\mathsf{x}}] = x_0, \tag{6.12}$$

$$\mathbb{V}[\hat{\mathsf{x}}] = \sum_{k \in \mathcal{K} \setminus \{0\}} \langle \Phi_k, \Phi_k \rangle x_k^2. \tag{6.13}$$

The propagation of the uncertainty from the reference load and generation to the nodal voltages and current involves mainly two operations, summation and multiplication. For three approximated random variables $\hat{\mathsf{x}}$, $\hat{\mathsf{y}}$ and $\hat{\mathsf{z}}$ which are defined on the same orthogonal basis, the Galerkin projection of the operations are:

$$\hat{\mathsf{z}} = \hat{\mathsf{x}} + \hat{\mathsf{y}} \overset{\text{gp}}{\rightarrow} z_k = x_k + y_k, \qquad \forall k \in \mathcal{K}, \tag{6.14}$$

$$\hat{\mathsf{z}} = \hat{\mathsf{x}} \cdot \hat{\mathsf{y}} \overset{\text{gp}}{\rightarrow} z_k = \sum_{k_1, k_2 \in \mathcal{K}} M(x_{k_1} \cdot y_{k_2}), \quad \forall k \in \mathcal{K}, \tag{6.15}$$

where M denotes the multiplication tensor,

$$M = \frac{\langle \Phi_{k_1}, \Phi_{k_2}, \Phi_k \rangle}{\langle \Phi_k, \Phi_k \rangle}, \tag{6.16}$$

and can be computed ahead-of-time.

In a stochastic context, enforcing deterministic bounds on continuous random variables is not possible. Alternatively, chance constraints ensure that the probability of a random variable x violating a deterministic bound $x^{\min}$ or $x^{\max}$ is below a certain stochastic bound ε:

$$\mathbb{P}(\mathsf{x} \geq x^{\min}) \geq (1 - \varepsilon) \text{ or } \mathbb{P}(\mathsf{x} \leq x^{\max}) \geq (1 - \varepsilon). \tag{6.17}$$

The choice of ε depends upon the nature of the solution required. If ε=0, the solution obtained is distributionally robust, which might lead to underestimating the actual HC of network [182]. This can be reformulated as a moment-based equation [35, 89],

$$x^{\min} \leq \mathbb{E}(\mathsf{x}) \pm \lambda(\varepsilon) \sqrt{\mathbb{V}(\mathsf{x})} \leq x^{\max}, \tag{6.18}$$

where $\lambda(\varepsilon) > 0$ is chosen based on knowledge of the random variable. In [89] and [35], $\lambda(\varepsilon)$ was taken assuming the output as Gaussian for all nodes and lines, i.e., $\lambda(\varepsilon) = 1.65$ if $\varepsilon = 0.05$. In this paper, an empirical tuning method is used where instead of taking the same $\lambda(\varepsilon)$ for all limits, each node is assigned a separate moment based reformulation factor $\lambda_i(\varepsilon^{\text{v}})$. Similarly, each line has a particular $\lambda_{lij}(\varepsilon^{\text{i}})$.

The following auxiliary variables are used in this formulation.

1. squared bus voltage magnitude: $\mathsf{W}_i = (\mathsf{U}_i^{\mathrm{re}})^2 + (\mathsf{U}_i^{\mathrm{im}})^2$,
2. squared branch current magnitude: $\mathsf{J}_{lij} = (\mathsf{I}_{lij}^{\mathrm{re}})^2 + (\mathsf{I}_{lij}^{\mathrm{im}})^2$,
3. active source power: $\mathsf{P}_s = \mathsf{U}_i^{\mathrm{re}}\mathsf{I}_s^{\mathrm{re}} + \mathsf{U}_i^{\mathrm{im}}\mathsf{I}_s^{\mathrm{im}}$, and,
4. reactive source power: $\mathsf{Q}_s = \mathsf{U}_i^{\mathrm{im}}\mathsf{I}_s^{\mathrm{re}} - \mathsf{U}_i^{\mathrm{re}}\mathsf{I}_s^{\mathrm{im}}$.

The reason was to eliminate the fourth order tensor product in gPC-CC-OPF formulation on IVR formulation space. Interested readers are referred to [35] on how the fourth-order permutation of $\mathcal{K}$ presented in [16] is reduced to second-order permutations by the use of these auxiliary variables.

Finally, the feasible set of the IVR formulation of the gPC-CC-OPF for chance constrained PV HC is:

gPC-CC-OPF HC

Reference Bus Constraints — $i \in \mathcal{I}^{\mathrm{ref}}$:

$$\mathsf{U}_{i,k}^{\mathrm{re}} = 0, \qquad \forall k \in \mathcal{K}_0, \tag{6.19a}$$

$$\mathsf{U}_{i,k}^{\mathrm{im}} = 0, \qquad \forall k \in \mathcal{K}, \tag{6.20}$$

Source generator Constraints — $si \in \mathcal{T}^{\mathrm{s}}$

$$\sum_{k_1,k_2\in\mathcal{K}} M(\mathsf{U}_{i,k_1}^{\mathrm{re}}\mathsf{I}_{s,k_2}^{\mathrm{re}} + \mathsf{U}_{i,k_1}^{\mathrm{im}}\mathsf{I}_{s,k_2}^{\mathrm{im}}) = \mathsf{P}_{s,k}, \qquad \forall k \in \mathcal{K}, \tag{6.20a}$$

$$\sum_{k_1,k_2\in\mathcal{K}} M(\mathsf{U}_{i,k_1}^{\mathrm{im}}\mathsf{I}_{s,k_2}^{\mathrm{re}} - \mathsf{U}_{i,k_1}^{\mathrm{re}}\mathsf{I}_{s,k_2}^{\mathrm{im}}) = \mathsf{Q}_{s,k}, \qquad \forall k \in \mathcal{K}, \tag{6.21}$$

$$-P_s^{\max} \leq \mathbb{E}(\overline{\mathsf{P}}_s) \pm \lambda(\varepsilon)\sqrt{\mathbb{V}(\overline{\mathsf{P}}_s)} \leq P_s^{\max}, \tag{6.21a}$$

$$-Q_s^{\max} \leq \mathbb{E}(\overline{\mathsf{Q}}_s) \pm \lambda(\varepsilon)\sqrt{\mathbb{V}(\overline{\mathsf{Q}}_s)} \leq Q_s^{\max}, \tag{6.21b}$$

Bus Constraints — $\forall i \in \mathcal{I}$:

$$\sum_{lij\in\mathcal{T}^{\mathrm{l}}} \mathsf{I}_{lij,k}^{\mathrm{re}} + \sum_{ui\in\mathcal{T}^{\mathrm{u}}} \mathsf{I}_{u,k}^{\mathrm{re}} + g_i^{\mathrm{sh}}\mathsf{U}_{i,k}^{\mathrm{re}} - b_i^{\mathrm{sh}}\mathsf{U}_{i,k}^{\mathrm{im}} = 0, \qquad \forall k \in \mathcal{K}, \tag{6.22a}$$

$$\sum_{lij\in\mathcal{T}^{\mathrm{l}}} \mathsf{I}_{lij,k}^{\mathrm{im}} + \sum_{ui\in\mathcal{T}^{\mathrm{u}}} \mathsf{I}_{u,k}^{\mathrm{im}} + g_i^{\mathrm{sh}}\mathsf{U}_{i,k}^{\mathrm{im}} + b_i^{\mathrm{sh}}\mathsf{U}_{i,k}^{\mathrm{re}} = 0, \qquad \forall k \in \mathcal{K}, \tag{6.22b}$$

$$\sum_{k_1,k_2\in\mathcal{K}} M(\mathsf{U}^{\mathrm{re}}_{i,k_1}\mathsf{U}^{\mathrm{re}}_{i,k_2}+\mathsf{U}^{\mathrm{im}}_{i,k_1}\mathsf{U}^{\mathrm{im}}_{i,k_2})=\mathsf{W}_{i,k}, \qquad \forall k\in\mathcal{K}, \tag{6.22c}$$

$$(U_i^{\min})^2 \le \mathbb{E}(\overline{\mathsf{W}}_i) \pm \lambda_i(\varepsilon^{\mathrm{v}})\sqrt{\mathbb{V}(\overline{\mathsf{W}}_i)} \le (U_i^{\max})^2, \tag{6.22d}$$

From Branch Constraints — $\forall lij \in \mathcal{T}^{\mathrm{l}\rightarrow}$:

$$\mathsf{U}^{\mathrm{re}}_{j,k}=\mathsf{U}^{\mathrm{re}}_{i,k}-r^{\mathrm{s}}_l\mathsf{I}^{\mathrm{s,re}}_{lij,k}+x^{\mathrm{s}}_l\mathsf{I}^{\mathrm{s,im}}_{lij,k}, \qquad \forall k\in\mathcal{K}, \tag{6.22e}$$

$$\mathsf{U}^{\mathrm{im}}_{j,k}=\mathsf{U}^{\mathrm{im}}_{i,k}-r^{\mathrm{s}}_l\mathsf{I}^{\mathrm{s,im}}_{lij,k}-x^{\mathrm{s}}_l\mathsf{I}^{\mathrm{s,re}}_{lij,k}, \qquad \forall k\in\mathcal{K}, \tag{6.22f}$$

Branch Constraints — $\forall lij \in \mathcal{T}^{\mathrm{l}}$:

$$\mathsf{I}^{\mathrm{re}}_{lij,k}=g^{\mathrm{sh}}_{lij}\mathsf{U}^{\mathrm{re}}_{i,k}-b^{\mathrm{sh}}_{lij}\mathsf{U}^{\mathrm{im}}_{i,k}+\mathsf{I}^{\mathrm{s,re}}_{lij,k}, \qquad \forall k\in\mathcal{K}, \tag{6.22g}$$

$$\mathsf{I}^{\mathrm{im}}_{lij,k}=g^{\mathrm{sh}}_{lij}\mathsf{U}^{\mathrm{im}}_{i,k}+b^{\mathrm{sh}}_{lij}\mathsf{U}^{\mathrm{re}}_{i,k}+\mathsf{I}^{\mathrm{s,im}}_{lij,k}, \qquad \forall k\in\mathcal{K}, \tag{6.22h}$$

$$\sum_{k_1,k_2\in\mathcal{K}} M(\mathsf{I}^{\mathrm{re}}_{lij,k_1}\mathsf{I}^{\mathrm{re}}_{lij,k_2}+\mathsf{I}^{\mathrm{im}}_{lij,k_1}\mathsf{I}^{\mathrm{im}}_{lij,k_2})=\mathsf{J}_{lij,k}, \qquad \forall k\in\mathcal{K}, \tag{6.22i}$$

$$\mathbb{E}(\overline{\mathsf{J}}_{lij}) \pm \lambda_{lij}(\varepsilon^{\mathrm{i}})\sqrt{\mathbb{V}(\overline{\mathsf{J}}_{lij})} \le (I_l^{\mathrm{rated}})^2, \tag{6.22j}$$

Demand Constraints — $\forall di \in \mathcal{T}^{\mathrm{d}}$:

$$\sum_{k_1,k_2\in\mathcal{K}} M(\mathsf{U}^{\mathrm{re}}_{i,k_1}\mathsf{I}^{\mathrm{re}}_{d,k_2}+\mathsf{U}^{\mathrm{im}}_{i,k_1}\mathsf{I}^{\mathrm{im}}_{d,k_2})=P^{\mathrm{ref}}_{d,k}, \qquad \forall k\in\mathcal{K}, \tag{6.22k}$$

$$\sum_{k_1,k_2\in\mathcal{K}} M(\mathsf{U}^{\mathrm{im}}_{i,k_1}\mathsf{I}^{\mathrm{re}}_{d,k_2}-\mathsf{U}^{\mathrm{re}}_{i,k_1}\mathsf{I}^{\mathrm{im}}_{d,k_2})=Q^{\mathrm{ref}}_{d,k}, \qquad \forall k\in\mathcal{K}, \tag{6.22l}$$

PV Constraints — $\forall pi \in \mathcal{T}^{\mathrm{p}}$:

$$\sum_{k_1,k_2\in\mathcal{K}} M(\mathsf{U}^{\mathrm{re}}_{i,k_1}\mathsf{I}^{\mathrm{re}}_{p,k_2}+\mathsf{U}^{\mathrm{im}}_{i,k_1}\mathsf{I}^{\mathrm{im}}_{p,k_2})=P^{\mathrm{kWp}}_{p}E^{\mathrm{ref}}_{k}, \qquad \forall k\in\mathcal{K}, \tag{6.22m}$$

$$\sum_{k_1,k_2\in\mathcal{K}} M(\mathsf{U}^{\mathrm{im}}_{i,k_1}\mathsf{I}^{\mathrm{re}}_{p,k_2}-\mathsf{U}^{\mathrm{re}}_{i,k_1}\mathsf{I}^{\mathrm{im}}_{p,k_2})=0, \qquad \forall k\in\mathcal{K}, \tag{6.22n}$$

$$0 \le P^{\mathrm{kWp}}_{p} \le P^{\max}_{p}. \tag{6.22o}$$

---

where M denotes the multiplication tensor for the Galerkin projection of uncertainties, and $\overline{\mathsf{X}}_{(\cdot)}=[\mathsf{X}_{(\cdot,k)}]_{k\in\mathcal{K}}$ is employed for convenience to represent the vector of coefficient variable of $\mathsf{X}_{(\cdot)}$. As in the deterministic case, the inverter

efficiency, soiling loss and other uncertainties involved with PV generation is neglected for this case. Finally, an objective is introduced to maximize total PV size :

$$\max \sum_{p \in \mathcal{P}} P_p^{\text{kWp}} \tag{6.23}$$

6.4 Numerical Illustration

In this section, the application of OPF and CC-OPF for HC calculation in a test feeder is shown. Then, the method is used to compute the PV HC of a set of actual LVDS feeders in [30]. The aim is to establish the method and discuss the limitations and applications of this tool. The code is available on github[1] along with the JSON format of dataset [2] [30].

All the simulations in this paper were done on a PC with a 2.11 GHz Intel i7-8650U CPU processor with 16 GB memory, using JULIA 1.6.5. No parallelization of any sort was used in all of the calculations. The non-linear Ipopt solver v0.7.0 is used for the calculations.

6.4.1 Test Feeder

A simplified test feeder, as shown in Fig. 6.3, is considered for a detailed study. The feeder has seven balanced three-phase loads, with each consumer having the possibility to have a three-phase PV installation between 0-15 kWp. The single time-period optimization is done on the timestamp with the maximum

[1] https://github.com/arpkoirala/StochasticPowerModels.jl/tree/PVhostingCap

[2] https://github.com/arpkoirala/StochasticPowerModels.jl/tree/PVhostingCap/test/data/Spanish

Table 6.1: Uncertainty PDF for load and irradiance for τ.

ω	Distribution	Consumer	Annual kWh	Connection kW
1	$\mathcal{B}([0, 3.98], 2.11, 20.90)$	1,2,4,7	1300-1612	4.6-5.75
2	$\mathcal{B}([0, 5.7], 1.17, 2.54)$	5	5308	17.32
3	$\mathcal{B}([0, 5.9], 1.29, 5.34)$	3	1226	4.6
4	$\mathcal{B}([0, 5.38], 1.75, 12.08)$	6	2022	4.6
5	$\mathcal{B}([748, 799], 0.24, 0.22)$	Irradiance	-	-

probability of overvoltage on a typical high irradiance spring day. The time-stamp chosen was 15:00-15:15 hrs as it was the time period with the highest p_t^{ov} in a priori PPF with all consumers having 10 kWp. The PV generation was considered perfectly correlated as the span of the feeder is 300 meters only. The consumers were classified into four categories based on their annual consumption and connected kVA. The consumers in the same category were considered to be linearly correlated. All the load and PV distributions obtained from the historical profiles were fitted as a beta distribution [103, 188]. In Table 6.1, the details of the uncertainty terms considered for the test case are given. In this work, all consumers in the same group are assumed to be perfectly correlated. However, if available, the correlation factor between the consumers can be readily endorsed in the model. The deterministic calculations were done assuming the maximum irradiance and minimal load of 0.1 kW for all consumers.

6.4.2 Deterministic OPF HC

This section compares the result obtained from OPF HC with the deterministic scenario based HC method. In the scenario based method, several planning scenarios were taken, and the best planning scenario within the operational limits was considered to be the PV HC of that feeder. The main drawback of this method is the number of scenarios to be considered. For this test case, for seven consumers with the possibility of having PV from 0-15 kWp in steps of 1 kWp the number of scenarios is 268 435 456 (16^7). These scenarios are only due to PV size, as other uncertainties, e.g., PV type and connected phase, are not

Table 6.2: Computation time of Scenario based MC, OPF HC, and gPC-CC-OPF HC for test feeder (7 consumer, each consumer can get 0-15 kWp in step of 1 kWh)

HC Method	No of power flow	computation	HC
Scenario based (Det., 3% scenarios)	2 million	≈3 days	50 kWp
Scenario based (Det., complete set)	68 435 456	≈190 days*	-
Scenario based (MC, 1000 samples)	68 435 456 000	≈520 yrs**	-
OPF HC	-	0.241 sec	50.4 kWp
gPC-CC-OPF HC	-	4.04 sec	52.4 kWp
gPC-CC-OPF HC (With equal PV)	-	0.51 sec	44.8 kWp

* scaled from time for 1% of the cases

** scaled from time for OPF with 1 sample

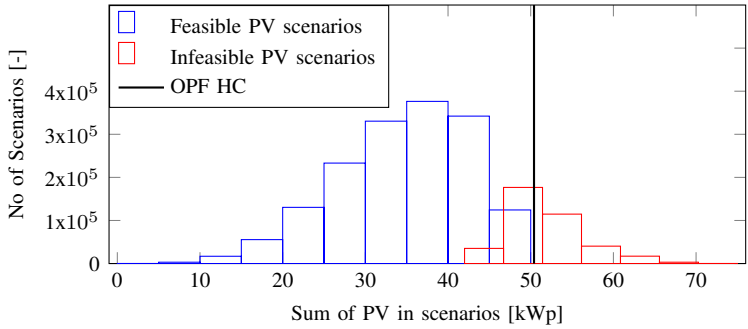


Figure 6.4: Histogram of feasible and infeasible scenarios compared with OPF-HC

even considered. A brute force deterministic evaluation of all these scenarios might take a few months, if not years. If the scenario based method is stochastic as proposed in [188], the time consumed will be multiplied by an order. For this practical reason, the OPF HC and deterministic scenario based HC methods are compared, assuming that a similar relationship will hold between stochastic OPF HC and stochastic scenario based HC methods.

A histogram is plotted for feasible and infeasible PV scenarios using only two million randomly selected possible scenarios (only 3% of the possible scenarios) using deterministic limits along with the OPF HC (Fig. 6.4). The time taken by the OPF HC was 0.241 sec, while it was ≈ 2 days for deterministic scenario based HC (Table 6.2). The OPF HC calculated was 50.4 kWp, while the scenario-based HC was 50 kWp, the closest possible integer value below OPF-HC. There are 41 621 scenarios where the sum of PV installed is equal to 50 kWp, of which 3 252 were feasible combinations ($\approx$8% of total scenarios with 50 kWp). In Fig. 6.4, all the feasible solutions are less than (on the left-hand side) the OPF HC. While the unfeasible solutions are on both sides of the OPF HC. The OPF HC is hence the theoretically possible upper limit as there are unfeasible scenarios where the total PV is less than this limit when the location of PV installations changes.

6.4.3 gPC-CC-OPF PV HC assuming Gaussian output

In Fig. 6.5, the OPF HC is compared with the gPC-CC-OPF HC. The gPC-CC-OPF considers that the voltage or current violation probability is not strict, but can be risk based, i.e., can be violated for some scenarios, which is denoted by the stochastic limit ϵ. The gPC-CC-OPF HC of the test network is calculated

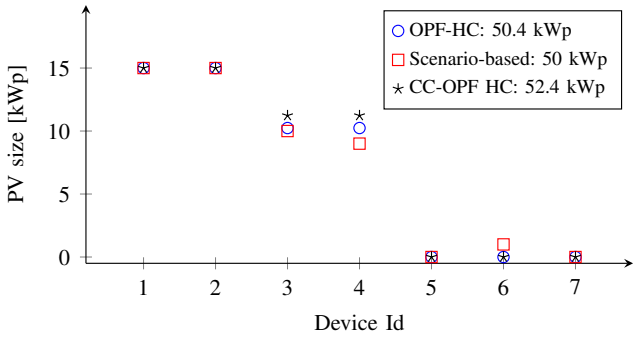


Figure 6.5: PV HC scenarios obtained from OPF and Scenario based method (PV HC for each scenario is given in the legend)

to be 52.4 kWp, 4% more than the result obtained with the deterministic limit. The computation time for gPC-CC-OPF HC was 4.04 sec, almost 16 times that taken by the OPF HC. In Table 6.2, the computation time for all methods is given.

PV size of each consumer is a variable in the above calculation. The advantage is that each PV installed can have different constraints for PV size, for example, all consumers can have equal PV. This means that the proposed CC-OPF HC calculations can be used when there is a limited knowledge of the scenario. The individual limits of the consumer PV size is an input for the method. In Fig. 6.6, four different scenarios of PV size limits are depicted, and the respective gPC-CC-OPF HC are as follows:

- **S1**: The PV installations of consumers are in between 0 to 15 kWp. The corresponding CC-OPF HC is 52.5 kWp.
- **S2**: The PV installations of consumers are equal, and in between 0 to 15 kWp. The corresponding CC-OPF HC of the test feeder is 44.8 kWp.
- **S3**: Consumer 1 and 2 in the test feeder can have PV installations between 0 and 10 kWp, consumer 3 has fixed PV installation of 5 kWp and remaining consumer can have PV installations between 0 and 15 kWp. The corresponding CC-OPF HC of the test feeder is 47.4 kWp.
- **S4**: All consumers have PV installations between 5 and 10 kWp. The corresponding CC-OPF HC is 47.4 kWp.

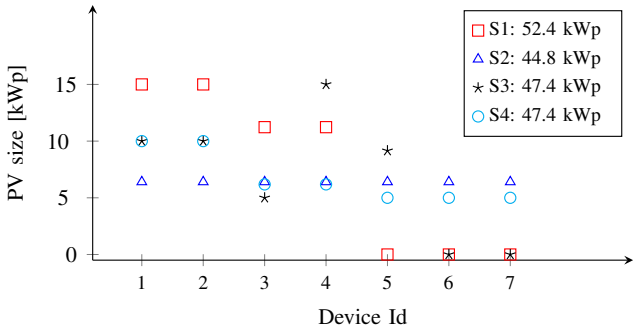


Figure 6.6: PV HC scenarios obtained from gPC-CC-OPF with different PV size limits with HC in the legend

6.4.4 Tuning of moment-based chance-constraints

One of the limitations of gPC-CC-OPF is that the moment based reformulation in (6.18), wherein we assumed the $\lambda_i(\varepsilon^{\mathrm{v}})$ to be based on the Gaussian distribution where its value is 1.65 for $\epsilon^{\mathrm{v}} = 0.05$ [35, 89]. But the input we assume is not Gaussian, and the non-Gaussian uncertainties propagated through non-linear power flow equations resulting in an output which is certainly not Gaussian. In Fig. 6.7, two distributions taken Normal and Beta with the same expectation and variance. However, the Beta distribution has a more skewed tail compared to the Normal distribution. The cyan region corresponds to the area for $\varepsilon^{\mathrm{v}} = 0.05$ for the Normal PDF, which is exact. For beta distribution, the area covered by PDF after the point given by Gaussian reformulation is higher than 0.05 (the space between two PDFs). For moment based reformulation of the such probability distribution, the value of $\lambda_i(\varepsilon^{\mathrm{v}})$ needs to be increased for getting $\varepsilon^{\mathrm{v}} = 0.05$.

An iterative method is proposed to correct the value of $\lambda_i(\varepsilon^{\mathrm{v}})$. In the first run of gPC-CC-OPF, the dual variable is recorded for W_i constraint in (6.22d). The $\lambda_i(\varepsilon^{\mathrm{v}})$ is tuned only for the bus with binding constraint, i.e., the constraints with non-zero dual variable. For the test case, it is bus 5. The voltage is sampled at those buses, and based on the 95% quantile of U_i, new $\lambda_i(\varepsilon^{\mathrm{v}})$ is decided. The second run of gPC-CC-OPF is done after the tuning of moment based formulations. The empirical tuning method is proposed, where a new $\lambda_i(\varepsilon^{\mathrm{v}})$ is assigned to the node with binding chance constraints to reduce the discrepancy created due to assuming the output as Gaussian.

a) If 95% quantile of $\mathsf{U}_i > 1.001U^{\max}$: $\lambda_i(\varepsilon^{\mathrm{v}})= 2.05$

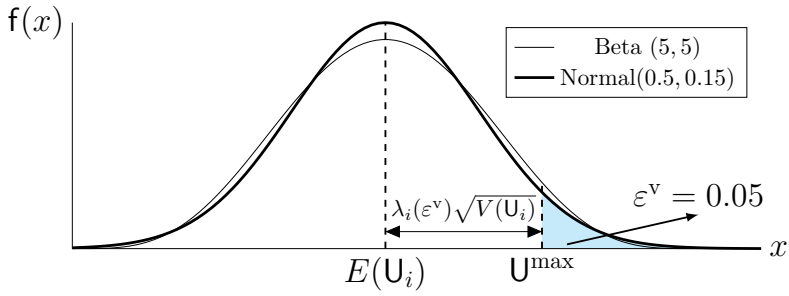


Figure 6.7: Moment based reformulation of stochastic limits. The cyan region is the area for $\varepsilon^{\mathrm{v}} = 0.05$ for Normal PDF

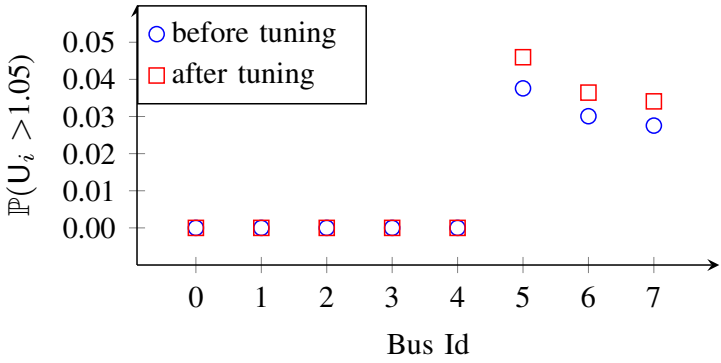


Figure 6.8: Probability of voltage in each node crossing 1.05 pu

b) If 95% quantile of $\mathsf{U}_i > 1.0001U^{\max}$: $\lambda_i(\varepsilon^{\mathrm{v}}) = 1.85$

c) If 95% quantile of $\mathsf{U}_i < 0.99U^{\max}$: $\lambda_i(\varepsilon^{\mathrm{v}}) = 1.05$

d) If 95% quantile of $\mathsf{U}_i < 0.099U^{\max}$: $\lambda_i(\varepsilon^{\mathrm{v}}) = 1.45$

After tuning, the probability $\mathbb{P}(\mathsf{U}_i < U^{\max}) \geq 0.95$ is satisfied. The new computation time is 9.26 sec, which is 2.3 times the original gPC-CC-OPF problem. The gPC-CC-OPF is solved twice, and the remaining time is due to the sampling of the voltage, finding quantiles and tuning λ. The initial and after-tuning probability of overvoltage for each nodes is shown in the Fig. 6.8. The new HC calculated is 52.7 kWp which is 0.38% higher than before tuning. As the name suggests, the tuning is done to improve the accuracy of gPC-CC-OPF HC and is a good-to-have feature to the original solution.

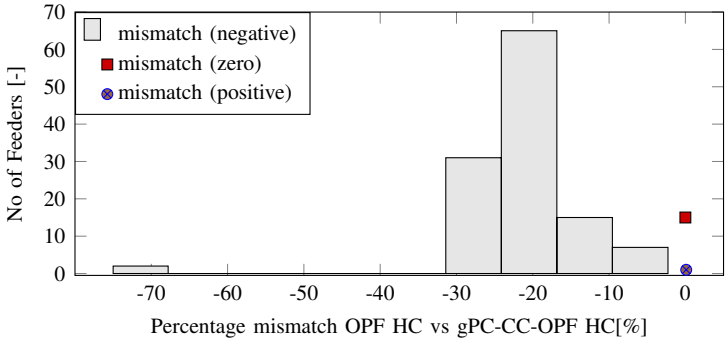


Figure 6.9: Histogram of mismatch between OPF HC and gPC-CC-OPF HC normalized by gPC-CC-OPF HC in percentage. The number of feeders with no mismatch or positive mismatch are plotted as dot.

6.4.5 OPF HC vs gPC-CC-OPF HC in actual network

In this section, the OPF HC is compared with gPC-CC-OPF HC with tuning for feeders presented in [30]. The OPF HC is computed assuming the maximum irradiance in the day and basic load of 0.1 kW for all consumers. The analysis is done feeder-wise instead of the whole network at once. For gPC-CC-OPF, the time-stamp of maximum congestion probability is calculated beforehand using probabilistic power flow when all consumers have 10 kWp PV installed. The single time-period gPC-CC-OPF solutions with tuning is obtained for each feeder with uncertainties in PV generation and consumer load.

In Fig. 6.9, the mismatch percentage of the HC for 136 feeders is shown where,

$$\text{mismatch} = \frac{(\text{OPF HC} - \text{gPC-CC-OPF HC})}{\text{gPC-CC-OPF HC}}. \tag{6.24}$$

The histogram shows that the deterministic OPF HC is below the HC of CC-OPF HC, concretely on 120 feeders. The lower value from OPF HC is attributed to the assumption of a constant load of all consumers and peak irradiance, which is only one snapshot of all possible operational scenarios due to load and irradiance. It shows that the assumption of low load and high irradiance gives us solution in more distributionally robust space (worst-case).

6.5 Conclusion

This paper introduces a SOPF method for stochastic HC calculation using gPC-CC-OPF. This method uses the stochastic load and irradiance input for a peak summer day to get the best PV placement scenario for a low voltage distribution feeders. Stochastic chance constraints in the operational variables, i.e., nodal voltage and branch current are considered. The use of gPC-CC-OPF enables us to solve the problem as a single-shot problem without any linearization, sampling, or relaxation. This method gives the stochastic HC in a few seconds, while solving a similar problem using MC-based methods would take a few days. The method is made openly available as a branch of StochasticPowerModels.jl along with the network data and the uncertainties definitions. The comparison between HC obtained from deterministic OPF and stochastic gPC-CC-OPF in actual LVDS feeders showed that the OPF HC calculated are underestimate of the CC-OPF HC. Finally, an iterative empirical tuning approach for moment-based reformulation of chance constraint is proposed to obtain a correct chance constraint limit when the inputs are not Gaussian distributed, and the distribution of nodal voltage and current are unknown.

For the planning application, such as calculation of the PV HC, a single-time period optimization in single phase equivalent system is sufficient. It is because the PV generation at a particular time period in the same street is usually correlated perfectly. Furthermore, three-phase installations are more congestion-proof compared to single-phase installations for future installations. However, for computing the PV-battery and electric vehicles HC, a multi-time period SOPF is recommended. Similarly, if the interest is in obtaining single-phase PV HC, an unbalanced SOPF is recommended. Hence, an unbalanced, multi-period SOPF is identified as future work.

Chapter 7

Scaling Hosting Capacity Using Representative Feeders

Koirala, A., Hashmi, Md. U., D'hulst, R., Van Hertem, D., *Representative feeders for spatial scaling of stochastic PV hosting capacity.* Proceedings of ISGT Europe, 2022

The first author is the main author of the article. The contributions of the first author include the conceptualization, methodology, visualization, and the writing of the article.

Abstract: The photovoltaics (PV) hosting capacity (HC) of the power system infrastructure is an important planning problem. The energy transition is happening now, resulting in the addition of new load types and generation in the low voltage distribution network. With these new loads and generation devices connected to the distribution network, a new planning approach to accomodate their stochastic nature is required. Computing the stochastic HC for individual low voltage distribution feeder is challenging, as even a small service area can have hundreds of feeders. This work aims to capture appropriate clustering schemes and the most relevant features of low voltage distribution feeders so that an accurate estimate of the hosting capacity of the full service area can be calculated by scaling up the hosting capacity of a small number of representative feeders. A case study using actual feeders from suburban Spain showed that representative feeders obtained from feature selection and using

appropriate cluster size enables to scale the stochastic PV HC of an extensive service area. The case study showed that using 3% of the total feeders enables to estimate the PV HC of the LVDS feeders in the large service area.

Keywords: distribution system, renewable hosting capacity, spatial scaling, representative feeders, low voltage, photovoltaics.

7.1 Introduction and Motivation

The hosting capacity (HC) of an Low Voltage Distribution System (LVDS) refers to "the amount of new generation or consumption that can be accommodated on a given feeder without impacting system operation under existing control and infrastructure configuration" [17]. In [182], it was seen that finding the LVDS hosting capacity for new photovoltaics (PV) is a multidimensional stochastic problem due to many uncertainties, which can be classified into three types: a) power generation and load consumption uncertainties, b) PV scenario uncertainties, i.e., size, connection phase and type of the installations, and c) feeder uncertainties, i.e., feeder type and feeder size. The first type of uncertainty can be regarded as an operational uncertainty that depends on weather and consumer behavior. The second is a planning level uncertainty, dealing with the size, type and location of PV installed. The third uncertainty emerges from the features of the feeder involved. The distribution network is the final capillary of the power grid. They are a significant part of the power network by length and spread. A small suburban town may have more than 100 such low voltage (LV) feeders [30], while the total number of LV feeders operated by a single operator can go up by tens of thousands. However, the variability in the feeder's features is usually neglected, and the HC is calculated only for a few selected feeders [182].

The easiest but computationally expensive solution for getting the total HC of a network, considering the uncertainty of the feeder features, is through individual analysis of each feeder. This feeder-wise calculation is an easy solution if the HC calculation method is deterministic or the service area is small. However, it might not be feasible if computationally expensive stochastic HC calculation methods [182] are used or the service area is quite large [21]. The limited availability of LVDS feeder data and load data also limits the feederwise analysis of the complete network. The alternative studied here is to estimate the HC of the whole network based on a smaller set of representative feeders [77].

PNNL has published 24 synthetic representative feeders [194] based on the statistical features of medium voltage (MV) grids. Similarly, EPRI has published six representative MV feeders sanitized for scientific use [195]. In [196], 9 MV

and 8 LV prototype feeders are presented to represent Australian feeders after the analysis of 204 members of the MV and 8858 members of the LV database. In [77], a set of 383 feeders from England were analyzed and clustered to get 11 representative feeders. One of such representative feeders was also published as the IEEE European LV test feeder [49]. In [77], the usefulness of the representative feeders was shown to identify the group of feeders which can host "100% PV penetration". It was shown that if a representative feeder has 100% PV penetration as HC capacity, then all the feeders in the subset represented by that feeder also have 100% PV HC. In [21], it was shown that the PV HC trend of an LVDS system with 50 000 feeders can be represented by a set of 2% of randomly chosen feeders. In [197], representative feeders for a set of 24,000 Austrian LVDS feeders were obtained to identify the limiting constraints of the group represented by those feeders, i.e., thermal or voltage, for a high PV penetration. These results were for simplified deterministic PV HC calculations with no other uncertainties considered [77], only the location of PV was considered as uncertain [21], or only the limiting constraint is observed [197]. So far, the existing studies do not consider scaling stochastic PV HC using representative feeders. Furthermore, there are no studies regarding the selection of features of low voltage feeders to reduce the number of representative feeders required to scale stochastic PV HC.

The main contributions of this paper are as follows:

a) a methodology is proposed to obtain the appropriate number of the representative feeders, whose stochastic PV HC can be used to estimate the cumulative PV HC of the network,

b) a method to identify the most relevant features of LVDS feeders for stochastic PV HC scaling is proposed (Section 7.2),

c) a case study is demonstrated for a) and b) using the set of European LV feeders from [30] (Section 7.3).

7.2 Methodology

7.2.1 PV Hosting Capacity

The main objective of this paper is to capture a small set of representative feeders for a given LVDS so that an accurate estimate of the PV HC of the full service area can be calculated by scaling up the HC of the representative feeders. The representative feeders are obtained through clustering schemes based on

the features of LVDS feeders. The PV HC calculation method used in this paper is the decoupled PV HC method proposed in [188]. This stochastic HC method splits the planning level uncertainties from the operational uncertainties (Fig. 5.1b), and the best planning scenario out of the considered scenarios is considered to be the PV HC of the feeder. Defining the PV HC itself is a well-debated topic [182], and not in the scope of this paper. Hence, a detailed discussion on the methodology to calculate the PV HC is skipped in this paper and interested readers are referred to [182] and [188].

Large scale LVDS can have hundreds of small feeders whose PV HC needs to be calculated independently. Calculating the HC is a computationally intensive process and can substantially be reduced by using representative feeders generated from the relevant features of the feeders in the network.

7.2.2 Representative Feeders

The first step for obtaining representative feeders is to create clusters of similar feeders, and crucial to this is how the *distance* between feeders is defined and which features of the feeders are taken into account to define this distance. In this paper, the following features are available for all feeders:

(1) number of consumers connected to a feeder,

(2) yearly active power consumption per customer in kWh,

(3) yearly reactive power consumption per customer in kVArh,

(4) total conductor length in km,

(5) main path length (longest path) in km,

(6) average length to the customer in km,

(7) total line impedance in Ohm, and

(8) average path impedance in Ohm.

The steps involved in finding appropriate representative feeders suitable for scaling PV HC can be summarized as follows:

1. Identify clusters of closely related LVDS feeders.
2. Choose one representative feeder per cluster.
3. Calculate stochastic PV HC of representative feeders.

4. Scale up the HC of representative feeders to the full set LVDS feeders.

The first step of the process is clustering. There are several clustering methods, of which k-means is most often used for LVDS feeder classification [77, 197, 198]. For k-means, an unsupervised clustering algorithm, the number of clusters and hence the number of representative feeders is an input.

The next step is then to find the representative feeder of each cluster based on statistical parameter, e.g., the feeder closest to the centroid. A parameter $x_{i,\psi}$ is in standard form when

$$x_{i,\psi} = \frac{\gamma_{i,\psi} - \bar{\gamma}_k}{\text{std}(\gamma_\psi)}, \tag{7.1}$$

where $\gamma_{i,\psi}$ is the measurement value of the feature ψ for feeder i, $\bar{\gamma}_\psi$ denotes the mean value of that feature within the cluster, and std(γ) denotes the standard deviation of the feature within the cluster. Consider, $x_k^{\min}$ and $x_k^{\max}$ denoting the minimum and maximum value of $x_{i,\psi}$ in a cluster. We define the normalized distance of a given feature and a feeder in that cluster as

$$P_{i,\psi} = \frac{x_{i,\psi} - x_\psi^{\min}}{x_\psi^{\max}}. \tag{7.2}$$

The representative feeders are selected by identifying ζ_{rep} given as (7.3) for all clusters.

$$\zeta_{\text{rep}} = \min_i \sum_\psi |P_{i,\psi}|. \tag{7.3}$$

The unknowns in these steps are: a) how many clusters (and representative feeders) do we need to get the scaled up PV HC as close to the fully calculated one? and b) which features are more relevant for PV HC scaling? This paper presents a case study in European LVDS feeders to answer these questions.

7.2.2.1 Appropriate number of representative feeders

The number of clusters required to represent the overall network is unclear. One possibility to define the required number of clusters is through verifying the goodness of the clustering measured using the *silhouette coefficient*. The silhouette value of a member i in a cluster is the confidence indicator of its association in the cluster [199] and is defined as:

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}, \tag{7.4}$$

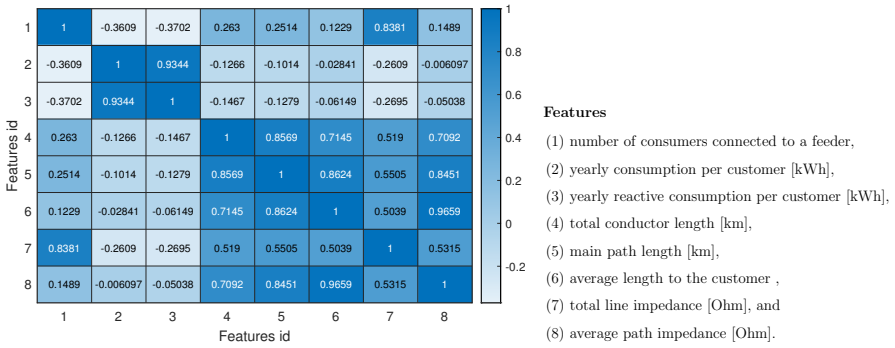


Figure 7.1: Covariance matrix of normalized to standard form used for feature correlationship.

where $a(i)$ is the mean distance between i and all other data points in the same cluster and $b(i)$ is the smallest mean distance of i to all points in any other cluster. Rousseeuw in [200] proposed an interpretation based on the maximum value of the mean $s(i)$ per cluster across the whole dataset termed as *silhoutte coefficient*: where 0.71 to 1.0 implies a strong structure. However, there is no guarantee that the clusters based on maximization of this coefficient will correctly represent the similarity in PV HC. Different cluster sizes are investigated to see if the *silhouette coefficient* based cluster size is an appropriate choice for getting the best representative feeder for *scaling* LVDS PV HC.

7.2.2.2 Choosing appropriate features

The covariance matrix of the eight features in standard form is shown in the heatmap shown in Figure 7.1. Note the high correlationship between (i) active and reactive power, (ii) total number of consumers and total line impedance, and (iii) features (4), (5), (6), and (8). This covariance matrix can be used to select the features taken for clustering. However, it does not say anything about if the feature selection improves the scaling of the PV HC. As seen in [77], the mapping of the PV HC may not be straightforward in all clusters. The clustering process is repeated with a reduced number of features and different combinations of features to fine-tune the result. The aim is to find the relevant set of features of LV feeders, which will assist the *scaling* of the stochastic PV HC.

7.2.3 Scaling PV HC

Once a set of representative feeders is obtained using the features and defined number of clusters, the PV HC of those feeders is compared with the cumulative PV HC of the respective cluster. The term *scaling* here refers to the use of representative feeders' PV HC to estimate the PV HC of bigger service area in distribution network. The stochastic decoupled PV HC [188] is used. Herein, two terminologies used in the remaining paper are introduced: *total* HC, meaning the sum of PV size that can be installed in the feeder and *normalized* HC, meaning the average PV size per consumer. Both total and normalized HC is compared while scaling using the representative feeders.

For this comparison, two indices are proposed:

Mean average absolute error (MAAE): In this index the absolute difference between HC of representative feeder (HC_{κ}^{rep}) and the mean HC of all feeders in the cluster κ ($\bar{HC}_{\kappa}$) is summed for all clusters and normalized by total number of clusters K:

$$\text{MAAE} = \sum_{\kappa} \frac{|HC_{\kappa}^{\text{rep}} - \bar{HC}_{\kappa}|}{K} \tag{7.5}$$

Difference from the representative feeder (DR): In this index, difference of HC of each feeder i in cluster c from their representative feeder is measured:

$$\text{DR}_{(i,\kappa)} = HC_{\kappa}^{\text{rep}} - HC_{(i,\kappa)} \tag{7.6}$$

The output will be in the form of distribution of error for a particular clustering scenario from which mean (DR_{μ}), and variance (DR_{σ}) of the difference can be obtained.

In the case-study following, increasing or decreasing the number of clusters and the use of different combination of features is done to observe the error on scaling using indices MAAE and DR. The objective of this study are:

a) can the representative feeders using all features be used to obtain the cumulative PV HC of an extensive distribution service area?

b) if not, what are the most relevant features of LVDS when scaling the stochastic PV HC spatially for a large service area?

c) what is the influence of increasing or decreasing the number of clusters in scaling the PV HC?

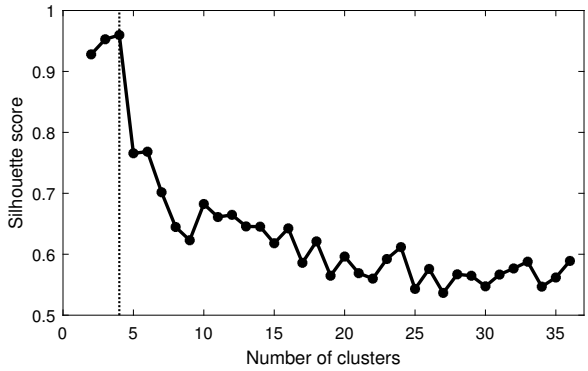


Figure 7.2: Silhouette coefficient for different sized feeder clusters for 160 Spanish distribution network feeders

7.3 Case Study

In this section, feeders from an actual European LVDS [30] are used to show the case study on *scaling* the stochastic PV HC to a large service area using representative feeders. The *silhouette coefficient* for varying numbers of the clusters is shown in Figure 7.2. It can be observed that the *silhouette coefficient* is maximum and exceeds 0.95 when four clusters are formed. The four representative feeders identified in this clustering scheme are shown in Table 7.1. Interestingly, the *silhouette coefficient* is above 0.51 for the cluster numbers above four, which means that increased cluster numbers still have a reasonable similarity between the members of the clusters. The *scaling* of stochastic PV HC using the cluster sizes on both sides of the one with maximum *silhouette coefficient* is investigated to check the significance of this coefficient in such a *scaling*. The aim is to obtain an appropriate number of representative feeders (or cluster numbers) for a given error threshold in *scaling*.

The first part of this section will deal with the impact of increasing the number of representative feeders on the scaling of HC. Secondly, the impact of scaling based on different features is shown. Finally, the impact of both feature selection and increasing the number of representative feeders is shown. All the experiments are repeated for 100 times to get a mean value to reduce the error due to uncertainty in cluster centroids.

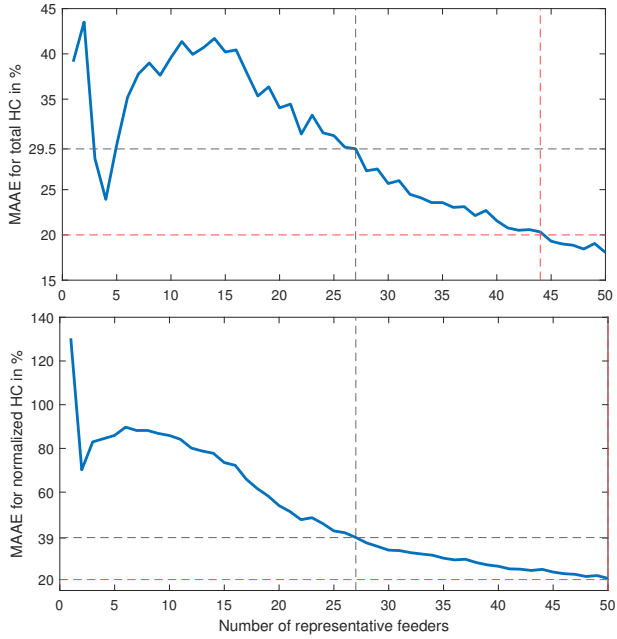


Figure 7.3: MAAE of stochastic PV HC scaling from representative feeders using all features

Table 7.1: Representative feeders selected

		Cluster 1	Cluster 2	Cluster 3	Cluster 4
	Feeder ID	144	52	27	23
(1)	No. of consumer	44	20	2	0
(2)	Avg active energy/year [kWh]	2068.79	4256.73	5424.98	0
(3)	Avg reactive energy/year [kVArh]	202.84	415.14	484.25	0
(4)	Total line length [km]	0.38	0.15	0.31	0.25
(5)	Main path [km]	0.21	0.13	0.29	0.22
(6)	Avg length to costumer [km]	0.15	0.13	0.16	0.0
(7)	Total line impedence [ohm]	4.60	2.52	0.30	0.67
(8)	Average path impedence [ohm]	0.10	0.13	0.15	0
	Indices	0.17	1.19	1.65	1.62

7.3.1 Role of number of representative feeders in scaling

Based on the maximum *silhouette coefficient*, the best suited number of clusters is chosen to be four. This clustering considers all eight features described in Section 7.2 (Fig. 7.2). On close observation, cluster 1 has the majority of residential feeders. In contrast, cluster 4 contains all the special feeders: such as feeders without any load on them or serving consumers like hospitals or sport-centres. For further calculations, the feeders of cluster 4 were removed as their PV HC was very high or zero, creating a bias in the residential PV HC scaling. The 136 feeders in clusters 1 to 3 were used for scaling PV HC.

In Fig 7.3, it is seen that the MAAE of the total HC of the network while using the representative feeders based on maximum *silhouette coefficient* is 24%, when all features are used. The MAAE increases when increasing the number of representative feeders till some point and then starts decreasing. The MAAE of the normalized HC is 80% for the clustering based on maximum *silhouette coefficient*. Similar to the total HC, it increases when increasing the number of representative feeders first and then decreases. This inconsistent behavior means we need to specify a threshold for MAAE of the total HC and normalized HC to get the required number of representative feeders. For a threshold of 20%, the number of needed representative feeders is 44 and 50, respectively. If the maximum of these two is taken, 50 representative feeders ($\approx$35% of total feeders) are required with an error threshold of 20% MAAE.

Alternatively, the MAAE is observed for a fixed number of representative feeders. In this case, 20% of the total feeders were selected, i.e., 27 representative feeders. The MAAE is 29.5% and 39% for the total and normalized HC, respectively. Hence, the conclusion obtained from this study is mixed. Yes, using representative feeders can estimate the PV HC of a larger network. However, a better estimation is obtained only by increasing the number of representative feeders. Furthermore, using all features to obtain representative feeders is not efficient enough, as we need a detailed analysis of 35% of feeders for a MAAE of less than 20%.

7.3.2 Feature Selection

The role of different features in *scaling* the stochastic PV HC is studied in this section. The correlation analysis in Fig. 7.1 shows the relation among the features, namely features (1) and (4), features (2) and (3), and features (4), (5), (6), and (8) are correlated but it is not straight forward how to use this correlation PV HC *scaling*. Three clustering scenarios are considered to get a better understanding of these features in PV HC *scaling*:

a) C1: Clustering using all features,

b) C2: Clustering using the individual features only, and

c) C3: Clustering using a combination of selected features.

A rank-based method is proposed to find better features for PV HC scaling. Eight different indices based on MAAE and DR are calculated:

a) **I1**: the number of representative feeders when an error threshold of 20% is fixed for the MAAE of total HC,

b) **I2**: the number of representative feeders when an error threshold of 20% is fixed for the MAAE of normalized HC

c) **I3**:MAAE of total HC when 20% of feeders are taken as representative feeders,

d) **I4**: MAAE of normalized HC when 20% of feeders are taken as representative feeders,

e) **I5**: absolute value of DR_{μ} of total HC when 20% of feeders are taken as representative feeders,

f) **I6**: absolute value of DR_{μ} of normalized HC when 20% of feeders are taken as representative feeders,

g) **I7**: DR_{σ} of total HC when 20% of feeders are taken as representative feeders,

Table 7.2: Rank of different indices when taking representative feeders, considering all features (C1) and individual features (C2)

Features	**I1**	**I2**	**I3**	**I4**	**I5**	**I6**	**I7**	**I8**	Rank sum
All	9	9	9	9	9	9	6	3	63
(1)	2	4	2	3	4	5	7	2	**29**
(2)	7	2	8	2	8	6	8	8	49
(3)	2	2	6	7	6	8	9	9	49
(4)	2	5	3	5	1	2	4	5	**27**
(5)	2	7	5	4	5	1	3	4	31
(6)	2	1	1	8	2	3	1	7	**25**
(7)	7	6	4	1	3	7	5	1	34
(8)	1	8	7	6	7	4	2	6	41

Table 7.3: Rank of different indices when taking representative feeders considering all features (C1), individual features (C2), and combinations of features (C3)

Features	**I1**	**I2**	**I3**	**I4**	**I5**	**I6**	**I7**	**I8**	Rank sum
All	13	13	13	12	13	11	10	5	90
(1)	3	6	2	4	6	6	11	3	41
(2)	8	2	10	3	10	7	12	12	64
(3)	3	2	8	9	8	9	13	13	65
(4)	3	7	3	7	3	3	8	8	42
(5)	3	9	7	5	7	2	5	6	44
(6)	3	1	1	10	4	4	3	11	**37**
(7)	8	8	5	1	5	8	9	1	45
(8)	1	10	9	8	9	5	4	10	56
(4), (6)	1	2	4	6	1	10	2	9	**35**
(1), (4)	11	11	12	11	11	12	6	4	78
(1), (6)	12	12	11	13	12	13	1	7	81
(1), (4), (6)	10	5	6	2	2	1	7	2	**35**

h) **I8**: DR_σ of normalized HC when 20% of feeders are taken as representative feeders,

At first, cases C1 and C2 are taken, and the features with the first three ranks are taken to form C3. Finally, a new ranking is done, taking C1, C2 and C3 together. The rank matrix for when taking only C1 and C2 is shown in Table 7.2, from which features (1), (4) and (6) had the top three sums of the rank. The C3 scenarios are built from the combination of (1), (4) and (6), and a new rank table is formed using all C1, C2, and C3.

Table 7.4: Comparison after feature selection

Features	no of rep. feeders MAAE ≤20%					MAAE for 20% total feeder			
	total HC	norm. HC	max HC	% of total feeder	% reduction	total HC	norm. HC	max HC	% reduction
All	44	50	50	37%	-	29.50	39.20	39.20	-
(6)	3	3	3	2.2%	94%	2.20	14.90	14.90	62%
(4),(6)	2	4	4	3%	92%	3.98	9.21	9.21	76.5%
(1), (4), (6)	7	5	7	5.22%	86%	5.20	5.29	5.29	86.5%

From Table 7.3, it is seen that feature (6) and the combinations (4),(6) and (1), (4), (6) had better performance based on the ranking of all eight indices.

We can reasonably conclude that taking all features with equal weight would not necessarily lead to better representative feeders. However, the features such as the number of consumers connected to a feeder (feature 1), total conductor length in km (feature 4), and average length to the consumer (feature 6) have a lower error on scaling PV HC when taken individually. Similarly, the combination of these three features had a better ranking among all other tested combinations.

7.3.3 Impact of feature selection

In this section, the three top ranking feature combination from the previous study is evaluated to compare the improvement from the first case study with all features in terms of MAAE. The two factors considered were the reduction in MAAE on taking 20% of total feeders as representative feeders and the number of representative feeders required for a threshold of 20% of MAAE of both the total and normalized HC.

From Table 7.4 and Fig. 7.4, two main conclusions can be drawn. First, the feature selection reduces the number of representative feeders required by 94% if the threshold of MAAE is fixed to 20%. On the other hand, if 20% of feeders were taken as representative feeders, the MAAE threshold for both normalized and total HC reduces to 5.29% compared to 39.2% when using all features (86.5% reduction).

Representative feeders based on average length to the consumer (Feature 6) have better PV HC scaling results alone and combined with total conductor length and the number of consumers connected (Features 1, 4 and 6). However, as seen in Fig. 7.4, the combination of Features 1 and 4 only does not lead to better scaling of PV HC. It can be fairly concluded that using the right features for clustering the LVDS feeders means 3% of the total feeders can be used as representative feeders to scale the HC of the whole network while having 20% MAAE.

7.4 Conclusion

The conclusion of the stochastic PV HC scaling exercise in this paper is two-fold. The first one is that representative feeders with specific cluster size and features can be used to scale the HC of the representative feeder to get a cumulative HC

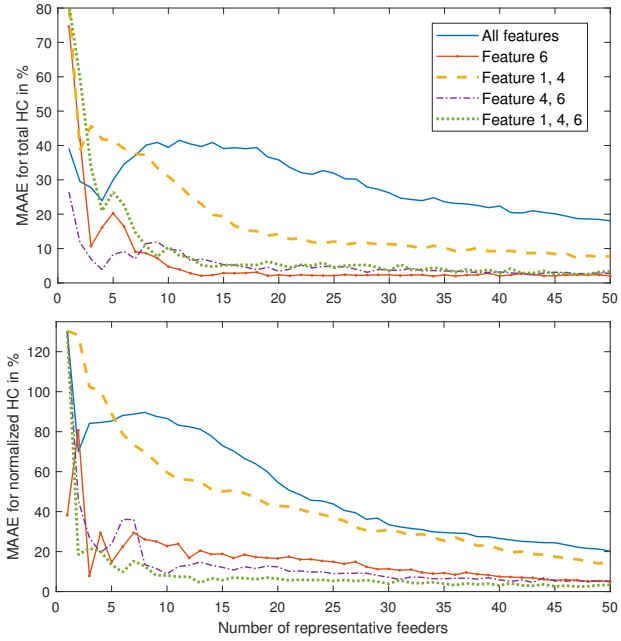


Figure 7.4: MAAE of stochastic PV HC from representative feeders after feature selection

of a larger service area, thereby reducing the need to calculate PV HC of all feeders. Proper feature selection can reduce the error associated with scaling or the number of representative feeders required to scale the PV HC of the LVDS. Concretely, the case study showed that a selection of only 3% of feeders could estimate the PV HC of the whole service area with 20% of error. On the other hand, this type of estimation neglects all the variation in different feeders, meaning the HC of any random feeder in a cluster cannot be extracted precisely. The only solution in such a case is to have HC of all feeders independently.

For the test LVDS, it was seen that maximum *silhouette coefficient* based clustering was not good enough to scale the PV HC. However, it could be used to separate the special feeders. On increasing the number of clusters (also the representative feeders), the MAAE starts saturating at one level, which can be used to select the appropriate number of the representative feeders. Similarly, for the feature selection, it was seen that using all features does not necessarily give better representative feeders. In contrast, the representative feeders based on individual features had better performance on scaling PV HC. The salient features identified are the number of consumers connected to the feeder, total

conductor length, and average length to the consumer. The selection of salient features reduces the error in scaling by 86.5% in the test case and the number of representative feeders required by 94% for the 20% threshold on MAAE.

Chapter 8

Conclusion and Future work

8.1 Summary and conclusions

The Low Voltage Distribution System (LVDS) is the part of the power system where the new distributed renewable energy resources (RES), i.e., photovoltaics (PV), and loads such as electric vehicles (EV) and heat pumps (HP) supporting the energy transition are connected. The traditional role of such an LVDS was to supply power generated by large power plants to passive consumers. The usual planning approach to such LVDS is "fit-and-forget", i.e., installing a substantial size cable assuming the worst case. However, with the revolution in PV, EV and HP installations, the loading of the grid increases as well as the uncertainties in LVDS. The traditional passive role of consumers and the standard approach to LVDS planning for the worst-case scenario is no longer efficient. Such planning results in extra investment in grid infrastructure to prepare the grid for rare events, the cost of which will eventually be reflected in the consumer's bill. New risk-based stochastic approaches for planning and active operation of LVDS are deemed essential to manage and design future LVDS. One of the critical measures is the hosting capacity (HC) of the LVDS to new PV devices. The many new uncertainties make this problem of computing HC computationally hard to solve. This thesis aims at understanding the stochastic approaches for LVDS, improving them and scale to a regional level. The focus is on the PV HC calculation of European LVDS.

8.1.1 LVDS PV HC

In Chapter 3, existing deterministic and stochastic PV HC methods are evaluated and benchmarked using the "Non-synthetic European low voltage (LV) test feeders" with actual network and smart meter data. A stochastic reference method is proposed for consistent benchmarking. The comparison of the PV HC obtained from different existing methods showed that the spread of HC obtained was huge. The factors mostly affecting the PV HC calculated are: the assumption on the *grid* and *stochastic limits*, size and number of the PV installations, and the granularity of PV deployment scenario. A sensitivity analysis of HC with varying *stochastic* and *grid limits* shows that small changes in those limits could give completely different HC for the same feeder under the same assumptions. Standardizing those limits is one major step in defining the actual HC of LVDS feeders. For HC calculation *stochastic limit* of 5% and voltage limit of 0.95-1.05 pu are recommended.

In Chapter 5, decoupled LVDS PV HC method is presented. The two types of uncertainties in LVDS: planning (i.e., size of PV, location of PV, etc.) and operational (i.e., load and PV generations) are decoupled to obtain the best planning scenario (i.e., HC) for given operational uncertainties. The planning scenario with the highest PV installation within the *stochastic limits* was the PV HC. Illustration on how the PV HC of a network increases with change in operational variables through PV curtailment is presented through this decoupled method. An impact study shows that three-phase PV installations are more resilient to congestion than single-phase installations. Another case study illustrates how the PV HC of an LVDS feeder increases if the maximum voltage violation is allowed for 25% of the time, concretely by three times in the test feeder considered. The HC method presented in this chapter opens an avenue to decouple planning level decision problems from operational decision problems.

In Chapter 6, the conclusions from Chapters 3 and 5 are used to propose stochastic optimization based PV HC calculation. The iterative PV deployment schemes in Chapters 3 and 5 are replaced by the constraints in the PV installations size, and the optimal planning scenario is obtained for stochastic constraints or chance-constraints (CC) in the nodal voltages and line currents. Comparison with scenario based methods shows that this method gives an upper bound of the PV HC capacity. The final stochastic PV HC is higher than HC calculated using simple deterministic Optimal Power Flow (OPF), with a mismatch of 20% in an average.

From this thesis, it was seen that using stochastic calculation tools gives more PV HC compared to deterministic methods. It can be fairly concluded that

using such a stochastic method in other planning applications helps to utilize the existing LVDS network more efficiently. In this work, only one planning problem is investigated in detail. However, the findings are the motivations to apply similar stochastic methods for efficient investment planning, reducing the cost of the energy transition.

8.1.2 Calculation tools

In this thesis, two types of stochastic calculation tools for LVDS are evaluated: Probabilistic Power Flow (PPF) and Stochastic Optimal Power Flow (SOPF).

Chapter 4 evaluates three PPF tools for an LVDS network, i.e., Monte Carlo (MC), quasi Monte Carlo (qMC) and general Polynomial Chaos (gPC). Using qMC methods improve the estimate of Quantities of Interest (QoIs) compared to MC but at the cost of losing the statistical properties. However, the gPC-based PPF outweighs MC and qMC methods clearly in terms of computational effort and accuracy in an LVDS case study. These advantages are more significant in feeders with more sources of uncertainty. For gPC based methods, the accuracy and computational time depend upon the choice of the testing point vector and the polynomial degree. The degree two gPC method based on Sobol sequences is recommended as the best choice with the same accuracy and ten times less computation time compared to MC methods. The evaluations are done for both mean and tail events for a real LVDS with demand and PV generation uncertainties represented by continuous univariate distributions. Decoupled LVDS PV HC method in Chapter 5 utilizes the gPC based PPF for calculation of the probability of congestions for each planning scenario. The method itself is a brute-force approach but using gPC based PPF reduces the computational time by at least ten times compared to MC method.

Chapter 6 presents a method for SOPF, i.e., gPC-CC-OPF, using an intrusive gPC for stochastic PV HC calculations. This method uses the stochastic load and irradiance input to get the best PV deployment scenario for LVDS feeders. Stochastic chance constraints in the operational variables, i.e., nodal voltage and branch current, are considered. The proposed SOPF enables us to solve the stochastic HC calculation as a single-shot problem without any linearization, sampling, or relaxation. This method gives the stochastic HC in a few seconds while solving a similar problem using MC-based methods would take a few days. The method is available publicly as a branch of StochasticPowerModels.jl along with network data and the uncertainties definitions. An iterative empirical tuning approach for the moment-based reformulation of chance constraints is proposed to obtain a correct chance constraint limit when the inputs are not Gaussian and the distributions of nodal voltage and current are unknown. This

SOPF tool can be easily modified to solve any other stochastic optimization problems.

In a nutshell, this thesis shows that SOPF could be conveniently used in the calculation of PV HC calculations. However, the publicly available modular SOPF framework can be utilised for other stochastic optimal investment planning with few modifications on the individual constraints and variables. The framework is analogous to PowerModels.jl for deterministic OPF and allows researchers to try different formulations and relaxations for SOPF.

8.1.3 Scaling of PV HC

Chapter 7 deals with scaling the PV HC to a large service area based on representative feeders. The representative feeders can be used to scale the HC of the representative feeder to get a cumulative HC of a larger service area. Features must be carefully selected to obtain feeders representative to scale the PV HC. Using only 3% of the feeders resulted in a scaled PV HC that approximates the service area HC with an error of 20%. The salient features that can be used for PV HC scaling are the number of consumers connected to the feeder, total conductor length, and average length to the consumer. On the other hand, this estimation neglects all the variations in different feeders, meaning that the HC of any random feeder in a cluster cannot be extracted precisely. The only solution in such a case is to have the HC of all feeders calculated independently.

8.2 Recommendations and Future work

Critical reflections on the work are presented in this section, with recommendations and identification of possible future lines of work.

8.2.1 LVDS uncertainties

This work uses a Beta distribution based on limited historical data to define the load and PV generation uncertainties. Presence of storage devices is not considered in this work. It might only be accurate for some consumers. The conclusion presented is based on the assumption that there will be more grid-connected PV in the future. However, if there are more PV-battery systems, the injection and offtake from the grid might be different. Although the optimization framework works fine when the uncertainties are well-defined, a

better understanding of these uncertain phenomena would contribute to a better design of the future grid. Furthermore, there are uncertainties regarding how technology evolves and how the behavior of consumers changes as they become more and more aware of the environmental impact of their consumption. As seen from the last two years, increasing energy prices might be another factor inducing drastic behavioral changes. These uncertainties are not modelled in this work. Furthermore, this work does not explore the uncertainties in the load due to added HP and EV. The specific challenge in modeling these loads is the introduction of new uncertainty, such as state of charge and mode of operation, which are discrete and time-dependent. This is an added complexity in the modeling.

In this line, the recommendation is to look at more extensive modelling of PV generation and consumer load. Stochastic modelling of future load based on consumer behavior, technology innovation, or price is an interesting area for future research.

8.2.2 LVDS power quality limits

The *grid limits* and *stochastic limits* for LVDS power quality are not in line with existing measurements. In Europe, EN50160 is taken as the standard for LVDS power quality. However, using them in a real application in LVDS is complicated due to the limitation of the measurement equipment. For e.g., the voltage quality of LVDS is defined as *the 10-minute average voltage should be within ±10% of base for 95% of the time.* However, smart meters (SMs) in Europe usually measure 15-minute averages. In this thesis, we "compensate" the impact of 15 minutes average measurement by reducing the voltage limits to ±5%. The stochastic limit on congestion is something still undefined by any available standards. The standard mentions that defined quality should be maintained for 95% of the time, which is not equivalent to what is an acceptable probability of congestion at any given time period. Furthermore, there are indices such as voltage unbalance factor, which cannot be measured with a SM without angular measurement possibility. In this work, the *grid limits* and *stochastic limits* are based on a simple sensitivity analysis concerning the HC. As seen in Chapter 3, these limits can be manipulated to get the desired planning scenario and investment requirements.

In this line, more research is required to define more practical limits of the power quality, which can be measured using the existing SMs. Defining a comparable industrial standard for LVDS power quality limits for stochastic planning and operations is identified as an important future work.

8.2.3 Hosting capacity definition and calculations

This thesis focused on understanding the PV hosting capacity of LVDS considering a passive deployment, i.e., without any active control of the PV inverter, deployment of PV with a battery system, or installation of smaller inverters. Hence, the HC is proposed to be in kWp of the PV installed. However, if an active control policy without a battery system is to be studied, minor modifications on PV generations would be enough to obtain a new hosting capacity in terms of apparent power (kVA) or active power (kW). For e.g., when the inverter size is 70% of the PV installations, PV HC is the sum of the active power ratings (in kW) of these inverters, but requires redefining uncertainty in PV generation. Additionally, for the PV battery systems, the HC calculation requires the introduction of new uncertain variables, such as the state of charge and the mode of operation (charging or discharging). This addition will also be useful for calculating EV HC.

Other planning problems, such as the location of new PV and the optimal way of installing shared PV or community PV can be studied using the proposed HC method with some improvements. The improvements could be minor, involving the change of the limits of some variables, or major involving the addition of new variables. Similarly, the redefinition of the input operational uncertainties is required for other cases, such as studying the impact of PV generation curtailments.

In Chapter 6, three-phase PV HC is computed using a balanced SOPF. However, for calculating the single-phase PV HC, an unbalanced SOPF is required and is identified as a future work.

8.2.4 Stochastic optimal power flow

This work presents a single phase equivalent, balanced SOPF for PV HC calculations using intrusive Polynomial Chaos Expansion (PCE) as a proof-of-concept (See Chapter 6). The natural extension of this work will be towards a three-phase unbalanced SOPF. The proposed Rectangular Current-Voltage (IVR) formulation in the format of multi-network `PowerModels.jl` has a computational gain compared to Rectangular Power-Voltage (ACR). The extension of this work should be towards new methods and strategies to reduce the computational burden, e.g., fixing the coefficients which are zero in degree-one PCE calculations while having a higher degree PCE. Furthermore, the current implementation considers decision variables of SOPF to be continuous. The SOPF formulations for discrete decision variables are identified as future work.

In this work, the QoIs of SOPF are assumed to be Gaussian while transforming them to moment based realization. To reduce the error, an empirical tuning method is proposed. However, this tuning mechanism is not perfect, though it reduces the error introduced by the assumption. Finding a better way to tune a moment based realization is an interesting line of work for improving this method.

8.2.5 Scaling of PV HC

A small network from Spain validates this paper's proposed method to scale PV HC. The result might differ from country to country due to the difference in their grid investment strategy. It will also depend upon the data availability, i.e., if all the features discussed are available in the new network. Future research should be towards validating the result in a larger network.

In this work, we aim to find the aggregated PV HC. However, if the requirement is to identify problematic feeders, or problematic points in a feeder, another approach might be required. Investigation to find such a tool is identified as an interesting future research direction.

8.2.6 Open LVDS data

Open data in LVDS is a barrier in research activities. The load and network data are scarce and not representative. Unlike transmission networks, where there are open libraries such as pglib-opf [201], distribution networks are rarely available for research use. Most of the available data are merely for the purpose of validating power flow and optimal power flow tools and do not reflect the real operational conditions. Furthermore, data on consumer load and EV charging are challenging to obtain even for research work due to privacy concerns. Future work will be towards releasing more open data suitable for research in solving planning and operational issues of LVDS.

8.2.7 LVDS modelling

The LVDS model used in this work is dependent upon the data provided by the DSO. However, those data might have uncertainties in topology, impedance value and consumer location. These parameters of LVDS can be improved if more measurement devices such as SMs are installed. Techniques to verify the data of DSO with the help of local measurements could be an important future work not covered in this thesis. Furthermore, the mutual impedance and shunt

capacitance of the LV network are usually missing. A detailed study on the impact of this mutual impedance and shunt capacitance could improve the quality of LVDS models.

Appendix A

LVDS Test network

One of the challenges of working with European LVDS is a limited test feeders or networks representing European-style distribution systems. Most of the existing test feeders and networks represent the North-American style distribution networks as listed below:

- IEEE PES test networks (1991, 2001) [202, 203],
- Neutral to earth voltage (NEV) test feeder [204],
- 8500 node test feeder [205],
- Comprehensive Distribution Test Feeder [206],
- 342-nodes low voltage network test system, [207].
- EPRI Green Circuit Project Database,
- PNNL taxonomy feeders [194],
- Western Australian taxonomic feeders [196].

However, there are a few publicly available test networks for European style LVDS:

- **Kerber Networks (2013)**: These are seven German LV feeders used by Georg Kerber in his dissertation [208]. However, they are single-phase equivalent network assuming all loads are balanced.

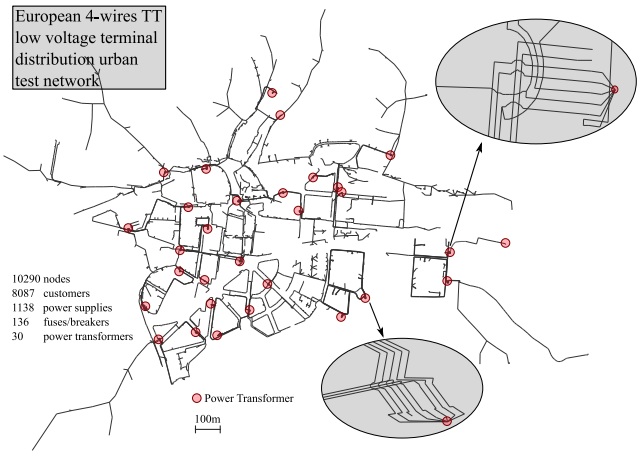


Figure A.1: Schematic of Non-synthetic European LVDS network

- **IEEE European LV test feeder (2014)**: IEEE currently has only one European LVDS test feeder [49]. However, it does not completely represent the European-style network apart from maintaining the base frequency of 50 Hz and LV maintained at 416 V.

- **Electricity North West (ENWL) network (2014)**: The 128 feeders from ENWL [209] are the extended version of the IEEE European test feeder. However, they lack complete cable description, and consumer loads are selected randomly from a pool of 100 profiles [210].

- **Non-synthetic European low voltage test system (2020)[1]:** The Non-synthetic European LVDS presented in [30] presents: a) raw-data from Geographical Information System (GIS), b) raw data of consumer SM installation and their phases obtained from field-survey, c) raw data of consumer SM as collected by the SM data management system of the DSO, c) a modular tool to process and convert the raw data into OpenDSS model, and d) an OpenDSS format final LVDS network with time-series load. The OpenDSS model is a 4-wire LVDS model with neutral grounded only at the substation. This thesis utilizes the feeders of this test network for all of the case studies. The 'json' version of feederwise data is made available in github[2]. A schematic representation of the non-synthetic network is given in Fig. A.1. The mentioned network has following features:

[1]https://data.mendeley.com/datasets/685vgp64sm/1

[2]https://github.com/arpkoirala/StochasticPowerModels.jl/tree/PVhostingCap/test/data/Spanish/All_feeder

- Rated voltage of 416V and the frequency 50 Hz,
- Three-phase four wire model with exclusive representation of neutral,
- 8087 loads and 30 distribution transformers,
- 10290 buses including the buses for geographical coordinates, and
- A time series hourly load-data of 20 days obtained from SM.

Bibliography

[1] European Commission, "Proposal for a regulation of the european parliament and of the council establishing the framework for achieving climate neutrality and amending regulation (eu) 2018/1999 (european climate law), march 4, 2020," *COM(2020) 563 final 2020/0036(COD).*

[2] European Commission, 23 and 24 October 2014, "2030 climate and energy policy framework."

[3] European Commission, 30 November 2016, "Renewable energy directive."

[4] Council of the EU, 27 June 2022, "Fit for 55."

[5] European Environment Agency, 18 November 2021, "New registrations of electric vehicles in europe."

[6] Boston Consultancy Group, A. Sahoo, K. Mistry, and T. Baker, "The costs of revving up the grid for electric vehicles," BCG Report, 2021.

[7] Department of Mobility and Public Works, Flanders, Belgium, June, 2021, "Electric driving an important pillar in the flemish climate plan."

[8] Eurelectric, EDSO, and Monitor Deloitte, "Connecting the dots: Distribution grid investment to power the energy transition," 2021.

[9] EDSO, "E.DSO strategic statement: European energy security needs grid investments," September 2022.

[10] Fluvius, Flanders, Belgium, November, 2022, "Investeringsplan 2023-2032."

[11] Y. Zhang, F. Li, Z. Hu, and G. Shaddick, "Quantification of low voltage network reinforcement costs: A statistical approach," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 810–818, 2013.

[12] S. Claeys, *Distribution Network Modeling: From Simulation Towards Optimization*. PhD thesis, KU Leuven, 2021.

[13] C. Masetti, "Revision of european standard en 50160 on power quality: Reasons and solutions," in *Proceedings of 14th International Conference on Harmonics and Quality of Power-ICHQP 2010*, pp. 1–7, IEEE, 2010.

[14] C. Gault, F. Watson, C. Breadon, C. Abbey, J. Miller, E. Davidson, and B. Currie, "Meeting emerging and future requirements for managing DER in highly active distribution network," in *4th International Conference on Electricity Distribution, Glasgow, 2017*, 2017.

[15] H. Al-Saadi, R. Zivanovic, and S. F. Al-Sarawi, "Probabilistic analysis of maximum allowable PV connections across bidirectional feeders within a distribution network," in *2017 Asian Conference on Energy, Power and Transportation Electrification (ACEPT)*, pp. 1–6, 2017.

[16] T. Mühlpfordt, T. Faulwasser, V. Hagenmeyer, L. Roald, and S. Misra, "On polynomial real-time control policies in stochastic AC optimal power flow," *Electric Power Systems Research*, vol. 189, p. 106792, 2020.

[17] M. H. J. Bollen and S. K. Rönnberg, "Hosting capacity of the power grid for renewable electricity production and new large consumption equipment," *Energies*, vol. 10, no. 9, 2017.

[18] S. M. Ismael, S. H. A. Aleem, A. Y. Abdelaziz, and A. F. Zobaa, "State-of-the-art of hosting capacity in modern power systems with distributed generation," *Renewable Energy*, vol. 130, pp. 1002 – 1020, 2019.

[19] E. Mulenga, M. H. Bollen, and N. Etherden, "A review of hosting capacity quantification methods for photovoltaics in low-voltage distribution grids," *International Journal of Electrical Power & Energy Systems*, vol. 115, p. 105445, 2020.

[20] M. Thomson and D. G. Infield, "Impact of widespread photovoltaics generation on distribution systems," *IET Renewable Power Generation*, vol. 1, pp. 33–40, March 2007.

[21] R. Torquato, D. Salles, C. Oriente Pereira, P. C. M. Meira, and W. Freitas, "A comprehensive assessment of PV hosting capacity on low-voltage distribution systems," *IEEE Transactions on Power Delivery*, vol. 33, pp. 1002–1012, April 2018.

[22] M. Rossi, G. Viganò, and D. Moneta, "Hosting capacity of distribution networks: Evaluation of the network congestion risk due to distributed generation," in *2015 International Conference on Clean Electrical Power (ICCEP)*, pp. 716–722, June 2015.

[23] C. Gaunt, E. Namanya, and R. Herman, "Voltage modelling of LV feeders with dispersed generation: Limits of penetration of randomly connected photovoltaic generation," *Electric Power Systems Research*, vol. 143, pp. 1 – 6, 2017.

[24] M. J. Chihota, B. Bekker, and T. Gaunt, "A stochastic analytic-probabilistic approach to distributed generation hosting capacity evaluation of active feeders," *International Journal of Electrical Power & Energy Systems*, vol. 136, p. 107598, 2022.

[25] E. Mulenga, M. H. Bollen, and N. Etherden, "Solar PV stochastic hosting capacity in distribution networks considering aleatory and epistemic uncertainties," *International Journal of Electrical Power & Energy Systems*, vol. 130, p. 106928, 2021.

[26] V. Klonari, J. Toubeau, J. Lobry, and F. Vallée, "Photovoltaic integration in smart city power distribution: A probabilistic photovoltaic hosting capacity assessment based on smart metering data," in *2016 5th International Conference on Smart Cities and Green ICT Systems (SMARTGREENS)*, pp. 1–13, April 2016.

[27] A. Navarro-Espinosa and L. F. Ochoa, "Probabilistic impact assessment of low carbon technologies in LV distribution systems," *IEEE Transactions on Power Systems*, vol. 31, pp. 2192–2203, May 2016.

[28] T. J. Sullivan, *Orthogonal Polynomials and Applications*, pp. 133–164. Springer International Publishing, 2015.

[29] T. Mühlpfordt, *Uncertainty Quantification via Polynomial Chaos Expansion-Methods and Applications for Optimization of Power Systems*. PhD thesis, Karlsruher Institut für Technologie (KIT), 2020.

[30] A. Koirala, L. Suárez-Ramón, B. Mohamed, and P. Arboleya, "Non-synthetic European low voltage test system," *International Journal of Electrical Power & Energy Systems*, vol. 118, 2020.

[31] EPOC Belgium, 25 Feb 2019, "Epoc and bregilab project unite belgian top experts to support belgian energy policy."

[32] A. Koirala, R. D'hulst, and D. Van Hertem, "Impedance modelling for European style distribution feeder," in *2019 International Conference on Smart Energy Systems and Technologies (SEST)*, pp. 1–6, Sep. 2019.

[33] F. Geth, R. Heidari, and A. Koirala, "Computational analysis of impedance transformations for four-wire power networks with sparse neutral grounding," in *Proceedings of the Thirteenth ACM International Conference on Future Energy Systems*, pp. 105–113, 2022.

[34] A. Koirala, T. Van Acker, D. Van Hertem, J. Camargo, and R. D'hulst, "General polynomial chaos vs crude monte carlo for probabilistic evaluation of distribution systems," in *2020 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)*, pp. 1–6, 2020.

[35] T. Van Acker, F. Geth, A. Koirala, and H. Ergun, "General polynomial chaos in the current–voltage formulation of the optimal power flow problem," *Electric Power Systems Research*, vol. 211, p. 108472, 2022.

[36] ESB Networks, 27 March 2021, "Company standard: The distribution system security and planning standards."

[37] W. H. Kersting, *Distribution system modeling and analysis*. CRC press, 2001.

[38] C.-K. Chang and M. A. F. Hassan, "Application of low voltage high resistance grounding in nuclear power plants," *Nuclear Engineering and Technology*, vol. 48, no. 1, pp. 211–217, 2016.

[39] C. S. Cheng and D. Shirmohammadi, "A three-phase power flow method for real-time distribution system analysis," *IEEE Transactions on Power Systems*, vol. 10, pp. 671–679, May 1995.

[40] T. Chen, M. Chen, K. Hwang, P. Kotas, and E. A. Chebli, "Distribution system power flow analysis-a rigid approach," *IEEE Transactions on Power Delivery*, vol. 6, pp. 1146–1152, July 1991.

[41] G. X. Luo and A. Semlyen, "Efficient load flow for large weakly meshed networks," *IEEE Transactions on Power Systems*, vol. 5, pp. 1309–1316, Nov 1990.

[42] M. F. AlHajri and M. E. El-Hawary, "Exploiting the radial distribution structure in developing a fast and flexible radial power flow for unbalanced three-phase networks," *IEEE Transactions on Power Delivery*, vol. 25, pp. 378–389, Jan 2010.

[43] W. H. Kersting and W. H. Phillips, "Distribution feeder line models," *IEEE Transactions on Industry Applications*, vol. 31, pp. 715–720, Jul 1995.

[44] J. Desmet, I. Sweertvaegher, G. Vanalme, K. Stockman, and R. Belmans, "Analysis of the neutral conductor current in a three phase supplied network with nonlinear single phase loads," in *IEMDC 2001. IEEE International Electric Machines and Drives Conference (Cat. No.01EX485)*, pp. 448–453, June 2001.

[45] R. M. Ciric, A. P. Feltrin, and L. F. Ochoa, "Power flow in four-wire distribution networks - general approach," *IEEE Trans. Power Syst.*, vol. 18, no. 4, pp. 1283–1290, 2003.

[46] F. Olivier, R. Fonteneau, and D. Ernst, "Modelling of three-phase four-wire low-voltage cables taking into account the neutral connection to the earth," in *Proc. of CIRED Workshop 2018*, 2018.

[47] C. Acosta, R. A. Hincapié, M. Granada, A. H. Escobar, and R. A. Gallego, "An efficient three phase four wire radial power flow including neutral-earth effect," *Journal of Control, Automation and Electrical Systems*, vol. 24, pp. 690–701, Oct 2013.

[48] P. Arboleya, A. Koirala, L. Suarez-Ramon, B. Mohamed, and C. Gonzalez-Moran, "Impact evaluation of the new self-consumption spanish scenario on the low voltage terminal distribution network," *IEEE Transactions on Industry Applications*, pp. 1–1, 2019.

[49] IEEE PES AMPS DSAS Test Feeder Working Group, "IEEE PES distribution systems analysis subcommittee radial test feeders.," 2016.

[50] J. R. Carson, "Wave propagation in overhead wires with ground return," *The Bell System Technical Journal*, vol. 5, pp. 539–554, Oct 1926.

[51] R. M. Ciric, L. F. Ochoa, and A. Padilha, "Power flow in distribution networks with earth return," *International Journal of Electrical Power & Energy Systems*, vol. 26, no. 5, pp. 373 – 380, 2004.

[52] B. Sudret, "Uncertainty propagation and sensitivity analysis in mechanical models – contributions to structural reliability and stochastic spectral methods," *Habilitationa diriger des recherches, Université Blaise Pascal, Clermont-Ferrand, France*, 01 2007.

[53] R. R. Booth, "Power system simulation model based on probability analysis," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-91, no. 1, pp. 62–69, 1972.

[54] A. Meliopoulos, G. Cokkinides, and X. Chao, "A new probabilistic power flow analysis method," *IEEE Transactions on Power Systems*, vol. 5, no. 1, pp. 182–190, 1990.

[55] K. N. Hasan, R. Preece, and J. V. Milanović, "Existing approaches and trends in uncertainty modelling and probabilistic stability analysis of power systems with renewable generation," *Renewable and Sustainable Energy Reviews*, vol. 101, pp. 168–180, 2019.

[56] T. Kloek and H. K. van Dijk, "Bayesian estimates of equation system parameters: An application of integration by monte carlo," *Econometrica*, vol. 46, no. 1, pp. 1–19, 1978.

[57] I. Sobol', "On the distribution of points in a cube and the approximate evaluation of integrals," *USSR Computational Mathematics and Mathematical Physics*, vol. 7, no. 4, pp. 86 – 112, 1967.

[58] J. H. Halton, "Algorithm 247: Radical-inverse quasi-random point sequence," *Commun. ACM*, vol. 7, p. 701–702, Dec. 1964.

[59] Y. Luo, X. Wang, and S. Yan, "Risk assessment of photovoltaic distribution network based on adaptive kernel density estimation and cumulant method," *Energy Reports*, vol. 8, pp. 1152–1159, 2022.

[60] R. E. Caflisch, "Monte carlo and quasi-monte carlo methods," *Acta Numerica*, vol. 7, p. 1–49, 1998.

[61] H. Huang, C. Chung, K. W. Chan, and H. Chen, "Quasi-monte carlo based probabilistic small signal stability analysis for power systems with plug-in electric vehicle and wind power integration," *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 3335–3343, 2013.

[62] W. C. Snyder, "Accuracy estimation for quasi-monte carlo simulations," *Mathematics and computers in simulation*, vol. 54, no. 1-3, pp. 131–143, 2000.

[63] S. Asmussen and P. Glynn, *Stochastic Simulation: Algorithms and Analysis*. Stochastic Modelling and Applied Probability, Springer New York, 2007.

[64] H. Ergun, *Grid planning for the future grid: optimizing topology and technology considering spatial and temporal effects*. PhD thesis, KU Leuven, 2015.

[65] C. A. Peñuela, G. E. Mauricio, and J. R. S. Mantovani, "Probabilistic analysis of the distributed power generation in weakly meshed distribution systems," in *2010 IEEE/PES Transmission and Distribution Conference and Exposition: Latin America (T D-LA)*, pp. 171–177, Nov 2010.

[66] A. Bracale, G. Carpinelli, P. Caramia, A. Russo, and P. Varilone, "Point estimate schemes for probabilistic load flow analysis of unbalanced electrical distribution systems with wind farms," in *Proceedings of 14th International Conference on Harmonics and Quality of Power - ICHQP 2010*, pp. 1–6, Sep. 2010.

[67] F. Ruiz-Rodriguez, J. Hernández, and F. Jurado, "Probabilistic load flow for photovoltaic distributed generation using the cornish–fisher expansion," *Electric Power Systems Research*, vol. 89, pp. 129 – 138, 2012.

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

[69] M. Fan, V. Vittal, G. T. Heydt, and R. Ayyanar, "Probabilistic power flow studies for transmission systems with photovoltaic generation using cumulants," *IEEE Transactions on Power Systems*, vol. 27, pp. 2251–2261, Nov 2012.

[70] M. Aien, M. Fotuhi-Firuzabad, and F. Aminifar, "Probabilistic load flow in correlated uncertain environment using unscented transformation," *IEEE Transactions on Power Systems*, vol. 27, pp. 2233–2241, Nov 2012.

[71] P. Amid and C. Crawford, "A cumulant-tensor-based probabilistic load flow method," *IEEE Transactions on Power Systems*, vol. 33, pp. 5648–5656, Sep. 2018.

[72] G. Gruosso, P. Maffezzoni, Z. Zhang, and L. Daniel, "Probabilistic load flow methodology for distribution networks including loads uncertainty," *International Journal of Electrical Power & Energy Systems*, vol. 106, pp. 392 – 400, 2019.

[73] J. Laowanitwattana and S. Uatrongjit, "Probabilistic power flow analysis based on arbitrary polynomial chaos expansion for networks with uncertain renewable sources," *IEEJ transactions on Electrical and Electronic Engineering*, vol. 13, no. 12, pp. 1754–1759, 2018.

[74] H. Sheng and X. Wang, "Probabilistic power flow calculation using non-intrusive low-rank approximation method," *IEEE Transactions on Power Systems*, vol. 34, pp. 3014–3025, July 2019.

[75] M. Eldred and J. Burkardt, "Comparison of non-intrusive polynomial chaos and stochastic collocation methods for uncertainty quantification," in *47th AIAA aerospace sciences meeting including the new horizons forum and aerospace exposition*, p. 976, 2009.

[76] H. Wu, Y. Zhou, S. Dong, and Y. Song, "Probabilistic load flow based on generalized polynomial chaos," *IEEE Transactions on Power Systems*, vol. 32, pp. 1–2, 01 2016.

[77] V. Rigoni, L. F. Ochoa, G. Chicco, A. Navarro-Espinosa, and T. Gozel, "Representative residential LV feeders: A case study for the north west of England," *IEEE Transactions on Power Systems*, vol. 31, pp. 348–360, Jan 2016.

[78] S. Gupta, S. Misra, D. Deka, and V. Kekatos, "Dnn-based policies for stochastic ac opf," *Electric Power Systems Research*, vol. 213, p. 108563, 2022.

[79] Y. Yang, Z. Yang, J. Yu, B. Zhang, Y. Zhang, and H. Yu, "Fast calculation of probabilistic power flow: A model-based deep learning approach," *IEEE Transactions on Smart Grid*, vol. 11, no. 3, pp. 2235–2244, 2020.

[80] X. A. Sun and Á. Lorca, *Robust Optimization in Electric Power Systems Operations*, pp. 227–258. Cham: Springer International Publishing, 2017.

[81] W. Gautschi, *Orthogonal polynomials: computation and approximation*. OUP Oxford, 2004.

[82] C.-T. Chan, A. Mironov, A. Morozov, and A. Sleptsov, "Orthogonal polynomials in mathematical physics," in *Ludwig Faddeev Memorial Volume: A Life in Mathematical Physics*, pp. 119–182, World Scientific, 2018.

[83] Y. Xu, M. Korkali, L. Mili, and X. Chen, "An efficient multifidelity model for assessing risk probabilities in power systems under rare events," in *53rd Hawaii International Conference on System Sciences 2020*, 2020.

[84] V. Goncharov, "The theory of best approximation of functions," *Journal of Approximation Theory*, vol. 106, no. 1, pp. 2–57, 2000.

[85] R. Askey and J. A. Wilson, *Some basic hypergeometric orthogonal polynomials that generalize Jacobi polynomials*, vol. 319. American Mathematical Soc., 1985.

[86] R. Koekoek, P. A. Lesky, and R. F. Swarttouw, *Hypergeometric orthogonal polynomials and their q-analogues*. Springer Science & Business Media, 2010.

[87] D. Xiu and G. E. Karniadakis, "Modeling uncertainty in flow simulations via generalized polynomial chaos," *Journal of Computational Physics*, vol. 187, no. 1, pp. 137–167, 2003.

[88] J. Feinberg and H. P. Langtangen, "Chaospy: An open source tool for designing methods of uncertainty quantification," *Journal of Computational Science*, vol. 11, pp. 46 – 57, 2015.

[89] T. Mühlpfordt, L. Roald, V. Hagenmeyer, T. Faulwasser, and S. Misra, "Chance-constrained ac optimal power flow: A polynomial chaos approach," *IEEE Transactions on Power Systems*, vol. 34, no. 6, pp. 4806–4816, 2019.

[90] Y. Xu, L. Mili, and J. Zhao, "Probabilistic power flow calculation and variance analysis based on hierarchical adaptive polynomial chaos-ANOVA method," *IEEE Transactions on Power Systems*, vol. 34, no. 5, pp. 3316–3325, 2019.

[91] M. Eldred, "Recent advances in non-intrusive polynomial chaos and stochastic collocation methods for uncertainty analysis and design," in *50th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference 17th AIAA/ASME/AHS Adaptive Structures Conference 11th AIAA No*, p. 2274, 2009.

[92] S. Hosder and R. Walters, "Non-intrusive polynomial chaos methods for uncertainty quantification in fluid dynamics," in *48th AIAA Aerospace Sciences Meeting Including the New Horizons Forum and Aerospace Exposition*, p. 129, 2010.

[93] X. Li, P. B. Nair, Z. Zhang, L. Gao, and C. Gao, "Aircraft robust trajectory optimization using nonintrusive polynomial chaos," *Journal of Aircraft*, vol. 51, no. 5, pp. 1592–1603, 2014.

[94] S. Bhonsale, P. Nimmegeers, D. Telen, J. A. Paulson, A. Mesbah, and J. Van Impe, "On the implementation of generalized polynomial chaos in dynamic optimization under stochastic uncertainty: a user perspective," in *29th European Symposium on Computer Aided Process Engineering* (A. A. Kiss, E. Zondervan, R. Lakerveld, and L. Özkan, eds.), vol. 46 of *Computer Aided Chemical Engineering*, pp. 541–546, Elsevier, 2019.

[95] A. K. Papageorgiou, E. M. Papoutsis-Kiachagias, and K. C. Giannakoglou, "Development and assessment of an intrusive polynomial chaos expansion-based continuous adjoint method for shape optimization under uncertainties," *International Journal for Numerical Methods in Fluids*, vol. 94, no. 1, pp. 59–75, 2022.

[96] M. Tootkaboni, A. Asadpoure, and J. K. Guest, "Topology optimization of continuum structures under uncertainty–a polynomial chaos approach," *Computer Methods in Applied Mechanics and Engineering*, vol. 201, pp. 263–275, 2012.

[97] N. Wiener, "The homogeneous chaos," *American Journal of Mathematics*, vol. 60, no. 4, pp. 897–936, 1938.

[98] A. Dubey and S. Santoso, "On estimation and sensitivity analysis of distribution circuit's photovoltaic hosting capacity," *IEEE Transactions on Power Systems*, vol. 32, pp. 2779–2789, July 2017.

[99] A. Seppälä, *Load research and load estimation in electricity distribution: Dissertation*. PhD thesis, Helsinki University of Technology, Finland, 1996. Project code: N6SU00026.

[100] J. Munkhammar, J. Rydén, and J. Widén, "Characterizing probability density distributions for household electricity load profiles from high-resolution electricity use data," *Applied Energy*, vol. 135, pp. 382–390, 2014.

[101] M. J. Chihota and B. Bekker, "Modelling and simulation of uncertainty in the placement of distributed energy resources for planning applications," in *2020 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)*, pp. 1–6, 2020.

[102] M. D. Coutto Filho, A. L. Da Silva, V. Arienti, and S. Ribeiro, "Probabilistic load modelling for power system expansion planning," in *1991 Third International Conference on Probabilistic Methods Applied to Electric Power Systems*, pp. 203–207, IET, 1991.

[103] M. J. Chihota, *Extending the Herman-Beta transform for probabilistic load flow analysis of radial feeders*. PhD thesis, Faculty of Engineering and the Built Environment, University of Cape Town, 2019.

[104] J. Dickert and P. Schegner, "Residential load models for network planning purposes," in *2010 Modern Electric Power Systems*, pp. 1–6, IEEE, 2010.

[105] B. Pillot, S. de Siqueira, and J. B. Dias, "Grid parity analysis of distributed PV generation using Monte Carlo approach: The Brazilian case," *Renewable Energy*, vol. 127, pp. 974 – 988, 2018.

[106] R. Herman and J. Kritzinger, "The statistical description of grouped domestic electrical load currents," *Electric Power Systems Research*, vol. 27, no. 1, pp. 43–48, 1993.

[107] B. Dupont, P. Vingerhoets, P. Tant, K. Vanthournout, W. Cardinaels, T. De Rybel, E. Peeters, and R. Belmans, "Linear breakthrough project: Large-scale implementation of smart grid technologies in distribution grids," in *2012 3rd IEEE PES Innovative Smart Grid Technologies Europe (ISGT Europe)*, pp. 1–8, 2012.

[108] A. Soroudi, M. Aien, and M. Ehsan, "A probabilistic modeling of photo voltaic modules and wind power generation impact on distribution networks," *IEEE Systems Journal*, vol. 6, no. 2, pp. 254–259, 2012.

[109] 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 vol.1, 1994.

[110] G. Tina, S. Gagliano, and S. Raiti, "Hybrid solar/wind power system probabilistic modelling for long-term performance assessment," *Solar Energy*, vol. 80, no. 5, pp. 578–588, 2006.

[111] D. Thevenard and S. Pelland, "Estimating the uncertainty in long-term photovoltaic yield predictions," *Solar Energy*, vol. 91, pp. 432–445, 2013.

[112] M. Anvari, G. Lohmann, M. Wächter, P. Milan, E. Lorenz, D. Heinemann, M. R. R. Tabar, and J. Peinke, "Short term fluctuations of wind and solar power systems," *New Journal of Physics*, vol. 18, no. 6, p. 063027, 2016.

[113] V. Graham and K. Hollands, "A method to generate synthetic hourly solar radiation globally," *Solar Energy*, vol. 44, no. 6, pp. 333–341, 1990.

[114] Z. Salameh, B. Borowy, and A. Amin, "Photovoltaic module-site matching based on the capacity factors," *IEEE Transactions on Energy Conversion*, vol. 10, no. 2, pp. 326–332, 1995.

[115] A. M. Dizqah, A. Maheri, and K. Busawon, "An assessment of solar irradiance stochastic model for the uk," in *2012 2nd International Symposium On Environment Friendly Energies And Applications*, pp. 670–675, 2012.

[116] A. S. Al-Sumaiti, M. H. Ahmed, S. Rivera, M. S. El Moursi, M. M. Salama, and T. Alsumaiti, "Stochastic pv model for power system planning applications," *IET Renewable Power Generation*, vol. 13, no. 16, pp. 3168–3179, 2019.

[117] H. Palahalli, P. Maffezzoni, and G. Gruosso, "Gaussian copula methodology to model photovoltaic generation uncertainty correlation in power distribution networks," *Energies*, vol. 14, no. 9, p. 2349, 2021.

[118] A. Dubey, S. Santoso, and A. Maitra, "Understanding photovoltaic hosting capacity of distribution circuits," in *2015 IEEE Power Energy Society General Meeting*, pp. 1–5, July 2015.

[119] S. S. Al Kaabi, H. Zeineldin, and V. Khadkikar, "Planning active distribution networks considering multi-dg configurations," *IEEE Transactions on Power Systems*, vol. 29, no. 2, pp. 785–793, 2013.

[120] X. Lin, J. Sun, S. Ai, X. Xiong, Y. Wan, and D. Yang, "Distribution network planning integrating charging stations of electric vehicle with v2g," *International Journal of Electrical Power & Energy Systems*, vol. 63, pp. 507–512, 2014.

[121] A. Ehsan and Q. Yang, "State-of-the-art techniques for modelling of uncertainties in active distribution network planning: A review," *Applied Energy*, vol. 239, pp. 1509–1523, 2019.

[122] A. Kharrazi, V. Sreeram, and Y. Mishra, "Assessment of voltage unbalance due to single phase rooftop photovoltaic panels in residential low voltage distribution network: A study on a real lv network in western australia," in *2017 Australasian Universities Power Engineering Conference (AUPEC)*, pp. 1–6, Nov 2017.

[123] F. Shahnia, R. Majumder, A. Ghosh, G. Ledwich, and F. Zare, "Voltage imbalance analysis in residential low voltage distribution networks with rooftop PVs," *Electric Power Systems Research*, vol. 81, no. 9, pp. 1805 – 1814, 2011.

[124] A. Navarro, L. F. Ochoa, and D. Randles, "Monte carlo-based assessment of pv impacts on real UK low voltage networks," in *2013 IEEE Power Energy Society General Meeting*, pp. 1–5, July 2013.

[125] G. Guerra and J. A. Martinez, "A monte carlo method for optimum placement of photovoltaic generation using a multicore computing environment," in *2014 IEEE PES General Meeting | Conference Exposition*, pp. 1–5, July 2014.

[126] M. Pau, F. Ponci, and A. Monti, "Impact of network parameters uncertainties on distribution grid power flow," in *2019 International Conference on Smart Energy Systems and Technologies (SEST)*, pp. 1–6, 2019.

[127] S. J. Pappu, N. Bhatt, R. Pasumarthy, and A. Rajeswaran, "Identifying topology of low voltage distribution networks based on smart meter data," *IEEE Transactions on Smart Grid*, vol. 9, no. 5, pp. 5113–5122, 2018.

[128] A. Hoogsteyn, M. Vanin, A. Koirala, and D. Van Hertem, "Low voltage customer phase identification methods based on smart meter data," *Electric Power Systems Research*, vol. 212, p. 108524, 2022.

[129] M. Vanin, T. Van Acker, R. D'hulst, and D. Van Hertem, "Phase identification of distribution system users through a milp extension of state estimation," *arXiv preprint arXiv:2206.08436*, 2022.

[130] M. H. Wen, R. Arghandeh, A. von Meier, K. Poolla, and V. O. Li, "Phase identification in distribution networks with micro-synchrophasors," in *2015 IEEE PES General Meeting*, pp. 1–5, 2015.

[131] S. Claeys, F. Geth, and G. Deconinck, "Line parameter estimation in multi-phase distribution networks without voltage angle measurements," in *CIRED 2021-The 26th International Conference and Exhibition on Electricity Distribution*, vol. 2021, pp. 1186–1190, IET, 2021.

[132] M. M. Haque and P. Wolfs, "A review of high PV penetrations in LV distribution networks: Present status, impacts and mitigation measures," *Renewable and Sustainable Energy Reviews*, vol. 62, pp. 1195 – 1208, 2016.

[133] C. L. Masters, "Voltage rise: the big issue when connecting embedded generation to long 11 kv overhead lines," *Power Engineering Journal*, vol. 16, pp. 5–12, Feb 2002.

[134] R. Tonkoski, D. Turcotte, and T. H. M. EL-Fouly, "Impact of high PV penetration on voltage profiles in residential neighborhoods," *IEEE Transactions on Sustainable Energy*, vol. 3, pp. 518–527, July 2012.

[135] D. Schwanz, F. Möller, S. K. Rönnberg, J. Meyer, and M. H. J. Bollen, "Stochastic assessment of voltage unbalance due to single-phase-connected solar power," *IEEE Transactions on Power Delivery*, vol. 32, pp. 852–861, April 2017.

[136] V. H. M. Quezada, J. R. Abbad, and T. G. S. Roman, "Assessment of energy distribution losses for increasing penetration of distributed generation," *IEEE Transactions on Power Systems*, vol. 21, pp. 533–540, May 2006.

[137] A. Nguyen, M. Velay, J. Schoene, V. Zheglov, B. Kurtz, K. Murray, B. Torre, and J. Kleissl, "High PV penetration impacts on five local distribution networks using high resolution solar resource assessment with sky imager and quasi-steady state distribution system simulations," *Solar Energy*, vol. 132, pp. 221 – 235, 2016.

[138] H. Pezeshki, P. J. Wolfs, and G. Ledwich, "Impact of high PV penetration on distribution transformer insulation life," *IEEE Transactions on Power Delivery*, vol. 29, pp. 1212–1220, June 2014.

[139] M. Z. ul Abideen, O. Ellabban, and L. Al-Fagih, "A review of the tools and methods for distribution networks' hosting capacity calculation," *Energies*, vol. 13, p. 2758, Jun 2020.

[140] A. Kharrazi, V. Sreeram, and Y. Mishra, "Assessment techniques of the impact of grid-tied rooftop photovoltaic generation on the power quality of low voltage distribution network - a review," *Renewable and Sustainable Energy Reviews*, vol. 120, p. 109643, 2020.

[141] M. Alturki, A. Khodaei, A. Paaso, and S. Bahramirad, "Optimization-based distribution grid hosting capacity calculations," *Applied Energy*, vol. 219, pp. 350 – 360, 2018.

[142] M. S. S. Abad, J. Ma, D. Zhang, A. S. Ahmadyar, and H. Marzooghi, "Probabilistic assessment of hosting capacity in radial distribution systems," *IEEE Transactions on Sustainable Energy*, vol. 9, no. 4, pp. 1935–1947, 2018.

[143] C. S. Cheng and D. Shirmohammadi, "A three-phase power flow method for real-time distribution system analysis," *IEEE Transactions on Power Systems*, vol. 10, pp. 671–679, May 1995.

[144] S. Kikuchi, M. Machida, J. Tamura, M. Imanaka, J. Baba, D. Iioka, K. Miura, M. Takagi, and H. Asano, "Hosting capacity analysis of many distributed photovoltaic systems in future distribution networks," in *2017 IEEE Innovative Smart Grid Technologies - Asia (ISGT-Asia)*, pp. 1–5, Dec 2017.

[145] D. Bertini, D. Falabretti, M. Merlo, D. Moneta, J. Silva de Assis Carneiro, and A. Silvestri, "Hosting capacity of italian lv distribution networks," in *CIRED 2011*, pp. 1–4, 2011.

[146] I. T. Papaioannou and A. Purvins, "A methodology to calculate maximum generation capacity in low voltage distribution feeders," *International Journal of Electrical Power & Energy Systems*, vol. 57, pp. 141 – 147, 2014.

[147] S. M. Mirbagheri, M. Moncecchi, D. Falabretti, and M. Merlo, "Hosting capacity evaluation in networks with parameter uncertainties," in *2018 18th International Conference on Harmonics and Quality of Power (ICHQP)*, pp. 1–6, May 2018.

[148] K. Coogan, M. J. Reno, S. Grijalva, and R. J. Broderick, "Locational dependence of pv hosting capacity correlated with feeder load," in *2014 IEEE PES T D Conference and Exposition*, pp. 1–5, 2014.

[149] T. Stetz, W. Yan, and M. Braun, "Voltage control in distribution systems with high level pv-penetration," in *25th European PV Solar Energy Conference*, pp. 5000–5006, 2010.

[150] P. McNutt, J. Hambrick, M. Keesee, and D. Brown, "Impact of solarsmart subdivisions on SMUD's distribution system," *NREL/TP-550-46093*, 2009.

[151] A. de Keizer, E. ter Horst, and W. van Sark, "Performance evaluation of the 1 MW building integrated PV project in Nieuwland, Amersfoort, the Netherlands, January 2001 – February 2006," 2008.

[152] P. B. Kitworawut, D. T. Azuatalam, and A. J. Collin, "An investigation into the technical impacts of microgeneration on uk-type lv distribution networks," in *2016 Australasian Universities Power Engineering Conference (AUPEC)*, pp. 1–5, Sep. 2016.

[153] S. Cobben, B. Gaiddon, and H. Laukamp, "Impact of photovoltaic generation on power quality in urban areas with high PV population," *EIE/05/171/SI2.420208*, 2008.

[154] B. Bletterie, A. Goršek, B. Uljanić, B. Blazic, A. Woyte, T. Vu Van, F. Truyens, and J. Jahn, "Enhancement of the network hosting capacity—clearing space for/with pv," in *25th European Photovoltaic Solar Energy Conference and Exhibition / 5th World Conference on Photovoltaic Energy Conversion, 6-10 September 2010, Valencia, Spain*, 09 2010.

[155] M. H. Coddington, D. Baca, B. D. Kroposki, and T. Basso, "Deploying high penetration photovoltaic systems — a case study," in *2011 37th IEEE Photovoltaic Specialists Conference*, pp. 002594–002599, June 2011.

[156] E. Stewart, D. Nakafuji, T. Aukai, S. vasilic, and J. MacPherson, "Analysis of high-penetration levels of photovoltaics into the distribution grid on oahu, hawaii," in *High Penetration Solar Forum, San Diego, CA*, 04 2013.

[157] T. Stetz, M. Kraiczy, K. Diwold, M. Braun, B. Bletterie, C. Mayr, R. Bründlinger, B. Noone, A. Bruce, I. MacGill, B. Mather, K. Ogimoto, K. Washihara, Y. Ueda, A. Iaria, A. Gatti, M. R. D. Cirio, I. Theologitis, K. de Brabandere, C. B. S. Tselepis, and Y. Wang, "High penetration of pv in local distribution grid," *IEA PVPS T14-02:2014*, 2014.

[158] M. Ebad and W. M. Grady, "An approach for assessing high-penetration PV impact on distribution feeders," *Electric Power Systems Research*, vol. 133, pp. 347 – 354, 2016.

[159] S. Hashemi, J. Østergaard, T. Degner, R. Brandl, and W. Heckmann, "Efficient control of active transformers for increasing the pv hosting capacity of lv grids," *IEEE Transactions on Industrial Informatics*, vol. 13, pp. 270–277, Feb 2017.

[160] S. Heslop, I. MacGill, J. Fletcher, and S. Lewis, "Method for determining a pv generation limit on low voltage feeders for evenly distributed pv and load," *Energy Procedia*, vol. 57, pp. 207 – 216, 2014. 2013 ISES Solar World Congress.

[161] M. Altin, E. U. Oguz, E. Bizkevelci, and B. Simsek, "Distributed generation hosting capacity calculation of mv distribution feeders in turkey," in *IEEE PES Innovative Smart Grid Technologies, Europe*, pp. 1–7, Oct 2014.

[162] V. R. Pandi, H. H. Zeineldin, and W. Xiao, "Determining optimal location and size of distributed generation resources considering harmonic and protection coordination limits," *IEEE Transactions on Power Systems*, vol. 28, pp. 1245–1254, May 2013.

[163] A. Arshad, M. Lindner, and M. Lehtonen, "An analysis of photo-voltaic hosting capacity in Finnish low voltage distribution networks," *Energies*, vol. 10, no. 11, 2017.

[164] S. Wang, S. Chen, L. Ge, and L. Wu, "Distributed generation hosting capacity evaluation for distribution systems considering the robust optimal operation of oltc and svc," *IEEE Transactions on Sustainable Energy*, vol. 7, no. 3, pp. 1111–1123, 2016.

[165] A. Rabiee and S. M. Mohseni-Bonab, "Maximizing hosting capacity of renewable energy sources in distribution networks: A multi-objective and scenario-based approach," *Energy*, vol. 120, pp. 417–430, 2017.

[166] E. J. Palacios-Garcia, A. Moreno-Muñoz, I. Santiago, I. M. Moreno-Garcia, and M. I. Milanés-Montero, "PV hosting capacity analysis and enhancement using high resolution stochastic modeling," *Energies*, vol. 10, no. 10, 2017.

[167] S. Munikoti, M. Abujubbeh, K. Jhala, and B. Natarajan, "A novel framework for hosting capacity analysis with spatio-temporal probabilistic voltage sensitivity analysis," *International Journal of Electrical Power & Energy Systems*, vol. 134, p. 107426, 2022.

[168] S. Ruester, S. Schwenen, C. Batlle, and I. Pérez-Arriaga, "From distribution networks to smart distribution systems: Rethinking the regulation of European electricity DSOs," *Utilities Policy*, vol. 31, pp. 229 – 237, 2014.

[169] T. 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)*, pp. 1–6, 2013.

[170] M. Abdelmalak and M. Benidris, "A polynomial chaos-based approach to quantify uncertainties of correlated renewable energy sources in voltage regulation," *IEEE Transactions on Industry Applications*, vol. 57, no. 3, pp. 2089–2097, 2021.

[171] R. R. Appino, T. Mühlpfordt, T. Faulwasser, and V. Hagenmeyer, "On solving probabilistic load flow for radial grids using polynomial chaos," in *2017 IEEE Manchester PowerTech*, pp. 1–6, 2017.

[172] Y. Ou and L. Goel, "Using monte carlo simulation for overall distribution system reliability worth assessment," *IEE Proceedings - Generation, Transmission and Distribution*, vol. 146, pp. 535–540, Sep. 1999.

[173] J. Toubeau, V. Klonari, J. Lobry, Z. De Grève, and F. Vallée, "Planning tools for the integration of renewable energy sources into low- and medium-voltage distribution grids," in *Renewable Energy - Utilisation and System Integration*, ch. 10, London, United Kingdom: IntechOpen, 2016.

[174] F. Ni, P. H. Nguyen, and J. F. G. Cobben, "Basis-adaptive sparse polynomial chaos expansion for probabilistic power flow," *IEEE Transactions on Power Systems*, vol. 32, no. 1, pp. 694–704, 2017.

[175] K. Ye, J. Zhao, Y. Zhang, X. Liu, and H. Zhang, "A generalized computationally efficient copula-polynomial chaos framework for probabilistic power flow considering nonlinear correlations of PV injections," *International Journal of Electrical Power & Energy Systems*, vol. 136, p. 107727, 2022.

[176] X. Wang, X. Wang, H. Sheng, and X. Lin, "A data-driven sparse polynomial chaos expansion method to assess probabilistic total transfer capability for power systems with renewables," *IEEE Transactions on Power Systems*, vol. 36, no. 3, pp. 2573–2583, 2021.

[177] A. Singhee and R. A. Rutenbar, "Why quasi-monte carlo is better than monte carlo or latin hypercube sampling for statistical circuit analysis," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 29, no. 11, pp. 1763–1776, 2010.

[178] R. Ghanem and J. Red-Horse, *Polynomial Chaos: Modeling, Estimation, and Approximation*, pp. 521–551. Springer International Publishing, 2017.

[179] S. Rahman, "Wiener–hermite polynomial expansion for multivariate gaussian probability measures," *Journal of Mathematical Analysis and Applications*, vol. 454, no. 1, pp. 303 – 334, 2017.

[180] K. E. Hagan, O. O. Oyebanjo, T. M. Masaud, and R. Challoo, "A probabilistic forecasting model for accurate estimation of PV solar and wind power generation," in *2016 IEEE Power and Energy Conference at Illinois (PECI)*, pp. 1–5, 2016.

[181] Z. Liu, D. Lesselier, B. Sudret, and J. Wiart, "Surrogate modeling based on resampled polynomial chaos expansions," *Reliability Engineering & System Safety*, vol. 202, p. 107008, 2020.

[182] A. Koirala, T. Van Acker, R. D'hulst, and D. Van Hertem, "Hosting capacity of photovoltaic systems in low voltage distribution systems: A benchmark of deterministic and stochastic approaches," *Renewable and Sustainable Energy Reviews*, vol. 155, p. 111899, 2022.

[183] A. Koirala, T. Van Acker, R. D'hulst, and D. Van Hertem, "Uncertainty quantification in low voltage distribution grids: Comparing monte carlo and general polynomial chaos approaches," *Sustainable Energy, Grids and Networks*, vol. 31, p. 100763, 2022.

[184] Y. Ueda, K. Kurokawa, T. Tanabe, K. Kitamura, and H. Sugihara, "Analysis results of output power loss due to the grid voltage rise in grid-connected photovoltaic power generation systems," *IEEE Transactions on Industrial Electronics*, vol. 55, no. 7, pp. 2744–2751, 2008.

[185] R. Tonkoski, L. A. C. Lopes, and T. H. M. El-Fouly, "Coordinated active power curtailment of grid connected pv inverters for overvoltage prevention," *IEEE Transactions on Sustainable Energy*, vol. 2, no. 2, pp. 139–147, 2011.

[186] L. Wang, R. Yan, and T. K. Saha, "Voltage regulation challenges with unbalanced pv integration in low voltage distribution systems and the corresponding solution," *Applied Energy*, vol. 256, p. 113927, 2019.

[187] M. Seydali Seyf Abad, J. Ma, H. Marzooghi, and A. S. Ahmadyar, "Probabilistic impact assessment of residential BESSs on the HC of LV distribution systems," *International Journal of Electrical Power & Energy Systems*, vol. 136, p. 107652, 2022.

[188] A. Koirala, M. U. Hashmi, R. D'hulst, and D. Van Hertem, "Decoupled probabilistic feeder hosting capacity calculations using general polynomial chaos," *Electric Power Systems Research*, vol. 211, p. 108535, 2022.

[189] M. Bollen and S. Rönnberg, "Hosting capacity of the power grid for renewable electricity production and new large consumption equipment," *Energies*, vol. 10, p. 1325, Sep 2017.

[190] H. Al-Saadi, R. Zivanovic, and S. F. Al-Sarawi, "Probabilistic hosting capacity for active distribution networks," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 5, pp. 2519–2532, 2017.

[191] H. Wu, R. Huang, Y. Wang, and Y. Yuan, "Chance-constrained dg hosting capacity assessment with spatially correlated wind speed," in *2019 IEEE Innovative Smart Grid Technologies - Asia (ISGT Asia)*, pp. 1546–1551, 2019.

[192] M. Deakin, C. Crozier, D. Apostolopoulou, T. Morstyn, and M. McCulloch, "Stochastic hosting capacity in lv distribution networks," in *2019 IEEE Power & Energy Society General Meeting (PESGM)*, pp. 1–5, 2019.

[193] D. Xiu, "Numerical methods for stochastic computations," in *Numerical Methods for Stochastic Computations*, Princeton university press, 2010.

[194] K. P. Schneider, Y. Chen, D. Engle, and D. Chassin, "A taxonomy of north american radial distribution feeders," in *2009 IEEE Power Energy Society General Meeting*, pp. 1–6, July 2009.

[195] EPRI, "EPRI test circuits," 2017.

[196] Y. Li and P. J. Wolfs, "Taxonomic description for western australian distribution medium-voltage and low-voltage feeders," *IET Generation, Transmission Distribution*, vol. 8, pp. 104–113, Jan 2014.

[197] B. Bletterie, S. Kadam, and H. Renner, "On the classification of low voltage feeders for network planning and hosting capacity studies," *Energies*, vol. 11, no. 3, p. 651, 2018.

[198] J. Dickert, M. Domagk, and P. Schegner, "Benchmark low voltage distribution networks based on cluster analysis of actual grid properties," in *2013 IEEE Grenoble Conference*, pp. 1–6, IEEE, 2013.

[199] J. Ding, Q. Zhang, S. Hu, Q. Wang, and Q. Ye, "Clusters partition and zonal voltage regulation for distribution networks with high penetration of pvs," *IET Generation, Transmission & Distribution*, vol. 12, no. 22, pp. 6041–6051, 2018.

[200] P. J. Rousseeuw, "Silhouettes: a graphical aid to the interpretation and validation of cluster analysis," *Journal of computational and applied mathematics*, vol. 20, pp. 53–65, 1987.

[201] S. Babaeinejadsarookolaee, A. Birchfield, R. D. Christie, C. Coffrin, C. DeMarco, R. Diao, M. Ferris, S. Fliscounakis, S. Greene, R. Huang, *et al.*, "The power grid library for benchmarking ac optimal power flow algorithms," *arXiv preprint arXiv:1908.02788*, 2019.

[202] W. H. Kersting, "Radial distribution test feeders," *IEEE Transactions on Power Systems*, vol. 6, pp. 975–985, Aug 1991.

[203] W. H. Kersting, "Radial distribution test feeders," in *2001 IEEE Power Engineering Society Winter Meeting. Conference Proceedings*, vol. 2, pp. 908–912 vol.2, 2001.

[204] W. G. Sunderman, R. C. Dugan, and D. S. Dorr, "The neutral-to-earth voltage (NEV) test case and distribution system analysis," in *2008 IEEE Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century*, pp. 1–6, July 2008.

[205] R. F. Arritt and R. C. Dugan, "The IEEE 8500-node test feeder," in *IEEE PES T D 2010*, pp. 1–6, April 2010.

[206] W. H. Kersting, "A comprehensive distribution test feeder," in *IEEE PES T&D 2010*, pp. 1–4, April 2010.

[207] K. Schneider, P. Phanivong, and J. S. Lacroix, "IEEE 342-node low voltage networked test system," in *2014 IEEE PES General Meeting | Conference Exposition*, pp. 1–5, July 2014.

[208] G. Kerber, *Aufnahmefähigkeit von Niederspannungsverteilnetzen für die Einspeisung aus Photovoltaikkleinanlagen.* PhD thesis, Technische Universität München, 2011.

[209] "Low voltage network solutions closedown report," Tech. Rep. June, Low Carbon Networks Fund, 2014. https://www.ofgem.gov.uk/sites/default/files/docs/2017/04/lvns_closedown_report.pdf.

[210] S. Claeys, M. Vanin, F. Geth, and G. Deconinck, "Applications of optimization models for electricity distribution networks," *Wiley Interdisciplinary Reviews: Energy and Environment*, vol. 10, no. 5, p. e401, 2021.

Curriculum vitae

Arpan Koirala was born on 20th March, 1991 in Gorkha, Nepal.

2018 - 2022	Ph.D. in Engineering Science
	KU Leuven, Leuven, Belgium
	Energyville, Genk, Belgium
2018 - 2018	Master thesis Internship
	EDP Spain
	Oviedo, Spain
2016 - 2018	M. Sc. in Sustainable Transport & Electric Power System
	University of Oviedo, Spain
	University of Nottingham, UK
	Politecnico de Coimbra, Portugal
2016 - 2016	Research Associate
	Indian Institute of Technology, Mumbai
	Mumbai, India
2013 - 2015	Senior Engineer
	Tata Power Delhi Distribution Limited
	Delhi, India
2009 - 2013	Bachelor of Technology (Electrical)
	MNNIT Allahabad
	Allahabad India
2006 - 2008	Secondary school
	St. Xavier's College, Maitighar
	Kathmandu, Nepal

List of publications

Journal Papers

1. **Koirala, A.**, Van Acker, T., D'hulst, R., Van Hertem, D., *Hosting capacity of photovoltaic systems in low voltage distribution systems: A benchmark of deterministic and stochastic approaches.* Renewable and Sustainable Energy Reviews, 155, Art.No. 111899.

2. **Koirala, A.**, Van Acker, T., D'hulst, R., & Van Hertem, D. (2022). *Uncertainty quantification in low voltage distribution grids: Comparing Monte Carlo and general polynomial chaos approaches.* Sustainable Energy, Grids and Networks, Art. number 100763.

3. **Koirala, A.**, Hashmi, Md. U., D'hulst, R., Van Hertem, D., *Decoupled Probabilistic Feeder Hosting Capacity Calculations Using General Polynomial Chaos.* Electric Power Systems Research, 211, Art.No. 108535.

4. **Koirala, A.**, Van Acker, T., Hashmi, Md. U., D'hulst, R., Van Hertem, D., *Chance-constrained optimization based PV hosting capacity calculation using general Polynomial Chaos.* Prepared to submit in IEEE Transactions on Power System.

5. Hashmi, M.U., **Koirala, A.**, Ergun, H., Van Hertem, D. (2022). *Flexible and curtailable resource activation in three-phase unbalanced distribution networks.* Electric Power Systems Research, 212, Art.No. 108608.

6. Van Acker, T., Geth, F., **Koirala, A.**, Ergun, H. (2022). *General polynomial chaos in the current–voltage formulation of the optimal power flow problem.* Electric Power Systems Research, 211, Art.No. 108472.

7. Hoogsteyn, A., Vanin, M., **Koirala, A.**, Van Hertem, D. (2022). *Low voltage customer phase identification methods based on smart meter data.* Electric Power Systems Research, 212, Art.No. 108524.

8. **Koirala, A.**, Suárez-Ramón, L., Mohamed, B., Arboleya, P. (2020). *Non-synthetic European low voltage test system.* International Journal Of Electrical Power & Energy Systems, Art.No. 105712.

9. Arboleya, P., **Koirala, A.**, Suarerez, L., Mohamed, B., Gonzalez-Moran, C. (2019). *Impact Evaluation of the New Self-Consumption Spanish Scenario on the Low Voltage Terminal Distribution Network.* IEEE Transactions on Industry Applications, vol. 55, no. 6, pp. 7230-7239, Nov.-Dec. 2019.

Conference papers

1. **Koirala, A.**, Hashmi, M.U., D'hulst, R., Van Hertem, D. (2022). *Representative feeders for spatial scaling of stochastic PV hosting capacity.* In: 2022 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe), 2022, pp. 1-6.

2. Wu, J., **Koirala, A.**, Van Hertem, D. (2022). *Review of statistics based coping mechanisms for Smart Meter Missing Data in Distribution Systems.* In: 2022 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe), 2022, pp. 1-6.

3. Hashmi, M.U., **Koirala, A.**, Lundholm, R., Ergun, H., Van Hertem, D. (2022). *Evaluation of voltage magnitude based unbalance metric for low voltage distribution networks.* In: IEEE Power and Energy Society General Meeting. Presented at the 2022 IEEE PES GM, Denver Colorado, 17 Jul 2022-21 Jul 2022.

4. **Koirala, A.**, Hashmi, M.U., D'hulst, R., Van Hertem, D. (2022). *Decoupled Probabilistic Feeder Hosting Capacity Calculations Using General Polynomial Chaos.* In: Proceedings of the Power Systems Computation Conference 2022. Presented at the PSCC 2022, Porto, Portugal, 27 Jun 2022-01 Jul 2022.

5. Hoogsteyn, A., Vanin, M., **Koirala, A.**, Van Hertem, D. (2022). *Low Voltage Customer Phase Identification Methods Based on Smart Meter Data.* In: Proceedings of the Power Systems Computation Conference 2022. Presented at the PSCC 2022, Porto, Portugal, 27 Jun 2022-01 Jul 2022.

6. Van Acker, T., Geth, F., **Koirala, A.**, Ergun, H. (2022). *General Polynomial Chaos in the Current-Voltage Formulation of the Optimal Power Flow Problem.* In: Proceedings of the Power Systems Computation Conference 2022. Presented at the PSCC 2022, Porto, Portugal, 27 Jun 2022-01 Jul 2022

7. Hashmi, M.U., **Koirala, A.**, Ergun, H., Van Hertem, D. (2022). *Flexible and curtailable resource activation in three-phase unbalanced distribution networks.* In: Proceedings of the Power Systems Computation Conference 2022. Presented at the PSCC 2022, Porto, Portugal, 27 Jun 2022-01 Jul 2022.

8. Geth, F., Heidari, R., **Koirala, A.** (2022). *Computational Analysis of Impedance Transformations for Four-Wire Power Networks with Sparse Neutral Grounding.* In: e-Energy '22: Proceedings of the Thirteenth ACM International Conference on Future Energy Systems, (105-113). Presented at the ACM e-Energy 2022, Virtual, 29 Jun 2022-01 Jul 2022. New York, NY, United States. ISBN: 978-1-4503-9397-3.

9. Hashmi, M.U., **Koirala, A.**, Ergun, H., Van Hertem, D. (2021). *Flexible and curtailable resource activation in a distribution network using nodal sensitivities.* In: 2021 International Conference on Smart Energy Systems and Technologies (SEST), (1-6). Presented at the SEST 2021, Vaasa, Finland, 06 Sep 2021-08 Sep 2021. ISBN: 978-1-7281-7660-4.

10. **Koirala, A.**, Van Acker, T., Van Hertem, D., Camargo, J., D'hulst, R. (2020). *General Polynomial Chaos vs Crude Monte Carlo for Probabilistic Evaluation of Distribution Systems.* In: 2020 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS), (1-6). Presented at PMAPS 2020, Liege, Belgium, 18 Aug 2020-21 Aug 2020. ISBN: 978-1-7281-2822-1.

11. **Koirala, A.**, D'hulst, R., Van Hertem, D. (2019). *Impedance modelling for European style Distribution Feeder.* In: 2019 International Conference on Smart Energy Systems and Technologies (SEST), (1-1). Presented at the SEST 2019, Porto, Portugal, 09 Sep 2019-11 Sep 2019. ISBN: 978-1-7281-1156-8



FACULTY OF ENGINEERING SCIENCE
DEPARTMENT OF ELECTRICAL ENGINEERING
ELECTRICAL ENERGY SYSTEMS AND APPLICATIONS (ELECTA)
Kasteelpark Arenberg 10, Bus 2445
B-3001 Leuven
arpan.koirala@kuleuven.be
https://www.esat.kuleuven.be/electa