

# **Uncertainty Quantification via Polynomial Chaos Expansion – Methods and Applications for Optimization of Power Systems**

Zur Erlangung des akademischen Grades eines

**Doktor-Ingenieurs**

von der KIT-Fakultät für Informatik  
des Karlsruher Instituts für Technologie (KIT)

genehmigte

**Dissertation**

von

**M. Sc. Tillmann Mühlpfordt**

Tag der mündlichen Prüfung:

10. Dezember 2019

Hauptreferent:  
Korreferenten:

Prof. Dr. Veit Hagenmeyer  
Prof. Dr.-Ing. Timm Faulwasser  
Prof. Daniel Bienstock, PhD





This document is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0):  
<https://creativecommons.org/licenses/by/4.0/deed.en>



# Acknowledgment

Any piece of work is the child of its context. I have spent the last four years at the Institute for Automation and Applied Informatics (IAI) at the Karlsruhe Institute of Technology. My motivation to join IAI was to work with Veit Hagenmeyer and Timm Faulwasser—my supervisors to whom I owe a great deal when it comes to practicing science from day to day. Thank you for the inspiration, the discussions, the insights, the advice. Though I may not see the full picture, I am aware that I had tremendous scientific and non-scientific freedoms thanks to your efforts. I also thank Daniel Bienstock for agreeing to be my external reviewer.

When I joined IAI in July 2015, the optimization and control group was a tuple consisting of Timm Faulwasser and myself. I was lucky enough to see this group grow: Alexander Engelmann was our first Master's student, eventually becoming my office mate pursuing a doctoral degree of his own. Not only did I inherit an apartment from Alex, but he also brought the Julia programming language and programming paradigms as such to my attention. In the office across the hall there sat my fellow doctoral candidates Riccardo Appino and Alexander Murray. Man, did we have a lot of discussions about TV series and sneak previews. Florian Becker was the first Master student I supervised. I was perplexed to see: there really are people who use vim to code. Frederik Zahn, first a research assistant then a Master's student, became not just a helpful contributor to our Julia package but an honorable friend.

But there is more to IAI than the group I was part of: first and foremost I have to thank Lutz Gröll. Without his support for mathematical rigor I am not sure I would have dared to open the door to Hilbert spaces. Throughout the years he has been a constant source of help and insight, and an inspiration. The Institute is lucky to have you. Other people from IAI whom I would like to thank are (alphabetical order): Eric Braun, Tom Brown, Claudia Greceanu, Andreas Hofmann, Bernadette Lehmann, Jianlei Liu, Jörg Matthes, Ralf Mikut, Fabian

Neumann, Tessina Scholl, Simon Waczowicz. You contribute to a culture at IAI that makes it a pleasant place to work. Keep it up.

Collaborations beyond the own and familiar institute are the salt of the PhD-endeavor. I had outstanding scientific collaborators from the USA (Line Roald, Sidhant Misra, Daniel Molzahn) and from China (Yuning Jiang, Boris Houska). These days I feel compelled to emphasize that a collaboration across national borders does not just mean working toward a publication, but fostering intercultural exchange, opening our eyes to unfamiliar yet perhaps interesting or insightful traditions and habits. My research stays at the Los Alamos National Laboratory and at the École Polytechnique Fédérale de Lausanne are prime examples that research is always a cultural exchange too.

Pursuing a doctoral degree is as much a professional effort as it is a personal endurance training. It is in these times I realized (once again) what a unique family I grew up in: thank you Mama, thank you Papa for equipping me with curiosity and perseverance. Regardless of whether my mood was euphoric or pessimistic I knew I could rely on you, Stefanie. Not only did we start a life as a married couple, you also gave life to our wonderful son Jonathan Paul. Yes, the last four years were remarkable.

September 2019

Tillmann Mühlpfordt

# Abstract

Fossil fuels paved the way to prosperity for modern societies, yet alarmingly, we can exploit our planet’s soil only so much. Renewable energy sources inherit the burden to quench our thirst for energy, and to reduce the impact on our environment simultaneously. However, renewables are inherently volatile; they introduce uncertainties. What is the effect of uncertainties on the operation and planning of power systems? What is a rigorous mathematical formulation of the problems at hand? What is a coherent methodology to approaching power system problems under uncertainty? These are among the questions that motivate the present thesis that provides a collection of methods for uncertainty quantification for (optimization of) power systems.

We cover power flow (PF) and optimal power flow (OPF) under uncertainty (as well as specific derivative problems). Under uncertainty—we view “uncertainty” as continuous random variables of finite variance—the state of the power system is no longer certain, but a random variable. We formulate PF and OPF problems in terms of random variables, thusly exposing the infinite-dimensional nature in terms of  $L^2$ -functions. For each problem formulation we discuss a solution methodology that renders the problem tractable: we view the problem as a mapping under uncertainty; uncertainties are propagated through a known mapping. The method we employ to propagate uncertainties is called polynomial chaos expansion (PCE), a Hilbert space technique that allows to represent random variables of finite variance in terms of real-valued coefficients.

The main contribution of this thesis is to provide a rigorous formulation of several PF and OPF problems under uncertainty in terms of infinite-dimensional problems of random variables, and to provide a coherent methodology to tackle these problems via PCE. As numerical methods are moot without numerical software another contribution of this thesis is to provide *PolyChaos.jl*: an open source software package for orthogonal polynomials, quadrature rules, and PCE written in the Julia programming language.



# Deutsche Kurzfassung

Fossile Brennstoffe ebneten modernen Gesellschaften den Weg zum Wohlstand; der Energievorrat unseres Planeten ist allerdings begrenzt. Den erneuerbaren Energien kommt nun die Bürde zu, den Durst nach Energie zu stillen und dabei den Einfluss auf die Umwelt zu verringern. Jedoch – erneuerbare Energien sind volatil: wann Sonne oder Wind zur Verfügung stehen, ist nie *genau* bekannt. Was ist der Einfluss erneuerbarer Energien auf den Betrieb des elektrischen Netzes? Was ist eine mathematische Beschreibung dieses Problems? Existiert eine strukturierte Vorgehensweise, um elektrische Netze unter Unsicherheiten zu berechnen? Diese Fragen bilden die Grundlage der vorgelegten Dissertation; sie umfasst Herangehensweisen und Methoden, um die Wirkung von Unsicherheiten auf den (optimalen) Betrieb elektrischer Netze zu quantifizieren.

Insbesondere werden (optimale) Lastflussprobleme unter Unsicherheiten untersucht. Unter Unsicherheiten lässt sich der Gesamtzustand des Netzes als Zufallsvariable endlicher Varianz modellieren. Die Dissertation zeigt, wie (optimale) Lastflussprobleme unter Unsicherheiten als Optimierungsprobleme über Zufallsvariable formuliert werden können, also als unendlichdimensionale Probleme in  $L^2$ -Räumen. Fernerhin wird ein einheitlicher Zugang vorgestellt, um jene Probleme in endlichdimensionale zu überführen: sämtliche Probleme werden als Abbildungen unter Unsicherheiten aufgefasst, welche durch eine polynomiale Chaoserweiterung (PCE) angegangen werden können. Besagte Hilbertraum-Methode erlaubt, Zufallsvariable endlicher Varianz durch reellwertige Koeffizienten darzustellen.

Der Hauptbeitrag der Dissertation besteht in der rigorosen Problemformulierung als auch der strukturierten PCE-basierten Methodik, die diverse (optimale) Lastflussprobleme unter Unsicherheiten zu lösen vermag. Überdies stellt die Dissertation mit *PolyChaos.jl* eine Software in der Programmiersprache Julia zur Verfügung, mit welcher orthogonale Polynome, Quadraturen und polynomiale Chaoserweiterungen berechnet werden können.



# Content

<b>Acknowledgment</b> . . . . .	<b>i</b>
<b>Abstract</b> . . . . .	<b>iii</b>
<b>Deutsche Kurzfassung</b> . . . . .	<b>v</b>
<b>Notation</b> . . . . .	<b>xi</b>
<b>Acronyms</b> . . . . .	<b>xiii</b>
<b>1 Introduction</b> . . . . .	<b>1</b>
<b>2 Mappings under uncertainty</b> . . . . .	<b>11</b>
2.1 Problem formulation – Idea . . . . .	11
2.2 Preliminaries . . . . .	15
2.2.1 Measure theory . . . . .	15
2.2.2 Random variables . . . . .	18
2.2.3 Hilbert space theory . . . . .	20
2.3 Problem formulation – Revisited . . . . .	24
2.4 Polynomial chaos . . . . .	25
2.4.1 Orthogonal polynomials . . . . .	25
2.4.2 Polynomial chaos expansion . . . . .	30
2.4.3 Advantages . . . . .	34
2.4.4 Disadvantages . . . . .	38
2.5 Quantifying truncation errors . . . . .	39
2.5.1 Polynomial mappings . . . . .	40
2.5.2 Quadratic programs . . . . .	42
2.5.3 Non-polynomial mappings . . . . .	45

2.5.4	Selected examples . . . . .	46
2.6	Problem formulation – Finalized . . . . .	49
<b>3</b>	<b>Power flow under uncertainty . . . . .</b>	<b>55</b>
3.1	Problem formulation . . . . .	57
3.2	Existing approaches . . . . .	63
3.3	Solution methodology . . . . .	64
3.3.1	Discussion . . . . .	68
3.4	Backward-forward sweep . . . . .	69
3.4.1	Problem formulation . . . . .	69
3.4.2	Existing approaches . . . . .	72
3.4.3	Solution methodology . . . . .	73
3.4.4	Discussion . . . . .	76
3.5	Optimal adaptive linearizations . . . . .	77
3.5.1	Problem formulation . . . . .	78
3.5.2	Existing approaches . . . . .	80
3.5.3	Solution methodology . . . . .	81
3.5.4	Discussion . . . . .	83
<b>4</b>	<b>Optimal power flow under uncertainty . . . . .</b>	<b>85</b>
4.1	Chance-constrained AC optimal power flow . . . . .	86
4.1.1	Problem formulation . . . . .	86
4.1.2	Existing approaches . . . . .	92
4.1.3	Solution methodology . . . . .	95
4.1.4	Discussion . . . . .	100
4.2	Chance-constrained DC optimal power flow . . . . .	101
4.2.1	Problem formulation . . . . .	101
4.2.2	Existing approaches . . . . .	105
4.2.3	Solution methodology . . . . .	107
4.2.4	Discussion . . . . .	109
4.3	Multi-stage chance-constrained DC optimal power flow . . . . .	110
4.3.1	Problem formulation . . . . .	110
4.3.2	Existing approaches . . . . .	115
4.3.3	Solution methodology . . . . .	116
4.3.4	Discussion . . . . .	120

---

<b>5 PolyChaos.jl</b> . . . . .	<b>123</b>
5.1 Existing software . . . . .	125
5.2 Type hierarchy . . . . .	126
5.3 Tutorial example . . . . .	134
<b>6 Case studies</b> . . . . .	<b>137</b>
6.1 Setup . . . . .	138
6.2 Probabilistic power flow . . . . .	139
6.3 Chance-constrained AC optimal power flow . . . . .	143
6.4 Chance-constrained DC optimal power flow . . . . .	147
6.5 Comparison . . . . .	150
<b>7 Summary and outlook</b> . . . . .	<b>151</b>
<b>A Appendix</b> . . . . .	<b>157</b>
A.1 Orthogonal polynomials . . . . .	157
A.2 Gauss quadrature . . . . .	160
A.3 Number of basis polynomials . . . . .	163
A.4 Polynomial chaos and stochastic processes . . . . .	163
A.5 Bus admittance matrix . . . . .	167
A.6 AC and DC power flow equations . . . . .	170
A.7 Backward-forward sweep method . . . . .	173
<b>References</b> . . . . .	<b>178</b>
<b>List of publications</b> . . . . .	<b>190</b>
<b>Curriculum vitæ</b> . . . . .	<b>191</b>



# Notation

$\mathbb{N}$	natural numbers
$\mathbb{R}, \mathbb{R}_{\geq 0}$	real numbers, non-negative real numbers
$\mathbb{R}^n$	$n$ -dimensional Euclidean space
$\mathbb{1}$	indicator function
$0_{n \times m}, 0_n$	zero matrix, zero vector of given dimension
$\delta_{ij}$	Kronecker-delta

## Specific distributions

$\mathbb{N}$	normal distribution
$\mathbb{B}$	beta distribution
$\mathbb{U}$	uniform distribution

## Measure theory, random variables

$\Omega$	sample space
$\mathfrak{F}$	sigma algebra
$\mathfrak{B}$	Borel sigma algebra
$\mu$	measure
$\rho$	Lebesgue density or probability density
$P, \mathbb{P}$	probability distribution, see (2.9)
$x$	random variable, vector-valued random variable
$\mathbb{E}(x)$	(element-wise) expected value of $x$
$\mathbb{V}(x)$	(element-wise) variance of $x$
$\mathbb{C}(x, y)$	covariance of $x$ and $y$

## Hilbert spaces

$\langle \cdot, \cdot \rangle$	scalar product
$\  \cdot \ $	2-norm
$\Pi$	projection operator
$L^2(\Omega, \mu; \mathbb{R})$	Hilbert space of random variables

Orthogonal polynomials

$k$	index
$\mathcal{K}$	set of indices
$\phi_k$	$k^{\text{th}}$ basis polynomial

Power systems

$N_b$	number of buses
$N_{br}$	number of lines
$\mathcal{N}$	index set of buses
$\mathcal{N}_{PV}$	index set of PV buses
$\mathcal{N}_{PQ}$	index set of PQ buses
$\mathcal{N}_{SL}$	index set of slack bus(es)
$\mathcal{L}, \mathcal{L}^{\text{lin}}$	index sets of lines
$p_i, p_i$	active power at bus $i$
$q_i, q_i$	reactive power at bus $i$
$v_i, v_i$	voltage magnitude at bus $i$
$\delta_i, \delta_i$	voltage angle at bus $i$
$e_i, e_i$	real voltage at bus $i$
$f_i, f_i$	imaginary voltage at bus $i$
$z_i, z_i$	state of bus $i$
$p^{\text{br}}, p^{\text{br}}$	active power branch flow
$q^{\text{br}}, q^{\text{br}}$	reactive power branch flow
$g(\cdot), g_{\text{PCE}}(\cdot)$	power flow equations
$h(\cdot), h_{\text{PCE}}(\cdot)$	bus specifications
$p_i^c$	controllable active power at bus $i$
$p_i^u$	uncontrollable active power at bus $i$
$q_i^c$	controllable reactive power at bus $i$
$q_i^u$	uncontrollable reactive power at bus $i$

# Acronyms

AC	alternating current
BFS	backward-forward sweep
CC	chance constraint
CC-OPF	chance-constrained optimal power flow
DC	direct current
KKT	Karush Kuhn Tucker
KL	Karhunen-Loève
LTI	linear time-invariant
MPC	model predictive control
NLP	nonlinear program
OPF	optimal power flow
PCE	polynomial chaos expansion
PDF	probability density function
PF	power flow
PPF	probabilistic power flow
PTDF	power transfer distribution factor
QP	quadratic program
SOCP	second-order cone program



# 1 Introduction

The world's supply with electrical power has come a long way since Faraday, Henry, and Zantedeschi independently discovered the effect of electromagnetic induction in the early 1830s. What the citizens of industrialized countries take for granted today—an uninterrupted, reliable, and cheap access to electricity—is the outcome of two world wars, unforeseen economic growth, and of course tremendous efforts of ever-curious scientists and engineers. The current power systems are impressive results of craftsmanship, planning, and engineering. However, there is a key player missing: their exceptional role for society makes power systems inadvertently political. For instance, here is how regulations changed in Germany in the last 20 years:

- On April 29, 1998, the new version of the *Gesetz über die Elektrizitäts- und Gasversorgung* enabled the so-called liberalization of the German electricity market: customers are free to choose their electricity provider [29].
- Since April 01, 2000, the *Gesetz für den Ausbau erneuerbarer Energien* guarantees feed-in tariffs for renewable energy sources [30].
- On August 06, 2011, the *13. Gesetz zur Änderung des Atomgesetzes* demanded the shut-down of all nuclear power plants by the end of 2022 [31].
- On January 26, 2019, the *Kommission für Wachstum, Strukturwandel und Beschäftigung* with its final report recommended to phase out of coal power by 2038 [93].

These policies have induced a paradigm change in the planning and operation of (German) power systems. For instance, in Germany the share of renewable energy production rose from 8.8 % in 2002 to 40.6 % in 2018. In the same time, the installed capacity for solar power increased from 0.3 GW to 45.9 GW in

Germany; for onshore wind it rose from 12.0 GW to 53.0 GW.<sup>1</sup> The installation of solar panels and wind parks affects different voltage levels of the power systems. In turn, this affects the transmission system operators and may induce re-dispatches, that is power plants have to account for possible shortages of electric power. In Germany, the re-dispatched energy grew from 4,956 GWh in 2012 to 20,438 GWh in 2017.<sup>2</sup> The installation of renewables is not a German but a global phenomenon: we observe similar endeavors—with similar consequences—across Europe [12], in the USA [1], in Mexico [13], in India [89], and in Japan [132].

There is an interesting side effect to the undeniable growth in the already complex challenges in power systems: diverse research fields discover their interest in the topic. For instance, the field of systems and control dedicates entire sessions to *power systems* at premier conferences such as the IEEE Conference on Decision and Control. Computer scientists have even coined a new term: *energy informatics* [68], highlighting that it takes an effort of a broad and interdisciplinary research community to enable the transition toward sustainable power systems.

Abstractly speaking, the installation of renewable energy sources and the liberalization of the electricity market *introduce uncertainty*: it is no longer known with certainty who will produce how much energy where and when exactly. Clearly, uncertainty was present all along in power systems, but the necessity to consider uncertainties in a structured and explicit manner has become more relevant. Exactly this is the seed around which this thesis agglomerates: power systems under uncertainty. We tackle questions such as: Given a power system, and given an uncertainty description of, for instance, a load, how does this affect the overall state of the power system? What is the probability that some lines are overloaded? How can we operate a power system cost-optimally despite the presence of uncertainties? How to compute and quantify the effect of uncertainties on power systems?

---

<sup>1</sup> Figures taken from [https://www.energy-charts.de/energy\\_pie\\_de.htm](https://www.energy-charts.de/energy_pie_de.htm) retrieved August 13, 2019, 16:15.

<sup>2</sup> Figures taken from [https://www.bundesnetzagentur.de/DE/Sachgebiete/ElektrizitaetundGas/Unternehmen\\_Institutionen/Versorgungssicherheit/Engpassmanagement/Redispatch/Redispatch\\_2017.jpg](https://www.bundesnetzagentur.de/DE/Sachgebiete/ElektrizitaetundGas/Unternehmen_Institutionen/Versorgungssicherheit/Engpassmanagement/Redispatch/Redispatch_2017.jpg), retrieved August 13, 2019, 16:25.

---

Motivated by the celebrated Don't Repeat Yourself (DRY) principle from software engineering [87] this thesis aims for a unifying methodology to analyze, tackle, and solve specific problems of power systems under uncertainty.<sup>3</sup> To realize DRY means to introduce abstraction, in our case leading to the study of mappings under uncertainty. That is, given a mapping, and given a random variable, how can we characterize the image? For us this given mapping resembles (optimal) power flow problems.

## Outline and contributions

The content of this thesis interconnects according to Figure 1.1. Chapter 2 forms the theoretical and methodological baseplate: starting from an infinite-dimensional problem formulation that represents a mapping under uncertainty in terms of random variables, we show how a method called polynomial chaos expansion (PCE) helps in rendering the problem finite-dimensional, irrespective of whether the uncertainties are Gaussian, non-Gaussian, or a combination thereof. Before we apply PCE we cover its surrounding theory: measure theory, Hilbert space theory, and orthogonal polynomials. On top of the baseplate provided by Chapter 2 we erect two columns, each standing for a specific class of problems from power systems: Chapter 3 covers power flow problems under uncertainty, and Chapter 4 covers optimal power flow problems under uncertainty. Both Chapter 3 and Chapter 4 are built bottom-up: starting from the generic problem (probabilistic power flow (PPF) for Chapter 3, respectively chance constraint (CC)-alternating current (AC)-OPF for Chapter 4), we cover two special cases (backward-forward sweep (BFS) and optimal adaptive linearizations for Chapter 3, respectively single and multi-stage direct current (DC)-OPF for Chapter 4). In both Chapter 3 as well as Chapter 4 we adhere to the formalism from the methodological Chapter 2. We show how the same principles and the same workflow can be applied to different problems, all of which lead to finite-dimensional problems in terms of PCE coefficients.

The case studies from Chapter 6 are to Chapter 3 and Chapter 4 what an arch is to the columns it rests upon: it connects, it bridges a gap, and it orna-

---

<sup>3</sup> The opposite of DRY being—of course—Write Everything Twice (WET) [87].

ments. Chapter 6 shows how power systems problems under uncertainty can be tackled with PCE. We study PPF, AC-OPF, and DC-OPF for the same grid, the same uncertainty, and the same constraints. What is the role of Chapter 5? It literally provides the tool to construct and implement our problem formulations: the software package *PolyChaos.jl* written in the Julia programming language. *PolyChaos.jl* facilitates the computation of orthogonal polynomials, quadrature rules, and PCEs. We conclude with Chapter 7, giving a summary and outlining open issues.

Let us examine the content and the individual contributions of the chapters more closely.

## **Chapter 2 – Mappings under uncertainty**

This chapter introduces the problem formulation that permeates the entire thesis: mappings under uncertainty. That means, given a mapping, and given a random variable, how to characterize the image of the random variable under the mapping? We study this problem by means of a Hilbert space technique called polynomial chaos expansion (PCE): PCE is to a random variable what a Fourier series is to a periodic signal, namely a representation of an infinite-dimensional mathematical object in terms of scalar coefficients. Chapter 2 introduces the relevant concepts from measure theory, Hilbert space theory, and orthogonal polynomials. These three fields are connected such that their relation to PCE becomes obvious. The amalgamation of established material to form a self-contained and concise problem formulation is a pedagogical and methodological contribution of Chapter 2. The main scientific contribution lies with Section 2.5, the quantification of ever-present truncation errors for PCE, specifically for polynomial and non-polynomial explicit mappings [122, 123]. Insightful examples complement the chapter.

## **Chapter 3 – Power flow under uncertainty**

The abstract problem formulation from Chapter 2 is applied to concrete power systems problems, namely power flow problems under uncertainty. Solving the probabilistic power flow (PPF) problem covered in Section 3.1 means to determine the state of an electrical grid in terms of random vari-

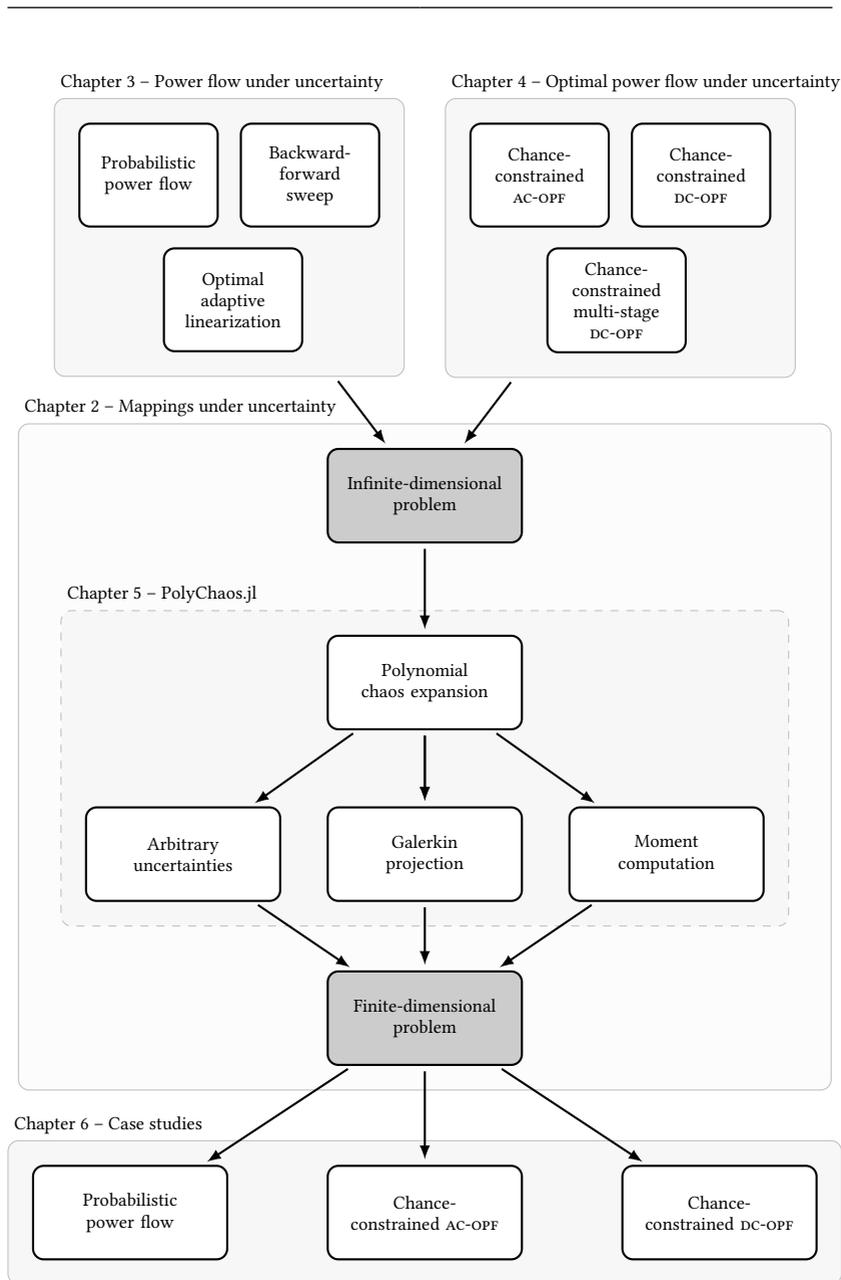


Figure 1.1: Content of this thesis as a block diagram. Chapter 5 has a special role, because it provides the numerical software *PolyChaos.jl*.

ables: given an electrical grid under AC conditions governed by Kirchhoff's laws, and given, for instance, a probabilistic description of the load at one bus of the power system, how does this affect the flows among all lines? How does this affect the voltage magnitudes at all buses? Mathematically, PPF leads to a system of nonlinear equations in terms of random variables, which can be rendered to an enlarged system of nonlinear equations in terms of the PCE coefficients. One set of equations is given by the so-called power flow equations that model the underlying physics. The other set of equations models equipment behavior, the so-called bus specifications.

For radial grids, i.e. power systems whose underlying graph is a tree, the PPF can be simplified, leading to the so-called backward-forward sweep (BFS) method that is covered in Section 3.4. The BFS method is an iterative method that switches between applying Kirchhoff's current law and Kirchhoff's voltage law. The iteration between nodal and branch quantities is possible mathematically, because the reduced incidence matrix<sup>4</sup> of a tree is invertible.

A slightly different kind of power flow under uncertainty is the subject of Section 3.5. Perhaps a full PPF problem is too costly to solve, and we are interested only in a sufficiently accurate proxy. In this context "sufficiently accurate" refers to a setting in which we model the operating range of a power system in terms of random variables and we seek an affine input-to-output mapping that is optimal with respect to a specific error metric. This is the idea of optimal adaptive linearizations that combine the computational prowess of affine mappings with optimization under uncertainty.

The advantages of PCE for all of the considered PPF problems are evident: our setting is not restricted to a specific class of uncertainties, such as Gaussian random variables. Furthermore, we require no sampling to obtain the solution. Instead, we formulate a modified problem in terms of the PCE coefficients that needs to be solved just once, then providing all the information we sought. Also, we do not need to linearize the power flow equations; we employ the full nonlinear equations.

To summarize our contributions:

---

<sup>4</sup> That is the quadratic incidence matrix of the tree with the root column removed.

- 
- rigorous formulation of power flow problem as a mapping under uncertainty in terms of arbitrary random variables of finite variance [119];
  - exposition of the infinite-dimensional character of the PPF problem [119];
  - one of the first applications of intrusive PCE to PPF [119];
  - first application of intrusive PCE to BFS [9];
  - first application of PCE to optimal adaptive linearizations [125].

## Chapter 4 – Optimal power flow under uncertainty

In this chapter we turn to optimal power flow (OPF) under uncertainty, the task of operating a power system optimally despite possible fluctuations in load and/or generation. Compared to the PPF problem from Chapter 2, OPF is a numerical optimization problem.<sup>5</sup> In Section 4.1 we formulate the generic problem under AC conditions, putting great emphasis on how to reformulate the deterministic OPF problem in the presence of uncertainties, leading to chance-constrained optimal power flow (CC-OPF) formulations. This is necessary because it is not clear per se how inequality constraints are to be treated in the presence of uncertainties. We also show how PCE renders the infinite-dimensional problem finite-dimensional. Under the simplifying DC conditions Section 4.2 derives a similar chance-constrained problem formulation, and shows how its solution simplifies when using PCE. The multi-stage setting, i.e. the problem of solving OPF over a time horizon, is covered in Section 4.3.

Similar to the PPF problems from Chapter 2 we highlight the advantages of PCE for chance-constrained OPF problems: Our setting is not restricted to a specific class of uncertainties, such as Gaussian random variables. No sampling is required to obtain the solution. Instead, we formulate a modified problem in terms of the PCE coefficients that needs to be solved just once, then providing all the information we sought. We do not rely on linearizations of the power flow equations or convex reformulations thereof; we propagate un-

---

<sup>5</sup> Strictly speaking PPF is a specific numerical optimization problem, namely a feasibility problem. The point is: OPF always involves numerical optimization.

certainties through the fully nonlinear power flow equations. Finally, PCE is a natural framework for moment-based reformulations of chance constraints, because moments can be computed directly from the PCE coefficients.

To summarize our contributions:

- rigorous formulation of chance-constrained OPF as a mapping under uncertainty in terms of random variables [55, 117–121, 125];
- exposition of the infinite-dimensional character of the OPF problem [55, 117–121, 126];
- first application of intrusive PCE to chance-constrained OPF for both AC and DC settings [55, 117–121, 125, 126];
- first solution of CC-OPF for the fully nonlinear AC power flow equations [55, 119, 126];
- rigorous formulation of multi-stage DC-OPF in the presence of Gaussian uncertainties in terms of first and second moments.

## Chapter 5 – PolyChaos.jl

Reliable and efficient software forms the backbone of numerics-driven research. Chapter 5 is about *PolyChaos.jl*, a software package written in the Julia programming language [19] for orthogonal polynomials, quadrature rules, and PCE. As there existed no PCE package in Julia it was up to the author of this thesis to co-create one. Clearly, we could have resorted to existing PCE packages in Matlab or Python. However, the Julia programming language offers too many intriguing features for scientific computing to not witness a PCE package in due time anyway. The focus of *PolyChaos.jl* is on intrusive PCE with an efficient computation of the tensors of scalar products. The open source code is available free of charge [127]. An example shows how to construct orthogonal polynomials for a Gaussian mixture with *PolyChaos.jl*.

To summarize our contributions:

- first open source Julia package for orthogonal polynomials, quadrature rules, and PCE [127];

- 
- given an absolutely continuous, non-negative measure the package *PolyChaos.jl* allows to
    - ▷ compute recurrence relations of underlying orthogonal polynomials via Stieltjes procedure, Lanczos procedure, or multiple discretization;
    - ▷ compute quadrature rules via Gauss, Gauss-Radau, or Gauss-Lobatto quadrature;
    - ▷ compute tensorized scalar products of basis functions for intrusive PCE;
  - comprehensible documentation.

## Chapter 6 – Case studies

Chapter 3 and Chapter 4 provide methodologies, Chapter 5 provides a numerical tool. The purpose of Chapter 6 is to demonstrate all of the above in action, applied to the same power system. Starting from a non-trivial uncertainty model (Gaussian mixture) we use *PolyChaos.jl* to compute the orthogonal basis, then solve the PPF problem, and then compare AC-OPF to DC-OPF under uncertainty. The contribution of Chapter 6 is to study three different (optimal) power flow problems under uncertainty for the same grid and the same uncertainty description, also providing a thorough comparison of the solution characteristics.



## 2 Mappings under uncertainty

The concept of a function in mathematics is as simple as it is profound: elements are mapped to other elements according to some rule. The usual vocabulary surrounding functions suggests a visual interpretation: the elements that a function takes as inputs are said to be *pre-images*, and the function renders them to *images*. Imagine we select a collection of pre-images such that certain elements of the collection appear more frequently than others. Mapping all the elements of the collection of pre-images, what can we say about the resulting collection of images—this is the question we attempt to answer.<sup>1</sup>

### 2.1 Problem formulation – Idea

Let  $\mathcal{X}$  be the pre-image (or domain), and let  $\mathcal{Y}$  be the image (or codomain). The function  $f$  associates elements of  $\mathcal{X}$  with elements of  $\mathcal{Y}$ , and we write

$$f : \mathcal{X} \rightarrow \mathcal{Y}, \quad x \mapsto y = f(x). \quad (2.1)$$

We are interested in applying  $f$  to specific sets

$$\tilde{\mathcal{X}} = \{x \in \mathcal{X} : x \text{ is a realization of the random variable } x\}, \quad (2.2)$$

where we assume the random variable  $x$ , and hence  $\tilde{\mathcal{X}}$ , are known to us. If we map these specific sets (2.2) through the function  $f$ , we obtain the images

$$\left\{ y \in \mathcal{Y} : y = f(x) \forall x \in \tilde{\mathcal{X}} \right\}. \quad (2.3)$$

---

<sup>1</sup> We shall use the terms *mapping*, *function*, and *map* synonymously [10]. Also, we exclude set-valued functions from our investigations.

These images depend on both the random variable  $x$  and the mapping  $f$ .

In lighter—yet for now imprecise and purely conceptual—notation this reads

$$y = f(x), \tag{2.4}$$

meaning that the random variable  $x$  is propagated via the function  $f$  to the random variable  $y$ . We emphasize that the mapping  $f$  itself is deterministic (or causal)—identical pre-images yield identical images—but it is uncertain which pre-images will be mapped at all. This problem at hand is well-known in the field of uncertainty quantification [95, 154, 172].

**Problem 2.1** (Mapping under uncertainty). *Given a function  $f$  that maps elements of the domain  $\mathcal{X}$  to the codomain  $\mathcal{Y}$ , and given the random variables  $x = [x_1, \dots, x_{n_x}]^\top$  for  $n_x \in \mathbb{N}$  with known  $\tilde{\mathcal{X}} \subseteq \mathcal{X}$ , find the random variables  $y = [y_1, \dots, y_{n_y}]^\top$  for  $n_y \in \mathbb{N}_y$  that are the image of  $x$  under  $f$ . We formally write  $y = f(x)$ .*

Let us study a concrete example.

**Example 2.1** (Chi-squared distribution). *Let  $\mathcal{X} = \tilde{\mathcal{X}} = \mathbb{R}^n$  for some finite  $n \in \mathbb{N}$ , and  $\mathcal{Y} = \mathbb{R}_{\geq 0}$  both of which are linked by the mapping*

$$y = f(x_1, \dots, x_n) = \sum_{i=0}^n x_i^2.$$

*We treat each  $x_i$  as the realization of a standard Gaussian random variable for which we write  $x_i \sim \mathcal{N}(0, 1)$  for  $i \in \{1, \dots, n\}$ . Hence, we are interested in the random variable  $y$  that is the image of the sum of the squared  $x_i$*

$$y = f(x_1, \dots, x_n) = \sum_{i=0}^n x_i^2.$$

In Example 2.1 we are fortunate enough to have an explicit expression for the function  $f$ . More often than not this will not be the case. The function  $f$  might, for example, be a black box to which we have no access: a proprietary piece of software for instance. The function  $f$  may not admit an explicit form at all: it may be an implicit function, it may be the solution to an ordinary

differential equation to which no analytic solution is known, it may be a discretized solution of a partial differential equation—all of the above scenarios are common in the field of uncertainty quantification [95, 154, 172]. In this thesis the function  $f$  often represents the solution to an optimization problem, and we are interested in solving a family of optimization problems for a range of uncertain parameter values. From here on we assume that the argmin of the optimization problems is not set-valued.

**Example 2.2** (Quadratic program with uncertain data). *Let  $\mathcal{X} = \mathbb{R}$  and  $\mathcal{Y} = \mathbb{R}^2$ . For a positive definite matrix  $A \in \mathbb{R}^{2 \times 2}$  consider the quadratic program*

$$\begin{aligned} y = f(x) := \operatorname{argmin}_{z \in \mathbb{R}^2} & \quad \frac{1}{2} z^\top A z \\ \text{s. t.} & \quad \begin{bmatrix} 1 & 1 \end{bmatrix} z = x, \end{aligned} \tag{2.5}$$

where  $x$  is the realization of a Beta random variable  $x \sim \mathbf{B}([\underline{x}, \bar{x}], \alpha, \beta)$  with so-called support  $\tilde{\mathcal{X}} = [\underline{x}, \bar{x}]$  where  $0 < \underline{x} < \bar{x} < \infty$  and positive shape parameters  $\alpha, \beta$ . Hence, find  $y = f(x)$ .<sup>2</sup>

Example 2.2 makes it clear that the conceptually simple “ $y = f(x)$ ” requires care: if we merely substituted  $x \leftarrow x$  in (2.5), then the very meaning of the equality constraint would be questionable; we would have “ $[1 \ 1]z = x$ ”, which is not meaningful as there is a real number on the left-hand side, and a random variable on the right-hand side. Hence, it is imperative to distinguish the function from the uncertainty: the optimization problem (2.5) is a function that maps a value  $x$  to an optimal value  $y$ , and it is this function that shall be evaluated for realizations of a random variable.

It is natural to ask: How can we solve Problem 2.1? What are existing methods?

Perhaps the most intuitive approach is sampling: evaluate the function  $f$  for each realization of the uncertainty  $x$ . Having solved a plethora of deterministic problems, the ensemble of the outputs provides statistical information. The simplicity of this so-called Monte-Carlo (MC) approach is both its strength and its weakness. It “just” requires a reliable sampler, the imple-

<sup>2</sup> From here on we no longer specify  $\tilde{\mathcal{X}}$  explicitly. It will be clear from the context.

mentation is then conceptually simple and amenable to parallelization. Existing code can be re-used, no additional massaging of the problem is required. However, the mean value of the output ensemble converges in proportion to  $1/\sqrt{N}$  with  $N$  being the number of samples [172, p. 3]; the same applies to the convergence of the variance [95, p. 9]. Hence, mc typically requires a lot of samples. Several sampling methods have been proposed in the literature to speed up convergence [95, 154, 173].

Perturbation methods are an alternative to sampling-based approaches [172, p. 3]: the function  $f$  is expanded around the mean value of  $x$  in terms of an (often second-order) Taylor series. The surrogate model may be used to derive analytic expressions of the mean and variance. Clearly, this technique is valid only in a neighborhood of the mean value of  $x$ , hence no large deviations should occur.

Moment-based approaches, instead, compute solutions in terms of moments: the input uncertainty  $x$  is given in terms of its moments, and we are interested in the moments of  $y$  [172, p. 4]. Hence, neither the input random variable  $x$  nor the desired output random variable  $y$  is modelled by its probability density function. The derivation of the moments, however, often exhibits a dependency on higher moments, leading to the so-called moment closure problem [94].

In this thesis we study mappings of the form  $y = f(x)$  by means of a Hilbert space technique called polynomial chaos expansion (pce). Besides having appealing convergence properties, pce is a powerful tool that allows to propagate uncertainties through mappings in a single step. The idea is to expand random variables as a linear combination of orthogonal basis functions weighted by deterministic coefficients. It is these coefficients that characterize the random variable, and it is these coefficients we need to compute. The canvas for pce is weaved of measure theory, Hilbert space theory, and orthogonal polynomials—all three of which we touch upon in the following.

## 2.2 Preliminaries

In view of polynomial chaos expansion (PCE) the basic ingredients we need to formulate Problem 2.1 are measure theory, Hilbert space theory, and orthogonal polynomials. The material is assembled from [82–84, 154].

### 2.2.1 Measure theory

Given a collection  $\Omega$ —a so-called sample space—we want to assign numbers to elements (or combinations of elements) of  $\Omega$ . These numbers may have a physical meaning such as length, area, volume or mass, but they may as well stand for the abstract idea of a probability. We call this act of assignment a *measurement*. But what does it mean *to measure*? The verb *to measure* is a transitive verb. In its active form *to measure* requires a subject that does the action of *measuring*, and it acts on an receiving object to which the act of measuring is done to. In mathematical terms the doer is the measure, and the receiving objects are elements of a *sigma algebra*.

**Definition 2.1** (Sigma algebra, measurable space). *Let  $\Omega$  be a sample space. Then, the family of sets  $\mathfrak{F} \subseteq \mathcal{P}(\Omega)$ , where  $\mathcal{P}(\Omega)$  denotes the power set of  $\Omega$ , is called a sigma algebra on  $\Omega$  if it satisfies*

- $\Omega \in \mathfrak{F}$ ,
- if  $\mathcal{A} \in \mathfrak{F}$ , then its complement  $\mathcal{A}^c \in \mathfrak{F}$  (closed under complement),
- if  $\mathcal{A}_n \in \mathfrak{F}$  for all  $n \in \mathbb{N}$ , then  $\bigcup_{n \in \mathbb{N}} \mathcal{A}_n \in \mathfrak{F}$  (closed under countable union).

*The pair  $(\Omega, \mathfrak{F})$  is called a measurable space. Every set  $\mathcal{A} \in \mathfrak{F}$  is called  $\mathfrak{F}$ -measurable.*

A specific sigma algebra is the so-called Borel sigma algebra [82, 83, 154]. Given a sample space  $\Omega$  its Borel sigma algebra  $\mathfrak{B}(\Omega)$  is the smallest sigma algebra on  $\Omega$  that contains all open sets—where “small” refers to the cardinality of the set. For the scope of this thesis the sample space  $\Omega$  associated with Borel sigma algebras is either  $\mathbb{R}$  or  $\mathbb{R}^n$  with  $\mathbb{N} \ni n \geq 2$ . A possible way

to define the Borel sigma algebra on  $\mathbb{R}$  is that it contains all left-bounded intervals, hence [34]

$$\mathfrak{B}(\mathbb{R}) = \bigcap \{ \mathfrak{F} : \mathfrak{F} \text{ is a sigma algebra containing all } (a, \infty) \forall a \in \mathbb{R} \}.$$

We can study Borel sigma algebras in higher dimensions too. To do so we introduce the following element-wise interval notation for some  $a, b \in \mathbb{R}^n$

$$[a, b] = \{ x \in \mathbb{R}^n : a_i \leq x_i \leq b_i \forall i \in \{1, \dots, n\} \}. \quad (2.6)$$

The definitions for the open interval  $(a, b)$ , and the half-open intervals  $(a, b]$ ,  $[a, b)$  are analogous. With this notation we find the Borel sigma algebra  $\mathfrak{B}(\mathbb{R}^n)$  on  $\mathbb{R}^n$  to have the following properties [82, p. 41]:

- $\{c\} \in \mathfrak{B}(\mathbb{R}^n)$  for all  $c \in \mathbb{R}^n$ ,
- for all  $a, b \in \mathbb{R}^n$

$$[a, b], [a, b), (a, b], (a, b) \in \mathfrak{B}(\mathbb{R}^n),$$

- for all  $a \in \mathbb{R}^n$  it holds that

$$(-\infty, a], (-\infty, a), [a, \infty), (a, \infty) \in \mathfrak{B}(\mathbb{R}^n).$$

Given a measurable set, a measure is a function that assigns a number to this set. Albeit an intuitive concept the mathematically rigorous definition of a measure requires some care.

**Definition 2.2** (Measure, measure space). *Let  $(\Omega, \mathfrak{F})$  be a measurable space. Then, the function  $\mu: \mathfrak{F} \rightarrow [0, \infty]$  is called a measure on the sigma algebra  $\mathfrak{F}$  if it satisfies*

- $\mu(\emptyset) = 0$ , and
- if  $\mathcal{A}_n \in \mathfrak{F}$  for all  $n \in \mathbb{N}$  and all  $\mathcal{A}_n$  being mutually disjoint, i.e.  $\mathcal{A}_n \cap \mathcal{A}_m = \emptyset$  for all  $n \neq m$ , then

$$\mu \left( \bigcup_{n \in \mathbb{N}} \mathcal{A}_n \right) = \sum_{n \in \mathbb{N}} \mu(\mathcal{A}_n). \quad (2.7)$$

The triple  $(\Omega, \mathfrak{F}, \mu)$  is called a measure space.

**Remark 2.1** (Lebesgue measure, integral). For the scope of this thesis the prevalent measure is the Lebesgue measure on  $\mathbb{R}^n$ . The Lebesgue measure is the mathematically precise definition of the common sense notion of volume. In fact, for the measure space  $(\mathbb{R}^n, \mathfrak{B}(\mathbb{R}^n))$  the Lebesgue measure is the unique measure such that<sup>3</sup>

$$\mu([a_1, b_1] \times \cdots \times [a_n, b_n]) = (b_1 - a_1) \cdots (b_n - a_n) \quad (2.8)$$

holds with  $a_j \leq b_j$  for all  $j \in \{1, \dots, n\}$  [82, 83]. All integrals that appear in the following are to be understood as Lebesgue integrals that employ the familiar notation from Riemann integrals.

**Definition 2.3** (Measurable function). Let  $(\Omega, \mathfrak{F})$  and  $(\Omega', \mathfrak{F}')$  be measurable spaces. A function  $f: \Omega \rightarrow \Omega'$  is called  $\mathfrak{F}/\mathfrak{F}'$ -measurable if

$$f^{-1}(\mathcal{A}') = \{x \in \Omega: f(x) \in \mathcal{A}'\} \in \mathfrak{F} \quad \forall \mathcal{A}' \in \mathfrak{F}'.$$

Hence every pre-image  $f^{-1}(\mathcal{A}')$  of every  $\mathfrak{F}'$ -measurable subset  $\mathcal{A}'$  is  $\mathfrak{F}$ -measurable. If  $f: \Omega \rightarrow \Omega'$  is  $\mathfrak{F}/\mathfrak{F}'$ -measurable we write  $f: (\Omega, \mathfrak{F}) \rightarrow (\Omega', \mathfrak{F}')$ .

**Definition 2.4** (Absolutely continuous measure). Let  $\mu$  be a measure on  $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ , and let  $\rho: \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}$  be a Lebesgue-measurable function such that for any  $\mathcal{A} \subseteq \mathfrak{B}(\mathbb{R})$

$$\mu(\mathcal{A}) = \int_{\mathcal{A}} \rho(\tau) d\tau.$$

Then, the measure  $\mu$  is said to be absolutely continuous w.r.t. the Lebesgue measure and non-negative, and  $\rho$  is called the Lebesgue density. We formally write  $d\mu = \rho(\tau) d\tau$  with  $d\tau$  as the Lebesgue measure. If  $\mu$  is a probability measure such that  $\mu(\mathbb{R}) = \int_{\mathbb{R}} \rho(\tau) d\tau = 1$ , then we call  $\rho$  a probability density.

---

<sup>3</sup> If the Lebesgue measure were defined on the power set of the real numbers, then whether the thusly defined Lebesgue measure truly exists is an instance of Gödel's incompleteness theorem—it can neither be proved nor disproved [84]. Technically, the domain of the Lebesgue measure is the Lebesgue sigma algebra, a sigma algebra which contains the Borel sigma algebra, but is still smaller than the power set in general. Hence, every Borel-measurable set is Lebesgue-measurable, while the converse is not true. In the following we assume that we are dealing with Borel-measurable sets exclusively.

## 2.2.2 Random variables

**Definition 2.5** (Probability space). *Let  $(\Omega, \mathfrak{F}, \mu)$  be a measure space with  $\mu(\Omega) = 1$ . Then, the triple  $(\Omega, \mathfrak{F}, \mu)$  is called a probability space.*

**Definition 2.6** (Real-valued random variable). *Let  $(\Omega, \mathfrak{F}, \mu)$  be a probability space. Then, any  $\mathfrak{F}/\mathfrak{B}(\mathbb{R})$ -measurable function is called a real-valued random variable, i.e.  $f: (\Omega, \mathfrak{F}) \rightarrow (\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ . We call  $P_f: (\mathbb{R}, \mathfrak{B}(\mathbb{R})) \rightarrow [0, 1]$  with*

$$P_f(\mathcal{A}') = \mu(f^{-1}(\mathcal{A}'))$$

*the probability distribution of the random variable  $f$  under  $\mu$ .*

It can be shown that the image space triple  $(\mathbb{R}, \mathfrak{B}(\mathbb{R}), P_f)$  is a probability space [82]. This means that the probability distribution from Definition 2.6 is itself a probability measure for the measure space  $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ . We introduce a tempting and engineering-motivated short-hand notation “ $\mathbb{P}(f \leq a)$ ” to describe “the probability of the real-valued random variable  $f$  being less than or equal to  $a$ .” This verbose definition can be made rigorous with the help of Definition 2.6

$$\mathbb{P}(f \leq a) := P_f((-\infty, a]) = \mu(\{\omega \in \Omega : f(\omega) \leq a\}). \quad (2.9)$$

**Definition 2.7** ( $\mathbb{R}^n$ -valued random vector). *Let  $(\Omega, \mathfrak{F}, \mu)$  be a probability space, and let  $f_i$  be a real-valued random variable for all  $i \in \{1, \dots, n\}$  with  $n \in \mathbb{N}$ . Then,*

$$f: \Omega \rightarrow \mathbb{R}^n, x \mapsto \begin{bmatrix} f_1(x) \\ \vdots \\ f_n(x) \end{bmatrix}$$

*is  $(\mathfrak{F}, \mathfrak{B}(\mathbb{R}^n))$ -measurable and called an  $\mathbb{R}^n$ -valued random vector.*

Real-valued random variables with absolutely continuous probability distributions allow to compute integrals as one would expect [34, Thm. 4.16, Cor. 4.2], namely

$$\int_{\Omega} g(f(\omega)) d\mu(\omega) = \int_{\mathbb{R}} g(\tau) dP_f(\tau) = \int_{\mathbb{R}} g(\tau) \rho(\tau) d\tau, \quad (2.10)$$

provided any of the integrals exist. Note that  $g$  is some  $\mathfrak{B}(\mathbb{R})/\mathfrak{B}(\mathbb{R})$ -measurable function. For the special cases  $g(x) = x$  and  $g(x) = x^2$  we obtain

$$\mathbb{E}(f) = \int_{\Omega} f d\mu = \int_{\mathbb{R}} \tau dP_f(\tau) = \int_{\mathbb{R}} \tau \rho(\tau) d\tau \quad (2.11a)$$

$$\mathbb{E}(f^2) = \int_{\Omega} f^2 d\mu = \int_{\mathbb{R}} \tau^2 dP_f(\tau) = \int_{\mathbb{R}} \tau^2 \rho(\tau) d\tau = \mathbb{V}(f) + \mathbb{E}(f)^2, \quad (2.11b)$$

where  $\mathbb{E}(\cdot)$  is called the expected value, and  $\mathbb{V}(\cdot)$  is called the variance.

**Definition 2.8** (Real-valued stochastic process). *Let  $\mathcal{T}$  be any set, and let  $(\Omega, \mathfrak{F}, \mu)$  be a probability space. We call the function  $f: \mathcal{T} \times \Omega \rightarrow \mathbb{R}$  an  $\mathbb{R}$ -valued stochastic process on  $\mathcal{T}$  if for every  $t \in \mathcal{T}$  we have that  $f(t, \cdot) = f(t)$  is a random variable with respect to the probability space  $(\Omega, \mathfrak{F}, \mu)$ . In other words, a stochastic process is a collection  $\{f(t, \cdot) \forall t \in \mathcal{T}\}$  of random variables. We say a stochastic process is discrete if the index set  $\mathcal{T}$  has finitely many elements; if  $\mathcal{T}$  has infinitely many elements we say it is a continuous stochastic process.*

Stochastic processes can be viewed in three equivalent ways [84, 154]:<sup>4</sup>

- as the map  $(t, \omega) \mapsto f(t, \omega)$  from  $\mathcal{T} \times \Omega$  to the real numbers;
- as the map  $t \mapsto f(t, \cdot)$  from  $\mathcal{T}$  to the set of all random variables with respect to the probability space  $(\Omega, \mathfrak{F}, \mu)$ ; or
- as the map  $\omega \mapsto f(\cdot, \omega)$  from the sample space  $\Omega$  to the set of all functions that map  $\mathcal{T}$  to the real line.

The last of the three views relates discrete stochastic processes to random vectors: Let's say that the index set is  $\mathcal{T} = \{1, \dots, n\}$ , and we consider a stochastic process  $f: \mathcal{T} \times \Omega$ . Then we can view the stochastic process equivalently as an  $\mathbb{R}^n$ -valued random vector  $f: \Omega \rightarrow \mathbb{R}^n$ , because every realization of the stochastic process is an element of  $\mathbb{R}^n$ . Consequently, for index sets  $\mathcal{T}$  with infinitely many elements the realization of the stochastic process becomes infinitely-valued, hence a “true” function from  $\mathcal{T}$  onto the real

<sup>4</sup> For the scope of this thesis, stochastic processes are functions defined on a domain  $\mathcal{T} \times \Omega$ . In contrast to the field of signal processing we do not put emphasis on the properties of a stochastic process, such as ergodicity, but study the effect of propagating stochastic processes through given mappings.

line. Loosely speaking, a random variable corresponds to a randomly chosen number, a random vector corresponds to a randomly chosen vector, and a stochastic process corresponds to a randomly chosen function.

### 2.2.3 Hilbert space theory

Hilbert spaces generalize the geometric intuitions we hold dear from Euclidean spaces: length, orthogonality, and orthogonal projections. Euclidean spaces are perhaps the most prominent example of Hilbert spaces. The core subject of study for this thesis are Lebesgue spaces of square-integrable measurable functions. These so-called  $L^2$ -spaces include continuous random variables of finite variance. Other notable examples of Hilbert spaces are the space of square-summable sequences, and the Sobolov space of  $L^2$ -functions whose weak derivatives (up to a specified order) are also  $L^2$ -functions. We present standard results for real (pre-)Hilbert spaces with a focus on orthonormal sequences and when they generate a Hilbert space.

A pre-Hilbert space is a linear vector space endowed with an inner product  $\langle \cdot, \cdot \rangle$  which induces the norm  $\| \cdot \| = \sqrt{\langle \cdot, \cdot \rangle}$ . Also, pre-Hilbert spaces allow for the concept of orthogonality.

**Definition 2.9** (Orthogonal elements, orthonormality). *Two elements  $f, g$  of a pre-Hilbert space are called orthogonal if  $\langle f, g \rangle = 0$  whenever  $f \neq g$ . If in addition the elements  $f, g$  have unit length  $\|f\| = \|g\| = 1$ , they are called orthonormal.*

If a pre-Hilbert space is complete—i.e. every Cauchy sequence of elements of the space attains its limit in the very space—we call it a Hilbert space. In Hilbert spaces the notion of orthogonality leads to the classical projection problem: Given an element of a Hilbert space, characterize and find the element from a closed subspace that is closest in the induced norm. The idea is shown in Figure 2.1.

**Theorem 2.1** (Projection theorem [104, 154]). *Let  $\mathcal{M}$  be a closed subspace of a Hilbert space  $\mathcal{H}$ . For any element  $f \in \mathcal{H}$  there is a unique element  $\Pi_{\mathcal{M}}f \in \mathcal{M}$  such that  $\|f - \Pi_{\mathcal{M}}f\| \leq \|f - g\|$  for all  $g \in \mathcal{M}$ . A necessary and sufficient condition for  $\Pi_{\mathcal{M}}f$  to be the unique minimizer is that  $f - \Pi_{\mathcal{M}}f$  be orthogonal to all elements of  $\mathcal{M}$ . This allows to write  $f$  uniquely as  $f = \Pi_{\mathcal{M}}f + e$  with*

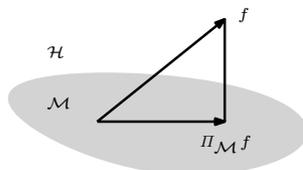


Figure 2.1: Idea of projection theorem (Theorem 2.1): a Hilbert space element  $f \in \mathcal{H}$  is projected orthogonally onto the closed subspace  $\mathcal{M}$ , yielding  $\Pi_{\mathcal{M}}f \in \mathcal{M}$ .

$e \in \mathcal{M}^\perp$  where  $\mathcal{M}^\perp$  is the orthogonal complement of  $\mathcal{M}$ . The element  $\Pi_{\mathcal{M}}f \in \mathcal{M}$  is called the orthogonal projection of  $f$  onto  $\mathcal{M}$ .

The projection theorem is a powerful lever. It gives an answer to the basic approximation problem: given a—now finite—sequence of elements  $\{t_n\}_{n=1}^N$  in a Hilbert space  $\mathcal{H}$ , find the best approximation  $\hat{f} \in \text{span}\{t_n\}_{n=1}^N$  such that the norm of the error  $\|f - \hat{f}\|$  is minimized. For all the computational and methodological advantages of orthogonality we focus not just on any sequence of elements in the following, but on orthogonal/orthonormal sequences. Orthonormal sequences are sequences of elements of a Hilbert space that are mutually orthonormal. The Gram-Schmidt procedure is one possibility to construct a (in-)finite orthonormal sequence of elements given an (in-)finite sequence of linearly independent elements in some pre-Hilbert space. The importance of Gram-Schmidt is twofold: it ensures the existence of an orthonormal sequence *and* it provides a constructive method to compute them.<sup>5</sup> If we combine the idea of the Projection Theorem 2.1 with the Gram-Schmidt procedure we find that for finite orthonormal sequences  $(e_n)_{n=1}^N$  the aforementioned approximation problem of an element  $f \in \mathcal{H}$  reduces to

$$\hat{f} = \sum_{n=1}^N \langle f, e_n \rangle e_n \quad (2.12)$$

with the error  $\|f - \hat{f}\|$  being minimized. This is known as (generalized) Fourier series, and the scalars  $\langle f, e_n \rangle$  are called (generalized) Fourier coefficients [104, 154]. The natural question to ask is whether this series con-

<sup>5</sup> However, the importance of Gram-Schmidt is more theoretical as it tends to be numerically unstable [70, p. 151].

verges given an infinite orthonormal sequence  $(e_n)_{n \in \mathbb{N}}$  in the Hilbert space.<sup>6</sup> Bessel's inequality and Parseval's identity help to affirm that  $\sum_{n \in \mathbb{N}} \langle f, e_n \rangle e_n$  does converge to an element of the closed subspace generated by the orthonormal sequence [104, 154]. Hence, for every element of the Hilbert space to admit a converging Fourier series we need the closure of the span of the orthonormal sequence to be the Hilbert space itself. This leads to the notion of a dense set.

**Definition 2.10** (Complete orthogonal/orthonormal sequence, dense set, separable space [104]). *An orthogonal/orthonormal sequence  $(e_n)_{n \in \mathbb{N}}$  with all elements  $e_n \in \mathcal{H}$  for  $n \in \mathbb{N}$  is called complete if*

$$\overline{\text{span}\{e_n\}_{n \in \mathbb{N}}} = \mathcal{H}.$$

*Hence the closed subspace generated by orthogonal/orthonormal sequence is the Hilbert space itself. Then, by definition, the set  $\text{span}\{e_n\}_{n \in \mathbb{N}}$  is dense in  $\mathcal{H}$ , and  $\mathcal{H}$  is separable.*

We remark that the use of the word “complete” in Definition 2.10—i.e. the entire Hilbert space is generated—is different from the notion of completeness in metric spaces. To highlight the importance of the entire space being generated, we added the notion of a dense set and a separable space to Definition 2.10. It remains to provide a criterion to check whether an orthonormal sequence is complete according to Definition 2.10. This criterion and its implications are given in the following Theorem 2.2.

**Theorem 2.2** (Complete orthonormal basis [104, 154]). *An orthonormal sequence  $(e_n)_{n \in \mathbb{N}}$  in a Hilbert space  $\mathcal{H}$  is complete if and only if its orthogonal complement is the set containing only the zero element  $0_{\mathcal{H}}$  of  $\mathcal{H}$ , thus  $\{e_n\}_{n \in \mathbb{N}}^{\perp} = \{0_{\mathcal{H}}\}$ . The following statements are then equivalent:*

- i)  $\{e_n\}_{n \in \mathbb{N}}^{\perp} = \{0_{\mathcal{H}}\}$ ;
- ii)  $\mathcal{H} = \overline{\text{span}\{e_n\}_{n \in \mathbb{N}}}$ ;

---

<sup>6</sup> Convergence refers to the sequence of partial sums being convergent.

iii) for all  $f \in \mathcal{H}$  it holds

$$\|f\|^2 = \sum_{n \in \mathbb{N}} \langle f, e_n \rangle^2;$$

iv) for all  $f \in \mathcal{H}$  it holds

$$f = \sum_{n \in \mathbb{N}} \langle f, e_n \rangle e_n.$$

**Corollary 2.1** (Complete orthogonal basis [104, 154]). *In case of an orthogonal sequence  $(e_n)_{n \in \mathbb{N}}$  the results from Theorem 2.2 are analogous, but Item iii) and Item iv) read:*

$$\|f\|^2 = \sum_{n \in \mathbb{N}} \frac{\langle f, e_n \rangle^2}{\langle e_n, e_n \rangle^2}, \quad f = \sum_{n \in \mathbb{N}} \frac{\langle f, e_n \rangle}{\langle e_n, e_n \rangle} e_n$$

for any  $f \in \mathcal{H}$ .

Hence, orthogonality and orthonormality differ only in the normalization constants; we use them interchangeably whenever no confusion is expected.

Let us turn to the Hilbert space prevalent in this thesis, the  $L^2$ -space whose definition is based on square-integrable measurable functions. If we restrict ourselves to those square-integrable measurable functions that are random variables we can define the Hilbert space of equivalence classes of random variables. The notion of equivalence classes is required because square-integrable measurable functions that differ only on a set of measure zero have the same norm, hence violating positive-definiteness of the norm.

**Definition 2.11** (Equivalence almost everywhere). *Let  $(\Omega, \mathfrak{F}, \mu)$  be a measure space. Then,*

$$\mathcal{E} = \{f: (\Omega, \mathfrak{F}) \rightarrow (\mathbb{R}, \mathfrak{B}(\mathbb{R})) \text{ with} \\ \exists \mathcal{A} \in \mathfrak{F}: \mu(\mathcal{A}) = 0 \text{ and } f(x) = 0 \forall x \in \Omega \setminus \mathcal{A}\}$$

*is the set of all  $\mathfrak{F}/\mathfrak{B}(\mathbb{R})$ -measurable functions from  $\Omega$  to  $\mathbb{R}$  that are zero almost everywhere. Any two  $\mathfrak{F}/\mathfrak{B}(\mathbb{R})$ -measurable functions  $f, g$  from  $\Omega$  to  $\mathbb{R}$  that satisfy  $f - g \in \mathcal{E}$  are called equivalent  $\mu$ -almost everywhere.*

**Definition 2.12** ( $L^2$ -space of random variables). *Let  $(\Omega, \mathfrak{F}, \mu)$  be a probability space, and let  $\mathcal{E}$  be defined according to Definition 2.11. The quotient space*

$$L^2(\Omega, \mu; \mathbb{R}) = \left\{ f: (\Omega, \mathfrak{F}) \rightarrow (\mathbb{R}, \mathfrak{B}(\mathbb{R})) \text{ with } \int_{\Omega} f(\tau)^2 d\mu(\tau) < \infty \right\} / \mathcal{E}$$

*of equivalence classes of  $\mathfrak{F}/\mathfrak{B}(\mathbb{R})$ -measurable square-integrable random variables from  $\Omega$  to  $\mathbb{R}$  is a Hilbert space with respect to the inner product*

$$\langle x, y \rangle_{L^2(\Omega, \mu; \mathbb{R})} := \mathbb{E}(xy) = \int_{\Omega} x(\tau)y(\tau)d\mu(\tau)$$

*for elements  $x, y \in L^2(\Omega, \mu; \mathbb{R})$ .*<sup>7</sup>

We yield to the convention and refer to the equivalence classes by their representatives [34, 104, 151, 154]. Saying that a random variable  $x \in L^2(\Omega, \mu; \mathbb{R})$  is square-integrable is equivalent to  $x$  having a finite variance.

## 2.3 Problem formulation – Revisited

Let us revisit Problem 2.1 in light of measure theory and Hilbert space theory.

**Problem 2.2** (Mapping under uncertainty). *Let  $(\Omega, \mathfrak{F}, \mu)$  be a probability space, for which  $L^2(\Omega, \mu; \mathbb{R})$  is the Hilbert space of all real-valued random variables of finite variance, and let  $x_i \in L^2(\Omega, \mu; \mathbb{R})$  be  $n_x \in \mathbb{N}$  given random variables for  $i \in \{1, \dots, n_x\}$ . Furthermore, let*

$$f: L^2(\Omega, \mu; \mathbb{R}) \times \dots \times L^2(\Omega, \mu; \mathbb{R}) \rightarrow L^2(\Omega, \mu; \mathbb{R}) \times \dots \times L^2(\Omega, \mu; \mathbb{R})$$

*be a  $\mathfrak{B}(\mathbb{R}^{n_x})/\mathfrak{B}(\mathbb{R}^{n_y})$ -measurable and component-wise square-integrable function.*

*Then, find the random variables  $y_i \in L^2(\Omega, \mu; \mathbb{R})$  for all  $i \in \{1, \dots, n_y\}$  such that*

$$y = f(x) \quad \text{with} \quad y = [y_1, \dots, y_{n_y}]^\top, \quad x = [x_1, \dots, x_{n_x}]^\top.$$

---

<sup>7</sup> Whenever convenient we drop the subscript in the scalar product specifying the Hilbert space.

A challenge intrinsic to Problem 2.2 is that real-valued random variables are infinite-dimensional mathematical objects, but practical computations are naturally restricted to finite evaluations. Hence, we are interested in finite-dimensional representations of infinite-dimensional objects. This idea dates back to Fourier series and has been addressed in Section 2.2.3 already. Fourier series allow to represent periodic signals of finite energy in terms of deterministic Fourier coefficients, see Theorem 2.2. In certain cases—surely in practical applications—only finitely many coefficients serve as a proxy for the periodic signal. The same idea can be applied to random variables of finite variance where it is then called polynomial chaos expansion.

## 2.4 Polynomial chaos

A polynomial chaos expansion (PCE) for a random variable is what a classic Fourier series is for a periodic signal, namely a Hilbert space method to represent an infinite-dimensional mathematical object in terms of (in-)finitely many coefficients. We study PCE for continuous random variables and extensions to random vectors. We further assume that all random variables considered admit an absolutely continuous probability distribution, hence can be visualized by their probability density functions.

### 2.4.1 Orthogonal polynomials

**Definition 2.13** (Ordered set of monic orthogonal polynomials [64]). *Let  $(\mathbb{R}, \mathfrak{B}(\mathbb{R}), \mu)$  be a measure space with the measure  $\mu$  absolutely continuous and non-negative, and let  $\mathcal{K}$  be the index set  $\mathbb{N}_0$  (or  $\{0, 1, \dots, d\}$  for some finite  $d \in \mathbb{N}$  in case we truncate). The elements  $\mathcal{P}_\mu = \{\phi_k\}_{k \in \mathcal{K}}$  are called an ordered set of monic orthogonal polynomials with respect to  $\mu$  if*

$$\begin{aligned}
 \phi_0 &= 1, \\
 \phi_k(t) &= t^k + a_{k-1}t^{k-1} + \dots + a_0, \quad \forall k \in \mathcal{K} \setminus \{0\}, \\
 \langle \phi_i, \phi_j \rangle &= \int_{\mathbb{R}} \phi_i(\tau)\phi_j(\tau)d\mu(\tau) = \gamma_i\delta_{ij}, \quad \forall i, j \in \mathcal{K}, \\
 \gamma_k &> 0, \quad \forall k \in \mathcal{K},
 \end{aligned} \tag{2.13}$$

where  $\delta_{ij}$  is the Kronecker-delta. The numbers  $\gamma_k$  are normalization constants.

We stick to the convention and use  $\phi_0 = 1$ , cf. [64, 154, 172]. Note that the index  $k$  is equivalent to the degree of the polynomial  $\phi_k$

$$\deg \phi_k = k. \quad (2.14)$$

The index  $k$  hence is a natural ordering of the elements of  $\mathcal{P}_\mu$ ; the greater  $k$ , the greater the degree. As the inner product from Definition 2.13 is positive definite, there exists an infinite sequence of orthogonal polynomials relative to the measure  $\mu$  [154, Thm. 8.5]. In light of the theory of Hilbert spaces from Section 2.2.3, it is fair to ask whether these ordered sets of orthogonal polynomials form complete orthogonal sequences according to Definition 2.10? Unfortunately, there is no affirmative answer in general. Take for example the log-normal density  $\rho(\tau) \propto 1/\tau \exp(-\ln^2(\tau)/2)$  for all real  $t > 0$ : there are infinitely many orthogonal polynomials relative to the measure, but their sequence is not complete, hence the closed subspace generated by these polynomials is *not* the underlying Hilbert space [154].

The interest in orthogonal polynomials arose in the middle of the 19<sup>th</sup> century when Chebyshev studied least-squares problems to fit experimental data [72]. This is a classic example of the Projection Theorem 2.1—which ensures convergence in the mean square sense. In fact, the smoother the function that is to be approximated, the better the rate of convergence is going to be; this is the notion of spectral convergence [154, 172]. Uniform convergence can be arbitrarily ludicrous, however, independent of the smoothness of the function to be approximated—this is the Gibbs phenomenon [95, 154, 172].

Let us now turn to multivariate polynomials. We write

$$t^\alpha := t_1^{\alpha_1} t_2^{\alpha_2} \cdots t_m^{\alpha_m}, \quad \deg t^\alpha = |\alpha| := \alpha_1 + \cdots + \alpha_m \quad (2.15)$$

for a multivariate monomial, where  $\alpha$  is a multi-index  $\alpha = [\alpha_1 \ \alpha_2 \ \dots \ \alpha_m]^\top \in \mathcal{K}^m$ . The degree of the multivariate monomial is defined to be the sum of the entries of the multi-index. Given some multivariate polynomial we call the largest degree among all its monomials the total degree.<sup>8</sup>

---

<sup>8</sup> Sometimes the total degree refers to the expression for the degree in (2.15), see [42, p. 2].

$ \alpha $	grevlex	grevlex	revgrevlex	#
0	[0 0 0 0]	[0 0 0 0]	[0 0 0 0]	0
1	[0 0 0 1]	[0 0 0 1]	[1 0 0 0]	1
	[0 0 1 0]	[0 0 1 0]	[0 1 0 0]	2
	[0 1 0 0]	[0 1 0 0]	[0 0 1 0]	3
	[1 0 0 0]	[1 0 0 0]	[0 0 0 1]	4
2	[0 0 0 2]	[0 0 0 2]	[2 0 0 0]	5
	[0 0 1 1]	[0 0 1 1]	[1 1 0 0]	6
	[0 0 2 0]	[0 1 0 1]	[1 0 1 0]	7
	[0 1 0 1]	[1 0 0 1]	[1 0 0 1]	8
	[0 1 1 0]	[0 0 2 0]	[0 2 0 0]	9
	[0 2 0 0]	[0 1 1 0]	[0 1 1 0]	10
	[1 0 0 1]	[1 0 1 0]	[0 1 0 1]	11
	[1 0 1 0]	[0 2 0 0]	[0 0 2 0]	12
	[1 1 0 0]	[1 1 0 0]	[0 0 1 1]	13
	[2 0 0 0]	[2 0 0 0]	[0 0 0 2]	14

Table 2.1: Different orderings on  $\mathbb{N}_0^4$  and their indexing in  $\mathbb{N}$  by #, see Examples 2.3-2.5.

In contrast to the univariate case there is no natural ordering for multivariate polynomials by their multi-indices  $\alpha$ , because there is no natural ordering on  $\mathcal{K}^m \subseteq \mathbb{N}_0^m$ . There exist several schemes for ordering multi-indices. We just mention a few representatives [41, 67, 95, 154, 172]:

**Example 2.3** (Graded lexicographic). For all  $\alpha, \beta \in \mathbb{N}_0^m$  we write

$$\alpha \succ_{\text{grevlex}} \beta$$

if  $|\alpha| > |\beta|$ , or if  $|\alpha| = |\beta|$  and the left-most non-zero entry of  $\alpha - \beta$  is positive.

**Example 2.4** (Graded reverse lexicographic). For all  $\alpha, \beta \in \mathbb{N}_0^m$  we write

$$\alpha \succ_{\text{grevlex}} \beta$$

if  $|\alpha| > |\beta|$ , or if  $|\alpha| = |\beta|$  and the right-most non-zero entry of  $\alpha - \beta$  is negative.

**Example 2.5** (Reversal of graded lexicographic). For all  $\alpha, \beta \in \mathbb{N}_0^m$  we write

$$\alpha \succ_{\text{revgrevlex}} \beta$$

if  $|\alpha| > |\beta|$ , or if  $|\alpha| = |\beta|$  and the left-most non-zero entry of  $\alpha - \beta$  is negative.

Note that the graded reverse lexicographic order is *not* the same as the reversal of the graded lexicographic order; they are mirrored. Table 2.1 shows the subtle differences on  $\mathbb{N}_0^4$ . Having mentioned these orderings we can define the multivariate extension of Definition 2.13.

**Definition 2.14** (Ordered set of multivariate monic orthogonal polynomials). *Consider some finite  $m \in \mathbb{N}$ . Let  $\mu = \mu_1 \otimes \cdots \otimes \mu_m$  be a product measure<sup>9</sup> on  $(\mathbb{R}^m, \mathfrak{B}(\mathbb{R}^m))$ , and let  $\mathcal{P}_{\mu_i} = \{\phi_k^{(i)}\}_{k \in \mathcal{K}}$  be a family of ordered sets of monic orthogonal polynomials relative to  $\mu_i$  with normalization constants  $\gamma_k^{(i)}$  for all  $i \in \{1, \dots, m\}$ . Then, the ordered set  $\mathcal{P}_\mu^m$  of  $m$ -variate monic orthogonal polynomials with respect to the measure  $\mu$  of total degree less than or equal to  $\sup \mathcal{K}$  is given by*

$$\mathcal{P}_\mu^m = \left\{ \phi_{\#(\alpha)} = \prod_{i=1}^m \phi_{\alpha_i}^{(i)} \text{ with } |\alpha| \leq \sup \mathcal{K} \right\}_{\alpha \in \mathcal{K}^m} \quad (2.16a)$$

$$= \left\{ \phi_{\#(\alpha)} = \prod_{i=1}^m (t^{\alpha_i} + a_{\alpha_i-1}^{(i)} t_i^{\alpha_i-1} + \cdots + a_0^{(i)}) \right. \\ \left. \text{with } |\alpha| \leq \sup \mathcal{K} \right\}_{\alpha \in \mathcal{K}^m} \quad (2.16b)$$

$$\langle \phi_{\#(\alpha)}, \phi_{\#(\beta)} \rangle = \prod_{i=1}^m \langle \phi_{\alpha_i}^{(i)} \phi_{\beta_i}^{(i)} \rangle_{L^2(\mathbb{R}, \mu_i; \mathbb{R})} \quad (2.16c)$$

$$= \prod_{i=1}^m \gamma_{\alpha_i}^{(i)} \delta_{\alpha_i \beta_i} \quad \forall \alpha, \beta \in \mathcal{K}^m, \quad (2.16d)$$

<sup>9</sup> The product measure  $\mu_1 \otimes \mu_2 : \mathfrak{F}_1 \times \mathfrak{F}_2 \rightarrow [0, \infty]$  is defined such that  $(\mu_1 \otimes \mu_2)(F_1 \times F_2) = \mu_1(F_1)\mu_2(F_2)$  for all  $F_1 \in \mathfrak{F}_1, F_2 \in \mathfrak{F}_2$ , see [154, Def. 2.33]. In light of random variables, product measures appear whenever random variables are independent. In that case, the overall probability density is the product of the individual probability densities.

where  $\# : \mathcal{K}^m \subseteq \mathbb{N}_0^m \rightarrow \mathbb{N}_0$  with  $\alpha \mapsto |\{\beta \in \mathbb{N}_0^m : \alpha \succ_{\text{ord}} \beta\}|$  enumerates the multi-index order  $\succ$  for  $\text{ord} \in \{\text{grlex}, \text{grevlex}, \text{revgrlex}\}$ .<sup>10</sup> In case of  $\mathcal{K} = \{0, 1, \dots, d\}$ , the cardinality of the set  $\mathcal{P}_\mu^m$  is (see Appendix A.3)

$$|\mathcal{P}_\mu^m| = \frac{(d+m)!}{d!m!} = \binom{d+m}{d}. \quad (2.16e)$$

The definition of the multivariate basis follows from Fubini's Theorem [154].

**Remark 2.2** (Fixed ordering). *In everything that follows we employ the reversal of the graded lexicographic ordering, see Example 2.5.*

**Remark 2.3** (Alternative construction of basis). *Often (physical) models are driven by low-order effects and low-order interactions [115]. This can help to tame the curse of dimensionality in Definition 2.14. Imagine we construct a multivariate basis of total degree less than  $\sup \mathcal{K}$  for which only products of at most  $j \leq \sup \mathcal{K}$  univariate bases form a multivariate element*

$$\mathcal{P}_\mu^m = \left\{ \phi_{\#(\alpha)} := \prod_{k=1}^m \phi_{\alpha_k}^{(k)} \text{ with } |\alpha| \leq \sup \mathcal{K} \quad \sum_{i=1}^m \mathbb{1}_{\alpha_i > 0} < j \right\}_{\alpha \in \mathcal{K}^m}. \quad (2.17)$$

The number  $\sum_{i=1}^m \mathbb{1}_{\alpha_i > 0}$  of non-zero elements of  $\alpha$  is called the rank of  $\alpha$  and represents the interactions among the univariate bases [7, 23]. It is shown in [23] that “decreasing  $j$  allows a relative decrease [...] of about 10 with respect to the usual index set  $[\mathcal{P}_\mu^m]$ .” Arguably, if the true model exhibits interactions  $\hat{j}$  with  $j < \hat{j} \leq \sup \mathcal{K}$ , then these interactions are not captured. Another approach to reducing the number of basis elements is to use hyperbolic index sets [23]. Both approaches are computationally superior than the ad-hoc definition of  $\mathcal{P}_\mu^m$  according to Definition 2.14 as they exploit sparsity.

All of the presented ways to create the multivariate basis are, so far, based on a-priori choices: the maximum degree  $\sup \mathcal{K}$  and, if applicable, the interaction  $j$  are chosen by the user before making any computation. If there is any data available, there exist other possibilities to create the multivariate basis, leading

<sup>10</sup> The cardinality of the empty set is zero.

to basis-adaptive and sparse polynomial bases. The idea is to add incrementally only those polynomials that have a significant numerical effect [24, 25].

## 2.4.2 Polynomial chaos expansion

Polynomial chaos expansions combine measure theory, Hilbert space theory, and orthogonal polynomials. The main idea is to let polynomials that are orthogonal relative to a given probability measure span a Hilbert space to which we can apply Theorem 2.1. To overcome technical details with whether a sequence of orthogonal polynomials is complete or not, we make the following technical assumption.

**Assumption 2.1** (Orthogonal sequence in  $L^2$ ). *Consider  $L^2(\Omega, \mu; \mathbb{R})$  according to Definition 2.12, let the sample space be the real numbers  $\Omega = \mathbb{R}$ , and let the probability measure  $\mu$  be absolutely continuous and non-negative. We assume  $\mathcal{P}_\mu$  is an ordered set of monic orthogonal polynomials in  $L^2(\Omega, \mu; \mathbb{R})$ , and that it forms a complete orthogonal sequence according to Definition 2.10.<sup>11</sup>*

In Assumption 2.1 the triple  $(\Omega, \mathfrak{F}, \mu) = (\mathbb{R}, \mathfrak{B}(\mathbb{R}), \mu)$  is a probability space, and the symbol “ $\mu$ ” takes over the meaning of the probability distribution. This allows to keep the notation somewhat lighter than the more technical notation in Section 2.2.2.

As the  $L^2$ -space from Definition 2.12 is a Hilbert space it inherits all the useful properties mentioned in Section 2.2.3. It was Norbert Wiener who studied the standard Gaussian measure

$$\mu_G(A) = \int_A \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx, \quad \forall A \in \mathfrak{B}(\mathbb{R}) \quad (2.18)$$

and who found the Hermite polynomials to be a complete orthogonal sequence in  $L^2(\mathbb{R}, \mu_G; \mathbb{R})$ , for which he coined the term *homogeneous chaos* [167].<sup>12</sup> Wiener’s homogeneous chaos is Fourier series for real-valued square-

---

<sup>11</sup> Hence, the elements  $\phi_k$  of  $\mathcal{P}_\mu$  are both random variables and “ordinary” polynomials that map real numbers to real numbers.

<sup>12</sup> Sometimes the Hermite polynomials orthogonal to (2.18) are called probabilists’ Hermite polynomials.

integrable random variables that are the image of functions of standard Gaussian random variables

$$f: L^2(\mathbb{R}, \mu_G; \mathbb{R}) \rightarrow L^2(\mathbb{R}, \mu_G; \mathbb{R}). \quad (2.19)$$

The extension to real-valued square-integrable random variables that are the image of functions of *several* random variables that are allowed to be non-Gaussian is called *generalized polynomial chaos*, or simply *polynomial chaos* [154, 172, 173]. The mappings of interest for PCE are thus

$$f: \bigotimes_{i=1}^m L^2(\Omega_i, \mu_i; \mathbb{R}) \rightarrow L^2(\Omega, \mu; \mathbb{R}). \quad (2.20)$$

It is fair to ask about the relation between the tensor product of Hilbert spaces and the image space in (2.20)? Before we can answer this question, let us assume that there is an ordered set of monic orthogonal polynomials for every Hilbert space associated with every Hilbert space  $L^2(\Omega_i, \mu_i; \mathbb{R})$ .

**Assumption 2.2** (Several  $L^2$ -spaces). *For a finite  $m \in \mathbb{N}$  let the Hilbert spaces  $L^2(\Omega_i, \mu_i; \mathbb{R})$  satisfy Assumption 2.1 for all  $i \in \{1, \dots, m\}$ . The random variables  $\Xi_1, \dots, \Xi_m$  with  $\Xi_i \in L^2(\Omega_i, \mu_i; \mathbb{R})$  for all  $i \in \{1, \dots, m\}$  are mutually independent.*

**Remark 2.4** (Stochastic germ [95, 154, 172]). *The  $m$ -valued random vector  $\Xi = [\Xi_1, \dots, \Xi_m]^\top$  from Assumption 2.2 is called the stochastic germ.<sup>13</sup>*

As all  $m$  spaces  $L^2(\Omega_i, \mu_i; \mathbb{R})$  are separable, see Assumption 2.1 and the reference to Definition 2.10 therein, the tensor product of the Hilbert spaces is isomorphic to the  $L^2$ -space over the product probability space [154, p. 237],

$$\begin{aligned} L^2(\Omega, \mu; \mathbb{R}) &= L^2(\Omega_1 \times \Omega_2 \times \dots \times \Omega_m, \mu_1 \otimes \mu_2 \otimes \dots \otimes \mu_m; \mathbb{R}) \\ &\cong \bigotimes_{i=1}^m L^2(\Omega_i, \mu_i; \mathbb{R}). \end{aligned} \quad (2.21)$$

<sup>13</sup> We use that term extensively in Section 4.3.

The scalar product in  $L^2(\Omega, \mu; \mathbb{R})$  is then the product of the  $m$  individual scalar products

$$\langle \cdot, \cdot \rangle_{L^2(\Omega, \mu; \mathbb{R})} = \prod_{i=1}^m \langle \cdot, \cdot \rangle_{L^2(\Omega_i, \mu_i; \mathbb{R})}. \quad (2.22)$$

It remains to characterize the elements that generate the tensor product space  $L^2(\Omega, \mu; \mathbb{R})$  in the sense of Theorem 2.2. These elements can be chosen as  $m$ -variate monic polynomials that are orthogonal relative to the product measure  $\mu = \mu_1 \otimes \mu_2 \otimes \dots \otimes \mu_m$ . Hence they can be constructed according to Definition 2.14 (for an unbounded index set). Thus by construction the ordered set of  $m$ -variate monic orthogonal polynomials  $\mathcal{P}_\mu^m$  for the tensor product space  $L^2(\Omega, \mu; \mathbb{R})$  forms a complete orthogonal sequence. Then, polynomial chaos expansion is the name of the manifestation of Theorem 2.1 for random variables that are elements of  $L^2(\Omega, \mu; \mathbb{R})$ , namely PCE is the expansion of a random variable of finite variance in a complete orthogonal basis. Importantly, the depiction of Theorem 2.1 from Figure 2.1 holds just the same: a given random variable is projected onto every basis function.<sup>14</sup>

**Definition 2.15** (PCE of random variables). *Let Assumption 2.2 hold such that  $\mathcal{P}_\mu^m = \{\phi_k\}_{k \in \mathbb{N}_0}$  is a complete orthogonal sequence of ordered  $m$ -variate monic orthogonal polynomials for the tensor product space  $L^2(\Omega, \mu; \mathbb{R})$ .<sup>15</sup> Then, the polynomial chaos expansion of the random variable  $x \in L^2(\Omega, \mu; \mathbb{R})$  is*

$$x = \sum_{k \in \mathbb{N}_0} x_k \phi_k \quad \text{with} \quad x_k = \frac{\langle x, \phi_k \rangle_{L^2(\Omega, \mu; \mathbb{R})}}{\langle \phi_k, \phi_k \rangle_{L^2(\Omega, \mu; \mathbb{R})}} \in \mathbb{R}, \quad (2.23)$$

where  $x_k$  is called the  $k^{\text{th}}$  PCE coefficient of  $x$ . The truncated polynomial chaos expansion of  $x$  for  $\mathcal{K} = \{0, \dots, \check{k}\}$  reads

$$\Pi_{\check{k}} x = \sum_{k \in \mathcal{K}} x_k \phi_k \quad \text{with} \quad x_k = \frac{\langle x, \phi_k \rangle_{L^2(\Omega, \mu; \mathbb{R})}}{\langle \phi_k, \phi_k \rangle_{L^2(\Omega, \mu; \mathbb{R})}} \in \mathbb{R}, \quad (2.24)$$

---

<sup>14</sup> With the modification that  $\mathcal{H} = L^2(\Omega, \mu; \mathbb{R})$ , and  $\mathcal{M} = \text{span}\{\phi_k\}_{k \in \mathcal{K}}$  (assuming  $\mathcal{K}$  has finitely many elements).

<sup>15</sup> Hence, the underlying measurable space is  $(\Omega, \mathfrak{F}) = (\mathbb{R}^m, \mathfrak{B}(\mathbb{R}^m))$ .

where the error  $x - \Pi_{\hat{k}}x$  is orthogonal to the subspace spanned by the elements of the sequence  $(\phi_k)_{k \in \mathcal{K}}$ .<sup>16</sup>

For the truncated PCE mentioned in Definition 2.12 it is common to let the subspace be spanned by all ordered  $m$ -variate orthogonal polynomials of degree at most  $d$ . This subspace has the dimension

$$\dim(\text{span}\{\phi_k\}_{k \in \mathcal{K}}) = \hat{k} + 1 = \binom{d+m}{d}, \quad (2.25)$$

We call the number  $\hat{k} + 1$  the PCE dimension.<sup>17</sup> In that case we identify the PCE of the random variable with its projection onto the subspace, i.e.  $x = \Pi_{\hat{k}}x$ .

**Definition 2.16** (PCE of random vectors). *Consider the setting from Definition 2.15. For a finite  $n \in \mathbb{N}$  consider the (not necessarily independent) random variables  $x_i \in L^2(\Omega, \mu; \mathbb{R})$  for all  $i \in \{1, \dots, n\}$  with their  $k^{\text{th}}$  PCE coefficient  $x_{i,k}$ . We call  $x = [x_1, \dots, x_n]^{\top}$  an  $\mathbb{R}^n$ -valued random vector and write its PCE*

$$x = \sum_{k \in \mathbb{N}_0} x_k \phi_k, \quad (2.26a)$$

where  $x_k = [x_{1,k}, \dots, x_{n,k}]^{\top} \in \mathbb{R}^n$ . Formally, we write

$$x \in L^2(\Omega, \mu; \mathbb{R}^n) \iff x_i \in L^2(\Omega, \mu; \mathbb{R}) \quad \forall i \in \{1, \dots, n\}. \quad (2.26b)$$

We see that the PCE of a random vector is defined as taking the PCE of each component. Clearly, the  $k^{\text{th}}$  PCE coefficient of a random vector is then vector-valued. Interestingly, the space  $L^2(\Omega, \mu; \mathbb{R}^n)$  can be understood as the following isomorphism [154]

$$L^2(\Omega, \mu; \mathbb{R}^n) \cong L^2(\Omega, \mu; \mathbb{R}) \otimes \mathbb{R}^n. \quad (2.27)$$

<sup>16</sup> To emphasize the dimension  $\hat{k} + 1$  of the subspace spanned by  $\{\phi_k\}_{k \in \mathcal{K}}$  rather than the fact that it is a subspace we slightly violate the notation of the projection operator, see its premier occurrence in Theorem 2.1.

<sup>17</sup> The number (2.25) appeared already in Definition 2.14. Note that the role of the sets “ $\mathcal{K}$ ” in both (2.25) and Definition 2.14 is the role of an ordered single-index set.

We can thus view the originally  $\mathbb{R}^n$ -valued random vector  $x$  instead as a real-valued discrete stochastic process  $x: \mathcal{T} \times \Omega \rightarrow \mathbb{R}$  with  $\mathcal{T} = \{1, \dots, n\}$ , hence for every instant  $t \in \mathcal{T}$  we have that  $x(t, \cdot) \in L^2(\Omega, \mu; \mathbb{R})$  is a real-valued random variable, and for every realization  $\omega \in \Omega$  we have that  $x(\cdot, \omega) \in \mathbb{R}^n$  is but an ordinary Euclidean vector. We call the real-valued discrete stochastic process we associate with  $x \in L^2(\Omega, \mu; \mathbb{R}^n)$  a square-integrable real-valued discrete stochastic process. For a more detailed investigation between PCE and stochastic processes we refer to Appendix A.4.

### 2.4.3 Advantages

We focus on four advantages of PCE for random variables: evaluation, applicability to non-Gaussian settings, moments of random variables/vectors, and uncertainty propagation.

#### Evaluation

Although more a property or point-of-view than an advantage, it is worth emphasizing that PCE allows for a straightforward evaluation of random variables to obtain realizations. Given a random variable  $x \in L^2(\Omega, \mu; \mathbb{R})$  and its PCE relative to some finite orthogonal basis  $\{\phi_k\}_{k \in \mathcal{K}}$  we write according to Definition 2.15

$$x = x(\tau) \cong \Pi_{\hat{k}} x(\tau) = \sum_{k \in \mathcal{K}} x_k \phi_k(\tau) \in \mathbb{R}. \quad (2.28)$$

The notation emphasizes that the random variable  $x$  is a function mapping values  $\tau \in \Omega$  to the real numbers  $x$ . It allows to leverage random variables that are outputs of mappings under uncertainty for decision-making: given a specific realization  $\tau$  of the uncertainty, the corresponding realization of the uncertainty of the output is obtained from (2.28). The computational cost associated with (2.28) is negligible: it requires to evaluate a polynomial basis and multiply with the respective PCE coefficients.

Table 2.2: Orthogonal bases for univariate random variables according to [173].

Type	Support	$\phi_k(\omega)$	Polynomial basis
Beta	$(0, 1)$	$P_k^{(\beta-1, \alpha-1)}(2\omega - 1)$	Jacobi
Gamma	$(0, \infty)$	$L_k^{(\alpha-1)}(\beta\omega)$	Generalized Laguerre
Gaussian	$(-\infty, \infty)$	$\text{He}_k(\omega)$	Hermite
Uniform	$[0, 1]$	$P_k^{(0,0)}(2\omega - 1)$	Legendre

### Non-Gaussian random variables

Polynomial chaos requires the random variables to have a finite variance, but there are no restrictions on the distribution of the random variable, e.g. Gaussian or uniform. Also, the random variables need not have a unimodal probability density. For instance, PCE works for (Gaussian) mixture models, i.e. continuous random variables whose probability density function is a convex combination of (Gaussian) densities [154]. Hence, PCE is applicable to Gaussian and non-Gaussian random variables alike—or combinations thereof. The Askey scheme of orthogonal polynomials characterizes polynomials that are orthogonal relative to measures reminiscent of often-employed random variables [173]. Table 2.2 shows some of the orthogonal polynomials that correspond to continuous random variables with absolutely continuous probability distributions. In case one has to deal with a non-standard probability density—because it has been fitted to experimental data for instance—not all hope is lost. There exist stable numerical methods to compute orthogonal polynomials relative to given probability density functions. We refer to Appendix A.1 for details; Chapter 5 introduces a Julia package written specifically for this purpose.

### Moments

Orthogonality of the polynomials admits straightforward formulæ for moments of random vectors in terms of their PCE coefficients. For example, the expected value of the random vector  $\mathbf{x} \in L^2(\Omega, \mu; \mathbb{R}^n)$  is

$$\mathbb{E}(\mathbf{x}) = \mathbb{E}(\mathbf{x}\phi_0) \quad (2.29a)$$

$$= \langle \mathbf{x}, \phi_0 \rangle_{L^2(\Omega, \mu; \mathbb{R})} \quad (2.29b)$$

$$= \sum_{k \in \mathbb{N}_0} x_k \langle \phi_k, \phi_0 \rangle_{L^2(\Omega, \mu; \mathbb{R})} \quad (2.29c)$$

$$= x_0, \quad (2.29d)$$

which follows from orthogonality and  $\mu$  being a probability measure, thus  $\langle \phi_0, \phi_0 \rangle_{L^2(\Omega, \mu; \mathbb{R})} = \mu(\Omega) = 1$ . Note that the notation in (2.29b) is to be understood as taking the scalar product component-wise with respect to  $\mathbf{x}$ . The covariance of two random vectors  $\mathbf{x}, \mathbf{y} \in L^2(\Omega, \mu; \mathbb{R}^n)$  becomes

$$\mathbb{C}(\mathbf{x}, \mathbf{y}) = \mathbb{E}((\mathbf{x} - \mathbb{E}(\mathbf{x}))(\mathbf{y} - \mathbb{E}(\mathbf{y}))^\top) \quad (2.30a)$$

$$= \mathbb{E} \left( \sum_{k_1 \in \mathbb{N}} x_{k_1} \phi_{k_1} \sum_{k_2 \in \mathbb{N}} y_{k_2}^\top \phi_{k_2} \right) \quad (2.30b)$$

$$= \sum_{k \in \mathbb{N}} x_k y_k^\top \mathbb{E}(\phi_k \phi_k) \quad (2.30c)$$

$$= \sum_{k \in \mathbb{N}} x_k y_k^\top \langle \phi_k, \phi_k \rangle_{L^2(\Omega, \mu; \mathbb{R})}, \quad (2.30d)$$

hence it is the sum of weighted squares of the PCE coefficients. Note that the summation in (2.30) is *not* over  $\mathbb{N}_0$  but  $\mathbb{N}$ . Similar expressions can be derived for higher-order moments of random variables such as skewness or kurtosis.

## Uncertainty propagation

Uncertainty propagation with PCE means to find the PCE coefficients of the image random variable  $\mathbf{y} = f(\mathbf{x})$ , given the mapping  $f$  and the PCE for  $\mathbf{x}$ . One way to approach this is via intrusive Galerkin projection. This means, given the mapping  $f : L^2(\Omega, \mu; \mathbb{R}^n) \rightarrow L^2(\Omega, \mu; \mathbb{R})$  according to Definition 2.16, and given the random vector  $\mathbf{x} \in L^2(\Omega, \mu; \mathbb{R}^n)$  with PCE coefficients  $x_k$  relative to the orthogonal polynomials  $\{\phi_k\}_{k \in \mathbb{N}_0}$ , we do the following to find the PCE coefficients of  $\mathbf{y} = f(\mathbf{x})$ :

1. Insert the PCE of  $\mathbf{x}$  and  $\mathbf{y}$

$$\sum_{k \in \mathbb{N}_0} y_k \phi_k = f \left( \sum_{k \in \mathbb{N}_0} x_k \phi_k \right). \quad (2.31)$$

2. For all  $m \in \mathbb{N}_0$  project onto the orthogonal element  $\phi_m$

$$\left\langle \sum_{k \in \mathbb{N}_0} y_k \phi_k, \phi_m \right\rangle_{L^2(\Omega, \mu; \mathbb{R})} = \left\langle f \left( \sum_{k \in \mathbb{N}_0} x_k \phi_k \right), \phi_m \right\rangle_{L^2(\Omega, \mu; \mathbb{R})}. \quad (2.32)$$

3. Solve

$$y_m = \frac{\left\langle f \left( \sum_{k \in \mathbb{N}_0} x_k \phi_k \right), \phi_m \right\rangle_{L^2(\Omega, \mu; \mathbb{R})}}{\left\langle \phi_m, \phi_m \right\rangle_{L^2(\Omega, \mu; \mathbb{R})}}. \quad (2.33)$$

The approach is said to be *intrusive* because we insert the pCE of  $x$  into the function  $f$  and derive new equations (2.33) that have to be evaluated. These new equations (2.33) are deterministic but there are  $\mathbb{N}_0$  of them. For practical considerations we restrict the investigations to the subspace generated by  $\{\phi_k\}_{k \in \mathcal{K}}$  see Definition 2.15, but still its dimension is known to grow rapidly. Intrusive approaches have the appealing property that the mathematical structure of the relations (2.33) may be equivalent to the structure of the mapping  $f$ . For example, if  $f$  is polynomial, then all equations (2.33) will remain polynomial. Hence, for selected classes of functions  $f$  the structure is preserved; this can be exploited both numerically and theoretically, see for instance Section 2.5.

Non-intrusive methods instead evaluate the function  $f$  for a fixed number of times (so-called deterministic model resolutions [95]) from which the pCE coefficients of the image of  $y$  are computed. Least-squares and its sparse counterparts can then be applied, but collocation methods are also possible [95, 172]. These methods compute the pCE to hold exactly at selected points, making the overall pCE of the image  $y$  an interpolatory rule. Non-intrusive methods may not be able exploit the structure of  $f$  as much as intrusive methods. However, non-intrusive methods have clear advantages in case there is no explicit formula for  $f$ . For instance, the evaluation of  $f$  might come from a proprietary piece of software, which in addition could be numerically expensive to run. In that case, non-intrusive methods may be favorable.

## 2.4.4 Disadvantages

As with anything PCE brings about advantages but also disadvantages. It is fair to say PCE is a method that is heavy in both abstraction and computation.

First, PCE requires to describe uncertainties in terms of probability density functions. This may be too much to ask for in practice as there might be—at best—histograms available. Hence, PCE might involve a pre-processing step that fits a continuous function to given histograms, yielding an analytic probability density function. Alternatively, the histograms can be used directly as a discrete measure for which we compute orthogonal polynomials. Mathematically speaking, however, we are then in the realm of discrete orthogonal polynomials. Second, for multivariate PCEs we assume independence among the random variables of each uni-variate basis, see Assumption 2.2.<sup>18</sup> It is possible to adapt PCE to the setting in which the random variables of each univariate are dependent, but the basis functions need then no longer be polynomials.<sup>19</sup> Third, PCE requires to *know* the orthogonal basis functions. Except for special cases, see Table 2.2, PCE relies on a pre-processing step to compute the basis functions relative to a given density. For this task there exist dedicated software packages, see Chapter 5 for details.

While the aforementioned disadvantages are generic to PCE, there may be issues specific to how we apply PCE. Specifically, this thesis employs intrusive PCE, see the previous Section 2.4.3: the governing equations are “intruded”, and we derive a larger set of deterministic equations. Although intrusive PCE may preserve the mathematical structure of the governing equations, it comes at the cost of having to derive the “intruded” governing equations, hence solving integrals. Clearly, non-intrusive PCE may be favorable for large systems with many uncertainties.

---

<sup>18</sup> Recall: this does not mean that the modeled random variables need to be independent, but it still imposes restrictions on the random variables that compose the orthogonal basis.

<sup>19</sup> An example is given in [154, Remark 11.16].

## 2.5 Quantifying truncation errors

A polynomial chaos expansion of a random variable as such is exact with infinitely many terms. In practice, the truncated PCE is used, meaning that only finitely many coefficients represent the random variable. This naturally leads to asking whether truncation errors are made and how they can be quantified. Is it possible to describe a random variable and its mapping precisely by a finite truncated PCE? And, if the PCE is truncated early, is it possible to establish an error bound on the image random variable?

It so appears that there exists but a small number of results that consider PCE truncation errors rigorously. For instance, in [60] illustrative examples evaluate the accuracy of PCE. Yet, the errors are not computed rigorously but rather are studied via extensive simulations. Similarly, the authors of [47] list several numerical challenges when using PCE, including the potential need for large PCE dimensions. The authors of [11] provide an upper bound on the truncation error using a univariate Hermitian basis based on differentiability assumptions of the mapping. Yet, these results do not easily carry over to other bases. The following material is based on [122, 123].

We begin with a definition.

**Definition 2.17** (Minimum expansion degree). *The minimum expansion degree of  $\mathbf{x} \in L^2(\Omega, \mu; \mathbb{R})$  is the number  $d_x \in \mathbb{N}_0$  such that all PCE coefficients associated with higher-degree polynomials are zero, i.e.  $x_j = 0$  for all  $j$  with  $\deg \phi_j > d_x$ .*

For multivariate polynomials the minimum expansion degree refers to the total degree, see Section 2.4.1.

**Assumption 2.3** (Exact PCE input). *For a given  $m$ -variate polynomial basis  $\mathcal{P}_\mu^m = \{\phi_k\}_{k=0}^{\hat{k}_x}$ , the PCE of the real-valued random variable  $\mathbf{x} \in L^2(\Omega, \mu; \mathbb{R})$  has the known and finite minimum degree  $d_x \in \mathbb{N}_0$  with*

$$\hat{k}_x \geq (m + d_x)! / (m! d_x!).$$

*Unless stated otherwise we assume that  $\hat{k}_x = (m + d_x)! / (m! d_x!)$ .*

In other words, for all sets  $\{\phi_k\}_{k=0}^{\hat{k}} \supseteq \{\phi_k\}_{k=0}^{\hat{k}_x}$  with  $\hat{k} \geq \hat{k}_x$  we have  $\|\mathbf{x} - \Pi_{\hat{k}} \mathbf{x}\|_{L^2(\Omega, \mu; \mathbb{R})} = 0$ .<sup>20</sup> Indeed, for many uncertainty models Assumption 2.3 is often satisfied for a minimum degree of  $d_x = 1$ , namely whenever Gaussian, Beta, Gamma, or uniform distributions are employed to model uncertainties. This is often done in the field of systems and control, see [57, 92, 103, 108, 109, 121].

Two questions arise:

- Q2.1 Choosing the PCE output dimension equal to the PCE input dimension what truncation error is made given a square-integrable nonlinear mapping?
- Q2.2 What is the minimum PCE dimension such that a zero truncation error is attained for?

We analyze and tackle these questions for polynomial mappings first.

### 2.5.1 Polynomial mappings

Let  $x_i, y \in \mathcal{H} = L^2(\Omega, \mu; \mathbb{R})$  with  $i \in \{1, \dots, n_x\}$  be real-valued random variables (with  $\Omega = \mathbb{R}$ ). Moreover, let the space generated by the orthogonal sequence of ordered  $m$ -variate monic orthogonal polynomials  $\mathcal{P}_\mu^m$  be a complete subspace of  $\mathcal{H}$ . For ease of presentation we consider that  $y$  and all  $x_i$  have the same PCE dimension.

**Theorem 2.3** (Error under polynomial mapping [122, 123]). *Suppose that all  $x_i$  satisfy Assumption 2.3 with minimum degree  $d_x$  relative to the orthogonal basis  $\mathcal{P}_\mu^m$ , and let  $f: \mathcal{H}^{n_x} \rightarrow \mathcal{H}$  be a square-integrable polynomial mapping*

---

<sup>20</sup> It might be the case that  $\|\mathbf{x} - \Pi_{\hat{k}} \mathbf{x}\|_{L^2(\Omega, \mu; \mathbb{R})} = 0$  for  $\hat{k} < \hat{k}_x$ , for instance  $\mathbf{x} = x_0 \phi_0 + x_1 \phi_1$  and  $\mathcal{P}_\mu^1 = \{\phi_k\}_{k=0}^2$ .

of total degree  $d_f$  such that  $\mathbf{y} = f(\mathbf{x}_1, \dots, \mathbf{x}_{n_x})$ . Then, the magnitude of the truncation error  $\mathbf{e}_n = \mathbf{y} - \Pi_n \mathbf{y}$  is

$$e_n := \|\mathbf{e}_n\| = \begin{cases} \sqrt{\sum_{k=n+1}^{\ell} y_k^2 \|\phi_k\|^2}, & n < \ell, \\ 0, & n \geq \ell, \end{cases} \quad (2.34a)$$

where

$$\ell + 1 = \frac{(m + d_x d_f)!}{m!(d_x d_f)!}, \quad (2.34b)$$

and  $y_k$  are the PCE coefficients of  $y$ .

*Proof.* From Assumption 2.3 the PCE for  $\mathbf{x}_i$  is given by

$$\mathbf{x}_i = \sum_{k=0}^{\hat{k}_x} x_{i,k} \phi_k \quad \forall i \in \{1, \dots, n_x\}, \quad (2.35)$$

where  $\hat{k}_x + 1 = (m + d_x)! / (m! d_x!)$ . Without loss of generality we assume the mapping  $f$  is of the form

$$f(\mathbf{x}_1, \dots, \mathbf{x}_{n_x}) = \mathbf{x}_1^{\alpha_1} \dots \mathbf{x}_{n_x}^{\alpha_{n_x}} + \dots, \quad (2.36)$$

where  $\alpha = [\alpha_1 \dots \alpha_{n_x}]^\top \in \mathbb{N}_0^{n_x}$  is a multi-index with  $|\alpha| = d_f$ . We insert the PCE for every  $\mathbf{x}_i$

$$\begin{aligned} f\left(\sum_{k=0}^{\hat{k}_x} x_{1,k} \phi_k, \dots, \sum_{k=0}^{\hat{k}_x} x_{n_x,k} \phi_k\right) = \\ \left(\sum_{k=0}^{\hat{k}_x} x_{1,k} \phi_k\right)^{\alpha_1} \cdots \left(\sum_{k=0}^{\hat{k}_x} x_{n_x,k} \phi_k\right)^{\alpha_{n_x}} + \dots = \\ \gamma t_1^{\beta_1} \cdots t_m^{\beta_m} + \dots, \end{aligned} \quad (2.37)$$

where  $\gamma$  is some constant. In (2.37) the multi-index  $\beta = [\beta_1 \dots \beta_m]^\top \in \mathbb{N}_0^m$  satisfies  $|\beta| = d_x |\alpha| = d_x d_f$ . Hence, the total degree of  $\mathbf{y}$  is  $d_x d_f$ , yielding a

total number  $\ell + 1$  of basis elements given by (2.16e), thus enlarging the basis by  $\ell - \hat{k}_x$  elements. The orthogonal projection of  $\mathbf{y}$  reads  $\Pi_n \mathbf{y} = \sum_{k=0}^n y_k \phi_k$ . Consequently, the truncation error  $\mathbf{e}_n$  becomes  $\mathbf{e}_n = \sum_{k=n+1}^{\ell} y_k \phi_k$ , which is zero in case of  $n \geq \ell$ . For  $n < \ell$ , apply Parseval's identity to obtain  $\|\mathbf{e}_n\|$ , cf. [28].  $\square$

With the help of Theorem 2.3 we provide answers to questions Q2.1 and Q2.2.

**Corollary 2.2** (Error for polynomial mapping [122]).

- A2.1 *Given a polynomial mapping  $f(\cdot)$  such that  $\mathbf{y} = f(x_1, \dots, x_{n_x})$  with  $y, x_1, \dots, x_{n_x} \in L^2(\Omega, \mu; \mathbb{R})$ , and choosing the PCE output dimension equal to the PCE input dimension, the truncation error is given by  $e_{\hat{k}_x}$  from (2.34a).*
- A2.2 *Furthermore, the minimum dimension to attain a zero truncation error is given by (2.34b).*

## 2.5.2 Quadratic programs

quadratic programs (QPs) are a special and often-encountered class of optimization problems. In the context of systems and control they frequently appear when applying model predictive control to discrete-time linear time-invariant systems with convex polytopic constraints and a convex quadratic cost function [105, 138]. Also, QPs are the basis for sequential quadratic programming methods for solving nonlinear programs that are encountered in nonlinear model predictive control. In power systems applications QPs occur when solving optimal power flow problems under so-called DC power flow conditions, see Section 4.2 and Appendix A.6. However, in many cases the data of the QP is uncertain—in these cases polynomial chaos can be of help.

**Problem 2.3** (Convex QP with uncertain data). *Let  $\mathbf{h} \in L^2(\Omega, \mu; \mathbb{R}^{n_x})$  be an  $\mathbb{R}^{n_x}$ -valued random vector, and let  $\mathbf{b} \in L^2(\Omega, \mu; \mathbb{R}^{n_{\text{con}}})$  be an  $\mathbb{R}^{n_{\text{con}}}$ -valued random vector. Set  $\mathbf{x} = [x_1^\top \ x_2^\top]^\top = [\mathbf{h}^\top \ \mathbf{b}^\top]^\top \in L^2(\Omega, \mu; \mathbb{R}^{n_x + n_{\text{con}}})$ . All these random variables satisfy Assumption 2.3, each with known and exact finite PCE dimension  $\hat{k} + 1$ .<sup>21</sup> Let the space generated by the orthogonal sequence of*

<sup>21</sup> For brevity of notation we demand the same dimension  $\hat{k} + 1$ .

ordered  $m$ -variate monic orthogonal polynomials  $\mathcal{P}_\mu^m$  be a complete subspace of  $L^2(\Omega, \mu; \mathbb{R})$ . For every realization  $x \in \mathbb{R}^{n_x + n_{\text{con}}}$  of  $\mathbf{x}$  consider the convex QP

$$\begin{aligned} y &:= \operatorname{argmin}_{\chi \in \mathbb{R}^{n_x}} \frac{1}{2} \chi^\top H \chi + x_1^\top \chi \\ \text{s. t.} \quad & A\chi + x_2 \leq 0, \end{aligned} \quad (2.38)$$

for some positive definite  $H \in \mathbb{R}^{n_x \times n_x}$  and a non-empty feasible set  $\{\chi \in \mathbb{R}^{n_x} : A\chi + x_2 \leq 0\}$ . Then, the problem is to find the  $\mathbb{R}^{n_x}$ -valued random variable  $y$  and quantify the element-wise truncation error  $\|y_i - \Pi_n y_i\|$  for all  $i \in \{1, \dots, n_x\}$ .

**Remark 2.5** (Overloading the argmin). In Problem 2.3 we interpret the argmin operator as a mapping from the realization  $x$  of the given random variable  $\mathbf{x}$  to the realization  $y$  of the sought random variable  $y$ . We shy away from directly overloading the argmin operator with the random-variable notation as that would lead to reformulating the inequality constraint: an inequality constraint in terms of random variables is per se not meaningful as there is no intuitive way of ordering random variables. A possible way out is to use chance constraints. We defer the discussion of chance constraints to the power systems applications in Chapter 4. The purpose of Problem 2.3 is to introduce a point of view: we are interested in evaluating the argmin of a convex QP for parameters that follow a known probability distribution.

**Remark 2.6** (QPs and model predictive control). In case of linear-quadratic model predictive control, Problem 2.3 is equivalent to considering uncertainty with respect to the initial condition at every time instant [105]. This uncertainty may be due to state estimation, or a lack of measurement precision/availability. The open-loop optimal control problem over the prediction horizon  $N \in \mathbb{N}$  for linear-quadratic model predictive control (MPC) can be written as

$$y := \operatorname{argmin}_{u \in \mathbb{R}^{N n_u}} \frac{1}{2} u^\top H u + x_k^\top F^\top u \quad (2.39a)$$

$$\text{s. t.} \quad Pu + Vx_k + v \leq 0. \quad (2.39b)$$

The construction of the QP ingredients  $H, F, P, V, v$  from the linear time-invariant system  $(A, B)$ , weight matrices  $Q, R$ , and constraint sets  $\mathcal{X}, \mathcal{U}$  is described for instance in [105].

We can leverage the results from polynomial mappings to tackle Problem 2.3.

**Theorem 2.4** (Uncertainty quantification for Problem 2.3 [122, 123]). *For all realizations of  $\mathbf{x}$ , let the active constraints in Problem 2.3 satisfy the linear inequality constraint qualification (LICQ) at the optimal solution  $y$ . If the set of active constraints  $\mathcal{A} = \{a_1, \dots, a_{n_{\text{act}}}\} \subseteq \{1, \dots, n_{\text{con}}\}$  is the same for all realizations of  $\mathbf{x} = [\mathbf{h}^\top \mathbf{b}^\top]^\top$ , then the element-wise truncation error becomes*

$$\|y_i - \Pi_n y_i\| = \begin{cases} \sqrt{\sum_{k=n+1}^{\hat{k}_x} (w_i^{h^\top} h_k + w_i^{b^\top} M_{\mathcal{A}} b_k)^2} \|\phi_k\|^2, & n < \hat{k}_x, \\ 0 & n \geq \hat{k}_x, \end{cases}$$

where  $w_i^{h^\top}$ ,  $w_i^{b^\top}$  are the  $i^{\text{th}}$  rows with  $i \in \{1, \dots, n_x\}$  of the matrices  $W^h$ ,  $W^b$  that satisfy

$$\begin{bmatrix} W^h & W^b \\ V^h & V^b \end{bmatrix} = - \begin{bmatrix} H & A^\top M_{\mathcal{A}}^\top \\ M_{\mathcal{A}} A & 0 \end{bmatrix}^{-1}. \quad (2.40)$$

The active constraint selection matrix  $M_{\mathcal{A}} \in \mathbb{N}_0^{n_{\text{act}} \times n_{\text{con}}}$  is constructed from the active set  $\mathcal{A}$  and has elements  $(M_{\mathcal{A}})_{i a_i} = 1$  for all  $i \in \{1, \dots, n_{\text{act}}\}$ , and it is zero elsewhere.

*Proof.* Because the set of active constraints is supposed to be  $\mathcal{A}$  for all realizations, the Karush Kuhn Tucker (KKT) conditions hold in terms of a function of random variables

$$\begin{bmatrix} y \\ \lambda^* \end{bmatrix} = \begin{bmatrix} W^h & W^b \\ V^h & V^b \end{bmatrix} \begin{bmatrix} \mathbf{h} \\ M_{\mathcal{A}} \mathbf{b} \end{bmatrix}, \quad (2.41)$$

where (2.40) and  $\mathbf{x} = [\mathbf{x}_1^\top \mathbf{x}_2^\top]^\top = [\mathbf{h}^\top \mathbf{b}^\top]^\top$  are used. Invertibility follows from LICQ. Consequently for all  $i \in \{1, \dots, n_x\}$ ,

$$y_i = w_i^{h^\top} \mathbf{h} + w_i^{b^\top} M_{\mathcal{A}} \mathbf{b} = \sum_{k=0}^{\hat{k}_x} (w_i^{h^\top} h_k + w_i^{b^\top} M_{\mathcal{A}} b_k) \phi_k,$$

and the result follows from Theorem 2.3 with  $d_f = 1$ .  $\square$

**Remark 2.7** (Extension to changes in the active set [122, 123]). *Note that even if the active set changes, Theorem 2.4 still holds locally on the current affine piece [18]. Furthermore, the error description can be turned into an upper bound by considering the worst case active set, which maximizes  $\|y_i - \Pi_n y_i\|$ .*

### 2.5.3 Non-polynomial mappings

We now turn to PCE truncation errors for non-polynomial mappings. Consider the real-valued random variables  $x_i, y \in \mathcal{H} = L^2(\Omega, \mu; \mathbb{R})$  for  $i \in \{1, \dots, n_x\}$ .

**Theorem 2.5** (Error for non-polynomial mapping [122]). *Let all  $x_i$  with  $i \in \{1, \dots, n_x\}$  satisfy Assumption 2.3 with minimum degree  $d_x$  relative to the orthogonal basis  $\mathcal{P}_\mu^m$ , and let  $f : \mathcal{H}^{n_x} \rightarrow \mathcal{H}$  be a square-integrable mapping such that  $y = f(x_1, \dots, x_{n_x})$ . Then, the magnitude of the truncation error  $e_n = y - \Pi_n y$  is*

$$e_n := \|e_n\| = \sqrt{\|y\|^2 - g^\top Q g}, \quad (2.42)$$

where  $Q = \text{diag}(1/\|\phi_0\|^2, \dots, 1/\|\phi_n\|^2) \in \mathbb{R}^{(n+1) \times (n+1)}$  is positive definite, and  $g = [g_1 \dots g_{n+1}]^\top \in \mathbb{R}^{n+1}$  with  $g_{j+1} = \langle y, \phi_j \rangle$  for all  $j \in \{0, \dots, n\}$ .

*Proof.* The PCE coefficients of  $\Pi_n y$  satisfy

$$\langle \phi_j, \phi_j \rangle y_j = \langle y, \phi_j \rangle \quad \forall j \in \{0, \dots, n\} \iff Q^{-1} \mathbf{y} = g, \quad (2.43)$$

which follows from orthogonality of the basis. The vector of PCE coefficients  $\mathbf{y} \in \mathbb{R}^{n+1}$  contains all PCE coefficients  $\mathbf{y} = [y_0 \dots y_n]^\top$ . The truncation error satisfies

$$\|e_n\|^2 = \langle y - \Pi_n y, y - \Pi_n y \rangle = \|y\|^2 - g^\top \mathbf{y}, \quad (2.44)$$

because  $(y - \Pi_n y) \perp \Pi_n y$ . Using (2.43), the result (2.42) follows.  $\square$

The error (2.42) can be computed efficiently using Gauss quadrature, see Appendix A.2. Recalling the questions Q2.1 and Q2.2 we asked at the beginning of this section we can now provide answers.

**Corollary 2.3** (Error for non-polynomial mappings [122]).

- A2.1 Given a non-polynomial mapping  $f(\cdot)$  such that  $y = f(x)$  with  $x, y \in L^2(\Omega, \mu; \mathbb{R})$ , and choosing the PCE output dimension equal to the PCE input dimension, the truncation error is given by  $e_{\hat{k}_x}$  from (2.42).
- A2.2 Unfortunately, a general statement remains open.<sup>22</sup> However, for a user-specified error threshold the according minimum PCE dimension is obtained from Theorem 2.5.  $\square$

## 2.5.4 Selected examples

We present some numerical examples from [122] that show how the presented results may be of use. The first example is more mathematical, while the other two examples are applications from systems and control.

**Example 2.6** (Quadratic mapping, Gaussian uncertainty [122]). Let  $n_x = 1$ ,  $m = 1$ , and let  $x$  be a Gaussian random variable with mean  $\mu$  and standard deviation  $\sigma > 0$ . Consider the mapping  $y = f(x) = x^2$ . If  $\mathcal{P}_\mu^m = \{\text{He}_0, \text{He}_1\}$ , where  $\text{He}_k$  is the  $k^{\text{th}}$  probabilists' Hermite polynomial, then Assumption 2.3 is satisfied with  $d_x = 1$ , and the PCE coefficients of  $x$  are  $x_0 = \mu$  and  $x_1 = \sigma$ . Direct inspection shows that  $y = f(x) = (\mu^2 + \sigma^2)\text{He}_0 + 2\sigma\mu\text{He}_1 + \sigma^2\text{He}_2$ . The error becomes  $e_1 = \sigma^2\text{He}_2$  with norm  $e_1 = \sqrt{2}\sigma^2 = \sqrt{2}x_1^2$ . The minimum exact PCE degree for  $y$  is  $d_x d_f = 2$ . Adding another basis function  $\mathcal{P}_\mu^m = \{\text{He}_0, \text{He}_1, \text{He}_2\}$ , the projection error becomes zero.

Table 2.3 shows the squared norm of  $e_{d_x} = y - \Pi_{d_x} y$  for an ascending input degree  $d_x$  and symbolic PCE input coefficients  $x_0, \dots, x_{d_x}$ ; the case in which we choose the PCE output dimension equal to the PCE input dimension. Note that higher-order coefficients tend to get weighted more heavily in the Hermitian basis.

**Example 2.7** (Linear-quadratic MPC [122]). Consider linear-quadratic MPC for an linear time-invariant (LTI) discrete-time model  $\chi(k+1) = A\chi(k) + Bu(k)$  of an aircraft. The open-loop optimal control problem can be cast as a QP [105]. The numerical values for the nominal system  $(A, B)$  and weights  $Q, R$  are taken

<sup>22</sup> Note that in the univariate case  $m = 1$  a zero truncation error in general requires an infinite PCE dimension, because a non-polynomial function cannot be represented exactly by a linear combination of a finite polynomial basis.

Table 2.3: Squared truncation errors  $e_{d_x}$  for Example 2.6.

$d_x$	1	2	3
$e_{d_x}^2$	$2x_1^4$	$24x_2^2(x_1^2 + x_2^2)$	$480x_2^2x_3^2 + 24(x_2^2 + 9x_3^2 + 2x_1x_3)^2 + 720x_3^4$

from [103]; the horizon length is  $N = 35$ . The input is the rate of change of the elevator angle, which introduces discrete-time integral action and an additional state. Uncertainty is introduced via the initial condition  $\chi(0) = \mathbf{x}$  for the altitude: it is modeled by the random variable  $x_4$  that follows a Beta distribution on  $[-402, -381]$  with shape parameters  $\alpha = 2$ ,  $\beta = 5$ , yielding the uncertain initial condition  $\mathbf{x} = [0 \ 0 \ 0 \ x_4 \ 0]^\top$ . Assumption 2.3 is satisfied with  $d_x = 1$  and the PCE coefficients are  $x_{4,0} = -396$ ,  $x_{4,1} = 3$  for a Jacobi polynomial basis. Figure 2.2 shows the evolution of the  $6\sigma$ -interval of the optimal input over time—note that the realization of the optimal random variable resembles the control input  $u$  to the system. For all realizations of the initial condition  $\mathbf{x}$  the constraints for the second state are active on the interval  $[0.5, 7.5]$  s, with no inequality constraints being active thereafter. Following Theorem 2.4, a Jacobi polynomial basis with  $n \geq 1$  allows a zero PCE truncation error in the decision variable  $y$ . The corresponding optimal control input trajectory over  $[0.0, 7.0]$  s is deterministic, as shown in Figure 2.2a. In terms of PCE coefficients, this is equivalent to all PCE coefficients of order greater than zero being zero, yielding a Dirac-delta as a probability density function. It is after the constraints become inactive that uncertainty plays a role; depicted in Figure 2.2b for  $t \in [7.5, 17.0]$  s.<sup>23</sup> Because the closed-loop system is asymptotically stable, the input uncertainty eventually fades out, resulting again in Dirac-deltas.

**Example 2.8** (Linear-quadratic regulator for continuous-time system [122]). Consider the continuous-time LTI dynamics  $\dot{\chi} = \mathbf{A}(\mathbf{x})\chi + \mathbf{B}u$  for a modified

<sup>23</sup> The histogram rather than the probability density is shown for sake of readability, because the peak of the densities for  $t \geq 12$  are orders of magnitude larger.

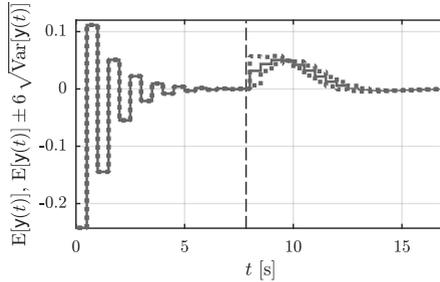


Figure 2.2a: Expected value and  $6\sigma$ -interval. Note that for  $t \leq 7$  s the input is deterministic as a state constraint is active.

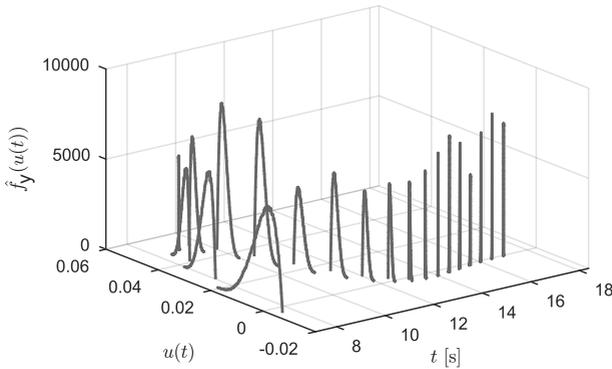


Figure 2.2b: Histogram of optimal input for  $t \geq 7.5$  s.

aircraft model from [105].<sup>24</sup> The initial condition is  $\chi(0) = [0 \ 0 \ 0 \ 40]^T$ , and the system dynamics are governed by

$$A = \begin{bmatrix} -1.2822+0.4x & 0 & 0.98 & 0 \\ 0 & 0 & 1 & 0 \\ -5.4293 & 0 & -1.8366 & 0 \\ -128.2 & 128.2 & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} -0.3 \\ 0 \\ -17 \\ 0 \end{bmatrix},$$

<sup>24</sup> The system matrix is a matrix-valued random variable.

where  $x \sim \mathcal{U}[-1, 1]$  follows a uniform distribution on  $[-1, 1]$ . The realization  $x = 0$  corresponds to the nominal system matrix  $A$ . The control  $u(t) = -K\chi(t)$  is a standard linear-quadratic regulator based on the weights  $Q = 0.001\mathbb{I}_4$  and  $R = 100$  for the nominal system  $(A, B)$ . Now we apply the above feedback to the uncertain system matrix. The closed-loop altitude trajectories  $\chi_4(t)$  are given in Figure 2.3 (left) for best case and worst case realizations, clearly showing the performance degradation under uncertainty. The uncertainty  $x$  is mapped to the state  $\chi(t)$  via the state transition map  $\chi(t) = \exp[(A(x) - BK)t]\chi_0$ . Figure 2.3 shows the altitude truncation error  $e_{4,n}(t)$  from (2.42) over time for increasing highest-degree  $n \in \{2, 3, 4\}$ . The basis consists of Legendre polynomials. The closed-loop system is asymptotically stable for all realizations of  $x$ , hence the truncation error decays to zero. However, it is clearly non-monotonic over time. Note how over- and undershooting of the deterministic solution, Figure 2.3 (left), carry over to the PCE error, Figure 2.3 (right).

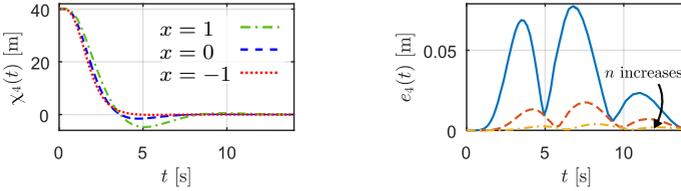


Figure 2.3: Closed-loop altitude trajectory for different  $x$ -realizations (left). Truncation error for altitude over time for difference PCE dimensions (right).

## 2.6 Problem formulation – Finalized

We have assembled all the material to formalize the verbose problem formulation from the beginning of this section, namely Problem 2.1.

**Problem 2.4** (Mapping under uncertainty). *Let  $(\Omega, \mathfrak{F}, \mu)$  be a probability space, for which  $L^2(\Omega, \mu; \mathbb{R})$  is the Hilbert space of all real-valued random variables of finite variance that is spanned by the uni- or multivariate orthogonal polynomials  $\{\phi_k\}_{k \in \mathbb{N}_0}$ , and let  $x_i \in L^2(\Omega, \mu; \mathbb{R})$  be  $n_x \in \mathbb{N}$  given random*

variables for  $i \in \{1, \dots, n_x\}$  that admit a PCE on the subspace spanned by  $\{\phi_k\}_{k \in \mathcal{K}}$  with  $\mathcal{K} = \{0, \dots, \hat{k}\}$ . Furthermore, let

$$f: L^2(\Omega, \mu; \mathbb{R}) \times \dots \times L^2(\Omega, \mu; \mathbb{R}) \rightarrow L^2(\Omega, \mu; \mathbb{R}) \times \dots \times L^2(\Omega, \mu; \mathbb{R})$$

be a  $\mathfrak{B}(\mathbb{R}^{n_x})/\mathfrak{B}(\mathbb{R}^{n_y})$ -measurable and square-integrable function. Then, find the random variables  $y_i \in L^2(\Omega, \mu; \mathbb{R})$  in terms of their PCE coefficients for all  $i \in \{1, \dots, n_y\}$  such that

$$y = f(x) \quad \text{with} \quad y = [y_1, \dots, y_{n_y}]^\top, x = [x_1, \dots, x_{n_x}]^\top.$$

Specifically, the non-set-valued mapping  $f$  may represent one of the following:

1. The solution to a system of nonlinear algebraic equations;
2. The argmin operator of an optimization problem.

In short, Problem 2.4 exploits polynomial chaos to represent mappings under uncertainty in terms of deterministic mappings of finitely many coefficients.

**Remark 2.8** (No set-valued argmin operators). *Implicit to Problem 2.4 is the assumption that the mapping  $f$  is not set-valued. This is especially important in light of argmin operators. For example, consider for all  $x > 0$*

$$f(x) = \underset{z \in \mathbb{R}}{\operatorname{argmin}} x(z - 1)^2(z + 1)^2 = \{-1, 1\},$$

which is a case we exclude by restricting  $f$  to be non-set-valued.

**Remark 2.9** (Interpretation of solution). *We emphasize that the nature of the mapping may be an explicit or implicit mapping, but it may also be the argmin operator of some optimization problem. That means the result of Problem 2.4 may be an optimal random variable. This does not mean that the argmin of the optimization problem takes on arbitrary values. Instead, from the engineering lens propagating uncertainties through an argmin operator means to solve many deterministic optimization problems—infinately many for continuous random variables—and studying the distribution of the resulting solutions. Problem 2.4 aims to perform the propagation in a single step. While we often think of continuous random variables in terms of their probability density functions and their moments, it is advisable to recall that mathematically a continuous random variable is merely a measurable function defined for a probability space,*

see Section 2.2.2. This function maps elements from its domain to the real numbers, see Section 2.4.3. Specifically, given a realization of the input uncertainty  $x$  we can evaluate the output uncertainty  $y$ , immediately obtaining a realization of the argmin; the solution provides a policy, see Example 2.10. This situation is typical in optimal control: a given optimal control problem is—among others—parameterized by the system’s initial condition. In case we are able to derive a closed-form solution of the optimal control problem, the optimal input trajectory is a function of this initial condition. Different initial conditions yield different optimal input trajectories. If we were to overload the initial condition with an uncertainty, we would be back in the setting from Problem 2.4.

Before we proceed let us revisit the very first examples we introduced at the beginning of this section.

**Example 2.9** (PCE applied to Example 2.1). *The problem is equivalent to finding the PCE coefficients of the random variable that follows a chi-squared distribution with  $m$  degrees of freedom.*

For every random variable  $x_i \sim N(0, 1)$  we introduce a distinct Hermitian basis  $\mathcal{P}_{\mu_i} = \{\text{He}_0, \text{He}_1^{(i)}, \text{He}_2^{(i)}\}$  for all  $i \in \{1, \dots, m\}$ , from which we construct the  $m$ -variate basis with total degree at most 2 as follows

$$\mathcal{P}_{\mu}^m = \{1, \text{He}_1^{(1)}, \text{He}_1^{(2)}, \dots, \text{He}_1^{(m)}, \dots\}$$

with a total of  $|\mathcal{P}_{\mu}^m| = 1 + \hat{k} = (m + 4)!/(m!4!)$  elements. Recall that we employ the reversal of lexicographic ordering, see Example 2.5. We set  $\mathcal{K} = \{0, 1, \dots, \hat{k}\}$ . With that, the PCE coefficients  $x_{i,k} \in \mathbb{R}$  of every  $x_i$  read

$$x_{i,k} = \begin{cases} 0, & k \in \mathcal{K} \setminus \{i + 1\}, \\ 1, & k \in \{i + 1\}, \end{cases}$$

for all  $i \in \{1, \dots, m\}$ . We obtain the PCE coefficients  $y \in \mathbb{R}^{\hat{k}+1}$  of  $y$  by Galerkin projection. This means to solve the following equations

$$y_k \langle \phi_k, \phi_k \rangle = \sum_{i=1}^m \sum_{k_1, k_2 \in \mathcal{K}} x_{i,k_1} x_{i,k_2} \langle \phi_{k_1} \phi_{k_2}, \phi_k \rangle$$

for all  $k \in \mathcal{K}$ .

**Example 2.10** (PCE applied to Example 2.2). *We are interested in the PCE coefficients of the argmin operator of an equality-constrained convex quadratic program. We know that we can express  $x$  exactly with two PCE coefficients in a Jacobi basis. Specifically, the PCE coefficients of  $x$  are*

$$x_0 = \frac{\alpha\bar{x} + \beta\underline{x}}{\alpha + \beta}, \quad x_1 = \bar{x} - \underline{x},$$

in the Jacobi basis  $\{\phi_k\}_{k=0}^1 = \{1, \xi - \frac{1}{2}(1 + \frac{(\alpha-1)^2(\beta-1)^2}{(\alpha+\beta-2)(\alpha+\beta)})\}$  which is orthogonal relative to

$$\rho(\xi) = \frac{\xi^{\alpha-1}(1-\xi)^{\beta-1}}{B(\alpha, \beta)},$$

where  $B(\cdot, \cdot)$  is the Beta function. In terms of random variables the solution to problem (2.5) reads

$$y_i = \nu_i x, \quad \nu_i = \frac{A_{3-i,3-i} - A_{1,2}}{A_{1,1} + A_{2,2} - 2A_{1,2}}, \quad \forall i \in \{1, 2\},$$

$$\text{where } A = \begin{bmatrix} A_{1,1} & A_{1,2} \\ A_{1,2} & A_{2,2} \end{bmatrix}$$

with  $A_{1,1} + A_{2,2} - 2A_{1,2} \neq 0$ . Given a specific realization  $x$  of the random variable  $x$  we immediately obtain the resulting argmin from  $y_i = \nu_i x$ .

We find the PCE coefficients of  $y$  to be

$$y_0 = \begin{bmatrix} \nu_1 \\ \nu_2 \end{bmatrix} \frac{\alpha\bar{x} + \beta\underline{x}}{\alpha + \beta}, \quad y_1 = \begin{bmatrix} \nu_1 \\ \nu_2 \end{bmatrix} (\bar{x} - \underline{x}).$$

Hence, we have an analytic expression for the desired PCE coefficients.

In terms of random variables the optimal cost becomes

$$z = \frac{1}{2} y^{*\top} A y^* = \frac{\det A}{2(A_{1,1} + A_{2,2} - 2A_{1,2})} \cdot x^2 =: \omega x^2,$$

and the PCE coefficients of  $z$  can be obtained from

$$z_k = \omega \sum_{k_1=0}^2 \sum_{k_2=0}^2 x_{k_1} x_{k_2} \frac{\langle \phi_{k_1} \phi_{k_2}, \phi_i \rangle_{L^2(\Omega, \mu; \mathbb{R})}}{\langle \phi_i, \phi_i \rangle_{L^2(\Omega, \mu; \mathbb{R})}}$$

for all  $k \in \{0, 1, 2\}$  with  $x_2 = 0$ . Note that we expanded the basis to attain a zero truncation error for the optimal cost  $z$ .

The problem is simple enough to allow for further insights. For example, we can compute the moments of  $x$

$$\begin{aligned} \mathbb{E}(x) &= x_0 = \frac{\alpha \bar{x} + \beta \underline{x}}{\alpha + \beta}, \\ \mathbb{V}(x) &= x_1^2 \|\phi_1\|_{L^2(\Omega, \mu; \mathbb{R})}^2 = (\bar{x} - \underline{x})^2 \frac{\alpha \beta}{(\alpha + \beta)^2 (\alpha + \beta + 1)}, \end{aligned}$$

and the moments of  $y$

$$\begin{aligned} \mathbb{E}(y) &= y_0 = \begin{bmatrix} \nu_1 \\ \nu_2 \end{bmatrix} \frac{\alpha \bar{x} + \beta \underline{x}}{\alpha + \beta}, \\ \mathbb{V}(y) &= \begin{bmatrix} y_{1,1}^2 \\ y_{2,1}^2 \end{bmatrix} \|\phi_1\|_{L^2(\Omega, \mu; \mathbb{R})}^2 = \begin{bmatrix} \nu_1^2 \\ \nu_2^2 \end{bmatrix} (\bar{x} - \underline{x})^2 \frac{\alpha \beta}{(\alpha + \beta)^2 (\alpha + \beta + 1)}. \end{aligned}$$

Furthermore, the probability density of each minimizer is

$$\rho(y_i) = \frac{(y_i - \nu_i \underline{x})^{\alpha-1} (\nu_i \bar{x} - y_i)^{\beta-1}}{(\nu_i \bar{x} - \nu_i \underline{x})^{\alpha+\beta-1} B(\alpha, \beta)}$$

for  $y_i \in [\nu_i \underline{x}, \nu_i \bar{x}]$  for  $i \in \{1, 2\}$ . Similarly, the probability density of the optimal cost is

$$\rho(z) = \frac{1}{2\sqrt{2}\omega z (\bar{x} - \underline{x})^{\alpha+\beta-1} B(\alpha, \beta)} \left( \sqrt{\frac{z}{\omega}} - \underline{x} \right)^{\alpha-1} \left( \bar{x} - \sqrt{\frac{z}{\omega}} \right)^{\beta-1}$$

for all  $z \in [\omega \underline{x}^2, \omega \bar{x}^2]$ .



## 3 Power flow under uncertainty

A power system is a large-scale alternating-current (AC) electrical circuit with considerable power consumption and generation that operates on different voltage levels (for example 380 kV, 220 kV, 110 kV, 20kV, 10 kV, 0.4 kV in Germany [149]). Typical components of a power system are generators, loads, transformers, shunt elements, and lines. Borrowing vocabulary from graph theory, we treat a power system as a connected, directed graph (without self-loops). The nodes of the graph correspond to generators, loads, or transformers; its edges are electrical lines. Thinking of the power system as a connected, directed graph without self-loops, we call its nodes the buses, and the edges are the lines. Given a power system with a known topology and known physical parameters, and given some known power demands and known power generations, we strive to determine all the quantities in the system such that Kirchhoff's laws and Ohm's law are satisfied—this is the traditional power flow study [8]. For the scope of this thesis the key word in that described task is *known*. What happens for instance if we are unable to specify the power demand in terms of a single real-valued number because we are unsure about what the precise value is going to be? Perhaps the power demand at a certain bus is modeled more adequately in terms of a continuous random variable with a probability density function.<sup>1</sup> The cornerstone assumption for the scope of this thesis is that we can adequately model any form of uncertainty in terms of continuous random variables. Hence, we are interested in the following power flow problem *under uncertainty*: given a power system with a known topology and known physical parameters, and

---

<sup>1</sup> In fact, the possible origins of uncertainty in power flow problems are numerous. For instance, the values for the physical parameters of the power systems, e.g. line admittances, may not be known precisely, and/or they might vary with changing weather conditions. However, we assume another origin to have a much more significant effect: the volatile and difficult-to-predict character of renewable energy sources, as well as all forms of uncertainties that root in forecasts. Forecasts, be they for renewable energies or for load/demand, are uncertain by nature. For specifics on probabilistic forecasts we refer to [73–75].

given some *unknown* power demands and *unknown* power generations, determine all the quantities in the system such that Kirchhoff's laws and Ohm's law are satisfied.

Before we can attempt to solve this task we need to make assumptions on how to model the power system.

**Assumption 3.1** (Power system model [8, 58, 61, 78, 168]). *We consider*

1. *a connected system/graph;*
2. *a lumped-parameter system;*
3. *steady-state conditions;*
4. *lines that are modeled by their  $\Pi$ -line equivalents (see Figure 3.1);*
5. *a balanced power system in terms of its single-phase equivalent.*

Under Assumption 3.1 the equations to model a power system become significantly simpler: we study merely a single problem (Item 1), partial differential equations become ordinary differential equations (Item 2) which become algebraic equations (Item 3, Item 4) that are reduced in number (Item 5).

We present a generic power flow problem under uncertainty (Section 3.1) and show how polynomial chaos facilitates its solution (Section 3.3). Thereafter, we study two specific applications of power flow under uncertainty: namely, how power flow can be simplified if the power system is radial (Section 3.4), and how linearizations can replace having to solve a full power flow (Section 3.5). We tackle each task by studying two consecutive questions:

Q3.1 *Problem formulation:* What is a mathematically sound formulation of the specific power flow problem under uncertainty?

Q3.2 *Solution methodology:* Having obtained a mathematical formulation, what is a solution approach?

**Remark 3.1** (Notation for random variables). *For everything to follow we model uncertainties as continuous random variables of finite variance. We leverage polynomial chaos expansion to represent all of the continuous random variables in terms of their deterministic PCE coefficients (to do so, we rely on the notations and definitions from Section 2.4). Specifically, we study probability*

spaces  $(\Omega, \mathfrak{F}, \mu) = (\mathbb{R}^m, \mathfrak{B}(\mathbb{R}^m), \mu)$  where  $m$  is the number of unique uncertainties, and the measure  $\mu$  is absolutely continuous and viewed as a probability distribution  $\mathbb{P}$  with respect to a known probability density.

### 3.1 Problem formulation

The following is inspired by [8, 58, 61, 78, 168]. We consider an  $N_b$ -bus electrical network that satisfies Assumption 3.1 and represent it by its set of bus indices  $\mathcal{N}$  with  $|\mathcal{N}| = N_b \in \mathbb{N}$ ; unless stated otherwise we set  $\mathcal{N} = \{1, \dots, N_b\}$ .<sup>2</sup> Each bus  $i \in \mathcal{N}$  is described by its complex power and its complex voltage: the complex power has real part  $p_i$  and imaginary part  $q_i$ , which are called active power and reactive power respectively; the complex voltage at bus  $i \in \mathcal{N}$  in rectangular coordinates has real part  $e_i$  and imaginary part  $f_i$ . Hence, we introduce

$$z_i = \begin{bmatrix} p_i & q_i & e_i & f_i \end{bmatrix}^\top \in \mathbb{R}^4 \quad (3.1)$$

as the state of bus  $i \in \mathcal{N}$ .<sup>3</sup> The standing assumption for what is to follow is that our knowledge about the state of every bus in the power system is uncertain. Thus we treat each bus state  $z_i$  as the realization of an  $\mathbb{R}^4$ -valued random variable. We formalize this as follows.

**Assumption 3.2** (Power system under uncertainty). *We study a power system under Assumption 3.1 in the presence of uncertainties. That implies we model the state of every bus  $i \in \mathcal{N}$  by an  $\mathbb{R}^4$ -valued random variable*

$$z_i = \begin{bmatrix} p_i & q_i & e_i & f_i \end{bmatrix}^\top \in L^2(\Omega, \mu; \mathbb{R}^4). \quad (3.2)$$

Assumption 3.2 addresses the *mathematical* model we choose to describe the state of the power system. What is the *physical* model of the power system,

<sup>2</sup> The backward-forward sweep method is an exception, see Section 3.4.

<sup>3</sup> A different representation of the complex-valued phasors leads to a different definition of the bus state. For DC power flow, for instance, it is more convenient to express the power in rectangular coordinates and the voltage phasor in polar coordinates, see Appendix A.6 and Section 3.5.

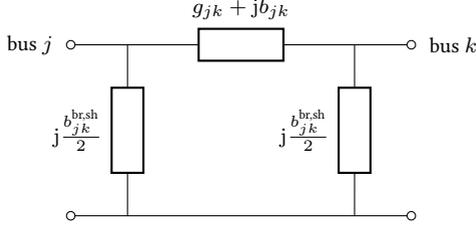


Figure 3.1:  $\Pi$ -line model of a transmission line between buses  $j$  and  $k$  [8, 78, 168].

hence how do the entries of the bus states relate to each other? The answer is given by the so-called power flow equations which we derive directly in terms of random variables.

We first relate the state  $\mathbf{z}_i$  of every bus  $i \in \mathcal{N}$  to power flows among the lines. All buses  $i \in \mathcal{N}$  are connected to each other by electrical lines (also called branches). We write  $\mathcal{L} \subseteq \mathcal{N} \times \mathcal{N}$  for the set of directed line indices

$$\mathcal{L} = \{(j, k) \in \mathcal{N} \times \mathcal{N} : \text{line begins in bus } i \text{ and ends in bus } j\}, \quad (3.3)$$

where  $|\mathcal{L}| = N_{\text{br}}$  is the number of lines. Each line  $(j, k) \in \mathcal{L}$  is modeled by a  $\Pi$ -circuit with shunt susceptance  $b_{jk}^{\text{br,sh}}$ , mutual conductance  $g_{jk}$ , and mutual susceptance  $b_{jk}$ , see Figure 3.1.<sup>4</sup> Kirchhoff's laws and Ohm's law relate the real and imaginary part of the random bus voltages to the random active power flow  $p_{jk}^{\text{br}}$  and the random reactive power flow  $q_{jk}^{\text{br}}$  across each line  $(j, k) \in \mathcal{L}$

$$p_{jk}^{\text{br}} = g_{jk}(e_j^2 + f_j^2 - e_j e_k - f_j f_k) + b_{jk}(e_j f_k - f_j e_k), \quad (3.4a)$$

$$q_{jk}^{\text{br}} = -(b_{jk} + b_{jk}^{\text{br,sh}}/2)(e_j^2 + f_j^2) - b_{jk}(e_j e_k - f_j f_k) + g_{jk}(e_j f_k - f_j e_k). \quad (3.4b)$$

We can use the branch flows from (3.4) to formulate energy balance at each

<sup>4</sup> For ease of presentation we do not consider transformers in the derivation to follow. Including transformers leads to the so-called unified power flows equations [8, 61].

Table 3.1: Common bus specifications.

Bus name	Fixed quantities
pQ	Active power & reactive power
pV	Active power & voltage magnitude
Slack	Voltage angle & voltage magnitude

bus  $i \in \mathcal{N}$ . If we assume that each bus connects not just to branches but also to some shunt, see Figure 3.2, the energy balance reads

$$p_i = \sum_{\substack{(j,k) \in \mathcal{L} \\ \text{s.t. } j=i}} p_{jk}^{\text{br}} + \sum_{\substack{(j,k) \in \mathcal{L} \\ \text{s.t. } k=i}} p_{kj}^{\text{br}} + g_i^{\text{sh}} (e_i^2 + f_i^2), \quad (3.5a)$$

$$q_i = \sum_{\substack{(j,k) \in \mathcal{L} \\ \text{s.t. } j=i}} q_{jk}^{\text{br}} + \sum_{\substack{(j,k) \in \mathcal{L} \\ \text{s.t. } k=i}} q_{kj}^{\text{br}} - b_i^{\text{sh}} (e_i^2 + f_i^2). \quad (3.5b)$$

These equations (3.5) are the power flow equations in terms of random variables. They are the physical model of the power system under uncertainty.<sup>5</sup> These  $2N_b$  equations relate all the bus states  $z_i$  for all  $i \in \mathcal{N}$ —hence a total of  $4N_b$  random variables—to each other.

Different buses represent different equipment, the behavior of which is typically modeled in one of three ways shown in Table 3.1. Hence, we decompose the bus index set  $\mathcal{N} = \mathcal{N}_{\text{pQ}} \cup \mathcal{N}_{\text{pV}} \cup \mathcal{N}_{\text{SL}}$  into the set of pQ buses  $\mathcal{N}_{\text{pQ}}$ , the set of pV buses  $\mathcal{N}_{\text{pV}}$  and the set of slacks  $\mathcal{N}_{\text{SL}}$ . Mathematically, each of these bus specifications fixes two values of the bus state  $z_i$ . The purpose of the slack bus is to provide an angle reference. In the absence of such an angle reference it would not be possible to entangle the voltages as they always appear in differences; the power flow equations exhibit rotational degeneracy [114]. For ease of presentation we assume the sets  $\mathcal{N}_{\text{pQ}}$ ,  $\mathcal{N}_{\text{pV}}$ , and  $\mathcal{N}_{\text{SL}}$  are mutually disjoint, and that there is just one slack bus such that  $|\mathcal{N}_{\text{SL}}| = 1$ .

To summarize: thus far we have presented the power flow equations, and the bus specifications. Let us now aim for a concise mathematical formulation of

<sup>5</sup> This form of the power flow equations is the so-called branch flow model [8, 61]. There also exists the formulation as the bus injection model, see Appendix A.6.

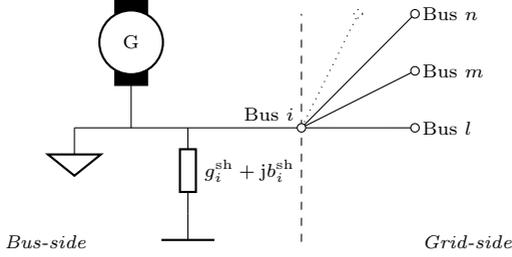


Figure 3.2: Power balance at bus  $i$  considers all the line flows from the grid-side and the nodal shunt flow from the bus-side.

the PPF problem in terms of random variables. We treat the power flow equations from (3.5) as a nonlinear mapping  $g: L^2(\Omega, \mu; \mathbb{R}^{4N_b}) \rightarrow L^2(\Omega, \mu; \mathbb{R}^{2N_b})$

$$\mathbf{z} = \left[ \mathbf{z}_1 \quad \dots \quad \mathbf{z}_{N_b} \right]^\top \mapsto g(\mathbf{z}). \quad (3.6)$$

We formulate the verbose bus specifications from Table 3.1 mathematically via  $h_i: L^2(\Omega, \mu; \mathbb{R}^4) \rightarrow L^2(\Omega, \mu; \mathbb{R}^2)$  for every bus  $i \in \mathcal{N} = \mathcal{N}_{\text{PQ}} \cup \mathcal{N}_{\text{PV}} \cup \mathcal{N}_{\text{SL}}$

$$\mathbf{z}_i \mapsto \mathbf{x}_i = h_i(\mathbf{z}_i) = \begin{cases} \begin{bmatrix} \mathbf{p}_i \\ \mathbf{q}_i \end{bmatrix}, & \text{if } i \in \mathcal{N}_{\text{PQ}}, \\ \begin{bmatrix} \mathbf{p}_i \\ \mathbf{e}_i^2 + \mathbf{f}_i^2 \end{bmatrix}, & \text{if } i \in \mathcal{N}_{\text{PV}}, \\ \begin{bmatrix} \mathbf{e}_i^2 + \mathbf{f}_i^2 \\ 0 \end{bmatrix}, & \text{if } i \in \mathcal{N}_{\text{SL}}. \end{cases} \quad (3.7)$$

We collect all bus specifications in the mapping  $h: L^2(\Omega, \mu; \mathbb{R}^{4N_b}) \rightarrow L^2(\Omega, \mu; \mathbb{R}^{2N_b})$  with

$$\left[ \mathbf{z}_1^\top \quad \dots \quad \mathbf{z}_{N_b}^\top \right]^\top \mapsto \mathbf{x} = \left[ h_1(\mathbf{z}_1)^\top \quad \dots \quad h_{N_b}(\mathbf{z}_{N_b})^\top \right]^\top. \quad (3.8)$$

In what follows the notion of “bus specification” refers to both the mappings  $h/h_i$  and their image. We are ready to formulate the probabilistic power flow problem mathematically.

**Problem 3.1** (Probabilistic power flow). *Let Assumption 3.1 and Assumption 3.2 hold. The probabilistic power flow problem is understood as mapping the bus specifications  $\mathbf{x} \in L^2(\Omega, \mu; \mathbb{R}^{2N_b})$  to a total bus state  $\mathbf{z} \in L^2(\Omega, \mu; \mathbb{R}^{4N_b})$  such that*

$$\begin{bmatrix} g(\mathbf{z}) \\ h(\mathbf{z}) \end{bmatrix} = \begin{bmatrix} \mathbf{0}_{2N_b} \\ \mathbf{x} \end{bmatrix}$$

*holds, where  $g$  are the power flow equations (3.6), and  $h$  are the bus specifications (3.8), both in terms of random variables.*

The idea of PPF is shown in Figure 3.3 for a four bus system. Bus 1 is the slack bus, bus 2 is an uncertain PQ bus, bus 3 is a certain PQ bus, and bus 4 is a PV bus. We visualize uncertainty in terms of the probability density function (PDF) of the active powers: at bus 2 there is a bi-modal density, at bus 4 there is Dirac-delta pulse as it is not uncertain. Figure 3.3a shows this setup. Figure 3.3b shows the solution to the PPF problem: the active power injection at the slack bus 1 mirrors the demand; the active power injection at the PV bus 3 is certain.

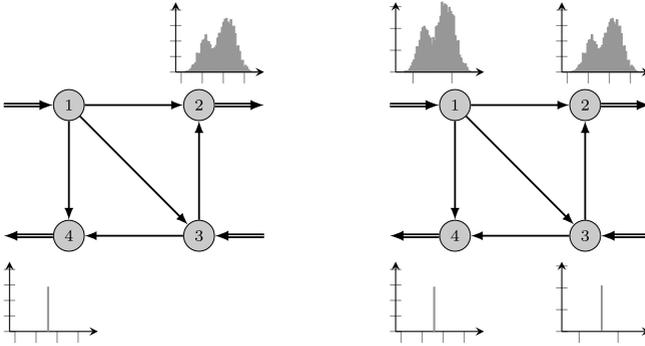
In light of Chapter 2 we can view PPF as an implicit mapping under uncertainties: the user-provided bus specifications  $\mathbf{x}$  are mapped through the power flow equations to an overall grid state in terms of random variables, namely  $\mathbf{z}$ . We loosely write

$$\mathbf{x} \xrightarrow{(3.6) \text{ and } (3.8)} \mathbf{z}, \quad (3.9)$$

highlighting that in general we are unable to provide a closed-form for the mapping (3.9).

The specific values of the bus specifications  $\mathbf{x}$  that occur in Problem 3.1 capture what the user knows about the power grid, for instance how the load at some PQ bus is expected to behave in terms of a random variable. The PPF according to Problem 3.1 is a system of nonlinear algebraic equations in terms of random variables, hence an infinite-dimensional problem that is challenging to solve in general. Before making an attempt to tackle the problem, let us investigate existing approaches from the literature.

**Remark 3.2** (Deterministic power flow). *If we imagine the random variables to have trivial probability density functions equal to Dirac-delta pulses centered at some nominal value, then the PPF problem reduces to the traditional determin-*



(a) Uncertainty specification at PQ buses: bus 2 is uncertain, bus 4 is certain (Dirac-delta pulse). (b) Having solved PPF we obtain the active power injection at the slack bus 1. Bus 3 is a PV bus (Dirac-delta pulse).

Figure 3.3: Idea of probabilistic power flow (PPF) for 4-bus system (bus 1 – slack bus, bus 2 – uncertain PQ bus, bus 3 – deterministic PQ bus, bus 4 – PV bus). The plots represent the probability density function (PDF) of the active power at every bus.

istic PF problem. Put differently, the PPF is a generalization of the deterministic PF problem in terms of probabilistic uncertainties.

**Remark 3.3** (Probabilistic DC power flow). *The AC power flow equations constitute a system of nonlinear algebraic equations. It is sometimes desirable to make additional assumptions in order to obtain a more pleasant computational formulation. A specific set of assumptions, often used for high-voltage transmission systems, is given by the so-called DC power flow conditions. Under these conditions the nonlinear AC power flow conditions reduce to a system of linear equations, see Appendix A.6 for a derivation. To study probabilistic DC power flow means to interpret and solve (A.47) from Appendix A.6 in terms of random variables. As DC power flow is linear, this is straightforward and can be done with zero truncation errors. The solution of (A.47) in terms of random variables gives the voltage angles in terms of random variables based on the active power bus specifications for all PQ and PV buses  $i \in \mathcal{N}_{\text{PQ}} \cup \mathcal{N}_{\text{PV}}$ . The net power at the slack bus is then computed from the line of (A.46) corresponding to  $\mathcal{N}_{\text{SL}}$ .*

## 3.2 Existing approaches

The deterministic PF problem is run-of-the-mill in the field of power systems engineering; its formulation is part of any given textbook on the subject such as [78, 134, 168]. The extension to PPF can be traced back to [27]: “[...] given a set of probable values of node loads, [...] the problem is to find the set of corresponding values of branch flows.” Ever since, the portfolio of methods to solve the PPF problem has been expanding. The existence of an 11-page bibliography [147] on power system analysis under uncertainties between the years 1962 and 1988 speaks for itself. We distinguish between analytical methods and sampling-based methods [36]. Sampling-based methods substitute the PPF problem by some (possibly large) number of deterministic PF problems. The ensemble of the solutions provides statistical information that is—hopefully—representative of the statistics of the true solution. The task of how to choose the deterministic PF problems is specific to the sampling-based method, e.g. (Markov chain) Monte-Carlo, (adaptive) rejection sampling, stratified random sampling, Latin hypercube sampling, or unscented transforms, see e.g. [4, 37, 40, 174].

Analytical methods analyze the PPF problem as an entity of mappings under uncertainty. The challenge is to propagate the uncertainties through the implicit nonlinear power flow equations. Analytical methods often modify the PPF problem to obtain a mathematical problem whose properties are easier to analyze and/or exploit. The earliest references on PPF may be classified as analytical methods. For instance, the already referenced paper [27] made the simplifying assumption that “active and reactive power flows are independent of each other.” The authors of [6, 49] linearized the power flow equations around a nominal operating point; propagating the uncertainties then becomes easy, because the mappings are affine after the linearization. Gram-Charlier expansions are another way to cope with PPF [178]. These expansions allow to write a series expansion of the probability density function. Moment-based methods may be viewed as a combination of analytical and sampling-based methods. These methods compute the statistics of the random variables by a fairly small number of deterministic power flow problems. The point-estimate method [86, 116, 153, 158] is a prominent example.

In this thesis we apply polynomial chaos expansion to PPF. Depending on how the PCE coefficients are computed we can classify polynomial chaos ei-

ther as a sampling-based method or as an analytical method. In 2016/2017 several independent works described how PCE can be applied to PPF problems [119, 131, 139, 169]. Interestingly, these works cover both sampling-based and analytical aspects of PCE. For example, [131, 139] present regression-based methods, and [119, 169] focus on the Galerkin-based approach instead. It is the hope that polynomial chaos can combine the advantages of both worlds: it is per se an analytical method, but may be implemented as a sampling-based method. Once the PCE coefficients are available, we have access to the moments, hence also entailing the advantages of moment-based approaches.

### 3.3 Solution methodology

The power flow problem under uncertainty according to Problem 3.1 is a problem of uncertainty propagation through implicit equations. We can employ PCE in combination with Galerkin projection to reformulate the infinite-dimensional problem as an enlarged deterministic problem. To apply PCE we need to assume that we know about the PCE of the bus specifications.

**Assumption 3.3** (PCE of bus specifications). *The overall bus specifications  $x \in L^2(\Omega, \mu; \mathbb{R}^{2N_b})$  from (3.8) in Problem 3.1 admit a finite and exact polynomial chaos expansion with respect to the polynomial basis  $\{\phi_k\}_{k \in \mathcal{K}}$  with  $\mathcal{K} = \{0, 1, \dots, \hat{k}\}$ . We write*

$$x = \sum_{k \in \mathcal{K}} \hat{x}_k \phi_k \quad \text{with} \quad \hat{x}_k \in \mathbb{R}^{2N_b}, \quad (3.10)$$

where the  $2N_b$  elements of  $\hat{x}_k$  are given by the individual PCE coefficients of all power system components

$$\left. \begin{aligned} p_i &= \sum_{k \in \mathcal{K}} \hat{p}_{i,k} \phi_k \\ q_i &= \sum_{k \in \mathcal{K}} \hat{q}_{i,k} \phi_k \end{aligned} \right\} \forall i \in \mathcal{N}_{\text{PQ}}, \quad (3.11a)$$

$$\left. \begin{aligned} \mathbf{p}_i &= \sum_{k \in \mathcal{K}} \hat{p}_{i,k} \phi_k \\ \mathbf{v}_i &= \sum_{k \in \mathcal{K}} \hat{v}_{i,k} \phi_k \end{aligned} \right\} \forall i \in \mathcal{N}_{\text{PV}}, \quad (3.11\text{b})$$

$$\left. \begin{aligned} \mathbf{e}_i &= \sum_{k \in \mathcal{K}} \hat{e}_{i,k} \phi_k \\ \mathbf{f}_i &= \sum_{k \in \mathcal{K}} \underbrace{\hat{f}_{i,k}}_0 \phi_k \end{aligned} \right\} \forall i \in \mathcal{N}_{\text{SL}}. \quad (3.11\text{c})$$

We use the notation  $\hat{\cdot}$  to indicate—and emphasize—known and given values.

Under Assumption 3.3 we can tackle the PPF Problem 3.1 by means of intrusive Galerkin projection, see Section 2.4.3. Let us begin by introducing polynomial chaos for every bus state  $\mathbf{z}_i$  with  $i \in \mathcal{N}$

$$\mathbf{z}_i = \begin{bmatrix} \mathbf{p}_i \\ \mathbf{q}_i \\ \mathbf{e}_i \\ \mathbf{f}_i \end{bmatrix} = \sum_{k \in \mathcal{K}} \begin{bmatrix} p_{i,k} \\ q_{i,k} \\ e_{i,k} \\ f_{i,k} \end{bmatrix} \phi_k = \sum_{k \in \mathcal{K}} z_{i,k} \phi_k \quad \text{with} \quad z_{i,k} \in \mathbb{R}^4. \quad (3.12)$$

Our goal is to compute all the PCE coefficients  $z_k = [z_{1,k}^\top \dots z_{N_b,k}^\top]^\top \in \mathbb{R}^{4N_b}$  for all  $k \in \mathcal{K}$ . To do so, we project the power flow across a transmission line given by (3.4) onto the orthogonal basis spanned by  $\phi_k$ , resulting in (3.17) on page 67. Then, we project the power flow equations given by (3.5) onto the basis polynomials, resulting in the system of equations from (3.18) on page 67. As we did before in Section 3.1 we introduce a shorthand notation for the collection of all Galerkin-projected power flow equations (3.18) in terms of the mapping  $g_{\text{PCE}}: \mathbb{R}^{4N_b(\hat{k}+1)} \rightarrow \mathbb{R}^{2N_b(\hat{k}+1)}$

$$\mathbf{z} = \left[ z_0^\top \quad \dots \quad z_{\hat{k}}^\top \right]^\top \mapsto g_{\text{PCE}}(\mathbf{z}). \quad (3.13)$$

Next, we apply Galerkin projection to the bus specifications (3.7), resulting in the equations

$$\text{PQ:} \quad p_{i,k} = \hat{p}_{i,k}, \quad q_{i,k} = \hat{q}_{i,k}, \quad \forall i \in \mathcal{N}_{\text{PQ}}, \quad (3.14\text{a})$$

$$\begin{aligned} \text{PV: } p_{i,k} &= \hat{p}_{i,k}, \quad \sum_{k_1, k_2 \in \mathcal{K}} \hat{v}_{i,k_1} \hat{v}_{i,k_2} \beta_{k_1 k_2 k} \\ &= \sum_{k_1, k_2 \in \mathcal{K}} (e_{i,k_1} e_{i,k_2} + f_{i,k_1} f_{i,k_2}) \beta_{k_1 k_2 k}, \quad \forall i \in \mathcal{N}_{\text{PV}}, \end{aligned} \quad (3.14\text{b})$$

$$\text{SL: } e_{i,k} = \hat{e}_{i,k}, \quad f_{i,k} = 0, \quad \forall i \in \mathcal{N}_{\text{SL}}, \quad (3.14\text{c})$$

for all  $k \in \mathcal{K}$  with  $\beta_{k_1 k_2 k} = \langle \phi_{k_1} \phi_{k_2}, \phi_k \rangle$ .<sup>6</sup> For (3.14) we introduce the mapping  $h_{\text{PCE}}: \mathbb{R}^{4N_b(\hat{k}+1)} \rightarrow \mathbb{R}^{2N_b(\hat{k}+1)}$  with

$$z = \begin{bmatrix} z_0^\top & \dots & z_{\hat{k}}^\top \end{bmatrix}^\top \mapsto h_{\text{PCE}}(z; \hat{x}), \quad (3.15\text{a})$$

which is parameterized by the collection of all PCE coefficients of the bus specifications

$$\hat{x} = \begin{bmatrix} \hat{x}_0^\top & \dots & \hat{x}_{\hat{k}}^\top \end{bmatrix}^\top. \quad (3.15\text{b})$$

We gather all the results and compose the PCE-overloaded PPF problem.

**Problem 3.2** (Probabilistic power flow using PCE). *Consider the probabilistic power flow Problem 3.1, and let Assumption 3.3 hold. Then, the PCE-overloaded probabilistic power flow problem is the solution to the system of nonlinear equations given by the PCE-overloaded power flow equations (3.18) and the PCE-overloaded bus specifications (3.14),*

$$\begin{bmatrix} g_{\text{PCE}}(z) \\ h_{\text{PCE}}(z; \hat{x}) \end{bmatrix} = \begin{bmatrix} 0_{2N_b(\hat{k}+1)} \\ 0_{2N_b(\hat{k}+1)} \end{bmatrix}. \quad (3.16)$$

*This way we map the user-given PCE coefficients  $\hat{x}_k \in \mathbb{R}^{2N_b}$  to the overall PCE coefficients  $z_k \in \mathbb{R}^{4N_b}$  for all  $k \in \mathcal{K}$ .*

<sup>6</sup> The numerical values for  $\beta_{k_1 k_2 k}$  can be computed by Gauss quadrature, see Appendix A.2.

$$p_{i,j,m}^{\text{br}} = \sum_{k_1, k_2 \in \mathcal{K}} \beta_{k_1 k_2 m} (g_{ij} (e_{i,k_1} e_{i,k_2} + f_{i,k_1} f_{i,k_2} - e_{i,k_1} e_{j,k_2} - f_{i,k_1} f_{j,k_2}) + b_{ij} (e_{i,k_1} f_{j,k_2} - f_{i,k_1} e_{j,k_2})) \quad (3.17a)$$

$$q_{i,j,m}^{\text{br}} = \sum_{k_1, k_2 \in \mathcal{K}} \beta_{k_1 k_2 m} \left( -(b_{ij} + b_{ij}^{\text{brsh}}/2) (e_{i,k_1} e_{i,k_2} + f_{i,k_1} f_{i,k_2}) - b_{ij} (e_{i,k_1} e_{j,k_2} - f_{i,k_1} f_{j,k_2}) + g_{ij} (e_{i,k_1} f_{j,k_2} - f_{i,k_1} e_{j,k_2}) \right) \quad (3.17b)$$

$$\forall (i, j) \in \mathcal{L}, \quad \forall m \in \mathcal{K}, \quad \text{and } \beta_{k_1 k_2 m} = \langle \phi_{k_1} \phi_{k_2}, \phi_m \rangle / \langle \phi_m, \phi_m \rangle$$

$$p_{i,m} = \sum_{\substack{(j,k) \in \mathcal{L} \\ \text{s.t. } j=i}}^{\text{br}} p_{jk,m} + \sum_{\substack{(j,k) \in \mathcal{L} \\ \text{s.t. } k=i}}^{\text{br}} p_{kj,m} + \sum_{k_1, k_2 \in \mathcal{K}}^{\text{br}} \beta_{k_1 k_2 m} g_i^{\text{sh}} (e_{i,k_1} e_{i,k_2} + f_{i,k_1} f_{i,k_2}) \quad (3.18a)$$

$$q_{i,m} = \sum_{\substack{(j,k) \in \mathcal{L} \\ \text{s.t. } j=i}}^{\text{br}} q_{jk,m} + \sum_{\substack{(j,k) \in \mathcal{L} \\ \text{s.t. } k=i}}^{\text{br}} q_{kj,m} - \sum_{k_1, k_2 \in \mathcal{K}}^{\text{br}} \beta_{k_1 k_2 m} b_i^{\text{sh}} (e_{i,k_1} e_{i,k_2} + f_{i,k_1} f_{i,k_2}) \quad (3.18b)$$

$$\forall i \in \mathcal{N}, \quad \forall m \in \mathcal{K}, \quad \text{and } \beta_{k_1 k_2 m} = \langle \phi_{k_1} \phi_{k_2}, \phi_m \rangle / \langle \phi_m, \phi_m \rangle$$

### 3.3.1 Discussion

#### Computational characteristics

The Galerkin-projected PF problem under uncertainty is rendered a deterministic system of nonlinear algebraic equations in terms of the PCE coefficients of the bus voltages and bus powers—in principle any established numerical technique for solving such problems such as the Newton-Raphson method can be used. The number of variables, however, grows to  $4N_b(\hat{k} + 1)$ : there is a power flow equation and a bus specification for every PCE coefficient for every entry of the total bus state. In case there is no uncertainty, the deterministic PF problem is recovered, and the problem size reduces back to  $4N_b$  because the PCE dimension reduces to  $\hat{k} + 1 = 1$ . In other words, the PCE-overloaded probabilistic power flow from Problem 3.2 is a generalization of the deterministic power flow problems for the specific case the uncertainties can be modeled adequately in terms of a PCE of a random variable. The Galerkin-projected power flow has the same mathematical structure as the deterministic power flow; namely, it is a system of quadratic equations in the real and imaginary voltage PCE coefficients, weighted by the scalars  $\beta_{k_1 k_2 m}$  that depend on the basis functions.

#### Interpretation of solution

Having solved the PCE-overloaded power flow problem, all vectors of PCE coefficients for all grid variables  $p_k$ ,  $q_k$ ,  $e_k$ , and  $f_k$  are known for all  $k \in \mathcal{K}$ . Recall from Remark 2.9 that the solution admits two different interpretations: we can compute stochastic moments of relevant quantities, see Section 2.4.3, or we can compute specific realizations of the grid state. The latter reduces merely to evaluating the polynomial basis at a point and multiplying with the respective PCE coefficients, see again Section 2.4.3.

#### Truncation errors

The PCE-overloaded PPF Problem 3.2 is an instance of the generic Problem 2.4 from page 49: random variables are mapped to other random variables, all

of which are described in terms of their finite PCE. Unfortunately, the implicit polynomial nature of the power flow equations makes it challenging to derive truncation errors. Whenever we employ a finite PCE to solve the PCE-overloaded PPF Problem 3.2 we have to accept a non-zero truncation error. Only for infinite PCEs we can expect a zero truncation error. However, numerical studies indicate that polynomial bases of low total degree such as 2 or 3 are sufficient to yield satisfactory numerical accuracy [55, 126].

## 3.4 Backward-forward sweep

In Section 3.1 we introduced the generic PPF problem; any PPF problem may be cast in this form. If, however, the problem at hand admits structure, we may be able to derive a formulation of PPF specific to that structure. An example is the backward-forward-sweep (BFS) method, a specific iterative solution method tailored to the power flow problem for radial grids (or tree networks) [150]. The BFS method exploits the tree structure of the underlying graph, and iterates successively between Kirchhoff's current law and Kirchhoff's voltage law [52]; the traditional nonlinear power flow equations need not be solved as a one-shot problem. The equations for the BFS-iterations are straightforward to derive, see Appendix A.7, they admit a physical interpretation, and compared to Newton-Raphson-based methods no Jacobian matrices need to be inverted.

The presented material is based on [9]; for a detailed derivation of the BFS scheme for generic load models and for generic constitutive laws we refer to Appendix A.7.

### 3.4.1 Problem formulation

We study a radial power grid with  $N_b + 1$  buses and  $N_b$  lines under uncertainty that satisfies Assumption 3.2. We define the set of bus indices by  $\mathcal{N} = \{0, 1, \dots, N_b\}$ , where we assume that the so-called root node 0 is the slack bus  $\mathcal{N}_{sl} = \{0\}$ . The BFS method iterates between Kirchhoff's current law and Kirchhoff's voltage law, i.e. between all nodal voltages, and

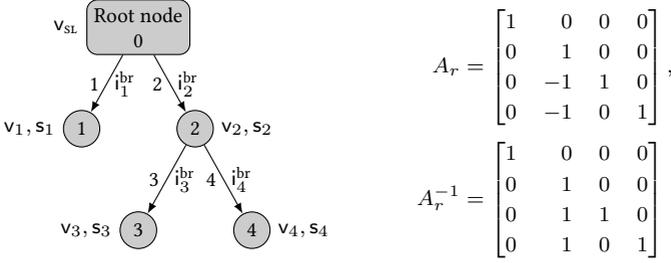


Figure 3.4: Depiction of bus and line numbering for a radial grid together with its nodal voltages and branch currents (left); reduced incidence matrix  $A_r$  and its inverse (right).

all branch currents. It is convenient for BFS to introduce a vectorized and complex-valued notation for all non-root nodes and all branches

$$\mathbf{v} = \begin{bmatrix} v_1 & \dots & v_{N_b} \end{bmatrix}^\top, \mathbf{s} = \begin{bmatrix} s_1 & \dots & s_{N_b} \end{bmatrix}^\top, \mathbf{i}^{br} = \begin{bmatrix} i_1^{br} & \dots & i_{N_b}^{br} \end{bmatrix}^\top, \quad (3.19a)$$

to represent the random-variable non-root node voltages  $\mathbf{v}$ , the random-variable non-root node apparent powers  $\mathbf{s}$ , and the random-variable branch currents  $\mathbf{i}^{br}$ , where

$$\text{Re}(w_i), \text{Im}(w_i) \in L^2(\Omega, \mu; \mathbb{R}) \quad \forall w \in \{\mathbf{v}, \mathbf{s}, \mathbf{i}^{br}\}, \forall i \in \mathcal{N}. \quad (3.19b)$$

The enumeration from (3.19) follows a simple pattern: the root node  $i \in \mathcal{N}_{sl}$  is number zero, and its leaves (i.e. non-root nodes)  $i \in \mathcal{N} \setminus \mathcal{N}_{sl}$  are numbered in ascending order from left to right. The lines are directed from the root to its leaves. The line number corresponds to the number of the terminating leaf node. For a 5-bus system Figure 3.4 shows the numbering of the buses and lines. The specific numbering allows to write the incidence matrix of the radial grid as

$$A = \begin{bmatrix} a_0 & A_r \end{bmatrix} \in \mathbb{Z}^{N_b \times (N_b+1)}, \quad a_0 \in \mathbb{Z}^{N_b}, \quad A_r \in \mathbb{Z}^{N_b \times N_b}, \quad (3.20)$$

and we call the quadratic matrix  $A_r$  the reduced incidence matrix. It is lower-triangular with ones on its diagonal, and has rank  $N_b$ , see Figure 3.4. The fact that the reduced incidence matrix is invertible is *the* key mathematical insight for the BFS method to work, because it allows to use Kirchhoff's laws directly



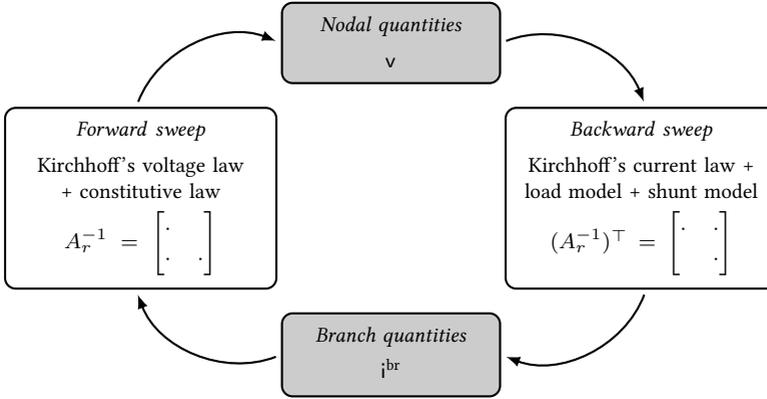


Figure 3.5: Scheme of BFS method as an iteration between Kirchhoff's current law and Kirchhoff's voltage law.

Algorithm 1 consists of the uncertain power  $s$  at all PQ buses, the shunt admittance matrix  $Y^{\text{sh}}$ , the branch impedances  $Z^{\text{br}}$ , the slack voltage  $v_{\text{SL}}$ , the reduced incidence matrix  $A_r$ , an initial voltage guess  $v_{\text{init}}$ , and termination tolerances  $\varepsilon_1, \varepsilon_2$ . If the first two moments between iterations do not differ by more than  $\varepsilon_1$ , respectively  $\varepsilon_2$ , the algorithm terminates. From Algorithm 1 we see how the uncertainties are propagated through the backward and the forward sweep, compare the graphical representation from Figure 3.5.

Besides the question about how to choose the initial random variable  $v_{\text{init}}$  the algorithm is not immediately amenable to computation; it is an iterative scheme among random variables. How has this problem been tackled before?

### 3.4.2 Existing approaches

The backward-forward sweep method dates back to [150]; extensions to three-phase power flow are presented in [38]. A review on backward-forward sweep methods is presented in [52]. In [26] the authors study analytic convergence criteria. Common to all of these works is that they apply BFS to deterministic power systems. To the best of the author's knowledge only the work [88] studies BFS under uncertainties, leveraging moment-based reformulations. Specifically, the authors of [88] reformulate BFS for the first two

moments, assuming that the first two moments characterize the uncertainties sufficiently accurately.

We clearly see that there is a considerable gap between the research on BFS for deterministic problems and problems under uncertainty. We attempt to close this gap to some degree by presenting an approach based on PCE that may be viewed as a generalization to the moment-based work [88]: while the authors of [88] restrict themselves to two moments, polynomial chaos does not need to make this assumption. By reformulating and solving BFS in terms of the PCE coefficients we can compute any moments of interest from the coefficients a posteriori, see Section 2.4.3. The resulting PCE-based approach is an iterative method reminiscent of the deterministic case. In fact, it is a generalization that contains deterministic BFS as a special case.

### 3.4.3 Solution methodology

Polynomial chaos can help render Algorithm 1 computationally tractable by rewriting it in terms of a deterministic algorithm that computes the PCE coefficients. To this end, let Assumption 3.3 hold. We write the PCE of the active power  $p$  and reactive power  $q$  for all PQ buses as follows

$$\begin{aligned} p &= \sum_{k \in \mathcal{K}} \begin{bmatrix} p_{1,k} \\ \vdots \\ p_{N_b,k} \end{bmatrix} \phi_k = \sum_{k \in \mathcal{K}} p_k \phi_k \in L^2(\Omega, \mu; \mathbb{R}^{N_b}), \\ q &= \sum_{k \in \mathcal{K}} \begin{bmatrix} q_{1,k} \\ \vdots \\ q_{N_b,k} \end{bmatrix} \phi_k = \sum_{k \in \mathcal{K}} q_k \phi_k \in L^2(\Omega, \mu; \mathbb{R}^{N_b}). \end{aligned} \quad (3.21)$$

This allows to introduce a complex-valued vector of PCE coefficients

$$\mathbb{C}^{N_b} \ni s_k = p_k + jq_k, \quad \forall k \in \mathcal{K}. \quad (3.22)$$

We employ the following formal notation for the PCE of the complex-valued apparent power among all non-root nodes  $\mathbf{s}$

$$\mathbf{s} = \sum_{k \in \mathcal{K}} s_k \phi_k, \quad (3.23a)$$

meaning that

$$\sum_{k \in \mathcal{K}} \operatorname{Re}(s_k) \phi_k, \sum_{k \in \mathcal{K}} \operatorname{Im}(s_k) \phi_k \in L^2(\Omega, \mu; \mathbb{R}^{N_b}). \quad (3.23b)$$

In the same manner we introduce the complex-valued PCE coefficients for the non-root node bus voltages and the branch currents

$$\mathbf{v} = \sum_{k \in \mathcal{K}} v_k \phi_k, \quad v_k \in \mathbb{C}^{N_b}, \quad \mathbf{i}^{\text{br}} = \sum_{k \in \mathcal{K}} i_k^{\text{br}} \phi_k, \quad i_k^{\text{br}} \in \mathbb{C}^{N_b}. \quad (3.24)$$

With this notation we substitute every random variable in Algorithm 1 by its PCE, and then apply Galerkin projection. This leads to Algorithm 2 (page 75): the PCE-overloaded BFS method that computes the PCE coefficients of all non-root node voltages, and the PCE coefficients of all branch currents. In Algorithm 2 the scalar  $\gamma_k$  is the normalization constant from Definition 2.13.

**Problem 3.4** (PCE-overloaded BFS under uncertainty). *Consider Problem 3.3, and let Assumption 3.3 hold. Then, the PCE-overloaded BFS under uncertainty means to solve a specific PCE-overloaded PPF problem according to Algorithm 2.*

In Algorithm 2 (page 75) we can still distinguish between the backward and the forward sweep, which are taken for every  $k^{\text{th}}$  PCE coefficient. Compared to the data provided to Algorithm 1, we need to feed Algorithm 2 the PCE coefficients of the uncertain power at the PQ buses. Note that the backward sweep requires the PCE coefficients  $i_{i,k_2}$ , i.e. all the  $k \in \mathcal{K}$  PCE coefficients of the current at buses  $i \in \mathcal{N}_{\text{PQ}}$ . Unfortunately, we obtain these coefficients only implicitly from the load model. Nonetheless, Algorithm 2 admits a straightforward and structurally equivalent implementation to the deterministic BFS according to Algorithm 3 from Appendix A.7. If the first two moments between iterations do not differ by more than  $\varepsilon_1$ , respectively  $\varepsilon_2$ , then the PCE-

**Algorithm 2:** PCE-overloaded BFS method.**Data:** PCE coefficients  $p_k, q_k \forall k \in \mathcal{K}$ ;  $Y^{\text{sh}}, Z^{\text{br}}, v_{\text{sl}}; A_r; \varepsilon_1, \varepsilon_2 > 0$ .**Result:** PCE coefficients  $v_k, i_k^{\text{br}} \forall k \in \mathcal{K}$ .Set  $\tau \leftarrow 0$ ;

Initialize

$$v_k(\tau) = \begin{cases} v_{\text{sl}} \mathbf{1}_{N_b}, & k = 0, \\ 0_{N_b}, & k \in \mathcal{K} \setminus \{0\}; \end{cases}$$

Compute  $s_k$  from (3.22); **do**

# Backward sweep

Evaluate  $\forall m \in \mathcal{K}$  and  $\forall i \in \mathcal{N}_{\text{PQ}}$ 

$$s_{i,m} \langle \phi_m, \phi_m \rangle = \sum_{k_1, k_2 \in \mathcal{K}} v_{i,k_1}(\tau) i_{i,k_2}^*(\tau) \langle \phi_{k_1} \phi_{k_2}, \phi_m \rangle;$$

Set  $\forall k \in \mathcal{K}$ 

$$i_k(\tau) = [i_{1,k}(\tau) \dots i_{N_b,k}(\tau)]^\top;$$

Evaluate  $\forall k \in \mathcal{K}$ 

$$i_k^{\text{br}}(\tau) = (A_r^\top)^{-1} (Y^{\text{sh}} v_k(\tau) - i_k(\tau));$$

# Forward sweep

Evaluate  $\forall k \in \mathcal{K}$ 

$$v_k(\tau + 1) = \begin{cases} -A_r^{-1} Z^{\text{br}} i_0^{\text{br}}(\tau) + v_{\text{sl}} \mathbf{1}_{N_b}, & k = 0, \\ -A_r^{-1} Z^{\text{br}} i_k^{\text{br}}(\tau), & k \in \mathcal{K} \setminus \{0\}; \end{cases}$$

Set  $\tau \leftarrow \tau + 1$ ;**while**  $\|v_0(\tau + 1) - v_0(\tau)\| > \varepsilon_1$  and  $\|\sum_{k \in \mathcal{K} \setminus \{0\}} (v_k(\tau + 1) - v_k(\tau)) \gamma_k\| > \varepsilon_2$ ;Return  $v_k \leftarrow v_k(\tau + 1), i_k^{\text{br}} \leftarrow i_k^{\text{br}}(\tau) \forall k \in \mathcal{K}$ ;

overloaded Algorithm 2 terminates. An alternative straightforward termination criterion is [9]:

$$\text{terminate if } \max_{k \in \mathcal{K}} \|v_k(\tau + 1) - v_k(\tau)\| < \varepsilon. \quad (3.25)$$

We refer the interested reader to [9] where Algorithm 2 is derived in terms of real and imaginary components of the PCE coefficients.

## 3.4.4 Discussion

### Computational characteristics

Algorithm 2, like its deterministic sibling, is an iterative algorithm, hence it creates no heavy computational burden. As usual with intrusive PCE, the scalar products need to be computed. For this, however, there exist dedicated software tools, see Chapter 5.

### Interpretation of solution

The situation is equivalent to Section 3.3.1: once BFS has terminated, the PCE coefficients are available from which the entire grid state can be derived in terms of random variables. Recalling Remark 2.9 we know that the random variable can either function as a policy: for a specific realization of the uncertainty we obtain a specific realization of the grid state. We may also use the random-variable solution to deduce statistical information about the grid state in terms of moments, for example. Via sampling we can also obtain the probability density function of desired grid states.

### Truncation errors

In light of truncation errors of PCE, see Section 2.5, the question arises how this applies to the PCE-overloaded BFS method. Recall the basic “ingredients” to the BFS method: Kirchhoff’s laws, a constitutive law, a load model, and a shunt model. Given the linearity of Kirchhoff’s laws, if we choose a linear relation for both the constitutive law (i.e. Ohm’s law), and a linear relation for the shunt model, then whether or not truncation errors can be quantified exactly will depend entirely on the load model. We presented the method for a constant-power load model from which the currents have to be determined. This leads to an implicit system of polynomial equations for which no general conclusions can be drawn. In case of a constant-impedance load model, however, the load model is also linear, hence the overall BFS scheme is linear, hence the theory from Section 2.5.1 applies; we can quantify the errors exactly. For this specific choice of constant-impedance loads we can also provide convergence guarantees, see Appendix A.7.

## 3.5 Optimal adaptive linearizations

In Section 3.1 we introduced the generic power flow problem under uncertainty. Mathematically, it is a nonlinear, non-convex, and implicit set of equations. For large-scale systems this may become numerically expensive to solve. It is common in power systems operations to determine proxies for the generic power flow problem that sacrifice numerical exactness for computational tractability. More often than not these proxies are linear and explicit sets of equations that can be evaluated quasi-instantaneously. Perhaps the most prominent example is power flow under so-called DC conditions, see Appendix A.6, leading to the so-called power transfer distribution factor (PTDF) matrix. Another approach is to linearize the power flow equations around some nominal operating point by means of first-order Taylor expansion; this leads to so-called AC-PTDF matrices [168]. In both cases we speak of power flow linearizations.

Both the PTDF and the AC-PTDF are general-purpose linearizations. By that we mean that they are applicable to any power system, but they do not exploit additional knowledge.<sup>7</sup> Instead, we strive for linearizations that are optimal and adaptive. Optimal refers to the linearizations minimizing a specific error metric, and adaptive means that the linearizations account for a specific operating range of the system. Specifically we choose to model the operating range in terms of random variables, and we aim to minimize the expected value of the squared error. Once more, polynomial chaos will be helpful to derive a concise and deterministic problem formulation.

Figure 3.6 visualizes the difference between general-purpose linearizations and adaptive linearizations. The solid black graph is the true correspondence between some bus power  $p$  and some branch power  $p^{\text{br}}$ —which we would like to approximate. The left plot in Figure 3.6 shows a generic linearization around some operating point (dashed grey graph); the linearization is sufficiently accurate within the dotted interval. If we know, however, that the bus injection  $p$  will fluctuate according to the density shown on the  $p$ -axis in the right plot of Figure 3.6, we also expect the branch flows to follow some

---

<sup>7</sup> This is similar to the generic power flow problem vs. backward-forward sweep; BFS exploits the additional knowledge about the radial structure of the grid, providing for a tailored power flow problem.

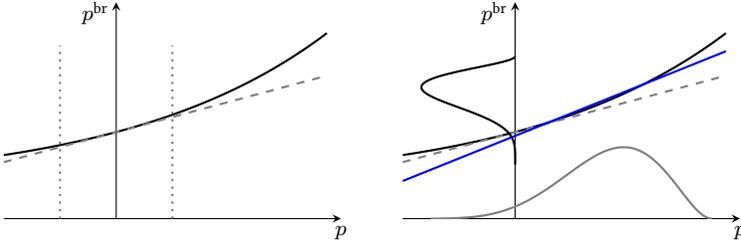


Figure 3.6: Graphical comparison between general-purpose linearizations (left) and adaptive linearizations (right). The curves on the bus power  $p$ -axis and the branch power  $p^{\text{br}}$ -axis in the right plot represent the probability density functions of  $p$  and  $p^{\text{br}}$ .

density, shown on the  $p^{\text{br}}$ -axis. We then see that the generic linearization will perform poorly in that case, because we did not design it to account for the information about how frequently certain bus injections  $p$  will occur. The output of the method we propose is the linearization shown in the right plot of Figure 3.6 as the solid graph—it provides a meaningful linearization given the density of the bus injections. The presented material is based on [125]. <sup>f</sup>

### 3.5.1 Problem formulation

We study a power grid under uncertainty that satisfies Assumption 3.2. The operating range is modeled by known random variables for the active and reactive power injections, denoted as  $p_i$  for all  $i \in \mathcal{N}_{\text{PQ}} \cup \mathcal{N}_{\text{PV}}$ , and  $q_i$  for all  $i \in \mathcal{N}_{\text{PQ}}$ . While the voltage magnitudes at PV and slack buses can also generally be represented by random variables  $v_i$  for all  $i \in \mathcal{N}_{\text{PV}} \cup \mathcal{N}_{\text{SL}}$ , we consider constant voltage magnitudes at these buses.<sup>8</sup> For the given power grid we seek a linearization that relates a single “output” quantity of interest to some “input” quantities.<sup>9</sup> The output is typically a quantity that is constrained or optimized but not directly controlled, such as voltage magnitudes at PQ buses, active and reactive line flows, or current flows. Although applicable to a variety of choices for input and output quantities we focus on outputs consisting

<sup>8</sup> In terms of random variables constant voltage magnitudes have a probability density equal to a Dirac-delta pulse centered at the value of the voltage magnitude.

<sup>9</sup> If we speak of *linearizations* (plural) in the following we mean the set of mappings that relates *several* outputs to inputs.

of active and reactive power flows on each line, and inputs consisting of active and reactive power injections at each non-slack bus  $i \in \mathcal{N} \setminus \mathcal{N}_{\text{SL}}$ . We want to identify an affine relation that maps random bus injections to some random active power branch flow via

$$\hat{\mathbf{p}}_{jk}^{\text{br}} = \ell_0 + \sum_{i \in \mathcal{N} \setminus \mathcal{N}_{\text{SL}}} (\ell_{p,i} \mathbf{p}_i + \ell_{q,i} \mathbf{q}_i) \quad (3.26)$$

for all lines  $(j, k) \in \mathcal{L}$  of interest. The coefficients  $\ell_0$  and  $\ell_{p,i}, \ell_{q,i}$  for all  $i \in \mathcal{N} \setminus \mathcal{N}_{\text{SL}}$  are the linearization parameters we need to compute. The ansatz (3.26) results in linearizations analogous to commonly used PTFD formulations.<sup>10</sup>

**Problem 3.5** (Optimal adaptive linearization with respect to expected squared error). *Consider a grid for which Assumption 3.1 and Assumption 3.2 hold. The operating range is modeled by known random variables for the active and reactive power injections, denoted as  $\mathbf{p}_i$  for all  $i \in \mathcal{N}_{\text{PQ}} \cup \mathcal{N}_{\text{PV}}$ , and  $\mathbf{q}_i$  for all  $i \in \mathcal{N}_{\text{PQ}}$ . We consider constant voltage magnitudes  $v_i$  for all buses  $i \in \mathcal{N}_{\text{PV}} \cup \mathcal{N}_{\text{SL}}$ . The optimal adaptive linearization with respect to the expected squared error is*

$$\min_{\substack{\ell_0, \ell_{p,i}, \ell_{q,i} \in \mathbb{R} \\ \forall i \in \mathcal{N} \setminus \mathcal{N}_{\text{SL}} \\ \mathbf{p}_i, \mathbf{q}_i, \mathbf{e}_i, \mathbf{f}_i \in L^2(\Omega, \mu; \mathbb{R}) \\ \forall i \in \mathcal{N}}} \mathbb{E}((\mathbf{p}_{jk}^{\text{br}} - \hat{\mathbf{p}}_{jk}^{\text{br}})^2) \quad (3.27a)$$

subject to

$$g(\mathbf{z}) = 0, \quad (3.27b)$$

$$h(\mathbf{z}) = \mathbf{x}, \quad (3.27c)$$

$$\mathbf{p}_{jk}^{\text{br}} = g_{jk}(\mathbf{e}_j^2 + \mathbf{f}_j^2 - \mathbf{e}_j \mathbf{e}_k - \mathbf{f}_j \mathbf{f}_k) + b_{jk}(\mathbf{e}_j \mathbf{f}_k - \mathbf{f}_j \mathbf{e}_k), \quad (3.27d)$$

$$\hat{\mathbf{p}}_{jk}^{\text{br}} = \ell_0 + \sum_{i \in \mathcal{N} \setminus \mathcal{N}_{\text{SL}}} (\ell_{p,i} \mathbf{p}_i + \ell_{q,i} \mathbf{q}_i), \quad (3.27e)$$

where  $\mathbf{p}_{jk}^{\text{br}}$  is the random variable corresponding to the active power flow on line  $(j, k) \in \mathcal{L}$ .

<sup>10</sup> In fact, if  $\ell_0 = 0$  and all  $\ell_{q_i} = 0$ , the ansatz is mathematically equivalent to the PTFD matrix, see (A.53) in Appendix A.6.

In Problem 3.5 the objective (3.27a) minimizes the expected squared error between the linearization and the nonlinear power flow equations for this line flow. Constraint (3.27b) denotes the power flow equations (3.5) with the bus specifications for the random variables being (3.27c), see (3.8). The constraints (3.27d) and (3.27e) model the nonlinear active power flow equation and its linearization, which is parameterized by the decision variables  $\ell_0, \ell_{p,i}, \ell_{q,i}$ . The solution to Problem 3.5 gives the coefficients of the respective optimal adaptive linearization. The optimal objective value bounds the linearization's expected squared error. Linearizations for the active power flows on each line  $(j, k) \in \mathcal{L}$  as well as variants of Problem 3.5 that consider other output quantities of interest such as reactive power flows, or voltage magnitudes can be computed in parallel.

### 3.5.2 Existing approaches

The idea of optimal adaptive linearizations originates in the recent works [107, 111]. In [111] the goal is to minimize the worst-case linearization error. To compute the linearization, a constraint-generation algorithm alternates between computing the power injections that result in the worst-case error for a candidate linearization, and updating the candidate linearization in order to reduce the error with respect to all previously calculated worst-case power injections. The works [85, 107] study adaptive linearizations using expected errors. Reference [107] (with extensions to three-phase systems in [3]) minimizes the least-square linearization error for a predefined set of evenly distributed points near a nominal operating point. While the resulting linearization is *adaptive* to the system and operating range of interest, the set of points is not selected in any *optimal* manner to minimize some objective. Using a so-called generalized moment approach, [85] also computes power flow linearizations that minimize the expected error. Specifically, [85] minimizes the error for the *solution* to a specified optimization problem (parameterized by the uncertain power injections). Another approach is presented in [98] where notions of regression analysis are applied to the linearization problem. Albeit being *adaptive* this approach does not explicitly consider optimality metrics that account for an uncertain operating range.

To summarize: linearizations have been studied for the expected error. They can be classified either as optimal *or* as adaptive, not satisfying both proper-

ties simultaneously. Our proposed solution is both optimal and adaptive, and allows for a decomposition that makes the idea intuitive and scalable.

### 3.5.3 Solution methodology

Recall that the linearization Problem 3.5 has finitely many decision variables: the linearization parameters  $\ell = [\ell_0, \ell_p^\top, \ell_q^\top]^\top \in \mathbb{R}^{1+2|\mathcal{N} \setminus \mathcal{N}_{sl}|}$ . Though finite-dimensional this problem is nevertheless challenging since the power flow equations in terms of random variables (3.27b) are infinite-dimensional. Additionally, the calculation of the expectation (3.27a) requires solving an integral. To address these challenges, we exploit the problem's structure, specifically the facts that

1. the active power line flow in (3.27d) is explicitly determined by the solution to the square system of power flow equations (3.27b), and
2. the decision variables  $\ell$  only appear in (3.27e).

Hence, Problem 3.5 can be decomposed:

1. *Feasibility problem:* Solve the probabilistic power flow problem (3.27b), i.e., compute the random variables  $e_i$  and  $f_i$  for all  $i \in \mathcal{N}$ ;  $q_i$  for all  $i \in \mathcal{N}_{pv} \cup \mathcal{N}_{sl}$ ; and  $p_i$  for  $i \in \mathcal{N}_{sl}$  given the random variables  $p_i$  for all  $i \in \mathcal{N}_{pv} \cup \mathcal{N}_{pq}$ , and  $q_i$  for all  $i \in \mathcal{N}_{pq}$ .
2. *Unconstrained optimization problem:* Determine the optimal linearization coefficients  $\ell$  by substituting the resulting random variables for  $e_i$ ,  $f_i$ ,  $p_i$ , and  $q_i$  into (3.27d) and (3.27e) and then minimizing (3.27a).

As the PPF problem (3.27b) is independent of the objective (3.27a), the solution from the first step can be used repeatedly (in parallel) to compute linearizations for multiple quantities of interest in the second step. Even though we can decompose the linearization Problem 3.5, the PF problem (3.27b) remains infinite-dimensional and is therefore challenging—but we can recycle the results from Section 3.3—hence, we expect Assumption 3.3 to hold.

However, the objective (3.27a) still requires evaluating an integral. Turning to the objective (3.27a), we next show how PCE facilitates the derivation of a closed-form solution in terms of PCE coefficients. We first substitute the

equalities (3.27d) and (3.27e) into the objective to obtain an unconstrained optimization problem. We then rewrite (3.27a) as

$$\mathbb{E}((\hat{p}_{jk}^{\text{br}} - \hat{p}_{jk}^{\text{br}})^2) = \mathbb{E}((\hat{p}_{jk}^{\text{br}})^2) - 2\mathbb{E}(p_{jk}^{\text{br}} \hat{p}_{jk}^{\text{br}}) + \mathbb{E}((p_{jk}^{\text{br}})^2). \quad (3.28)$$

For the first term  $\mathbb{E}((\hat{p}_{jk}^{\text{br}})^2)$  in (3.28), we use the ansatz (3.27e) for the active power flow  $\hat{p}_{jk}^{\text{br}}$  and compute the moments according to Section 2.4.3. We obtain (recall from Problem 3.5 that  $\ell = [\ell_0, \ell_p^\top, \ell_q^\top]^\top$ )

$$\mathbb{E}((\hat{p}_{jk}^{\text{br}})^2) = \ell^\top W \ell, \quad \text{where} \quad (3.29)$$

$$W = \begin{bmatrix} 1 & t^\top \\ t & T \end{bmatrix}, \quad t = \begin{bmatrix} p_0 \\ q_0 \end{bmatrix}, \quad T = \sum_{k \in \mathcal{K}} \begin{bmatrix} p_k p_k^\top & p_k q_k^\top \\ q_k p_k^\top & q_k q_k^\top \end{bmatrix}.$$

The vectors  $p_k, q_k$  are the  $k^{\text{th}}$  pCE coefficient of the active power and reactive power, respectively. Notice that the matrix  $W$  is positive (semi)definite. For the second term  $\mathbb{E}(p_{jk}^{\text{br}} \hat{p}_{jk}^{\text{br}})$  in (3.28), we substitute (3.27d) and (3.27e) and then apply the moment relation from Section 2.4.3 to obtain

$$\mathbb{E}(p_{jk}^{\text{br}} \hat{p}_{jk}^{\text{br}}) = w^\top \ell, \quad (3.30a)$$

where the vector  $w = [w_0 \ w_{pq}^\top]^\top$  has components

$$w_0 = \sum_{k \in \mathcal{K}} \gamma_k (g_{jk} (e_{j,k}^2 + f_{j,k}^2 - e_{j,k} e_{k,k} - f_{j,k} f_{k,k}) + b_{jk} (e_{j,k} f_{k,k} - f_{j,k} e_{k,k})) \quad (3.30b)$$

$$w_{pq} = \sum_{k_1, k_2, k_3 \in \mathcal{K}} \beta_{k_1 k_2 k_3} \left( g_{jk} \left( e_{j,k_1} e_{j,k_2} \begin{bmatrix} p_{k_3} \\ q_{k_3} \end{bmatrix} + f_{j,k_1} f_{j,k_2} \begin{bmatrix} p_{k_3} \\ q_{k_3} \end{bmatrix} - e_{j,k_1} e_{k,k_2} \begin{bmatrix} p_{k_3} \\ q_{k_3} \end{bmatrix} - f_{j,k_1} f_{k,k_2} \begin{bmatrix} p_{k_3} \\ q_{k_3} \end{bmatrix} \right) + b_{jk} \left( e_{j,k_1} f_{k,k_2} \begin{bmatrix} p_{k_3} \\ q_{k_3} \end{bmatrix} - f_{j,k_1} e_{k,k_2} p_{k_3} \right) \right) \quad (3.30c)$$

with  $\beta_{k_1 k_2 k_3} = \langle \phi_{k_1} \phi_{k_2}, \phi_{k_3} \rangle$ , and normalization constants  $\gamma_k = \langle \phi_k, \phi_k \rangle$ .<sup>11</sup>

Finally, the third term  $\mathbb{E}((p_{jk}^{\text{br}})^2)$  becomes

$$\mathbb{E}((p_{jk}^{\text{br}})^2) = \sum_{k \in \mathcal{K}} \gamma_k (p_{jk,k}^{\text{br}})^2 =: w_0, \quad (3.31)$$

where  $p_{jk,k}^{\text{br}}$  is given in (3.17a). To summarize, the objective (3.27a) subject to the equality constraints (3.27d) and (3.27e) can be written as an unconstrained convex quadratic program in the linearization parameters  $\ell$ ,

$$\left. \begin{array}{l} \min_{\substack{\ell = [\ell_0, \ell_p^\top, \ell_q^\top]^\top \\ \in \mathbb{R}^{1+2|\mathcal{N} \setminus \mathcal{N}_{\text{sl}}|}}} \mathbb{E}((p_{jk}^{\text{br}} - \hat{p}_{jk}^{\text{br}})^2) \\ \text{subject to (3.27d), (3.27e)} \end{array} \right\} = \min_{\substack{\ell = [\ell_0, \ell_p^\top, \ell_q^\top]^\top \\ \in \mathbb{R}^{1+2|\mathcal{N} \setminus \mathcal{N}_{\text{sl}}|}}} \ell^\top W \ell - 2w^\top \ell + w_0. \quad (3.13)$$

The “ingredients” ( $W, w, w_0$ ) are computed *after* the PPF problem has been solved, as they require knowledge of all vectors of PCE coefficients  $p_k, q_k, e_k, f_k$  for  $k \in \mathcal{K}$ . Notice that  $w_0$  is constant with respect to the linearization parameters  $\ell$  and therefore only affects the optimal value. We summarize the procedure as follows.

**Problem 3.6** (PCE-overloaded optimal adaptive linearization). *Consider Problem 3.5, and let Assumption 3.3 hold. The PCE-overloaded optimal adaptive linearization means to tackle Problem 3.5 in two steps:*

1. Solve probabilistic power flow (3.27b) according to Problem 3.2.
2. Solve an unconstrained QP according to (3.13).

## 3.5.4 Discussion

### Computational characteristics

The feasibility problem inherits all the computational characteristics we discussed in Section 3.3.1. The optimization problem is an unconstrained

<sup>11</sup> These numbers can be computed by Gauss quadrature, see Appendix A.2.

quadratic program for which the analytical solution is straightforward. Hence, the computational characteristics are dominated mainly by the ability to solve the PPF problem. Once that is solved, a quadratic program can be defined for every quantity of interest, which is straightforward to parallelize.

### **Interpretation of solution**

Just like there are two steps to solving the problem, there are two kinds of results we obtain from it: first and foremost we obtain the vector of parameters for the affine relation (3.26). The numerical values of the parameters indicate the sensitivity of the output quantity on the desired input quantities. As a byproduct from solving Problem 3.6 we also obtain information about the state of the overall grid in terms of random variables as we solve a PPF problem at the same time. Strictly speaking, we obtain the PCE coefficients from which we then recover the grid state.

### **Truncation errors**

The unconstrained optimization problem (3.13) does not cause any troubles. Hence, if the numerical values of the PCE coefficients were the exact solution to the probabilistic PF problem, then the obtained linearization coefficients were in fact the true solution to Problem 3.5. Unfortunately, as we have discussed previously in Section 3.3 we cannot solve the probabilistic power flow problem with finite PCE to exactness.

## 4 Optimal power flow under uncertainty

Optimal power flow is a cornerstone optimization problem for the planning and operation of power systems. Its solution is an optimal configuration of the power system in which the power flow equations are satisfied and engineering limits such as generation, voltage, and/or line flow limits are respected. The notion of “optimality” may represent, for instance, a minimum of monetary operation costs, or a minimum of transmission losses. The original formulation of OPF dates back to [35]. Ever since the research community has gone to tremendous lengths to develop tailored solution algorithms that are efficient, accurate, and computationally tractable [62, 63, 114]. To study OPF problems under uncertainty is a logical consequence in view of the steadily increasing importance of renewable energy sources.

Similarly to how we approached power flow under uncertainty in Section 3.1, we study optimal power flow under uncertainty in light of two questions:

- Q4.1 *Problem formulation*: What is a mathematically sound formulation of optimal power flow under uncertainty?
- Q4.2 *Solution methodology*: Having obtained a mathematical formulation, what is a solution approach?

We study three different OPF problems under uncertainty: the generic AC-OPF problem, the simplified DC-OPF problem, and the simplified DC-OPF problem for multiple time steps.

Before we begin, a note on the role of time for OPF: The solutions from OPF problems are applied for power system operation and operational planning. That includes time spans from about 15 minutes to several days [8, 78]. It is a standing assumption for the remainder that during that time span no new equipment is being installed, and that no maintenance is undertaken [146].

That is, the power system we study is not subjected to substantial structural changes.

## 4.1 Chance-constrained AC optimal power flow

The following material is based on [119, 126].

### 4.1.1 Problem formulation

We study a power system under Assumption 3.2. How does that affect the formulation of the optimization problem that aims to minimize a total cost of the power system subject to the power flow equations and engineering limits? A consequence of the presence of uncertainties according to Assumption 3.2 is that all variables describing the network become random variables. That means we need to consider probabilistic power flow according to Section 3.1 to model the physical behavior of the power system. With the notation from Section 3.1 we concisely write

$$g(\mathbf{z}) = \mathbf{0}_{2N_b}. \quad (4.1)$$

We make a technical assumption to exclude pathological grid configurations.

**Assumption 4.1** (Existence of high-voltage solution). *We assume there exists a (high-voltage) solution of the power flow equations for all realizations of the uncertainty, and that (4.1) models this (high-voltage) solution.*

Just like in Section 3.1 we model component behavior by declaring every bus  $i \in \mathcal{N} = \mathcal{N}_{\text{PQ}} \cup \mathcal{N}_{\text{PV}} \cup \mathcal{N}_{\text{SL}}$  a PQ bus ( $i \in \mathcal{N}_{\text{PQ}}$ ), a PV bus ( $i \in \mathcal{N}_{\text{PV}}$ ), or a slack bus ( $i \in \mathcal{N}_{\text{SL}}$ ). Further, we assume that each bus  $i \in \mathcal{N}$  connects to a single controllable (generating) unit and a single uncontrollable unit such that the net active and reactive power injection at bus  $i \in \mathcal{N}$  can be written as

$$\mathbf{p}_i = \mathbf{p}_i^c + \mathbf{p}_i^u, \quad (4.2a)$$

$$\mathbf{q}_i = \mathbf{q}_i^c + \mathbf{q}_i^u, \quad (4.2b)$$

Table 4.1: Engineering limits for OPF problems.

Quantity	Lower bound	Upper bound
Active power generation	$p^{\min}$	$p^{\max}$
Reactive power generation	$q^{\min}$	$q^{\max}$
Voltage magnitude	$v^{\min}$	$v^{\max}$
Branch flow	$-i^{\text{br,max}}, -p^{\text{br,max}}$	$i^{\text{br,max}}, p^{\text{br,max}}$

where the superscript “c” denotes controllable generation, and the superscript “u” denotes uncontrollable generation/demand. We subsume the entire uncontrollable generation/demand in  $x \in L^2(\Omega, \mu; \mathbb{R}^{2N_b})$ ,

$$x = \left[ p_1^u \quad q_1^u \quad \dots \quad p_{N_b}^u \quad q_{N_b}^u \right]^\top \in L^2(\Omega, \mu; \mathbb{R}^{2N_b}), \quad (4.3)$$

and we call  $x$  the (vector of) *bus parameterizations*. Given the bus parameterizations, it is our task to get the optimal configuration of the power grid.

With every bus we associate a cost of power generation. Since the power generation is not known with certainty, we introduce the random-variable cost per bus  $c_i : L^2(\Omega, \mu; \mathbb{R}) \rightarrow L^2(\Omega, \mu; \mathbb{R})$  via

$$p_i^c \mapsto c_i(p_i^c). \quad (4.4)$$

This cost is not directly amenable to (numerical) minimization. Often the expected value  $\mathbb{E}(c_i(p_i^c))$ —which is a scalar—is minimized.<sup>1</sup> This specific choice is motivated from stochastic programming [90, 91]. Other choices that include the variance are possible too.

Engineering limits require the power system to remain within reasonable technical bounds. This applies to active power generation, reactive power generation, voltage magnitudes and branch flows, see Table 4.1.<sup>2</sup> The branch

<sup>1</sup> To minimize a random variable would require to define a proper order relation for random variables, allowing us to characterize one random variable as “larger” than another one with respect to that notion.

<sup>2</sup> It is a simplification to assume that the limits for active power generation and reactive power generation are independent of each other, hence modeled as box constraints. They are actually coupled by the so-called reactive capability curve whose shape is specific to the type and the operating conditions of the generator [8]. It is also possible to consider engineering limits

flow limits follow from the thermal limits of the cable such that the sag does not exceed regulatory requirements.<sup>3</sup> In a deterministic setting the engineering limits have to be strictly satisfied in terms of inequality constraints in the optimization problem. In a setting under uncertainty, it is not per se meaningful and/or clear how to formulate inequality constraints for random variables (excluding pathological cases). So how can we formulate the inequality constraints given that we need to constrain random variables? A so-called chance constraint (CC) is one option. When we formulate an inequality constraint in the presence of uncertainties as a chance constraint we satisfy the constraint up to a pre-specified (high) level. This level is a tuning parameter and may be adjusted as desired. Let us consider an example: in a deterministic setting the upper generation limit at bus  $i$  reads

$$p_i^c \leq p_i^{\max}. \quad (4.5a)$$

We treat  $p_i^c$  as the realization of the random variable  $p_i^c$  and reformulate (4.5a) as the following chance constraint

$$\mathbb{P}(p_i^c \leq p_i^{\max}) \geq 1 - \varepsilon_p, \quad (4.5b)$$

for some  $\varepsilon_p \in (0, 1)$  and some bus  $i \in \mathcal{N}$ .<sup>4</sup> Hence, the probability of the upper generation bound being satisfied should exceed  $1 - \varepsilon$ . Likewise, we could model the converse statement, namely the probability of violating the constraint being below a level  $\varepsilon$ .

As an answer to Q4.1 from above, we can formulate OPF under uncertainty as the following chance-constrained optimization problem

**Problem 4.1** (Chance-constrained AC optimal power flow). *Let Assumption 3.2 and Assumption 4.1 hold for a given power system, and let the bus*

---

related to the absolute value of the voltage angles as well as their differences. For simplicity we do not consider these angle limits in what follows.

<sup>3</sup> In the dc setting the branch flow limits correspond to active power branch flow limits.

<sup>4</sup> We introduced this engineering-motivated notation in (2.9). Note that the symbol “ $\mu$ ” has a different meaning in the contexts of Section 2.2.2 and the present section, see also Remark 3.1.

parameterizations (4.3) be given. We call the following optimization problem a chance-constrained AC optimal power flow problem

$$\min_{\substack{\mathbf{p}_i^c, \mathbf{q}_i^c, \mathbf{p}_i^u, \mathbf{q}_i^u, \\ \mathbf{e}_i, \mathbf{f}_i \in L^2(\Omega, \mu; \mathbb{R}) \\ \forall i \in \mathcal{N}}} \sum_{i \in \mathcal{N}} \mathbb{E}(c_i(\mathbf{p}_i^c)) \quad (4.6a)$$

subject to

$$g(\mathbf{z}) = \mathbf{0}_{2N_b}, \quad (4.6b)$$

$$\mathbf{p}_i = \mathbf{p}_i^c + \mathbf{p}_i^u, \quad \forall i \in \mathcal{N}, \quad (4.6c)$$

$$\mathbf{q}_i = \mathbf{q}_i^c + \mathbf{q}_i^u, \quad \forall i \in \mathcal{N}, \quad (4.6d)$$

$$\mathbb{P}(\mathbf{p}_i^c \geq \mathbf{p}_i^{\min}) \geq 1 - \varepsilon_{p_i}, \quad \forall i \in \mathcal{N}, \quad (4.6e)$$

$$\mathbb{P}(\mathbf{p}_i^c \leq \mathbf{p}_i^{\max}) \geq 1 - \varepsilon_{p_i}, \quad \forall i \in \mathcal{N}, \quad (4.6f)$$

$$\mathbb{P}(\mathbf{q}_i^c \geq \mathbf{q}_i^{\min}) \geq 1 - \varepsilon_{q_i}, \quad \forall i \in \mathcal{N}, \quad (4.6g)$$

$$\mathbb{P}(\mathbf{q}_i^c \leq \mathbf{q}_i^{\max}) \geq 1 - \varepsilon_{q_i}, \quad \forall i \in \mathcal{N}, \quad (4.6h)$$

$$\mathbb{P}(v_i \geq v_i^{\min}) \geq 1 - \varepsilon_{v_i}, \quad \forall i \in \mathcal{N}, \quad (4.6i)$$

$$\mathbb{P}(v_i \leq v_i^{\max}) \geq 1 - \varepsilon_{v_i}, \quad \forall i \in \mathcal{N}, \quad (4.6j)$$

$$\mathbb{P}(l_{ij}^{\text{br}} \leq l_{ij}^{\text{br,max}}) \geq 1 - \varepsilon_{i,j}^{\text{br}}, \quad \forall (i, j) \in \mathcal{L}, \quad (4.6k)$$

$$\mathbf{f}_i = \mathbf{0}, \quad \forall i \in \mathcal{N}_{\text{SL}}, \quad (4.6l)$$

where  $g$  denotes the power flow equations (3.5) in terms of random variables.

Problem 4.1 minimizes the sum of the expected costs of active power generation (4.6a). The constraint (4.6b) stands for the power flow equations in terms of random variables; in other words, the AC power flow equalities hold for all realizations of the uncertainties. The single-sided chance constraints (4.6e)-(4.6k) are reformulations of the engineering limits for active power limits, reactive power limits, voltage magnitude limits, and branch flow limits, respectively, with acceptable violation probabilities  $\varepsilon_{p_i}, \varepsilon_{q_i}, \varepsilon_{v_i}, \varepsilon_{i,j}^{\text{br}} \in (0, 1)$ .<sup>5</sup> The voltage angle reference is set to zero for all realizations of the uncertainty by (4.6l). We can view Problem 4.1 as a mapping under uncertainty: the

<sup>5</sup> For ease of presentation we do not distinguish between different risk levels for upper and lower bounds.

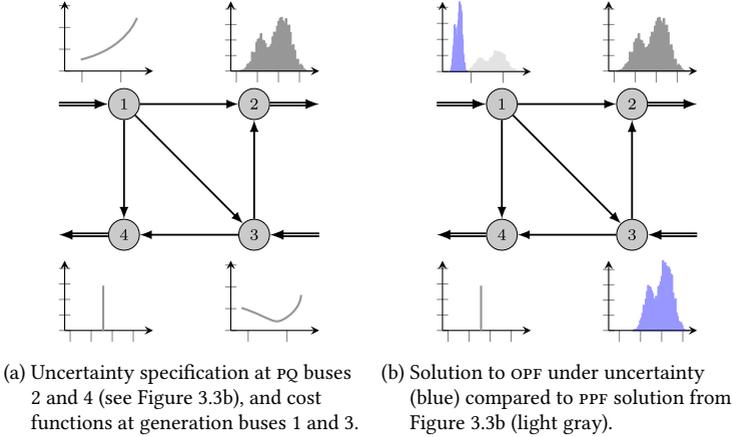


Figure 4.1: Idea of optimal power flow (OPF) under uncertainty for 4-bus system from Figure 3.3.

random-variable bus parameterizations are mapped to the random-variable overall grid state. We loosely write

$$“z^* = \operatorname{argmin} \text{Problem 4.1}(x),” \quad (4.7)$$

where  $z^* \in L^2(\Omega, \mu; \mathbb{R}^{4N_b})$  is the overall grid state, and the vector  $x$  contains the bus parameterizations that occur in the OPF problem (4.6).

Figure 4.1 visualizes the idea of OPF under uncertainty. The power system setup is equivalent to Figure 3.3 which demonstrates probabilistic power flow (PPF). Compared to the PPF setup from Figure 3.3a the OPF setup from Figure 4.1a contains not just uncertainty specifications but cost functions for the generator buses 1 and 3. The solution to OPF under uncertainty is shown in Figure 4.1b, plotted against the PPF solution from Figure 3.3b. As the generator at bus 3 is no longer injecting at a fixed value, its injection follows a PDF. The active power injection at bus 1 is shifted to lower values compared to the PPF solution; the degrees of freedom from OPF allow to assign power injections to the other generator at bus 3.

**Remark 4.1** (Joint chance constraint formulations). *In Problem 4.1 we chose to reformulate inequality constraints as individual chance constraints. Joint chance constraints are an alternative.*

Consider the following two inequalities

$$x^{\min} \leq x \leq x^{\max}, \quad y^{\min} \leq y \leq y^{\max};$$

treating  $x$  and  $y$  as realizations of the random variables  $\mathbf{x}$  and  $\mathbf{y}$  we would like to reformulate these constraints as chance constraints. Explicitly allowing for separate risk levels for upper and lower bounds, the approach from Problem 4.1 leads to

$$\begin{aligned} \mathbb{P}(\mathbf{x} \leq x^{\max}) &\geq 1 - \varepsilon_x^{\max}, \quad \mathbb{P}(x^{\min} \leq \mathbf{x}) \geq 1 - \varepsilon_x^{\min}, \\ \mathbb{P}(\mathbf{y} \leq y^{\max}) &\geq 1 - \varepsilon_y^{\max}, \quad \mathbb{P}(y^{\min} \leq \mathbf{y}) \geq 1 - \varepsilon_y^{\min}. \end{aligned}$$

A double joint chance constraint of the form

$$\mathbb{P}(x^{\min} \leq \mathbf{x} \leq x^{\max}) \geq 1 - \varepsilon_x, \quad \mathbb{P}(y^{\min} \leq \mathbf{y} \leq y^{\max}) \geq 1 - \varepsilon_y,$$

is an alternative; or a fully joint formulation

$$\mathbb{P}(x^{\min} \leq \mathbf{x} \leq x^{\max} \wedge y^{\min} \leq \mathbf{y} \leq y^{\max}) \geq 1 - \varepsilon.$$

We favor single-sided chance constraints over joint chance-constrained formulations, because if they are violated we have information about which quantity and which bound is not being satisfied. If a fully joint chance constraint is violated, it is in general difficult to track down the root cause. Also, single-sided chance constraints allow for a greater flexibility: for instance, if the lower bound  $x^{\min}$  is more critical than the upper bound  $x^{\max}$ , then  $\varepsilon_x^{\max}$  can be chosen smaller than  $\varepsilon_x^{\min}$ . For joint chance-constrained formulations, the risk level  $\varepsilon$  affects all constraints to the same degree. We refer to Table 4.2 for some references that apply joint chance constraints to OPF.

Problem 4.1 poses numerous challenges:

- c4.1 It is an infinite-dimensional problem because the decision variables are random variables of finite variance.
- c4.2 It entails a probabilistic power flow problem.

Table 4.2: Comparison of approaches for chance-constrained OPF in chronological order.

	Problem formulation			Solution methodology	
	PF	CC	PF	CC	
[177]	Full AC	Double	Monte Carlo	Gaussian uncertainties, Monte-Carlo	
[176]	Full AC	Single	Linearization	Back-mapping	
[175]	Full AC	Single	Linearization	Sparse grid integration, Gaussian	
[161]	Full AC	Joint	Semi-definite relaxation	Scenarios	
[14]	Approx. AC	Single	-	Linearization, convex approximation	
[44]	Lin. AC	Single	-	Convex approximation	
[148]	Full AC	Single	Iterative linearization	Moment-based	
[43]	Lin. AC	Single	-	Convex approximation	
[140]	Full AC	Single	Iterative linearization	Moment-based	
[157]	Convex Rel.	Joint	-	Gaussian uncertainties	
[156]	Convex Rel.	Joint	-	Scenarios uncertainties	
[155]	Convex Rel.	Joint	-	Scenarios	
[119]	Full AC	Single	Full AC, PCE	Moment-based, PCE	
[126]	Full AC	Single	Full AC, PCE	Moment-based, PCE	

c4.3 The cost function requires to compute the expected value of the individual random variable cost functions.

c4.4 The chance constraints require to compute the probability of constraint satisfaction in terms of decision variables.

Before we present our solution methodology to Problem 4.1, let us investigate how this problem has been tackled by previous works.

### 4.1.2 Existing approaches

We have presented *one* possibility to formulate and solve AC-OPF in the presence of uncertainties. Clearly, other works have proposed to formulate and solve AC-OPF in the presence of uncertainties as a chance-constrained optimization problem. Table 4.2 collects several works in chronological order, distinguishing between *problem formulation* and *solution methodology*.

Table 4.2 shows that early works on chance-constrained OPF formulated the problem with the full nonlinear AC power flow equations. Later works replaced the nonlinear and nonconvex AC equations by linear or convex approximations. There is a simple reason for that: it is a notoriously difficult task to propagate uncertainties through the original formulation of the power flow equations. As we can see from Table 4.2, the early works [161, 175–177] that did formulate the problem with the AC power flow equations were not able to solve the problem with these very equations, using Monte Carlo, linearization or relaxations instead. Formulating the problem itself with proxies such as linearizations or convex relaxations allows to also solve the problem directly. Albeit more appealing from a theoretical and computational side these proxies bear the danger of producing solutions that may not satisfy the original fully nonlinear AC power flow equations; for a survey on relaxation-based methods see [114].

For chance constraints the research landscape is more diverse: double, joint, and single chance constraint formulations exist along with a variety of methods to solve them. Scenario-based approaches replace chance constraints by a specific number of deterministic inequality constraints such that satisfying the many deterministic constraints ensures satisfying the original chance constraint [33]; scenario-based approaches for OPF are more often encountered in multi-period settings, see Section 4.2. Linearization-based methods for chance constraints replace the original nonlinear dependency by a linearization for which the first and second moment can then be computed. Hence, we can identify two mild trends: moment-based reformulations of chance constraints tend to be popular with linearization-based techniques, because the moments can be computed; and scenario-based approaches tend to be popular with convex relaxations, because they lead to large yet convex problems that can (hopefully) be solved reliably and efficiently.

The works [119, 126] provide a polynomial chaos-based approach to tackle the fully nonlinear AC power flow equations together with moment-based reformulations of chance constraints. Hence, PCE appears currently to be the only method that is able to propagate uncertainties through the power flow equations in a structured manner *and* providing the moments directly without having to linearize or sample.

Chance-constrained OPF is but one method to formulate OPF in the presence of uncertainties. There exist also robust formulations [100, 101, 113, 129]

of OPF. In that case, the inequality constraints are satisfied for all possible realizations of the random variable. This approach may be problematic as the information about the random variable—given by its probability density—is ignored; it is agnostic about what realizations will occur how often. Hence, robust constraints tend to be conservative. There is another reason that lets us be wary about robust constraints: if the support of the probability density function is the real axis, for instance, then a robust constraint formulation trivially induces infeasibility for any meaningful engineering limits.

Recently, so-called distributionally robust formulations have become popular. The idea is to introduce ambiguity sets that characterize a family of probability distributions that share certain properties [56]. For instance, a moment-based ambiguity set represents a set of distributions that all have the same statistical moments (up to a specified degree). Alternatively, ambiguity sets can be defined as all distributions that are close to some nominal distribution. The notion of “close” refers to a distance specific to probability distributions such as Wasserstein metrics, or Kullback-Leibler divergence; these ambiguity sets are then said to be metric-based [56]. The authors of [50] have applied Wasserstein-metric-based distributionally robust approaches to AC-OPF problems. However, in stark contrast to our proposed approach, the method from [50] is not able to consider the fully nonlinear AC power flow equations for all realizations of the uncertainty. Instead, the AC power flow equations are considered only for a nominal operating point; linearizations around that nominal operating point are employed to reformulate the chance constraints. The works [79–81] also consider Wasserstein-metric-based distributionally robust OPF problems, albeit in the multi-period setting. As the authors point out: “as long as an accurate linear model exists, the proposed technical approach can be utilized to formulate and solve a distributionally-robust chance-constrained AC-OPF problem.” Hence, the work [80] requires a linear model of the power flow equations, and is not able to propagate the uncertainties through the fully nonlinear AC power flow equations. The work [110] is an example for moment-based distributionally robust formulations in conjunction with a linearized dist-flow model.<sup>6</sup>

---

<sup>6</sup> The dist-flow model is an iterative formulation of the power flow equations for radial grids [15]. In contrast to backward-forward sweep (BFS) that alternates between voltages and currents, dist-flow is formulated with active powers, reactive powers, and voltage magnitudes.

To summarize: chance-constrained OPF is notoriously difficult owing to two reasons, namely the nonlinear AC power flow equations, and the chance constraints. There appears to be no structured approach that can deal with these challenges simultaneously. We aim to close that gap. Specifically, we propose to use polynomial chaos expansion to solve Problem 4.1 as it provides a natural framework to consider the full AC power flow equations and to treat moment-based reformulations of the chance constraints.

### 4.1.3 Solution methodology

We tackle the challenges c<sub>4.1</sub>-c<sub>4.4</sub> listed at the end of Section 4.1 subsequently by leveraging the advantages of polynomial chaos expansion mentioned in Section 2.4. Polynomial chaos directly addresses issue c<sub>4.1</sub>, because it allows to reformulate infinite-dimensional problems as finite-dimensional problems in terms of the PCE coefficients. As we would like to apply PCE, we need to make the following assumption.

**Assumption 4.2** (PCE of bus parameterizations). *The bus parameterizations  $\mathbf{x} \in L^2(\Omega, \mu; \mathbb{R}^{2N_b})$  from (4.3) in (4.6) admit a finite and exact polynomial chaos expansion with respect to the polynomial basis  $\{\phi_k\}_{k \in \mathcal{K}}$  with  $\mathcal{K} = \{0, 1, \dots, \hat{k}\}$ . We write*

$$\mathbf{x} = \sum_{k \in \mathcal{K}} \hat{\mathbf{x}}_k \phi_k \quad \text{with} \quad \hat{\mathbf{x}}_k \in \mathbb{R}^{2N_b}, \quad (4.8)$$

where

$$\hat{\mathbf{x}}_k = \left[ \hat{p}_{1,k}^u \quad \hat{q}_{1,k}^u \quad \dots \quad \hat{p}_{N_b,k}^u \quad \hat{q}_{N_b,k}^u \right]^\top. \quad (4.9)$$

We use the notation  $\hat{\cdot}$  to indicate and emphasize known and given values.

As a consequence of Assumption 4.2 we introduce PCE for every bus state  $\mathbf{z}_i$  with  $i \in \mathcal{N}$  according to (3.12). Just like for the PPF problem, our goal is to compute all the PCE coefficients  $\mathbf{z}_k = [z_{1,k}^\top \dots z_{N_b,k}^\top]^\top \in \mathbb{R}^{4N_b}$  for all  $k \in \mathcal{K}$ . Also, we introduce PCE for the random-variable controllable generation

$$\mathbf{p}_i^c = \sum_{k \in \mathcal{K}} p_{i,k}^c \phi_k. \quad (4.10)$$

We addressed c4.2, probabilistic power flow, in great detail in Chapter 3. We adopt the notation from (3.13) to represent the PCE-overloaded power flow equations in terms of the overall grid state PCE coefficients  $z = [z_0^\top \dots z_{\hat{k}}^\top]^\top$ , and write

$$g_{\text{PCE}}(z) = 0_{2N_b(\hat{k}+1)}. \quad (4.11)$$

We consider the PCE coefficients of the parameterizations by applying Galerkin projection to (4.2). This yields

$$p_{i,k} = p_{i,k}^c + \hat{p}_{i,k}^u, \quad \forall k \in \mathcal{K}, \forall i \in \mathcal{N}, \quad (4.12a)$$

$$q_{i,k} = q_{i,k}^c + \hat{q}_{i,k}^u, \quad \forall k \in \mathcal{K}, \forall i \in \mathcal{N}. \quad (4.12b)$$

The constraints for the slack bus angle references follow (3.14c) except that we do not specify a value for the real part,<sup>7</sup>

$$f_{i,k} = 0, \quad \forall k \in \mathcal{K}, \forall i \in \mathcal{N}_{\text{sl}}. \quad (4.13)$$

Let us turn to the cost function mentioned in c4.3. We consider convex quadratic costs

$$c_i(p_i^c) = c_{2,i}(p_i^c)^2 + c_{1,i}p_i^c, \quad (4.14a)$$

with  $c_{2,i} > 0$  for every bus  $i \in \mathcal{N}$ . The expected cost  $\mathbb{E}(c_i(p_i^c))$  per bus from (4.6a) written in terms of PCE coefficients becomes (see Section 2.4.3)

$$\begin{aligned} \mathbb{E}(c_i(p_i^c)) &= \underbrace{c_{2,i}\mathbb{E}(p_i^c)^2 + c_{1,i}\mathbb{E}(p_i^c)}_{=:c_i(\mathbb{E}(p_i^c))} + c_{2,i}\mathbb{V}(\mathbb{E}(p_i^c)) \\ &= c_{2,i} \sum_{k \in \mathcal{K}} \gamma_k (p_{i,k}^c)^2 + c_{1,i}p_{i,0}^c \\ &=: \tilde{c}_i(p_{i,k}^c), \end{aligned} \quad (4.14b)$$

with  $\gamma_k = \langle \phi_k, \phi_k \rangle$ . Notice that the cost function  $\tilde{c}_i(p_{i,k}^c)$  remains quadratic, but with respect to the PCE coefficients.

---

<sup>7</sup> For ease of presentation we assume every bus  $i \in \mathcal{N}_{\text{sl}}$  to have a deterministic imaginary voltage at zero.

It remains to address c4.4, the chance constraints. We reformulate the chance constraints from Problem 4.1 using the first two moments [32, 140, 145]. For example, the generation constraint in Problem 4.1 becomes

$$p_i^{\min} \leq \mathbb{E}(p_i^c) \pm \lambda(\varepsilon_{p_i}) \sqrt{\mathbb{V}(p_i^c)} \leq p_i^{\max}, \quad (4.15a)$$

where  $\lambda(\varepsilon_p) > 0$  is chosen based on the knowledge about and the confidence in the random variable  $p_i^c$ . The moment-based reformulation (4.15a) is particularly suitable with PCF as moments can be directly obtained from the PCF coefficients, see Section 2.4.3. Thus, the constraint (4.15a) becomes

$$p_i^{\min} \leq p_{i,0}^c \pm \lambda(\varepsilon_{p_i}) \sqrt{\mathbb{V}(p_i^c)} \leq p_i^{\max} \quad (4.15b)$$

with  $\mathbb{V}(p_i^c) = \sum_{k \in \mathcal{K} \setminus \{0\}} \gamma_k (p_{i,k}^c)^2$ . The reformulation of the other chance constraints for the generator (re)active powers from Problem 4.1 follows the same procedure. The chance constraints for voltage magnitudes  $v_i$  and line current magnitudes  $i_{ij}^{\text{br}}$  are replaced by constraints on their squared magnitudes and the corresponding first and second moment. The magnitude chance constraints become

$$(v_i^{\min})^2 \leq \mathbb{E}(v_i^2) \pm \lambda(\varepsilon_{v_i}) \sqrt{\mathbb{V}(v_i^2)} \leq (v_i^{\max})^2, \quad \forall i \in \mathcal{N}, \quad (4.16a)$$

$$\mathbb{E}((i_{ij}^{\text{br}})^2) + \lambda(\varepsilon_{i_{ij}^{\text{br}}}) \sqrt{\mathbb{V}((i_{ij}^{\text{br}})^2)} \leq (i_{ij}^{\text{br,max}})^2, \quad \forall (i, j) \in \mathcal{L}. \quad (4.16b)$$

The expressions for the moments are given in (4.18). The reason for using the moment-based reformulation on  $v_i^2$  and  $(i_{ij}^{\text{br}})^2$  instead of  $v_i$  and  $i_{ij}^{\text{br}}$  is that for the former, the moments can be obtained directly as an analytic function of the moments of  $e$  and  $f$  (see (4.17)), whereas for the latter, obtaining the moments would require additional equality constraints.

**Remark 4.2** (Choice of  $\lambda$ ). *In case  $p_i^c$  is Gaussian, the reformulation (4.15a) is exact with  $\lambda(\varepsilon_{p_i}) = \lambda_\Phi(\varepsilon_{p_i}) := \Phi^{-1}(1 - \varepsilon_{p_i})$ , where  $\Phi(\cdot)$  is the cumulative distribution function of a standard Gaussian random variable [20, 144]. Owing to the nonlinearity of the AC power flow, the resulting propagated random variables for CC-OPF (4.6) are non-Gaussian in general. Regardless, the distribution of those variables is often close to a Gaussian in practice. This is due to a concentration phenomenon similar to the central limit theorem [45, 140,*

145], making  $\lambda_\Phi$  a good heuristic that we employ in the following. In case the Gaussian heuristic is unsatisfactory, other choices of  $\lambda$  can be used to enforce distributionally robust chance constraints in terms of a moment-based ambiguity set. As these choices require weaker assumptions (such as symmetry and/or unimodality of the distribution) they become more conservative [32, 145]. Alternatively, the parameter  $\lambda$  can be chosen numerically via cross-validation or through online adaptive methods [133]. Finally, the Gaussian case is not the only one allowing exact reformulations for  $\lambda$ . Other examples include multivariate truncated Gaussians and uniform distributions on ellipsoidal supports [21, 22].

We are ready to cast CC-OPF according to Problem 4.1 as a finite-dimensional nonlinear program with the PCE coefficients as decision variables:

**Problem 4.2** (PCE-overloaded chance-constrained AC-OPF). *Consider Problem 4.1 and let Assumption 4.2 hold. Assuming a moment-based reformulation of the chance constraints, the PCE-overloaded reformulation of Problem 4.1 reads*

$$\min_{\substack{p_{i,k}^c, q_{i,k}^c, p_{i,k}^u, q_{i,k}^u, \\ e_{i,k}, f_{i,k} \in \mathbb{R} \\ \forall i \in \mathcal{N}, \forall k \in \mathcal{K}}} \sum_{i \in \mathcal{N}} \tilde{c}_i(p_{i,k}^c) \quad (4.18a)$$

subject to

$$g_{\text{PCE}}(z) = 0_{2N_b(\hat{k}+1)}, \quad (4.18b)$$

$$p_{i,k} = p_{i,k}^c + p_{i,k}^u, \quad \forall i \in \mathcal{N}, \forall k \in \mathcal{K}, \quad (4.18c)$$

$$q_{i,k} = q_{i,k}^c + q_{i,k}^u, \quad \forall i \in \mathcal{N}, \forall k \in \mathcal{K}, \quad (4.18d)$$

$$p_i^{\min} \leq p_{i,0} \pm \lambda(\varepsilon_{p_i}) \sqrt{\mathbb{V}(\mathbf{p}_i^c)} \leq p_i^{\max}, \quad \forall i \in \mathcal{N}, \quad (4.18e)$$

$$q_i^{\min} \leq q_{i,0} \pm \lambda(\varepsilon_{q_i}) \sqrt{\mathbb{V}(\mathbf{q}_i^c)} \leq q_i^{\max}, \quad \forall i \in \mathcal{N}, \quad (4.18f)$$

$$(v_i^{\min})^2 \leq \mathbb{E}(v_i^2) \pm \lambda(\varepsilon_{v_i}) \sqrt{\mathbb{V}(v_i^2)} \leq (v_i^{\max})^2, \quad \forall i \in \mathcal{N}, \quad (4.18g)$$

$$\mathbb{E}((i_{ij}^{\text{br}})^2) + \lambda(\varepsilon_{i_{ij}^{\text{br}}}) \sqrt{\mathbb{V}((i_{ij}^{\text{br}})^2)} \leq (i_{ij}^{\text{br,max}})^2, \quad \forall (i, j) \in \mathcal{L}, \quad (4.18h)$$

$$f_{i,k} = 0, \quad \forall i \in \mathcal{N}_{\text{SL}}, \forall k \in \mathcal{K}. \quad (4.18i)$$

Moments of squared line current magnitudes with  $(i, j) \in \mathcal{L}$ ,  $e_{ij,k} = e_{j,k}$ ,  $f_{ij,k} = f_{i,k} - f_{j,k}$ , and  $y_{ij} = g_{ij} + jb_{ij}$

$$\mathbb{E}(\left(\overset{\text{br}}{i_{ij}}\right)^2) = |y_{ij}|^2 \sum_{k \in \mathcal{K}} \langle \phi_k, \phi_k \rangle (e_{ij,k}^2 + f_{ij,k}^2) \quad (4.17a)$$

$$\mathbb{V}(\left(\overset{\text{br}}{i_{ij}}\right)^2) = |y_{ij}|^4 \sum_{k_1, k_2, k_3, k_4 \in \mathcal{K}} \langle \phi_{k_1}, \phi_{k_2}, \phi_{k_3}, \phi_{k_4} \rangle (e_{i, k_1} e_{ij, k_2} e_{i, k_3} e_{ij, k_4} + 2e_{ij, k_1} e_{ij, k_2} f_{ij, k_3} f_{ij, k_4} + f_{ij, k_1} f_{ij, k_2} f_{ij, k_3} f_{ij, k_4}) - \mathbb{E}(\left(\overset{\text{br}}{i_{ij}}\right)^2)^2 \quad (4.17b)$$

$$(4.17c)$$

Moments of squared voltage magnitudes with  $i \in \mathcal{N}$

$$\mathbb{E}(v_i^2) = \sum_{k \in \mathcal{K}} \langle \phi_k, \phi_k \rangle (e_{i,k}^2 + f_{i,k}^2) \quad (4.17d)$$

$$\mathbb{V}(v_i^2) = \sum_{k_1, k_2, k_3, k_4 \in \mathcal{K}} \langle \phi_{k_1}, \phi_{k_2}, \phi_{k_3}, \phi_{k_4} \rangle (e_{i, k_1} e_{i, k_2} e_{i, k_3} e_{i, k_4} + 2e_{i, k_1} e_{i, k_2} f_{i, k_3} f_{i, k_4} + f_{i, k_1} f_{i, k_2} f_{i, k_3} f_{i, k_4}) - \mathbb{E}(v_i^2)^2 \quad (4.17e)$$

The structural similarity between the PCE-overloaded Problem 4.2 and the original Problem 4.1 is obvious. Viewing Problem 4.2 as a mapping we loosely write

$$\left[ z_0^{*\top} \quad \dots \quad z_k^{*\top} \right]^\top = \operatorname{argmin} \text{ Problem 4.2}(\hat{x}_0, \dots, \hat{x}_k), \quad (4.19)$$

thus emphasizing that the PCE coefficients  $x_k$  of the bus parameterizations get mapped to optimal PCE coefficients of the overall grid state  $z_k^*$  for all  $k \in \mathcal{K}$ .

## 4.1.4 Discussion

### Computational characteristics

Problem 4.2 means to solve a nonconvex nonlinear program (NLP) in terms of the  $4N_b(\hat{k} + 1)$  PCE coefficients. This is a challenging task, given that the NLP contains already a PCE-overloaded probabilistic power flow problem via the equality constraints. It is advisable to warm-start the zero-order coefficients from the expected-value-solution, i.e. the solution obtained from solving a deterministic OPF problem for the expected value of the uncertainties. Also, constraint generation methods are of use for OPF problems: first neglect all inequality constraints, and solve the equality-constrained NLP. Then check whether there are violated inequality constraints. If there are, add them to the NLP and repeat the procedure until convergence. Clearly, this heuristic has merit only if the set of active inequality constraints is believed to have a few elements only. At least for the case studies from [127] this is the case. Other works also support this claim [130, 143].

### Interpretation of solution

The result of the NLP from Problem 4.2 is the collection of optimal PCE coefficients for the overall grid state. From these coefficients we can construct the optimal random variables that solve Problem 4.1. Recalling Remark 2.9 we can view interpret the solution from the angle of probability density functions and moments, or we can interpret the solution as policies that map specific realizations to specific inputs. Which view to choose depends on the

problem: are we solving a planning or an operational problem? For planning purposes the former frequentist point of view is perhaps more insightful. It allows to study the overall probability densities, including their (possible) skewness and/or kurtosis. For operational purposes the user is interested in *applicable* solutions, this means deterministic set points for the generators. This requires the policy point of view. In that case, however, the user has to bear in mind the meaning of the realization: inequality constraints cannot be expected to hold with certainty owing to the chance constraint formulation.

### Truncation errors

Problem 4.2 contains a PPF problem via the equality constraints. Hence, all considerations about truncation errors for PPF apply, see Section 3.3.1.

## 4.2 Chance-constrained DC optimal power flow

In Section 4.1 we studied chance-constrained AC optimal power flow. Its major methodological and computational challenge originates in the nonlinearity of the AC power flow equations. For high-voltage transmission systems we can impose additional assumptions that greatly simplify the mathematical structure of the power flow equations. This leads to the so-called DC power flow equations, see Remark 3.3 and Appendix A.6. Consequently, the DC power flow conditions lead to a simpler formulation of optimal power flow problems under uncertainty both mathematically and notationally. The main computational advantages of DC power flow compared to the fully nonlinear AC power flow are its linear structure and the fact that we can express line power flows as a linear combination of the net bus power injections by using the power transfer distribution factor matrix, see Appendix A.6.

The following material is based on [117, 121, 124].

### 4.2.1 Problem formulation

We consider a power system subject to DC power flow conditions. We restate the DC conditions from Appendix A.6.

**Assumption 4.3** (DC power flow). *Consider a power system under Assumption 3.1. Let the state of every bus  $i \in \mathcal{N}$  be given by (A.40) from Appendix A.6,*

$$z_i^{AC} = \begin{bmatrix} p_i & q_i & v_i & \delta_i \end{bmatrix}^\top \in \mathbb{R}^4, \quad (4.20)$$

and let the following conditions hold:

A4.1 *the Ohmic losses across each line  $(j, k) \in \mathcal{L}$  are negligible such that  $g_{jk} = 0$  holds for the  $\Pi$ -line model from Figure 3.1;*

A4.2 *the voltage angle differences are small across all lines  $(j, k) \in \mathcal{L}$ ;*

A4.3 *the voltage magnitudes are constant at one per unit for all buses  $i \in \mathcal{N}$ .*

A modeling consequence of Assumption 4.3 is that we need not consider either the reactive power or the voltage magnitude at any bus  $i \in \mathcal{N}$ , immediately halving the number of unknowns in the power system to  $2N_b$ ; only the active power and the voltage angle are unknowns. With this we can tailor Assumption 3.2 to the DC conditions.

**Assumption 4.4** (Power system under uncertainty and DC conditions). *We study a power system under Assumption 3.1 in the presence of uncertainties. Additionally we consider DC conditions according to Assumption 4.3. That means we can model the overall grid state by two  $\mathbb{R}^{N_b}$ -valued random variables*

$$\mathbf{p} = \begin{bmatrix} p_1 & \dots & p_{N_b} \end{bmatrix}^\top \in L^2(\Omega, \mu; \mathbb{R}^{N_b}), \quad (4.21a)$$

$$\delta = \begin{bmatrix} \delta_1 & \dots & \delta_{N_b} \end{bmatrix}^\top \in L^2(\Omega, \mu; \mathbb{R}^{N_b}), \quad (4.21b)$$

where  $p_i$  is the random-variable active power, and  $\delta_i$  is the random-variable voltage angle, each at bus  $i \in \mathcal{N}$ .

With the notation from Assumption 4.4 the DC power flow equations in terms of random variables read

$$\mathbf{p} = -B\delta, \quad (4.22)$$

where  $B \in \mathbb{R}^{N_b \times N_b}$  is the imaginary part of the bus admittance matrix under DC conditions, see Appendix A.5. In case we are not interested in the voltage angles we can sum up (4.22)

$$\sum_{i \in \mathcal{N}} p_i = \mathbf{1}_{N_b}^\top \mathbf{p} = 0, \quad (4.23)$$

which holds because  $\mathbf{1}_{N_b}$  is a left-eigenvector of  $B$  for the eigenvalue 0, see Appendix A.6. Under DC conditions we can leverage the so-called power transfer distribution factor (PTDF) matrix  $\Psi \in \mathbb{R}^{N_{br} \times N_b}$  to map the random-variable bus powers directly to the random-variable line flows<sup>8</sup>

$$\mathbf{p}^{br} = \Psi \mathbf{p}. \quad (4.24)$$

The entries of the random-variable line flows are

$$\mathbf{p}^{br} = \left[ p_1^{br} \quad \dots \quad p_{N_{br}}^{br} \right]^\top \in L^2(\Omega, \mu; \mathbb{R}^{N_{br}}), \quad (4.25)$$

where  $p_i^{br}$  corresponds to the pair  $(j, k) \in \mathcal{L}$  such that the  $i^{\text{th}}$  row of the incidence matrix  $A$  has two non-zero entries, one at position  $j$  and one at position  $k$ , i.e.  $A_{ij}, A_{ik} \in \{-1, 1\}$ . For what follows it is more convenient to have a linear indexing of the lines by  $\mathcal{L}^{\text{lin}} = \{1, \dots, N_{br}\}$  rather than the graph-theoretic indexing by edges.

In what follows we borrow largely from Section 4.1 and tailor it to the DC setting. We assume each bus  $i \in \mathcal{N}$  connects to a single controllable unit and a single uncontrollable unit, see (4.2)

$$\mathbf{p}_i = \mathbf{p}_i^c + \mathbf{p}_i^u, \quad (4.26a)$$

and we subsume the entire uncontrollable generation/demand in the bus parameterizations  $\mathbf{x} \in L^2(\Omega, \mu; \mathbb{R}^{N_b})$

$$\mathbf{x} = \left[ p_1^u \quad \dots \quad p_{N_b}^u \right]^\top \in L^2(\Omega, \mu; \mathbb{R}^{N_b}). \quad (4.27)$$

<sup>8</sup> We refer to Appendix A.6 for a concise derivation of the PTDF matrix.

The considerations from Section 4.1 about the cost functions per bus apply to the DC setting without changes. It remains to address the formulation of the engineering limits under uncertainty for DC conditions. Of the engineering limits from Table 4.1, only two apply to the DC setting: the active power generation limits, and the branch flow limits in terms of active power line flows. In the deterministic case we would have to enforce the inequality constraints

$$p_i^{\min} \leq p_i^c \leq p_i^{\max}, \quad (4.28a)$$

$$-p_j^{\text{br,max}} \leq p_j^{\text{br}} \leq p_j^{\text{br,max}}, \quad (4.28b)$$

for all buses  $i \in \mathcal{N}$  and all lines  $j \in \mathcal{L}^{\text{lin}} = \{1, \dots, N_{\text{br}}\}$ . Just like in Section 4.1 we reformulate the inequality constraints (4.28a) as individual chance constraints

$$\mathbb{P}(p_i^c \leq p_i^{\max}) \geq 1 - \varepsilon_{p_i}, \quad \mathbb{P}(p_i^c \geq p_i^{\min}) \geq 1 - \varepsilon_{p_i}, \quad (4.28c)$$

$$\mathbb{P}(p_j^{\text{br}} \geq -p_j^{\text{br,max}}) \geq 1 - \varepsilon_{p_j^{\text{br}}}, \quad \mathbb{P}(p_j^{\text{br}} \leq p_j^{\text{br,max}}) \geq 1 - \varepsilon_{p_j^{\text{br}}}, \quad (4.28d)$$

for some  $\varepsilon_{p_i}, \varepsilon_{p_j^{\text{br}}} \in (0, 1)$  and for all buses  $i \in \mathcal{N}$  and all lines  $j \in \mathcal{L}^{\text{lin}}$ . We are ready to formulate a chance-constrained DC-OPF problem.

**Problem 4.3** (Chance-constrained DC-OPF). *Let Assumption 4.4 hold for a given power system, and let the bus parameterizations (4.27) be given. We call the following optimization problem a chance-constrained DC optimal power flow problem*

$$\min_{\substack{p_i^c, p_i \in L^2(\Omega, \mu; \mathbb{R}) \\ \forall i \in \mathcal{N}}} \sum_{i \in \mathcal{N}} \mathbb{E}(c_i(p_i^c)) \quad (4.29a)$$

subject to

$$\sum_{i \in \mathcal{N}} p_i = 0, \quad (4.29b)$$

$$p_i = p_i^c + p_i^u, \quad \forall i \in \mathcal{N}, \quad (4.29c)$$

$$\mathbb{P}(p_i^c \geq p_i^{\min}) \geq 1 - \varepsilon_{p_i}, \quad \forall i \in \mathcal{N}, \quad (4.29d)$$

$$\mathbb{P}(p_i^c \leq p_i^{\max}) \geq 1 - \varepsilon_{p_i}, \quad \forall i \in \mathcal{N}, \quad (4.29e)$$

$$\mathbb{P}(\Psi_j^\top p \leq p_j^{\text{br,max}}) \geq 1 - \varepsilon_{p_j^{\text{br}}}, \quad \forall j \in \mathcal{L}^{\text{lin}}, \quad (4.29f)$$

$$\mathbb{P}(\Psi_j^\top \mathbf{p} \geq -p_j^{\text{br,max}}) \geq 1 - \varepsilon_{p_j^{\text{br}}}, \quad \forall j \in \mathcal{L}^{\text{lin}}, \quad (4.29\text{g})$$

where  $\Psi_j^\top \in \mathbb{R}^{1 \times N_b}$  is the  $j^{\text{th}}$  row of the PTDF matrix.<sup>9</sup>

Problem 4.3 minimizes the sum of the expected costs of active power generation subject to the DC power flow equations and individual chance constraints for the active power generation and line flow limits. Having solved Problem 4.3 we can compute the random-variable voltage angles from (4.22).

With  $\mathbf{x}$  being the bus parameterizations according to (4.27), we can view Problem 4.3 as mapping the bus parameterizations  $\mathbf{x}$  to the overall active power  $\mathbf{p} \in L^2(\Omega, \mu; \mathbb{R}^{N_b})$  via

$$\mathbf{p}^* = \text{argmin Problem 4.3}(\mathbf{x}), \quad (4.30)$$

from which we can determine the overall grid state by solving for the random-variable voltage angles  $\delta$  according to the DC power flow equations (4.22).

We are in the same situation we faced in Section 4.1.3: having formulated the chance-constrained DC-OPF problem we see that the same challenges C4.1-C4.4 appear. How has this problem been tackled in the existing literature?

## 4.2.2 Existing approaches

The chance-constrained OPF problem under DC conditions has been studied extensively in the literature. From Table 4.3 we immediately see that moment-based approaches appear to be most popular. This is not too surprising, because the linear equations of the DC power flow equations allow to derive analytical equations for the moments of all quantities of interest, especially the branch flows. In combination with the assumption of Gaussian uncertainties, this allows for an elegant, exact, and tractable reformulation of chance constraints, see also Remark 4.2. The scenario-based approaches [159, 160] make use of results from multi-parametric programming to compute the solution to the chance-constrained OPF problem as a piece-wise affine policy.

<sup>9</sup> If we were to consider constraints for the angles, they would appear as decision variables, and we would have to replace (4.29b) by (4.22).

Table 4.3: Approaches for chance-constrained OPF under DC conditions.

Approach	Work	Characteristics
Chance constraint	[20, 141, 144, 145]	Moment-based, Gaussian
	[21, 22]	Moment-based, Gaussian
	[142]	Weighted chance constraints
	[159, 160]	Scenario-based
	[117, 121, 124]	Moment-based, PCE
Distributionally robust	[96, 102, 145, 171, 179, 180]	Moment-based
	[97]	Ellipsoid-based

Just like in the AC case, distributionally robust approaches are an alternative to chance constraints; Table 4.3 lists several works. Again, this is to be expected because linear mappings integrate nicely with distributionally robust methods [56].

Just like their AC counterparts OPF problems under uncertainty and DC conditions are infinite-dimensional problems. To render the problem finite-dimensional affine response policies have been proposed as a potential remedy, one of the earliest works being [163] for the multi-stage setting. These affine policies ensure power balance despite the presence of stochastic fluctuations. Albeit intuitive from an engineering point of view it was not until [20] that a rigorous mathematical investigation was carried out, showing why affine policies *do* ensure power balance, and how. However, [20] focused on the Gaussian setting.

To summarize: the existing literature on chance-constrained OPF under DC conditions has focused to a large degree on affine policies and specific uncertainties, often Gaussian. With the help of polynomial chaos it is possible to propose a framework for chance-constrained OPF under DC conditions that naturally computes affine policies that satisfy power balance in the presence of arbitrary uncertainties of finite variance [117, 121, 124]. Hence, the PCE framework considers existing approaches as special cases, and provides extensions thereof.

### 4.2.3 Solution methodology

Chance-constrained OPF under DC conditions faces the same conceptual challenges C4.1-C4.4 as its AC counterpart. We address them step by step, keeping the presentation concise as it mirrors largely Section 4.1.3.

We tackle C4.1 by using PCE: it renders the problem finite-dimensional. To apply PCE means to introduce PCE for the bus specifications.

**Assumption 4.5** (PCE of bus parameterizations for Problem 4.3). *The bus parameterizations  $x \in L^2(\Omega, \mu; \mathbb{R}^{N_b})$  from (4.27) in Problem 4.3 admit a finite and exact polynomial chaos expansion with respect to the polynomial basis  $\{\phi_k\}_{k \in \mathcal{K}}$  with  $\mathcal{K} = \{0, 1, \dots, \hat{k}\}$ . We write*

$$x = \sum_{k \in \mathcal{K}} \hat{x}_k \phi_k \quad \text{with} \quad \hat{x}_k \in \mathbb{R}^{2N_b}, \quad (4.31)$$

where

$$\hat{x}_k = \left[ \hat{p}_{1,k}^u \quad \dots \quad \hat{p}_{N_b,k}^u \right]^\top. \quad (4.32)$$

We use the notation  $\hat{\cdot}$  to indicate and emphasize known and given values.

As a consequence of Assumption 4.5 we introduce PCE for the random-variable bus power and the random-variable generated power alike

$$p_i = \sum_{k \in \mathcal{K}} p_{i,k} \phi_k, \quad (4.33a)$$

$$p_i^c = \sum_{k \in \mathcal{K}} p_{i,k}^c \phi_k, \quad (4.33b)$$

for all buses  $i \in \mathcal{N}$ . Our goal is to reformulate Problem 4.3 as an optimization problem in terms of the PCE coefficients. Under DC conditions, the probabilistic power flow problem, i.e. challenge C4.2, becomes straightforward: ensuring power balance (4.29b) in terms of random variables becomes identical to ensuring “power balance in terms of the PCE coefficients.” This follows from applying Galerkin projection to (4.29b), giving

$$\sum_{i \in \mathcal{N}} p_{i,k} = 0, \quad (4.34)$$

for all  $k \in \mathcal{K}$ . We apply the same idea to the bus parameterizations (4.29c)

$$p_{i,k} = p_{i,k}^c + \hat{p}_{i,k}^u, \quad (4.35)$$

for all buses  $i \in \mathcal{N}$  and all PCE coefficients indices  $k \in \mathcal{K}$ . The reformulation of the cost, i.e. challenge c4.3, is analogous to Section 4.1.3. The same applies to the final challenge c4.4, the moment-based reformulation of the single-sided chance constraints, leading to

$$p_i^{\min} \leq p_{i,0} \pm \lambda(\varepsilon_{p_i}) \sqrt{\mathbb{V}(\mathbf{p}_i^c)} \leq p_i^{\max}, \quad (4.36a)$$

$$-p_j^{\text{br,max}} \leq p_{j,0}^{\text{br}} \pm \lambda(\varepsilon_{p_j^{\text{br}}}) \sqrt{\mathbb{V}(\mathbf{p}_j^{\text{br}})} \leq p_j^{\text{br,max}}, \quad (4.36b)$$

where

$$\mathbb{V}(\mathbf{p}_i^c) = \sum_{k \in \mathcal{K} \setminus \{0\}} \gamma_k (p_{i,k}^c)^2, \quad (4.36c)$$

$$\mathbb{V}(\mathbf{p}_j^{\text{br}}) = \mathbb{V}(\Psi_j^\top \mathbf{p}) = \sum_{i_1, i_2 \in \mathcal{N}} \Psi_{ji_1} \Psi_{ji_2} \sum_{k \in \mathcal{K} \setminus \{0\}} \gamma_k p_{i_1,k} p_{i_2,k}, \quad (4.36d)$$

with  $\gamma_k = \langle \phi_k, \phi_k \rangle$ . These results follow from orthogonality of the basis functions, see Section 2.4.3. Having addressed all the challenges c4.1-c4.4 we are ready to cast Problem 4.3 as a finite-dimensional optimization problem in terms of the PCE coefficients.

**Problem 4.4** (PCE-overloaded chance-constrained DC-OPF). *Consider Problem 4.3 and let Assumption 4.2 hold. Assuming a moment-based reformulation of the chance constraints, the PCE-overloaded reformulation of Problem 4.3 reads*

$$\min_{\substack{p_{i,k}^c, p_{i,k} \in \mathbb{R} \\ \forall i \in \mathcal{N}, \forall k \in \mathcal{K}}} \sum_{i \in \mathcal{N}} \tilde{c}_i(p_{i,k}^c) \quad (4.37a)$$

subject to

$$\sum_{i \in \mathcal{N}} p_{i,k} = 0, \quad \forall i \in \mathcal{N}, \forall k \in \mathcal{K}, \quad (4.37b)$$

$$p_{i,k} = p_{i,k}^c + \hat{p}_{i,k}^u, \quad \forall i \in \mathcal{N}, \forall k \in \mathcal{K}, \quad (4.37c)$$

$$p_i^{\min} \leq p_{i,0} \pm \lambda(\varepsilon_{p_i}) \sqrt{\mathbb{V}(\mathbf{p}_i^c)} \leq p_i^{\max}, \quad \forall i \in \mathcal{N}, \quad (4.37d)$$

$$-p_j^{\text{br,max}} \leq p_{j,0}^{\text{br}} \pm \lambda(\varepsilon_{p_j^{\text{br}}}) \sqrt{\mathbb{V}(p_j^{\text{br}})} \leq p_j^{\text{br,max}}, \quad \forall j \in \mathcal{L}^{\text{lin}}. \quad (4.37\text{e})$$

## 4.2.4 Discussion

### Computational characteristics

The reformulated Problem 4.3 is a second-order cone program (SOCP); the interested reader is referred to [117, 121] for a detailed derivation of the second-order cone formulation based on [99]. Hence we need to solve a convex problem, for which there are dedicated off-the-shelf solvers that solve SOCPs, e.g. Mosek or Gurobi. In principle this can be done efficiently with certified optimality, i.e. zero duality gap. Nonetheless, the combination of large grids with many uncertainties may lead to SOCPs that can be numerically challenging requiring tailored algorithms [20]. The SOCP from Problem 4.4 has a total of  $N_b(\hat{k} + 1)$  decision variables: for every bus  $i \in N_b$  the  $(\hat{k} + 1)$  PCE coefficients  $p_{i,k}^c$  of the controllable power need to be computed for all indices  $k \in \mathcal{K}$ .

### Interpretation of solution

This is analogous to the interpretation of the solution for the AC case from Section 4.1.4: according to Remark 2.9 we can use the computed PCE coefficients to derive either statistical information about the solution such as mean values or expected constraint violation, or we might use the PCE coefficients and the corresponding PCE as a policy that maps realizations of the uncertainty to realizations of the decision variables. Clearly, the latter interpretation is of interest to practitioners, while the former one is more advantageous for planning purposes.

### Truncation errors

As the DC power flow equations are linear, the results from Section 2.5 apply: given a finite and exact PCE representation of the bus parameterizations, power balance can be achieved for all realizations of the uncertainty

by expanding the decision variables in the same basis. This result is formalized in [117], independently of the underlying distribution. Previous results [20] restricted the analysis to Gaussian uncertainties. If in addition the moment-based reformulation of the chance constraints is exact (e.g. Gaussian uncertainties)—or conversely, if we start from a moment-based formulation of the problem—the PCE re-formulation is exact [121]. These properties are a direct consequence of the linearity of DC power flow, hence a prime example for how structure can be exploited in favor of theoretical findings.

## 4.3 Multi-stage chance-constrained DC optimal power flow

The chance-constrained OPF problems we considered thus far studied power systems under uncertainty at a single time instant; time did not appear explicitly. We now turn to the multi-stage setting where we study chance-constrained OPF problems over a time horizon.<sup>10</sup> In contrast to the previous single-stage settings from Section 4.1 or Section 4.2 the multi-stage setting allows to consider equipment ramp constraints and/or energy storage systems. For instance, generators cannot change their set points arbitrarily fast, which imposes additional inequality constraints. Including energy storage systems allows to determine schedules that obey their dynamics and guarantee, for example, a desired storage level at the end of the optimization horizon.

### 4.3.1 Problem formulation

In the following we borrow largely from the problem formulation for the single-stage problem from Section 4.2. It remains mainly to introduce a notation for the time instants, and to model the storages and ramp constraints.

**Assumption 4.6** (Power system under uncertainty and DC conditions; multi-stage). *We study a power system under Assumption 3.1 in the presence of uncertainties at time instants  $\mathcal{T} = \{1, \dots, T\}$ .*<sup>11</sup> *Additionally we consider*

---

<sup>10</sup> Also called multi-period setting.

<sup>11</sup> Beginning with time instant  $t = 1$  allows for a lighter notation.

DC conditions according to Assumption 4.3. We model the overall grid state by square-integrable discrete stochastic processes for the active power and the phase angle

$$\mathbf{p}_i = \left[ p_i(1) \quad p_i(2) \quad \dots \quad p_i(T) \right]^\top \in L^2(\Omega, \mu; \mathbb{R}^T), \quad (4.38a)$$

$$\delta_i = \left[ \delta_i(1) \quad \delta_i(2) \quad \dots \quad \delta_i(T) \right]^\top \in L^2(\Omega, \mu; \mathbb{R}^T), \quad (4.38b)$$

for every bus  $i \in \mathcal{N}$ . The overall grid state at every time instant  $t \in \mathcal{T}$  is modeled by the  $\mathbb{R}^{N_b}$ -valued random vectors

$$\mathbf{p}(t) = \left[ \mathbf{p}_1(t) \quad \dots \quad \mathbf{p}_{N_b}(t) \right]^\top \in L^2(\Omega, \mu; \mathbb{R}^{N_b}), \quad (4.38c)$$

$$\delta(t) = \left[ \delta_1(t) \quad \dots \quad \delta_{N_b}(t) \right]^\top \in L^2(\Omega, \mu; \mathbb{R}^{N_b}). \quad (4.38d)$$

In Assumption 4.6 we make explicit use of the dichotomy of discrete stochastic processes and random vectors, see Section 2.2.2: though mathematically equivalent, their interpretation is different. We treat the active power and phase at every bus  $i \in \mathcal{N}$  as a discrete process, but the overall grid state at a specific time instant  $t \in \mathcal{T}$  is a random vector, a snapshot in time across the power system. This covers the mathematical description of the power system variables; let us turn to the physical model.

The DC power flow conditions hold for every time instance  $t \in \mathcal{T}$ . Hence, the DC power flow equations, the energy balance, and the PPDF mapping that we introduced in Section 4.2 now read

$$\mathbf{p}(t) = -B\delta(t), \quad (4.39a)$$

$$\sum_{i \in \mathcal{N}} \mathbf{p}_i(t) = \mathbf{1}_{N_b}^\top \mathbf{p}(t) = 0, \quad (4.39b)$$

$$\mathbf{p}^{\text{br}}(t) = \Psi \mathbf{p}(t), \quad (4.39c)$$

for all time instants  $t \in \mathcal{T}$ . From (4.39c) we see that also every line flow is now a discrete stochastic process in the sense that

$$\mathbf{p}_i^{\text{br}} = \left[ \mathbf{p}_i^{\text{br}}(1) \quad \mathbf{p}_i^{\text{br}}(2) \quad \dots \quad \mathbf{p}_i^{\text{br}}(T) \right]^\top \in L^2(\Omega, \mu; \mathbb{R}^T) \quad (4.40a)$$

holds for all lines  $i \in \mathcal{L}^{\text{lin}}$ , and

$$\mathbf{p}^{\text{br}}(t) = \left[ \mathbf{p}_1^{\text{br}}(t) \quad \dots \quad \mathbf{p}_{N_{\text{br}}}^{\text{br}}(t) \right]^{\top} \in L^2(\Omega, \mu; \mathbb{R}^{N_{\text{br}}}) \quad (4.40b)$$

is the snapshot of all random-variable line flows at all time instants  $t \in \mathcal{T}$ .

Just like in Section 4.1 and Section 4.2 we suppose that each bus  $i \in \mathcal{N}$  connects to a single controllable unit and a single uncontrollable unit. Additionally we assume each bus connects to a single energy storage system, hence the net power at every bus  $i \in \mathcal{N}$  reads

$$\mathbf{p}_i(t) = \mathbf{p}_i^{\text{c}}(t) + \mathbf{p}_i^{\text{u}}(t) + \mathbf{p}_i^{\text{s}}(t), \quad (4.41)$$

for all time instants  $t \in \mathcal{T}$ , where  $\mathbf{p}_i^{\text{u}}(t)$  is the uncontrollable random-variable power and  $\mathbf{p}_i^{\text{s}}(t)$  is the random-variable power drawn from/injected into the storage at time instant  $t$ . We subsume the entire uncontrollable generation/demand as the discrete stochastic process

$$\forall i \in \mathcal{N}: \quad \mathbf{x}_i = \left[ \mathbf{p}_i^{\text{u}}(1) \quad \mathbf{p}_i^{\text{u}}(2) \quad \dots \quad \mathbf{p}_i^{\text{u}}(T) \right]^{\top} \in L^2(\Omega, \mu; \mathbb{R}^T) \quad (4.42)$$

with known and given characteristics for all buses. Hence, the bus parameterizations are given by discrete stochastic processes.

We consider integrator dynamics for the storage system

$$\mathbf{e}_i(t+1) = \mathbf{e}_i(t) - h \mathbf{p}_i^{\text{s}}(t), \quad \mathbf{e}_i(1) = \mathbf{e}_i^{\text{ic}}, \quad (4.43)$$

for all buses  $i \in \mathcal{N}$ . We call  $\mathbf{e}_i(t)$  the random-variable energy of the storage at bus  $i$  at time instant  $t$ , and  $\mathbf{e}_i^{\text{ic}} \in L^2(\Omega, \mu; \mathbb{R})$  is some given yet random-variable initial condition.<sup>12</sup> The value of  $h > 0$  subsumes the discretization time and a potential loss factor. We thus see that both the state of the storage and its input are discrete stochastic processes

$$\mathbf{e}_i = \left[ \mathbf{e}_i(1) \quad \mathbf{e}_i(2) \quad \dots \quad \mathbf{e}_i(T+1) \right]^{\top} \in L^2(\Omega, \mu; \mathbb{R}^{T+1}), \quad (4.44a)$$

$$\mathbf{p}_i^{\text{s}} = \left[ \mathbf{p}_i^{\text{s}}(1) \quad \mathbf{p}_i^{\text{s}}(2) \quad \dots \quad \mathbf{p}_i^{\text{s}}(T) \right]^{\top} \in L^2(\Omega, \mu; \mathbb{R}^T), \quad (4.44b)$$

---

<sup>12</sup> We accept the slight clash of notation with the real part of the voltage  $\mathbf{e}_i$ .

related by the storage dynamics. Just like the rest of the equipment, the energy storage system is constrained. We model its engineering limits in terms of individual chance constraints for every time instant, yielding

$$\mathbb{P}(e_i(t+1) \leq \bar{e}_i) \geq 1 - \varepsilon, \quad \mathbb{P}(e_i(t+1) \geq \underline{e}_i) \geq 1 - \varepsilon, \quad (4.45a)$$

$$\mathbb{P}(p_i^s(t) \leq \bar{s}_i) \geq 1 - \varepsilon, \quad \mathbb{P}(p_i^s(t) \geq \underline{s}_i) \geq 1 - \varepsilon, \quad (4.45b)$$

for all buses  $i \in \mathcal{N}$  with (4.45a) being the limits for the energy of the storage, and (4.45b) being the limits for the storage input. For ease of presentation we choose an identical risk level  $\varepsilon$  for all quantities.

Generator ramp constraints are considered as individual chance constraints

$$\mathbb{P}(p_i^c(t) - p_i^c(t-1) \leq \Delta p_i^{\max}) \geq 1 - \varepsilon, \quad (4.46a)$$

$$\mathbb{P}(p_i^c(t) - p_i^c(t-1) \geq \Delta p_i^{\min}) \geq 1 - \varepsilon, \quad (4.46b)$$

for all time instants  $t \in \mathcal{T} \setminus \{1\}$  and all buses  $i \in \mathcal{N}$ , thus limiting the probability of abrupt set point changes.

For the cost function we modify the formulation from Problem 4.3 by additionally summing over all time instants  $t \in \mathcal{T}$ .

**Problem 4.5** (Chance-constrained multi-stage DC-OPF). *Let Assumption 4.6 hold for a given power system, and let the bus parameterizations (4.42) be given. We call the following optimization problem a chance-constrained multi-stage DC optimal power flow problem*

$$\min_{\substack{p_i^c, p_i^s, e_i, p_i \in L^2(\Omega, \mu; \mathbb{R}^T) \\ \forall i \in \mathcal{N}}} \sum_{t \in \mathcal{T}} \sum_{i \in \mathcal{N}} \mathbb{E}(c_i(p_i^c(t))) \quad (4.47a)$$

subject to

$$p_i(t) = p_i^c(t) + p_i^u(t) + p_i^s(t), \quad \forall i \in \mathcal{N}, \forall t \in \mathcal{T}, \quad (4.47b)$$

$$\sum_{i \in \mathcal{N}} p_i^u(t) + p_i^c(t) + p_i^s(t) = 0 \quad \forall i \in \mathcal{N}, \forall t \in \mathcal{T}, \quad (4.47c)$$

$$e_i(t+1) = e_i(t) - h p_i^s(t), \quad \forall i \in \mathcal{N}, \forall t \in \mathcal{T} \setminus \{T\}, \quad (4.47d)$$

$$e_i(1) = e_i^{\text{ic}}, \quad \forall i \in \mathcal{N}, \quad (4.47e)$$

$$\mathbb{P}(p_i^c(t) \leq p_i^{\max}) \geq 1 - \varepsilon, \quad \forall i \in \mathcal{N}, \forall t \in \mathcal{T}, \quad (4.47f)$$

$$\mathbb{P}(\mathbf{p}_i^c(t) \geq p_i^{\min}) \geq 1 - \varepsilon, \quad \forall i \in \mathcal{N}, \forall t \in \mathcal{T}, \quad (4.47g)$$

$$\mathbb{P}(\Psi_j^\top \mathbf{p}(t) \leq p_j^{\text{br,max}}) \geq 1 - \varepsilon, \quad \forall i \in \mathcal{N}, \forall t \in \mathcal{T}, \quad (4.47h)$$

$$\mathbb{P}(\Psi_j^\top \mathbf{p}(t) \geq -p_j^{\text{br,max}}) \geq 1 - \varepsilon, \quad \forall i \in \mathcal{N}, \forall t \in \mathcal{T}, \quad (4.47i)$$

$$\mathbb{P}(\Delta \mathbf{p}_i^c(t) \leq \Delta p_i^{\max}) \geq 1 - \varepsilon, \quad \forall i \in \mathcal{N}, \forall t \in \mathcal{T} \setminus \{0\}, \quad (4.47j)$$

$$\mathbb{P}(\Delta \mathbf{p}_i^c(t) \geq \Delta p_i^{\min}) \geq 1 - \varepsilon, \quad \forall i \in \mathcal{N}, \forall t \in \mathcal{T} \setminus \{0\}, \quad (4.47k)$$

$$\mathbb{P}(\mathbf{e}_i(t+1) \leq \bar{\mathbf{e}}_i) \geq 1 - \varepsilon, \quad \forall i \in \mathcal{N}, \forall t \in \mathcal{T} \setminus \{T\}, \quad (4.47l)$$

$$\mathbb{P}(\mathbf{e}_i(t+1) \geq \underline{\mathbf{e}}_i) \geq 1 - \varepsilon, \quad \forall i \in \mathcal{N}, \forall t \in \mathcal{T} \setminus \{T\}, \quad (4.47m)$$

$$\mathbb{P}(\mathbf{e}_i(T) \leq e_i^{T,\max}) \geq 1 - \varepsilon, \quad \forall i \in \mathcal{N}, \quad (4.47n)$$

$$\mathbb{P}(\mathbf{e}_i(T) \geq e_i^{T,\min}) \geq 1 - \varepsilon, \quad \forall i \in \mathcal{N}, \quad (4.47o)$$

$$\mathbb{P}(\mathbf{p}_i^s(t) \leq s_i^{T,\max}) \geq 1 - \varepsilon, \quad \forall i \in \mathcal{N}, \forall t \in \mathcal{T}, \quad (4.47p)$$

$$\mathbb{P}(\mathbf{p}_i^s(t) \geq s_i^{T,\min}) \geq 1 - \varepsilon, \quad \forall i \in \mathcal{N}, \forall t \in \mathcal{T}, \quad (4.47q)$$

with  $\Delta \mathbf{p}_i^c(t) = \mathbf{p}_i^c(t) - \mathbf{p}_i^c(t-1)$ , and  $\varepsilon \in [0.5, 1)$ .<sup>13</sup> Note that  $\Psi_j^\top \in \mathbb{R}^{1 \times N_b}$  is the  $j^{\text{th}}$  row of the PTDF matrix.

Problem 4.5 minimizes over stochastic processes. Specifically, we minimize the expected cost of thermal generation (4.47a) over time while satisfying the power balance (4.47c) for every realization of the uncertain disturbance, and the storage dynamics (4.47d) in terms of stochastic processes, see (4.43). All engineering limits are formulated in terms of single-sided chance constraints. We added chance constraints (4.47n) and (4.47o) for the terminal state of the storage to allow for the storage to be at a prescribed level (with high probability) at the end of the horizon.

It is possible to add another constraint to restrict the standard deviation of all occurring random variables. For example, this enables to restrict the variation of certain generation units to be small. Another application would be to constrain the variation of line flows across areas to provide a more certain power exchange.

<sup>13</sup> It is straightforward to modify Problem (4.47) to consider time-varying and quantity-dependent risk levels  $\varepsilon$ .

We can view Problem 4.5 as a mapping under uncertainty. Specifically, the bus parameterizations (4.42) are mapped to the optimal random-variable power injections

$$"p_1^{c*}(t), \dots, p_{N_b}^{c*}(t), p_1^{s*}(t), \dots, p_{N_b}^{s*}(t) = \operatorname{argmin} \text{ Problem 4.5}(x_1, \dots, x_{N_b}),"$$

from which we can then determine the stochastic processes for the voltage angles via the DC power flow conditions (4.39a).

### 4.3.2 Existing approaches

Operating a system optimally subject to DC power flow conditions and uncertainties has been and continues to be an active area of research. Affine control policies<sup>14</sup> and scenario-based approaches seem predominant [5, 48, 76, 128, 135, 162–164, 166, 170]. What is common to all the listed works is that the choice of affine policies is motivated mostly by engineering intuition rather than mathematical rigor. Also, the mathematical treatment of uncertainties as discrete stochastic processes is not made explicit; the use of scenarios is rather ad-hoc. In other words, the abstract problem formulation according to Problem 4.5 is often bypassed in favor of a formulation that has a specific application in mind.

Our presented formulation according to Problem 4.5 is inspired mostly by the collection of works [162–164] as well as [166]. A notable difference to [162–164] is that we consider single chance constraints instead of joint chance constraints that are then reformulated deterministically via scenario-based approaches. And in contrast to [166] we consider not just line flow limits but also constraints for generation, storage injection, and the storage state.

To summarize: there is a rich suite of existing works, but these approaches do not present the problem formulation in terms of random variables as done in Problem 4.5. Explicitly formulating the problem with random variables makes it clear that we are—once again—dealing with an infinite-dimensional problem. The question is how to render it finite-dimensional; we tackle this

<sup>14</sup> Sometimes also referred to as linear decision rules [66].

by deriving explicit formulæ for the first two moments that allow for a structured problem formulation.

### 4.3.3 Solution methodology

We are in the familiar situation from Section 4.1.3 and Section 4.2.3: having formulated the chance-constrained multi-stage DC-OPF problem the same challenges c4.1-c4.4 appear; in addition, there is the time dependency.

We tackle c4.1 by restricting our attention to second-order stochastic processes that are fully described by their first and second moments. This allows to rewrite the infinite-dimensional Problem 4.5 as a finite-dimensional problem in terms of the moments—an approach that is strongly motivated by our previous use of PCE. Let us specify the discrete stochastic processes we are dealing with.

**Assumption 4.7** (Bus parameterizations as discrete stochastic process). *The bus parameterizations from Problem 4.5 are modeled by a discrete second-order stochastic process (4.42). Specifically, we model the stochastic process by*

$$\begin{bmatrix} p_i^u(1) \\ p_i^u(2) \\ \vdots \\ p_i^u(T) \end{bmatrix} = \underbrace{\begin{bmatrix} [\hat{p}_i^u]_1 \\ [\hat{p}_i^u]_2 \\ \vdots \\ [\hat{p}_i^u]_T \end{bmatrix}}_{=: \hat{p}_i^u} + \underbrace{\begin{bmatrix} [P_i^u]_{11} & 0 & \dots & 0 \\ [P_i^u]_{21} & [P_i^u]_{22} & & 0 \\ \vdots & & \ddots & \\ [P_i^u]_{T1} & & & [P_i^u]_{TT} \end{bmatrix}}_{=: P_i^u} \underbrace{\begin{bmatrix} [\Xi_i]_1 \\ [\Xi_i]_2 \\ \vdots \\ [\Xi_i]_T \end{bmatrix}}_{=: \Xi_i} \quad (4.48a)$$

$$= \hat{p}_i^u + P_i^u \Xi_i \quad (4.48b)$$

for all buses  $i \in \mathcal{N}$ , where  $\hat{p}_i^u \in \mathbb{R}^T$ , and  $P_i^u \in \mathbb{R}^{T \times T}$  is lower-triangular and non-singular.<sup>15</sup> The so-called stochastic germ  $\Xi_i \in L^2(\Omega, \mu; \mathbb{R}^T)$ , see Re-

<sup>15</sup> We use this dedicated bracket notation to refer to elements. This simplifies to distinguish the bus index  $i$  from the element indices.

mark 2.4 is an  $\mathbb{R}^T$ -valued random vector composed of independent random variables with mean zero and unit variance,

$$\forall i, j \in \mathcal{N}: \quad \mathbb{E}([\Xi_i]_k [\Xi_j]_l) = \delta_{ij} \delta_{kl} \quad \forall k, l \in \mathcal{T}, \quad (4.48c)$$

where  $\delta_{ab}$  is the Kronecker-delta.<sup>16</sup> Specifically we assume that  $\Xi_i$  is a  $T$ -variate Gaussian random vector. The uncertain disturbance (4.48) is then fully described by its mean  $\mathbb{E}(\mathbf{p}_i^u(t))$  and variance  $\mathbb{V}(\mathbf{p}_i^u(t))$

$$\mathbb{E}(\mathbf{p}_i^u(t)) = [\hat{p}_i^u]_t, \quad \mathbb{V}(\mathbf{p}_i^u(t)) = \sum_{k=1}^t [P_i^u]_{tk}^2 \quad (4.48d)$$

for all time instants  $t \in \mathcal{T}$ .

Note that the lower-triangularity of  $P_i^u$  means that the uncertain disturbance  $\mathbf{p}_i^u(t)$  is causal, i.e.

$$\mathbf{p}_i^u(t) = [\hat{p}_i^u]_t + \sum_{k=1}^t [P_i^u]_{tk} [\Xi_i]_k \quad (4.49)$$

depends only on past and present time instants  $k \in \{1, \dots, t\}$  for all  $t \in \mathcal{T}$ , but not on future ones.

**Remark 4.3** (Non-Gaussian uncertainties). *Condition (4.48c) may be satisfied by other (centralized) random variables too. Hence, it is restrictive to consider only Gaussian uncertainties in Assumption 4.7. This specific restriction is made with a user background in mind: the full parameterization (4.48a) must be provided as an input to the proposed method. For Gaussian uncertainties, such a parameterization can be obtained, for instance, via Gaussian process regression [137, Ch. 2]. For other uncertainties that satisfy the technical condition (4.48c) the parameterization (4.48a) may be more difficult to obtain.*

From Assumption 4.7 we see that the stochastic germ  $\Xi_i$  parameterizes each uncertain disturbance affinely. From [117] it is known that affine parameter-

<sup>16</sup> Notice that non-singularity of  $P_i^u$  means that (4.48a) is a one-to-one mapping between  $[\mathbf{p}_i^u(1), \dots, \mathbf{p}_i^u(T)]^\top$  and the stochastic germ  $\Xi_i$ . The lower-triangularity of  $P_i^u$  allows to create this mapping first for time instant  $t = 1$ , then  $t = 2$ , etc.

izations of the disturbance can be accounted for by affine parameterizations of the inputs to satisfy DC power balance. Hence, we introduce a discrete stochastic process of the form

$$L^2(\Omega, \mu; \mathbb{R}^T) \ni \begin{bmatrix} \mathbf{p}_i^c(1) \\ \vdots \\ \mathbf{p}_i^c(T) \end{bmatrix} = \hat{p}_i^c + \sum_{j \in \mathcal{N}} P_{i,j}^c \Xi_j, \quad \forall i \in \mathcal{N}, \quad (4.50a)$$

where  $\hat{p}_i^c \in \mathbb{R}^T$ , and every  $P_{i,j}^c \in \mathbb{R}^{T \times T}$  for all  $j \in \mathcal{N}$  is lower-triangular; notice that the notation is structurally equivalent to (4.48). Again, the requirement of lower-triangularity enforces the stochastic process to be causal. Explicitly, the components of (4.50a) are

$$\mathbf{p}_i^c(t) = [\hat{p}_i^c]_t + \sum_{j \in \mathcal{N}} \sum_{k=1}^t [P_{i,j}^c]_{tk} [\Xi_j]_k \quad \forall i \in \mathcal{N}, \quad (4.50b)$$

for all time instants  $t \in \mathcal{T}$ .<sup>17</sup>

**Remark 4.4** (Flexibility of policies). *The structure of the matrices  $P_{i,j}^c$  determines the “character” of the generation policy. For example, the generic policy from (4.50b) allows the generator at bus  $i$  to react to every possible source of uncertainty  $\Xi_j$ . In turn, if the generator at bus  $i$  is allowed to react only to the sum of the deviations, then we would set  $P_{i,1}^c = \dots = P_{i,N_b}^c$ . Also, according to [166] “[e]xisting reserve mechanisms could be modeled by matrices  $[P_{i,j}^c]$  for which only the main diagonal is populated.” Clearly, combinations of all of the above are possible.*

We introduce the same kind of policy (4.50a) for the storages

$$L^2(\Omega, \mu; \mathbb{R}^T) \ni \begin{bmatrix} \mathbf{p}_i^s(1) \\ \vdots \\ \mathbf{p}_i^s(T) \end{bmatrix} = \hat{p}_i^s + \sum_{j \in \mathcal{N}} P_{i,j}^s \Xi_j, \quad (4.51)$$

<sup>17</sup> Notice that the control policy (4.50a) is written in terms of the stochastic germs  $\Xi_j$  for  $j \in \mathcal{N}$ ; but in practice it is the realization of the uncertain disturbances  $\mathbf{p}_i^c(t)$  that can be measured. In our setting it is always possible to get the realization of the stochastic germ from the realization of the uncertain disturbance, and vice versa, see Footnote 16.

where every  $P_{i,j}^s \in \mathbb{R}^{T \times T}$  with  $j \in \mathcal{N}$  is lower-triangular.

Having introduced the stochastic processes for the uncertainty model (4.48), the generation (4.50b), and the storage (4.51), this allows to derive closed-form expressions for the line flow  $\mathbf{p}_j^{\text{br}}(t)$  at line  $j \in \mathcal{L}^{\text{lin}}$ , the change of inputs  $\Delta \mathbf{p}_i^c(t) = \mathbf{p}_i^c(t) - \mathbf{p}_i^c(t-1)$  at bus  $i \in \mathcal{N}$ , and the state of the storage  $\mathbf{e}_i(t+1)$  bus  $i$ —all of which are required to solve Problem (4.47). These closed-form expressions are listed in (4.54). All random variables, namely the controls ( $\mathbf{p}_i^c(t)$ ,  $\mathbf{p}_i^s(t)$ ) and the dependent quantities ( $\mathbf{e}_i(t+1)$ ,  $\Delta \mathbf{p}_i^c(t)$ ,  $\mathbf{p}_j^{\text{br}}(t)$ ) are Gaussian random variables for all respective time instants. This follows from linearity/affinity of all equations (uncertainty model (4.49), power balance (4.23), PTDF matrix, storage dynamics (4.43), control policies (4.50b), (4.51)), in combination with the fact that a Gaussian random variable remains Gaussian under affine mappings [83]. Hence, all stochastic processes from Problem (4.47) are fully described by their mean and variance, the closed-form expressions for which are listed in (4.55). This completes the solution to challenge c4.1.

Next, we address challenge c4.2: probabilistic power flow. Consider the power balance (4.47c), and substitute the uncertainty model (4.48) and the process for generation/storage control (4.50)/(4.51). Then, the power balance is satisfied for all realizations if [122]

$$\sum_{i \in \mathcal{N}} \hat{p}_i^u + \hat{p}_i^c + \hat{p}_i^s = 0_T, \quad (4.52a)$$

$$P_j^u + \sum_{i \in \mathcal{N}} P_{i,j}^c + P_{i,j}^s = 0_{T \times T}, \quad \forall j \in \mathcal{N}. \quad (4.52b)$$

The reformulation of the cost, i.e. challenge c4.3, is analogous to Section 4.1.3.

The final challenge c4.4, the reformulation of the chance constraints, becomes straightforward for our setting. As all random variables occurring in Problem (4.47) are Gaussian random variables, the chance constraints can be reformulated exactly using the first two moments: Let  $x$  be a Gaussian random variable with mean  $\mu$  and variance  $\sigma^2$ . Then for  $\varepsilon \in [0.5, 1)$ ,

$$\mathbb{P}(x \leq x^{\max}) \geq 1 - \varepsilon \iff \mu + \lambda(\varepsilon)\sqrt{\sigma^2} \leq x^{\max}, \quad (4.53a)$$

$$\mathbb{P}(x^{\min} \leq x) \geq 1 - \varepsilon \iff x^{\min} \leq \mu - \lambda(\varepsilon)\sqrt{\sigma^2}, \quad (4.53b)$$

where  $\lambda(\varepsilon) = \Phi^{-1}(1-\varepsilon)$ , and  $\Phi$  is the cumulative distribution function of a standard Gaussian random variable [20], see also Remark 4.2. Hence, all chance constraints from Problem (4.47) can be reformulated by applying relation (4.53) with the moments from (4.55).

Having addressed all the challenges C4.1-C4.4 we are ready to cast Problem 4.5 as a finite-dimensional optimization problem.

**Problem 4.6** (Reformulated chance-constrained multi-stage DC-OPF). *Consider Problem 4.5 and let Assumption 4.7 hold. For a moment-based reformulation of the chance constraints, the finite-dimensional reformulation of Problem 4.5 reads*

$$\min_{\substack{\hat{p}_i^c, \hat{p}_i^s \in \mathbb{R}^T, \\ P_{i,j}^c, P_{i,j}^s \in \mathbb{R}^{T \times T} \\ \forall i, j \in \mathcal{N}}} \sum_{t \in \mathcal{T}} \sum_{i \in \mathcal{N}} c_i(\mathbb{E}(u_i(t))) + c_{i,2} \mathbb{V}(\mathbf{p}_i^c(t)) \quad (4.56a)$$

subject to

$$\sum_{i \in \mathcal{N}} \hat{p}_i^u + \hat{p}_i^c + \hat{p}_i^s = 0_T, \quad (4.56b)$$

$$P_j^u + \sum_{i \in \mathcal{N}} P_{i,j}^c + P_{i,j}^s = 0_{T \times T},$$

$$\mathbf{e}_i(t+1) = \{\text{see (4.54)}\}, \quad (4.56c)$$

$$\mathbf{e}_i(1) = \mathbf{e}_i^{\text{IC}}, \quad (4.56d)$$

$$x^{\min} \leq \mathbb{E}(\mathbf{x}) \pm \lambda(\varepsilon) \sqrt{\mathbb{V}(\mathbf{x})} \leq x^{\max}, \quad (4.56e)$$

$$\forall \mathbf{x} \in \{\mathbf{p}_i^{\text{br}}(t), \mathbf{p}_i^c(t), \Delta \mathbf{p}_i^c(\tau), \mathbf{e}_i(t+1), \mathbf{e}_i(T+1), \mathbf{p}_i^s(t)\}$$

$$\forall i, j \in \mathcal{N}, t \in \mathcal{T}, \tau \in \mathcal{T} \setminus \{1\}, l \in \mathcal{L}^{\text{lin}}.$$

## 4.3.4 Discussion

### Computational characteristics

The reformulated Problem 4.5 is a second-order cone program (SOCP), hence all contemplations from Section 4.2.4 apply. It is generally advisable to introduce the minimum number of scalar decision variables in the SOCP (4.56). Specifically, assume that the uncertain disturbances are connected to buses

Closed-form expressions for line flows  $\mathbf{p}_i^{\text{br}}(t)$ , change of inputs  $\Delta \mathbf{p}_i^c(t)$ , and state of storage  $\mathbf{e}_i(t+1)$ .

$$\mathbf{p}_i^{\text{br}}(t) = \sum_{i \in \mathcal{N}} [\Psi]_{i,i}([\hat{p}_i^{\text{u}}]_t + [\hat{p}_i^c]_t) + \sum_{i \in \mathcal{N}} \sum_{k=1}^t \left( [\Psi]_{i,i} [P_i^{\text{u}}]_{tk} + \sum_{j \in \mathcal{N}} [\Psi]_{ij} ([P_{j,i}^c]_{tk} + [P_{j,i}^s]_{tk}) \right) [\Xi_i]_k \quad (4.54a)$$

$$\Delta \mathbf{p}_i^c(t) = [\hat{p}_i^c]_t - [\hat{p}_i^c]_{t-1} + \sum_{j \in \mathcal{N}} ([P_{i,j}^c]_{tt} [\Xi_j]_t + \sum_{k=1}^{t-1} ([P_{i,j}^c]_{tk} - [P_{i,j}^c]_{(t-1)k}) [\Xi_j]_k) \quad (4.54b)$$

$$\mathbf{e}_i(t+1) = \mathbf{e}_i^{\text{u}} - h \sum_{j \in \mathcal{N}} [\hat{p}_j^c]_k - h \sum_{j \in \mathcal{N}} \sum_{k=1}^t \left( \sum_{l=k}^t [P_{i,j}^c]_{lk} \right) [\Xi_j]_k \quad (4.54c)$$

Expected value and variance of random variables from Problem (4.47) under affine policies.

$$\mathbb{E}(\mathbf{p}_i^{\text{u}}(t)) = [\hat{p}_i^{\text{u}}]_t \quad \mathbb{V}(\mathbf{p}_i^{\text{u}}(t)) = \sum_{k=1}^t [P_{i,i}^{\text{u}}]_{tk}^2 \quad (4.55a)$$

$$\mathbb{E}(\mathbf{p}_i^c(t)) = [\hat{p}_i^c]_t \quad \mathbb{V}(\mathbf{p}_i^c(t)) = \sum_{j \in \mathcal{N}} \sum_{k=1}^t [P_{i,j}^c]_{tk}^2 \quad (4.55b)$$

$$\mathbb{E}(\mathbf{p}_i^s(t)) = [\hat{p}_i^s]_t \quad \mathbb{V}(\mathbf{p}_i^s(t)) = \sum_{j \in \mathcal{N}} \sum_{k=1}^t [P_{i,j}^s]_{tk}^2 \quad (4.55c)$$

$$\mathbb{E}(\mathbf{p}_i^{\text{br}}(t)) = \sum_{i \in \mathcal{N}} [\Psi]_{it} ([\hat{p}_i^{\text{u}}]_t + [\hat{p}_i^c]_t + [\hat{p}_i^s]_t) \quad \mathbb{V}(\mathbf{p}_i^{\text{br}}(t)) = \sum_{i \in \mathcal{N}} \sum_{k=1}^t \left( [\Psi]_{it} [P_{i,i}^{\text{u}}]_{tk} + \sum_{j \in \mathcal{N}} [\Psi]_{ij} ([P_{j,i}^c]_{tk} + ([P_{j,i}^s]_{tk})^2) \right) \quad (4.55d)$$

$$\mathbb{E}(\Delta \mathbf{p}_i^c(t)) = [\hat{p}_i^c]_t - [\hat{p}_i^c]_{t-1} \quad \mathbb{V}(\Delta \mathbf{p}_i^c(t)) = \sum_{i \in \mathcal{N}} \left( [P_{i,i}^c]_{tt}^2 + \sum_{k=1}^{t-1} ([P_{i,i}^c]_{tk} - [P_{i,i}^c]_{(t-1)k})^2 \right) \quad (4.55e)$$

$$\mathbb{E}(\mathbf{e}_i(t+1)) = \mathbb{E}(\mathbf{e}_i^{\text{u}}) - h \sum_{k=1}^t [\hat{p}_k^c]_k \quad \mathbb{V}(\mathbf{e}_i(t+1)) = \mathbb{V}(\mathbf{e}_i^{\text{u}}) + h^2 \sum_{j \in \mathcal{N}} \sum_{k=1}^t \left( \sum_{l=k}^t [P_{i,j}^c]_{lk} \right)^2 \quad (4.55f)$$

$\mathcal{N}_u \subseteq \mathcal{N}$  with  $|\mathcal{N}_u| = N_u$ , that the generators are connected to buses  $\mathcal{N}_c \subseteq \mathcal{N}$  with  $|\mathcal{N}_c| = N_c$ , and that the storages are connected to buses  $\mathcal{N}_s \subseteq \mathcal{N}$  with  $|\mathcal{N}_s| = N_s$ , where  $\mathcal{N}_c \cap \mathcal{N}_s = \emptyset$ , i.e. no bus has a generator *and* a storage. Then, the number of scalar decision variables in the SOCP (4.56) is

$$(N_c + N_s) \left( T + N_u \frac{T(T+1)}{2} \right) \quad (4.57)$$

for the generation/storage policies (4.50)/(4.51); in contrast to [166] we exploit lower-triangularity of the matrices  $P_{i,j}^c, P_{i,j}^s$ . For a global balancing policy for both generation and storage, see Remark 4.4, the number of scalar decision variables reduces to

$$(N_c + N_s) \left( T + \frac{T(T+1)}{2} \right), \quad (4.58)$$

hence it is independent of the number of uncertainties in the grid. The difference between the numbers (4.57) and (4.58) reflects the usual trade-off between computational tractability and complexity of the solution.

### Interpretation of solution

The considerations from Section 4.2.4 apply: having obtained the optimal PCE coefficients from Problem 4.6, we can construct the overall grid state in terms of random variables. According to Remark 2.9 this solution can either be used in terms of a policy that maps realizations of the uncertainty to realizations of the decision variables. Or it may be used to deduce statistics of the random-variable solutions.

### Truncation errors

The considerations from Section 4.2.4 apply: for the DC power flow equations the truncation error can be brought down to zero. In other words, the energy balance can be satisfied for all realizations of the uncertainty. This result is formalized in [20, 117], independently of the underlying distribution. Also, the moment-based reformulation of the chance constraints is exact, hence the PCE re-formulation is exact [121].

## 5 PolyChaos.jl

Thus far we presented problem formulations of (optimal) power flow problems under uncertainty based on intrusive pCE. Tracing back every problem formulation we presented in either Chapter 3 or Chapter 4 we see that there appear pCE-specific numerical ingredients that until now we assumed given; Table 5.1 lists all of them. We summarize the requirements from Table 5.1:

1. Given a probability density (or, more generally, an absolutely continuous, non-negative measure  $d\mu(\tau) = \rho(\tau)d\tau$ ), what are the respective orthogonal polynomials  $\phi_k$ ?
2. How can one compute the numbers

$$\langle \phi_{i_1} \phi_{i_2} \cdots \phi_{i_{m-1}}, \phi_{i_m} \rangle = \int \phi_{i_1}(\tau) \phi_{i_2}(\tau) \cdots \phi_{i_m}(\tau) d\mu(\tau) \quad (5.1)$$

for  $m \in \{2, 3, 4\}$  efficiently and accurately?

3. How to do both of the above in a multivariate setting?

Clearly, we would appreciate to have a software answer these questions for us: that is why the author of this dissertation co-created the package *Poly-*

Table 5.1: pCE-related quantities for (optimal) power flow problems under uncertainty.

Name	Problem	pCE-related quantities			
PPF	3.2	$\phi_k$	$\langle \phi_k, \phi_k \rangle$	$\langle \phi_{k_1} \phi_{k_2}, \phi_{k_3} \rangle$	
BFS	3.4	$\phi_k$	$\langle \phi_k, \phi_k \rangle$	$\langle \phi_{k_1} \phi_{k_2}, \phi_{k_3} \rangle$	
Optimal adaptive linearization	3.6	$\phi_k$	$\langle \phi_k, \phi_k \rangle$	$\langle \phi_{k_1} \phi_{k_2}, \phi_{k_3} \rangle$	
CC-AC-OPF	4.2	$\phi_k$	$\langle \phi_k, \phi_k \rangle$	$\langle \phi_{k_1} \phi_{k_2}, \phi_{k_3} \rangle$	$\langle \phi_{k_1} \phi_{k_2} \phi_{k_3}, \phi_{k_4} \rangle$
CC-DC-OPF	4.4	$\phi_k$	$\langle \phi_k, \phi_k \rangle$		

*Chaos.jl* written in the Julia programming language that provides a user-friendly interface for making computations with PCE [127].<sup>1</sup> *PolyChaos.jl* is a collection of numerical routines for orthogonal polynomials, quadrature rules, and (intrusive) polynomial chaos expansions. It is built around the above questions: allow for arbitrary probability densities (or product measures) in an intrusive PCE setting, and compute and store arrays of scalar products of the basis polynomials. *PolyChaos.jl* allows

- to compute the coefficients for the monic three-term recurrence relation numerically by the Lanczos or Stieltjes procedure, or multiple discretization [64], see Section 2.4.1 or Appendix A.1,
- to evaluate the orthogonal polynomials at arbitrary points,
- to compute the quadrature rule (Gauss, Gauss-Radau, Gauss-Lobatto, see Appendix A.2; Fejér’s first/second rule, Clenshaw-Curtis, see [165]),
- to compute arrays of scalar products (5.1),
- to do all of the above in a multivariate setting, i.e. product measures.

If the weight function of the measure corresponds to a probability density function, *PolyChaos.jl* further provides routines to compute polynomial chaos expansions of random variables. These routines allow

- to compute affine PCE coefficients for arbitrary densities,
- to compute moments,
- to compute the tensors of scalar products.

Why Julia? Julia is a just-in-time compiled programming language for scientific computing [19]. It provides a command line interface (the so-called Julia REPL) with a built-in package manager, unicode support, and easy access to the shell. Julia is committed to the paradigm of multiple dispatch, the ability of functions to act differently depending on their signatures. Macros and other concepts from metaprogramming are supported, too. Perhaps Julia’s greatest advantage—at least for the scientific programmer—is its abil-

---

<sup>1</sup> The documentation is available at <https://timueh.github.io/PolyChaos.jl/stable/>.

ity to solve the so-called two-language problem. This problem refers to the undesirable situation in which a programmer creates prototypes in one language (often based on easy-to-use and easy-to-read scripts or notebooks, e.g. Matlab or Python), then having to switch to a different language (often compiled, e.g. C/C++) to achieve fast execution times. Julia is a platform for both: rapid prototypes with intuitive code design can be morphed into type-specific, high-performance code. Julia is entirely based on types. The fact that types of values need not be declared explicitly is one reason why Julia solves the aforementioned two-language problem: users may never feel the need to declare types, yielding code reminiscent of scripted languages, yet users *can* leverage the full power and expressiveness of Julia’s type system to write cleaner code. The advantage of type-declared code is that it asserts the code behaves as expected, and it provides information to the user and the compiler which improves readability and performance in some cases.

The base library of Julia is slim by design: among others it provides collections and data structures, mathematical functions, arrays, and interfaces to several foreign languages such as C/C++, Fortran, Python, R, or Java. Additional functionalities are provided by the rich Julia package ecosystem.<sup>2</sup> Every Julia package is hosted open source on GitHub, built and deployed via continuous integration, and referenced in the Julia package manager—which is itself a package written in Julia. There exists, for instance, a diverse suite of solvers for differential equations [136], optimization tools [51, 112], iterative solvers for linear systems,<sup>3</sup> and plotting tools,<sup>4</sup> to name a few. However, there existed no previous Julia implementation that combined routines for orthogonal polynomials, quadrature rules, and polynomial chaos expansions. *PolyChaos.jl* aims to close that gap.

## 5.1 Existing software

Table 5.2 lists existing software packages for PCE. Except for *Chaospy* and *PolyChaos.jl* these packages are full-fledged libraries for uncertainty quan-

---

<sup>2</sup> See <https://juliaobserver.com/packages> for all publicly registered packages.

<sup>3</sup> See <https://github.com/JuliaMath/IterativeSolvers.jl>.

<sup>4</sup> See <https://github.com/JuliaPlots/Plots.jl/>.

tification; PCE comprises just one module of many, and is used mostly for non-intrusive applications. Amongst the software from Table 5.2 *UQLab* and *Dakota* provide the richest functionality, each coming with a superb documentation. While the core functions of *UQLab* are closed source, the scientific methods surrounding PCE are all available under the BSD 3-clause license. Furthermore, *UQLab* provides methods for basis-adaptive PCE based on [24, 25]. *Dakota* is a mature framework: currently at version 6.0, version 3.0 beta, for instance, dates back to 2001. The functionality of *MUQ* and *UQToolkit* is comparable; unfortunately they do not allow to compute orthogonal polynomials for arbitrary probability densities. *OpenTURNS* is a full-fledged uncertainty quantification framework that comes with rich and mathematically detailed documentation. The solely PCE-centered Python package *Chaospy* comes with the least restrictive MIT-license whilst providing the core PCE functionality that includes the computation of orthogonal polynomials for arbitrary probability densities.

Table 5.2 positions *PolyChaos.jl* in the landscape of software packages for PCE: its premise is to support arbitrary probability densities, for which it provides not just the Stieltjes but also the Lanczos procedure based on [64], see Appendix A.1. In case the density can be composed as a sum of individual densities—as is common for (Gaussian) mixture models—*PolyChaos.jl* provides a specific method for multiple discretization based on [64, p. 99]. Furthermore, *PolyChaos.jl* supports a range of quadrature rules. Most importantly though—as *PolyChaos.jl* is currently tailored to the intrusive PCE setting—it allows to compute arrays of scalar products of the form (5.1). These numbers are needed in every problem from Chapter 3 or Chapter 4. Future versions of *PolyChaos.jl* might include non-intrusive PCE, i.e. collocation-based methods [154, 172], or basis-adaptive sparse methods [23].

## 5.2 Type hierarchy

Every value in Julia has a type. The conceptual foundation of the type system relies on *abstract* types. Abstract types serve but a single purpose: to form a type hierarchy. It is neither desired nor possible to instantiate abstract types. The type hierarchy remains independent from functions that operate on types. To get what in other languages is called a struct or an ob-

Table 5.2: Existing software packages for PCE.

Name	Language	Features for PCE	License	Ref.
<i>UQLab</i>	Matlab	- Classic and arbitrary distributions - Stieltjes procedure - Gauss and sparse quadrature - Basis-adaptive sparse PCE - Least-angle regression	BSD 3-clause	[106]
<i>Chaospy</i>	Python	- Classic and arbitrary distributions - Gram-Schmidt, Stieltjes procedure - Gauss quadrature, Clenshaw-Curtis	MIT	[59]
<i>OpenTURNS</i>	Python	- Classic and arbitrary distributions - Stieltjes procedure - Gauss quadrature	GNU LGPL	[16]
<i>Dakota</i>	C++	- Classic and arbitrary distributions - Stieltjes, Gram-Schmidt, Chebyshev - Gauss and sparse quadrature - Stochastic collocation	GNU LGPL	[2]
<i>MUQ</i>	C++, Python	- Classic distributions - Gauss quadrature	n/a	[39]
<i>UQToolkit</i>	C++, Python	- Classic distributions - Gauss quadrature	GNU LGPL	[46]
<i>PolyChaos.jl</i>	Julia	- Classic and arbitrary distributions - Stieltjes and Lanczos procedure - Multiple discretization - Gauss quadrature, Fejér, Clenshaw-Curtis - Scalar products	MIT	[127]

ject Julia provides *composite* types. A composite type has fields,<sup>5</sup> it can be instantiated, and it can be declared a subtype of abstract types.

For *PolyChaos.jl* we devise our own type hierarchy. Figure 5.1 shows the two bread-and-butter type trees we need; abstract types carry the prefix “Abstract.” Take the abstract type *AbstractMeasure*: it has two composite subtypes: *Measure* and *ProductMeasure* with obvious meanings. There exist, however, well-studied canonical measures such as Gaussian or uniform measures for which we introduce the abstract subtype *AbstractCanonicalMeasure*. All subtypes of *AbstractCanonicalMeasure* are shown in Figure 5.1. The

<sup>5</sup> Methods can be fields too. In that case, the type of the field is *Function*.

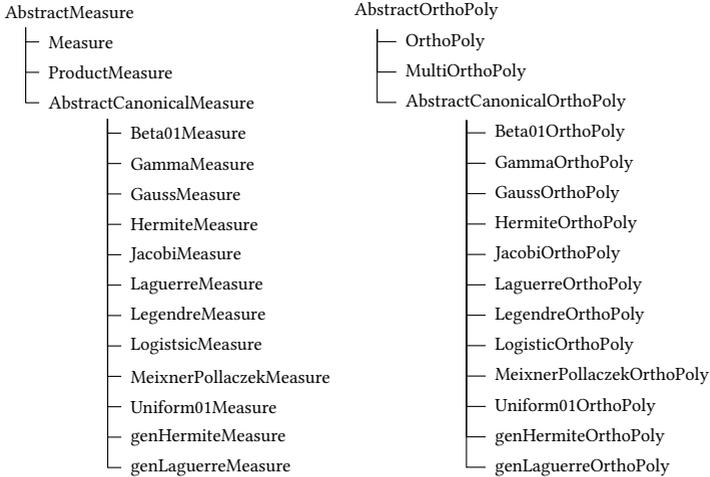


Figure 5.1: Type hierarchy for measures and orthogonal polynomials.



Figure 5.2: Type hierarchy for quadrature rules and tensors.

type hierarchy for orthogonal polynomials mirrors that of measures: there are generic composite types for univariate polynomials, namely *OrthoPoly*, and multivariate polynomials, namely *MultiOrthoPoly*, and there are canonical orthogonal polynomials. Figure 5.2 adds to the overall *PolyChaos.jl* type system quadrature rules via *AbstractQuad* and tensors of scalar products via *AbstractTensor*. We emphasize once more that the type hierarchies from Figure 5.1 and Figure 5.2 describe a *concept* and no implementation.

Let us declare fields for the composite types.

*AbstractMeasure*—We begin with the composite subtypes of *AbstractMeasure* which we assume to be absolutely continuous measures specified in terms of their domain and weight function; Table 5.3 (page 131) lists their fields. In Julia the operator “*::*” links a type annotation to an expression. For instance *name::String* means that the value of the variable *name* is of the type *String*.

The composite type *Measure* has a *name* and a weight function  $w : \Omega \subseteq \mathbb{R} \rightarrow \mathbb{R}$  with domain  $\Omega$  (*dom*).<sup>6</sup> If the weight function is symmetric relative to some  $m \in \Omega$ , the field *symmetric* is set to *true*; symmetry relative to  $m$  means that

$$\forall x \in \Omega : w(m - x) = w(m + x). \quad (5.2)$$

For example, the probability density of a Gaussian random variable is symmetric relative to the origin. If the weight function has any parameters, then they are stored in the dictionary *pars*. All subtypes of *AbstractCanonicalMeasure* follow that pattern, and all fields are conveniently pre-set. The composite type *ProductMeasure* has a weight function which is the product of the weight functions of the elements of *measures*, a vector whose elements are subtypes of *AbstractMeasure*.

*AbstractOrthoPoly*—Given an absolutely continuous measure in terms of its weight (respectively Lebesgue density) what are the monic polynomials  $\phi_i : \Omega \rightarrow \mathbb{R}$  that are orthogonal relative to this very measure? The subtypes of *AbstractOrthoPoly* store the system of monic orthogonal polynomials in terms of the three-term recurrence-coefficients, see Appendix A.1. We use the composite types from Table 5.3 as fields of the subtypes of *AbstractOrthoPoly*. Table 5.4 shows the fields of the composite types *OrthoPoly*, *MultiOrthoPoly*, and all the subtypes of *AbstractCanonicalOrthoPoly*. The purpose of *name* is obvious. The integer *deg* stands for the maximum degree of the polynomials. Rather than storing the polynomials  $\phi_i$  themselves, we store the recurrence coefficients  $\alpha$ ,  $\beta$  that characterize the system of orthogonal polynomials. These recurrence coefficients are the single most important piece of information for the orthogonal polynomials [64]. For several common measures, there exist analytic formulæ. These are built-in to *PolyChaos.jl* and should be preferred to numerical approximations thereof. Nonetheless, *PolyChaos.jl* can construct the recurrence coefficients for arbitrary densities via the Stieltjes or Lanczos procedure, see Appendix A.1. We use the vectors of recurrence coefficients to model the orthogonal polynomials, see Appendix A.1. Also, each orthogonal polynomial has an underlying measure. Notice how we use the specific canonical measure for the respective canonical orthogonal polynomial. A quadrature may be assigned in the field *quad*. By default, a Gauss

<sup>6</sup> The weight  $w$  corresponds to the Lebesgue density  $\rho$  from Definition 2.4.

quadrature rule is computed based on the recurrence coefficients, see Appendix A.1.

*AbstractQuad*—What are the fields of the composite type *Quad* from Figure 5.2? These are straightforward, see Table 5.3: the composite type *Quad* carries a name, the number of quadrature points, and the nodes and weights that are vectors of reals. The type *EmptyQuad* is assigned whenever we wish to assign no quadrature rule.

*AbstractTensor*—The composite type *Tensor* is used to store the results of scalar products (5.1). The “dimension”  $m$  of the tensor is the number of terms that appear in the scalar product. Let’s assume we set  $m = 3$ , hence (5.1) becomes  $\langle \phi_{i_1} \phi_{i_2}, \phi_{i_3} \rangle$ , then the concrete entry is obtained as  $Tensor.get([i_1, j_2, k_3])$ . For computing and storing the results in *Tensor* we rely on results from [17]. Specifically, we exploit the commutative property of the scalar products. For example for the scalar products with  $m = 3$  we have

$$\begin{aligned} \langle \phi_{i_1}, \phi_{i_2} \phi_{i_3} \rangle &= \langle \phi_{i_1}, \phi_{i_3} \phi_{i_2} \rangle = \langle \phi_{i_2}, \phi_{i_1} \phi_{i_3} \rangle \\ &= \langle \phi_{i_2}, \phi_{i_3} \phi_{i_1} \rangle = \langle \phi_{i_3}, \phi_{i_2} \phi_{i_1} \rangle = \langle \phi_{i_3}, \phi_{i_1} \phi_{i_2} \rangle \end{aligned} \quad (5.3)$$

for all  $i, i_2, i_3 \in \mathcal{K} = \{0, 1, \dots, \hat{k}\}$ . This yields a total of

$$\binom{\hat{k} + m}{m}$$

distinct computations.<sup>7</sup> The distinct computations of  $\langle \phi_{i_1} \phi_{i_2}, \phi_{i_3} \rangle$  exploit Fubini’s theorem and symmetry of the probability density functions to detect trivial zeros. All distinct scalar products are stored in sparse arrays. The  $j^{\text{th}}$  entry  $T_j$  of this array holds the entry

$$T_j = \langle \phi_{i_1}, \phi_{i_2} \cdots \phi_{i_m} \rangle,$$

with

$$j = 1 + \sum_{n=1}^m i_n \hat{k}^{m-n}. \quad (5.4)$$

---

<sup>7</sup> A naïve computation would yield  $(\hat{k} + 1)^m$  distinct computations.

With (5.4) we introduce a system of numbers to the basis  $\hat{k}$  to which we assign linear one-based indexing.

Table 5.3: Composite subtypes of *AbstractMeasure* and their fields.

<b>Measure</b>	
<i>name</i> ::String	Name of measure
<i>w</i> ::Function	Weight function $w : \Omega \rightarrow \mathbb{R}$
<i>dom</i> ::Tuple{<:Real,<:Real}	Domain $\Omega$
<i>symmetric</i> ::Bool	Is $w$ symmetric relative to some $m \in \Omega$ , hence $w(m - x) = w(m + x)$ for all $x \in \Omega$ ?
<i>pars</i> ::Dict	Additional parameters
<b>ProductMeasure</b>	
<i>w</i> ::Function	Weight function
<i>measures</i> ::Vector{<:AbstractMeasure}	Vector of univariate measures
<b>Beta01Measure</b>	
<i>w</i> ::Function	$\frac{1}{B(\alpha, \beta)} t^{\alpha-1} (1-t)^{\beta-1}$
<i>dom</i> ::Tuple{<:Real,<:Real}	(0, 1)
<i>symmetric</i> ::Bool	true if $\alpha = \beta$
<i>ashapeParameter</i> ::Real	$\alpha > 0$
<i>bshapeParameter</i> ::Real	$\beta > 0$
<b>GammaMeasure</b>	
<i>w</i> ::Function	$\frac{\beta^\alpha}{\Gamma(\alpha)} t^{\alpha-1} \exp(-\beta t)$
<i>dom</i> ::Tuple{<:Real,<:Real}	(0, $\infty$ )
<i>symmetric</i> ::Bool	false
<i>shapeParameter</i> ::Real	$\alpha > 0$
<i>rateParameter</i> ::Real	1
<b>GaussMeasure</b>	
<i>w</i> ::Function	$\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right)$
<i>dom</i> ::Tuple{<:Real,<:Real}	( $-\infty$ , $\infty$ )
<i>symmetric</i> ::Bool	true
<b>HermiteMeasure</b>	
<i>w</i> ::Function	$\exp(-t^2)$
<i>dom</i> ::Tuple{<:Real,<:Real}	( $-\infty$ , $\infty$ )
<i>symmetric</i> ::Bool	true
<b>JacobiMeasure</b>	
<i>w</i> ::Function	$(1-t)^\alpha (1+t)^\beta$
<i>dom</i> ::Tuple{<:Real,<:Real}	(-1, 1)
<i>symmetric</i> ::Bool	true if $\alpha = \beta$

<i>ashapeParameter::Real</i>	$\alpha > -1$
<i>bshapeParameter::Real</i>	$\beta > -1$
<b>LaguerreMeasure</b>	
<i>w::Function</i>	$\exp(-t)$
<i>dom::Tuple{&lt;:Real,&lt;:Real}</i>	$(0, \infty)$
<i>symmetric::Bool</i>	true
<b>LegendreMeasure</b>	
<i>w::Function</i>	1
<i>dom::Tuple{&lt;:Real,&lt;:Real}</i>	$(-1, 1)$
<i>symmetric::Bool</i>	true
<b>LogisticMeasure</b>	
<i>w::Function</i>	$\frac{\exp(-t)}{(1+\exp(-t))^2}$
<i>dom::Tuple{&lt;:Real,&lt;:Real}</i>	$(-\infty, \infty)$
<i>symmetric::Bool</i>	true
<b>MeixnerPollaczekMeasure</b>	
<i>w::Function</i>	$\frac{1}{2\pi} \exp((2\phi - \pi)t)  \Gamma(\lambda + it) ^2$
<i>dom::Tuple{&lt;:Real,&lt;:Real}</i>	$(-\infty, \infty)$
<i>symmetric::Bool</i>	false
<i>lambdaParameter::Real</i>	$\lambda > 0$
<i>phiParameter::Real</i>	$0 < \phi < \pi$
<b>Uniform01Measure</b>	
<i>w::Function</i>	1
<i>dom::Tuple{&lt;:Real,&lt;:Real}</i>	$(0, 1)$
<i>symmetric::Bool</i>	true
<b>genHermiteMeasure</b>	
<i>w::Function</i>	$ t ^{2\mu} \exp(-t^2)$
<i>dom::Tuple{&lt;:Real,&lt;:Real}</i>	$(-\infty, \infty)$
<i>symmetric::Bool</i>	true
<i>muParameter::Real</i>	$\mu > -0.5$
<b>genLaguerreMeasure</b>	
<i>w::Function</i>	$t^\alpha \exp(-t)$
<i>dom::Tuple{&lt;:Real,&lt;:Real}</i>	$(0, \infty)$
<i>symmetric::Bool</i>	false
<i>shapeParameter::Real</i>	$\alpha > -1$

---

Table 5.4: Composite subtypes of *AbstractOrthoPoly* and their fields.

<i>OrthoPoly</i>	
<i>name</i> ::String	Name
<i>deg</i> ::Int	Maximum degree
$\alpha$ ::Vector{<-Real}	Vector of recurrence coefficients
$\beta$ ::Vector{<-Real}	Vector of recurrence coefficients
<i>measure</i> ::AbstractMeasure	Underlying measure
<i>quad</i> ::AbstractQuad	Quadrature rule
<i>MultiOrthoPoly</i>	
<i>name</i> ::Vector{String}	Vector of names
<i>deg</i> ::Int	Maximum degree
<i>dim</i> ::Int	Dimension
<i>ind</i> ::Matrix{<-Int}	Array of multi-indices
<i>measure</i> ::ProductMeasure	Underlying product measure
<i>uni</i> ::Vector{<-AbstractOrthoPoly}	Vector of univariate orthogonal polynomials
<i>CanonicalOrthoPoly</i>	
<i>deg</i> ::Int	Maximum degree
$\alpha$ ::Vector{<-Real}	Vector of recurrence coefficients
$\beta$ ::Vector{<-Real}	Vector of recurrence coefficients
<i>measure</i> ::CanonicalMeasure	Underlying canonical measure
<i>quad</i> ::AbstractQuad	Quadrature rule
For all <i>CanonicalOrthoPoly</i> $\in$ { <i>Beta01OrthoPoly</i> , <i>GammaOrthoPoly</i> , <i>GaussOrthoPoly</i> , <i>HermiteOrthoPoly</i> , <i>JacobiOrthoPoly</i> , <i>LaguerreOrthoPoly</i> , <i>LegendreOrthoPoly</i> , <i>LogisticOrthoPoly</i> , <i>MeixnerPollaczekOrthoPoly</i> , <i>Uniform01OrthoPoly</i> , <i>genHermiteOrthoPoly</i> , <i>genLaguerreOrthoPoly</i> }	

Table 5.5: Composite subtypes of *AbstractQuad* and their fields.

<i>Quad</i>	
<i>name</i> ::String	Name
<i>Nquad</i> ::Int	Number of quadrature points
<i>nodes</i> ::Vector{<-Real}	Nodes
<i>weights</i> ::Vector{<-Real}	Weights
<i>EmptyQuad</i>	

Table 5.6: Composite subtype of *AbstractTensor* and its fields.

<i>Tensor</i>	
<i>dim::Int</i>	“Dimension” $m$ of tensor $\langle \phi_{i_1} \phi_{i_2} \cdots \phi_{i_{m-1}}, \phi_{i_m} \rangle$
<i>T::SparseVector{Float64,Int}</i>	Entries of tensor
<i>get::Function</i>	Function to get entries from $T$
<i>op::AbstractOrthoPoly</i>	Underlying univariate orthogonal polynomials

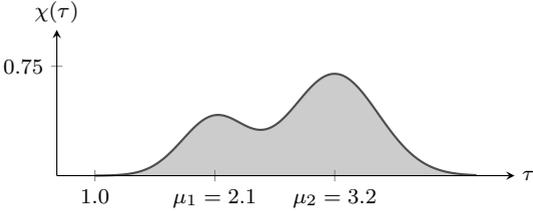


Figure 5.3: Probability density function of Gaussian mixture.

### 5.3 Tutorial example

In order to see *PolyChaos.jl* in action we consider an example.<sup>8</sup> We would like to compute the polynomials that are orthogonal relative to a so-called Gaussian mixture on the real axis

$$\chi(\tau) = \sum_{i=1}^n w_i f(\tau; \mu_i, \sigma_i), \tag{5.5a}$$

where

$$f(\tau; \mu, \sigma) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(\tau - \mu)^2}{2\sigma^2}\right) \tag{5.5b}$$

$$\sum_{i=1}^n w_i = 1, \quad w_i > 0. \tag{5.5c}$$

<sup>8</sup> Inspired by [https://timueh.github.io/PolyChaos.jl/stable/gaussian\\_mixture\\_model/](https://timueh.github.io/PolyChaos.jl/stable/gaussian_mixture_model/).

is the probability density of a Gaussian random variable with mean  $\mu$  and standard deviation  $\sigma$ . We use the generic Julia code to define a Gaussian mixture of  $n = 2$  components:

```
f(x, μ, σ) = 1/sqrt(2*π*σ^2)*exp(-(x-μ)^2/(2σ^2))
μ, σ, w = [2.1, 3.2], [0.3, 0.4], [0.3, 0.7]
χ(τ) = sum(w[i]*f(τ, μ[i], σ[i]) for i in 1:length(w))
```

Note how close Julia code is to the mathematical syntax, especially by supporting unicode. A plot of the probability density is shown in Figure 5.3. We use *PolyChaos.jl* to create the *Measure* associated with the density (5.5).

```
using PolyChaos
meas = Measure("myMeas", χ, (-∞, ∞), false, Dict(:μ=>μ, :σ=>σ, :w=>w))
```

So far we have just declared and defined the fields; no computations have been done. Let us change that now by creating the family of orthogonal polynomials up to degree 4.

```
deg = 4
op = OrthoPoly("myOP", deg, meas; Nquad=150, Nrec=5deg)
```

The keywords *Nquad* and *Nred* specify how many *Nquad* quadrature points are used internally to compute *Nred* recurrence relation coefficients. The optional keyword *discretization* specifies whether to use the Stieltjes or the Lanczos procedure. What are now the polynomials  $\{\phi_k\}_{k=0}^4$  that are orthogonal relative to the specific density  $\chi$  we defined?

```
showbasis(op, sym="τ", digits=2)
1
τ - 2.87
τ^2 - 5.58τ + 7.38
τ^3 - 8.5τ^2 + 23.2τ - 20.22
τ^4 - 11.33τ^3 + 46.6τ^2 - 82.25τ + 52.42
```

A quadrature rule is added automatically. By default *PolyChaos.jl* uses an  $(Nrec-1)$ -point Gauss quadrature rule based on the recurrence relation coefficients. The nodes and weights can be displayed by calling `nw(opq)`. We can use the quadrature rule to compute scalar products of the form  $\langle \phi_i, \phi_j \rangle$ .

```
T2 = Tensor(2, opq)
[ T2.get([i, j]) for i in 0:deg, j in 0:deg ]
```

```
5x5 Array{Float64, 2}:
 1.0  0.0  0.0  0.0  0.0
 0.0  0.3931  0.0  0.0  0.0
 0.0  0.0  0.179322  0.0  0.0
 0.0  0.0  0.0  0.117193  0.0
 0.0  0.0  0.0  0.0  0.0954813
```

The polynomials are indeed orthogonal, because  $\langle \phi_i, \phi_j \rangle = \gamma_i \delta_{ij}$  holds, where  $\delta_{ij}$  is the Kronecker-delta. The collection of the scalar products  $\langle \phi_i, \phi_j \phi_k \rangle$  can be computed by `Tensor(3, opq)`, and so on.

**Remark 5.1** (Multiple discretization). *We presented but one possibility to generate the orthogonal polynomials. In case the underlying measure can be written as a sum of individual measures for which dedicated quadrature rules exist, [64, p. 99] provides a tailored algorithm. This algorithm—which is implemented in PolyChaos.jl as `mcdiscretization`—allows to solve integrals of the form*

$$\int_a^b g(\tau) \rho(\tau) d\tau = \sum_{i=1}^n \int_{a_i}^{b_i} g_i(\tau) \rho_i(\tau) d\tau.$$

For our specific case of a Gaussian bi-mixture, `mcdiscretization` requires two Gauss-Hermite quadrature rules (with the quadrature weights weighted by the entries in  $w$ ) together with

$$\begin{aligned} n &= 2, \\ a_1 &= a_2 = -\infty, \\ b_1 &= b_2 = \infty, \\ \rho_1(\tau) &= w_1 f(\tau; \mu_1, \sigma_1), \\ \rho_2(\tau) &= w_2 f(\tau; \mu_2, \sigma_2). \end{aligned}$$

## 6 Case studies

How can uncertainties be propagated through (optimal) power flow problems? This was the question permeating Chapter 3 and Chapter 4. Let us now turn to three case studies that show the presented methods in action. We demonstrate PCE-overloaded probabilistic power flow according to Section 3.3 (Problem 3.2), PCE-overloaded optimal power flow according to Section 4.1 (Problem 4.2), and its DC counterpart from Section 4.2 (Problem 4.4). For all problems we study the same grid under the same uncertainty: a non-standard uncertainty in terms of a Gaussian mixture. The focus of our presentation is in line with the main theme from Chapter 2: mappings under uncertainty. We focus on showing the probability density functions of the obtained solutions, see Remark 2.9. The main purpose of our case study is not to drown in details, but to solve different power flow problems under uncertainty in an equivalent setting, thus allowing to show how their solutions differ.

After having specified the power system and the uncertainty in the following Section 6.1, we adhere to the same template for the three case studies:

- construct the orthogonal basis functions;
- derive the PCE coefficients for bus specifications;
- solve the PCE-overloaded problem, and visualize the solution.

Clearly, the scope of our case studies could be extended to further interesting questions: What about exactness of the power flow equations for all realizations of the uncertainties [126]? What about exactness of the moments obtained from PCE [117, 125]? What about larger systems [117, 121]? What about more uncertainties [117, 121]? The interested reader is kindly referred to the respective references [117, 121, 125, 126].

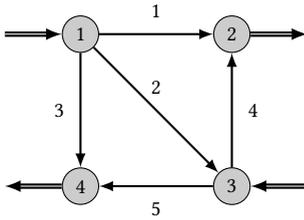


Figure 6.1: 4-bus system inspired by [78].

Table 6.1: Branch parameters in per-unit for 4-bus system from Figure 6.1.

Line	Impedance		Admittance	
	$r$	$x$	$g$	$b$
1 1-2	0.01008	0.0504	3.8156	-19.0781
2 1-3	0.00744	0.0372	5.1696	-25.8478
3 1-4	0.00744	0.0372	5.1696	-25.8478
4 3-2	0.01272	0.0636	3.0237	-15.1185
5 3-4	0.01004	0.0601	2.7129	-16.2127

Base: 100 MVA, 230 kV

The Julia code for this case study is available at <https://github.com/timueh/PowerFlowUnderUncertainty>. All units are normalized to per-unit values; angles are given in radians; the cost function has no physical/monetary unit.

## 6.1 Setup

We study the 4-bus system from Figure 6.1 with generators located at buses 1 and 3, and loads located at buses 2 and 4.<sup>1</sup> We assign bus 1 to be the slack bus; it is controlled to have a unit voltage magnitude and a zero voltage angle. There is no power demand at any of the generators, and conversely there is no power injection at any of the loads. The branch parameters and base values are listed in Table 6.1; their choice is inspired by [78]. We introduce a single source of uncertainty at the load at bus 2: the negative active power  $-p_2$  is assumed to have a PDF equal to the density (5.5) from Section 5.3, i.e.  $\rho_{-p_2}(\tau) = \chi(\tau)$ .<sup>2</sup> Assuming a constant power factor of 0.85 we have that  $q_2 = 0.85p_2$ , hence the reactive power has the same PDF modulo stretching and normalization, i.e.  $\rho_{-q_2}(\tau) = \chi(\tau/0.85)/0.85$ . Figure 6.2 shows the PDF for the uncertain load at bus 2 based on 5,000 samples.<sup>3</sup>

<sup>1</sup> It is the same power system we used to demonstrate PPF (see Figure 3.3) and OPF under uncertainty (see Figure 4.1).

<sup>2</sup> Note that we count power demand negative; it is a sink term.

<sup>3</sup> Although we have an analytic expression for both PDFs we choose to display the PDF created from samples, because we cannot expect to have closed-form expressions for the PDFs for the results of PPF or CC-OPF.

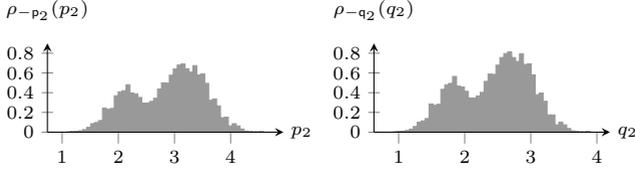


Figure 6.2: probability density function (PDF) of active and reactive power at bus 2.

All deterministic problems are formulated in Julia using JuMP [51]. We choose a flat-start as an initial condition which means setting all voltage phasors to real unity. We present the numerical values of the PCE coefficients for the active and reactive power for both generators, and then plot histograms of the PDFs of quantities of interest based on the 5,000 samples we used to generate the PDFs in Figure 6.2.

## 6.2 Probabilistic power flow

We first need to assign further bus specifications: the generator at bus 4 is a PV bus, and the loads at buses 2 and 4 are PQ buses, hence  $\mathcal{N} = \{1, 2, 3, 4\} = \mathcal{N}_{\text{SL}} \cup \mathcal{N}_{\text{PQ}} \cup \mathcal{N}_{\text{PV}} = \{1\} \cup \{3\} \cup \{2, 4\}$ . The PV bus injects an active power of 0.84 whilst being controlled to a voltage magnitude of 1.04. The load at bus 4 consumes an active power of 1.2, and a reactive power of 1.02; hence the power factor is 0.85. Table 6.2 summarizes the bus specifications and lists the known PDFs; deterministic values are modeled to have a PDF equal to a shifted Dirac-delta.

In order to apply the solution methodology from Problem 3.2 we need to satisfy Assumption 3.3: hence find the orthogonal basis, and determine the PCE coefficients.

### Construction of orthogonal basis

The driving uncertainty is the uncertain load at bus 2, see Table 6.2. Its density follows the Gaussian mixture density from (5.5). For this density we have computed the respective orthogonal polynomials in Section 5.3, hence the ba-

Table 6.2: Uncertainty description for Figure 6.1;  $\delta$  stands for a Dirac-delta pulse.

Bus	Type	Probability density function for all $\tau \in (-\infty, \infty)$			
1	Slack	$\rho_{e_1}(\tau) = \delta(\tau - 1)$	$\rho_{f_1}(\tau) = \delta(\tau)$		
2	PQ	$\rho_{-p_2}(\tau) = \chi(\tau)$	$\rho_{-q_2}(\tau) = \chi(\tau/0.85)/0.85$		
3	PV	$\rho_{p_3}(\tau) = \delta(\tau - 0.84)$	$\rho_{v_3}(\tau) = \delta(\tau - 1.04)$		
4	PQ	$\rho_{-p_4}(\tau) = \delta(\tau - 1.2)$	$\rho_{-q_4}(\tau) = \delta(\tau - 1.02)$		

sis functions  $\{\phi_k\}_{k=0}^{\hat{k}}$  are known. For a maximum total degree of  $d = 4$  the dimension of the orthogonal polynomial basis is  $\hat{k} + 1 = d + 1$ , see (2.25).

### PCE coefficients for bus specifications

Let us continue with the PCE coefficients for the bus specifications. This is straightforward for all deterministic quantities: all coefficients except for the zero-order coefficient are zero. The only non-zero first-order PCE coefficients appear for the uncertain load at bus 2: by construction of the basis the first-order PCE coefficient is unity for the active power.<sup>4</sup> The PCE coefficients of the reactive power  $q_2$  are the ones from the active power  $p_2$  scaled by the power factor of 0.85. The PCE coefficients of all higher-order coefficients  $k \in \{2, \dots, \hat{k}\} = \{2, \dots, 4\}$  are zero. We summarize the PCE coefficients for the bus specifications as follows

$$\begin{bmatrix} \hat{e}_{1,0} & \hat{e}_{1,1} & \hat{e}_{1,k} & \hat{f}_{1,0} & \hat{f}_{1,1} & \hat{f}_{1,k} \\ \hat{p}_{2,0} & \hat{p}_{2,1} & \hat{p}_{2,k} & \hat{q}_{2,0} & \hat{q}_{2,1} & \hat{q}_{2,k} \\ \hat{p}_{3,0} & \hat{p}_{3,1} & \hat{p}_{3,k} & \hat{v}_{3,0} & \hat{v}_{3,1} & \hat{v}_{3,k} \\ \hat{p}_{4,0} & \hat{p}_{4,1} & \hat{p}_{4,k} & \hat{q}_{4,0} & \hat{q}_{4,1} & \hat{q}_{4,k} \end{bmatrix} = \begin{bmatrix} 1.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ -2.87 & -1.00 & 0.00 & -2.44 & -0.85 & 0.00 \\ 0.84 & 0.00 & 0.00 & 1.04 & 0.00 & 0.00 \\ -1.20 & 0.00 & 0.00 & -1.02 & 0.00 & 0.00 \end{bmatrix},$$

where  $k \in \{2, \dots, \hat{k}\}$  are the indices relative to degrees greater than one.

<sup>4</sup> *PolyChaos.jl* provides the function `convert2affinePCE` to compute the zero- and first-order PCE coefficient for a random variable relative to a given orthogonal basis.

Table 6.3: PCE-based solution for generators for PPF.

x	$k^{\text{th}}$ PCE coefficient					Moments	
	0	1	2	3	4	$\mathbb{E}(x)$	$\sqrt{\mathbb{V}(x)}$
$p_1$	3.364	1.077	0.019	0.002	0.00023	3.364	0.676
$q_1$	0.276	0.597	0.053	0.005	0.00064	0.276	0.375
$p_3$	0.84	0.0	0.0	0.0	0.0	0.84	0.0
$q_3$	3.859	0.641	0.042	0.004	0.00055	3.859	0.402

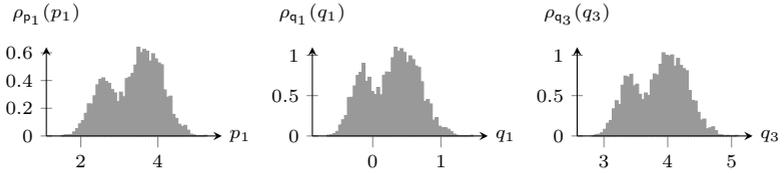
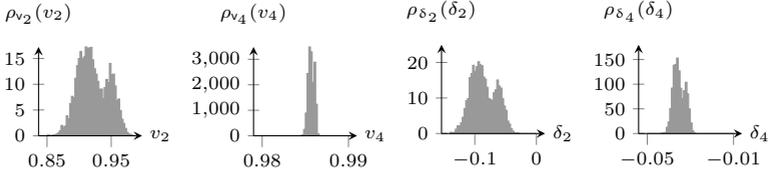
Figure 6.3: PDFs of active power at slack bus 1 and reactive power at buses  $\{1, 4\}$ .

Figure 6.4: (PPF) PDF of voltage magnitude and angle at PV buses.

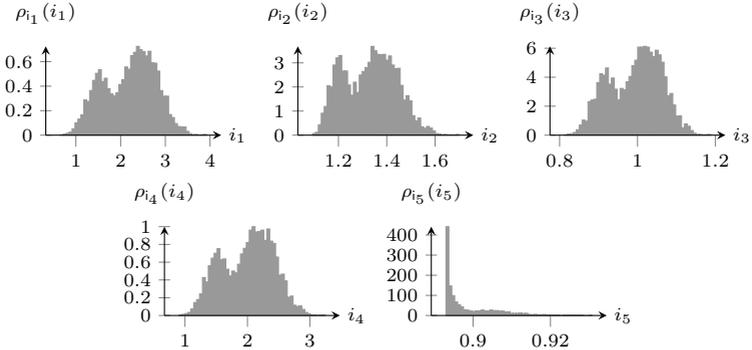


Figure 6.5: (PPF) PDFs of current magnitude for all lines.

## Numerical solution

The PCE-overloaded PPF problem according to Problem 3.2 has a total of  $4N_b(\hat{k} + 1) = 80$  decision variables. Ipopt is able to solve the feasibility problem in about 20 ms. Table 6.3 lists the resulting numerical values of the PCE coefficients for the active and reactive power at the generators: there are non-zero higher-order coefficients due to nonlinearity of the AC power flow equations. Based on the numerical values—especially the moments—we see that bus 1 accounts for the larger share of active power generation while bus 3 takes over the reactive power share. To get a better visual representation of the solution we study the PDFs of quantities of interest based on the 5,000 samples we used to generate the PDFs in Figure 6.2.

Figure 6.3 shows the PDFs of the generated active power at the slack bus, and the generated reactive power at both generation buses.<sup>5</sup> In terms of active power the generator at the slack bus 1 has to accommodate the entire uncertainty originating at the load at bus 2. The reactive power injections appear similar in terms of their distributions. We see that the bi-modal PDF of the uncertainty at bus 2 induces a bi-modal PDF of the generated powers.

What is the effect on the voltages at the PQ buses? This is shown in Figure 6.4: corresponding to the peak in demand at 3.2 there is a peak in the voltage at around 0.91. The other peak in the demand at 2.1 corresponds to the voltage peak at around 0.95. The phase is negative because the load is drawing power. The deterministic load at bus 4 shows a qualitatively similar behavior: two peaks in the voltage magnitudes and the angle are induced by the uncertain demand at bus 2. However, the voltage magnitude variation at bus 4 is much smaller—essentially constant—compared to bus 2. For both PQ buses the phase angles are negative as they are drawing power from the grid. Figure 6.5 shows the PDFs of the current magnitudes for all branches. We see the bi-modality in all lines; however line 5 has an extreme first peak at around 0.89. It is the line that connects the PV bus 3 to the deterministic load 4. There are two effects: first the voltage magnitude and the power injection are constant for the generator at bus 4, and second the voltage magnitude variation at bus 4 is also small because it has a deterministic load. As

---

<sup>5</sup> The power injection for the generator at bus 3 is not displayed as it is deterministic at 0.84, see Table 6.2 and Table 6.3.

Table 6.4: PCE-based solution for generators for OPF.

x	$k^{\text{th}}$ PCE coefficient					Moments	
	0	1	2	3	4	$\mathbb{E}(x)$	$\sqrt{\mathbb{V}(x)}$
p <sub>1</sub>	1.22	0.3114	0.005694	0.0004417	4.51e-5	1.22	0.1953
q <sub>1</sub>	1.883	0.4084	-0.03221	-0.009274	-0.00179	1.883	0.2565
p <sub>3</sub>	2.971	0.7651	0.01283	0.001229	0.0001456	2.971	0.4797
q <sub>3</sub>	2.189	0.8255	0.1251	0.01767	0.002753	2.189	0.5203

Table 6.5: PCE-based solution for generators for CC-OPF.

x	$k^{\text{th}}$ PCE coefficient					Moments		Constraint satisfaction	
	0	1	2	3	4	$\mathbb{E}(x)$	$\sqrt{\mathbb{V}(x)}$	Required	Achieved
p <sub>1</sub>	1.411	0.801	0.01346	0.001128	0.0001986	1.411	0.5022	-	-
q <sub>1</sub>	1.9	0.2419	0.05474	-0.03452	0.007012	1.9	0.1539	99.0 %	99.9 %
p <sub>3</sub>	2.778	0.2713	0.005252	0.0004625	8.105e-5	2.778	0.1701	95.0 %	96.7 %
q <sub>3</sub>	2.163	0.9697	0.03892	0.04255	-0.00561	2.163	0.6084	-	-

the current magnitude is proportional to the magnitude of the difference of the voltages—which is small for line 5—there is just a bit of fluctuation.

## 6.3 Chance-constrained AC optimal power flow

For PPF there are no degrees of freedom. In contrast, OPF introduces degrees of freedom such that operational costs can be minimized and inequality constraints can be considered. We keep the p<sub>q</sub> bus specifications and the slack bus specification from the previous Section 6.2, but drop the bus specification for the p<sub>v</sub> bus 3. In order to solve chance-constrained OPF according to Problem 4.2 we need to satisfy Assumption 4.2: we need to compute the orthogonal basis, and to parameterize the buses in terms of given PCE coefficients.

### Construction of orthogonal basis

We leave the uncertainty model untouched; the basis from PPF from Section 6.2 is used.

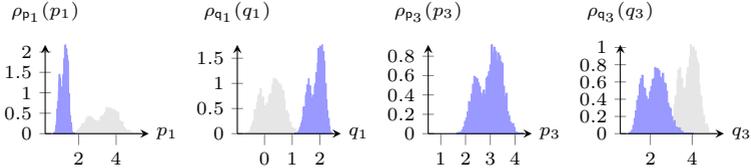


Figure 6.6: (OPF) PDFs of active power and reactive power at buses  $\{1, 4\}$ . PF solution in light gray.

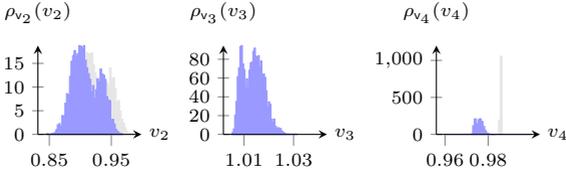


Figure 6.7: (OPF) PDFs of voltage magnitude at buses  $\{2, 3, 4\}$ . PF solution in light gray.

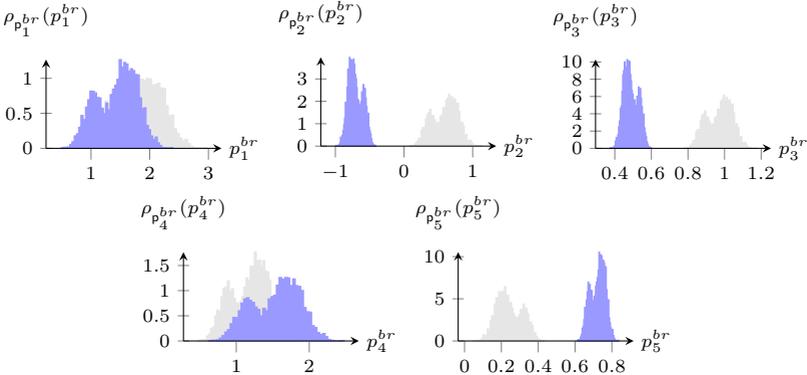


Figure 6.8: (OPF) PDFs of active power for all lines. PF solution in light gray.

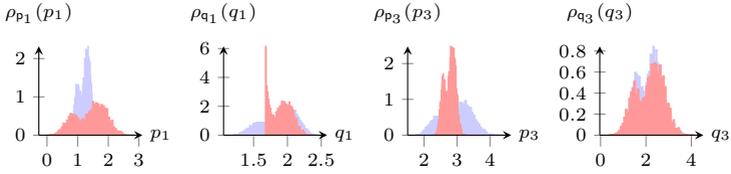


Figure 6.9: (CC-OPF) PDFs of active power and reactive power at buses  $\{1, 4\}$ . Unconstrained OPF solution in light blue.

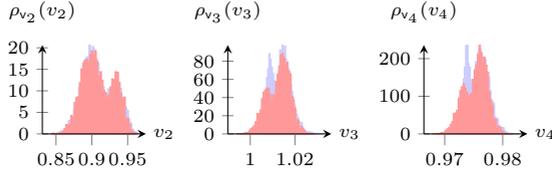


Figure 6.10: (CC-OPF) PDFs of voltage magnitude at buses  $\{2, 3, 4\}$ . Unconstrained OPF solution in light blue.

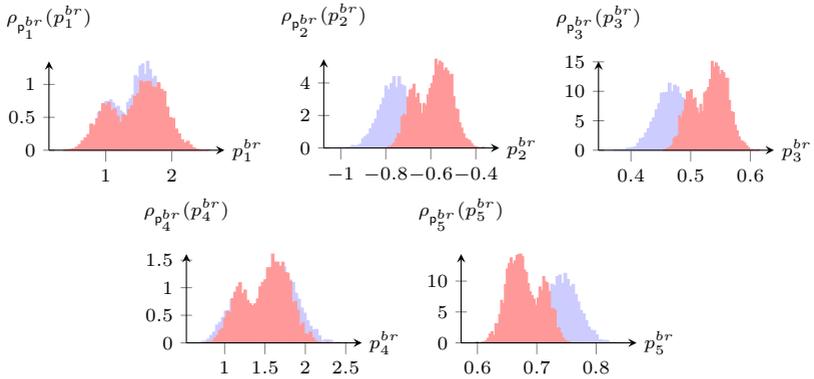


Figure 6.11: (CC-OPF) PDFs of active power for all lines. Unconstrained OPF solution in light blue.

## PCE coefficients for bus specifications

We recycle the results from PPF from Section 6.2:<sup>6</sup>

$$\begin{bmatrix} \hat{e}_{1,0} & \hat{e}_{1,1} & \hat{e}_{1,k} & \hat{f}_{1,0} & \hat{f}_{1,1} & \hat{f}_{1,k} \\ \hat{p}_{2,0} & \hat{p}_{2,1} & \hat{p}_{2,k} & \hat{q}_{2,0} & \hat{q}_{2,1} & \hat{q}_{2,k} \\ \hat{p}_{4,0} & \hat{p}_{4,1} & \hat{p}_{4,k} & \hat{q}_{4,0} & \hat{q}_{4,1} & \hat{q}_{4,k} \end{bmatrix} = \begin{bmatrix} 1.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ -2.87 & -1.00 & 0.00 & -2.44 & -0.85 & 0.00 \\ -1.20 & 0.00 & 0.00 & -1.02 & 0.00 & 0.00 \end{bmatrix},$$

where  $k \in \{2, \dots, \hat{k}\}$  are the indices relative to degrees greater than one. We keep only the bus specifications that leave no degrees of freedom.

## Numerical solution

We consider quadratic costs (4.14) for the generators with coefficients<sup>7</sup>

$$\begin{bmatrix} c_{2,1} & c_{1,1} \\ c_{2,3} & c_{1,3} \end{bmatrix} = \begin{bmatrix} 2500 & 100 \\ 1000 & 200 \end{bmatrix}. \quad (6.1)$$

The PCE-overloaded nonlinear program (NLP) from Problem 4.2 has a total of  $4N_b(\hat{k} + 1) = 80$  unknowns. Ipopt is able to solve the NLP in about 25 ms.

Table 6.4 shows the numerical values of the PCE coefficients for the active and reactive power at the generators. Based on the moments we see that bus 1 injects less active power compared to bus 2 because it is more expensive. The optimal value, hence the expected total cost, is 13,588.37; in comparison, the PPF solution results in a total expected cost of 30,664.30. We are interested in how the PDFs of the generators have changed compared to the case of plain PPF. This is shown in Figure 6.6: the light gray PDFs correspond to the PDFs from PPF from Section 6.2, see Figure 6.3. Clearly, the active power injection at bus 1 is reduced compared to PPF; and it is ramped up at the other generator that was previously set to a fixed active power injection. For the reactive power we see that bus 1 ramps up and bus 3 ramps down in relation to PPF.

<sup>6</sup> Note that we do not need to distinguish the buses according to (4.2), because each bus is either a generator or a load.

<sup>7</sup> Recall that we consider the cost to be dimensionless.

The effect on the voltage magnitudes is shown in Figure 6.7: the magnitude at bus 3 is no longer constant. A consequence of bus 3 now being flexible in its power injections—and being cheap—is that the (directed) power flow across line 2—which connects bus 1 to bus 3—is reversed compared to PPF, see Figure 6.8. Also, more power is delivered across line 4 from bus 3 to the uncertain load at bus 2.

Let us now consider two specific engineering limits in terms of the reformulated chance constraints from Problem 4.2

$$\mathbb{E}(q_1) + \lambda(\varepsilon_q) \sqrt{\mathbb{V}(q_1)} \leq q_1^{\max} = 2.30, \quad (6.2a)$$

$$\mathbb{E}(p_3) + \lambda(\varepsilon_p) \sqrt{\mathbb{V}(p_3)} \leq p_3^{\max} = 3.05. \quad (6.2b)$$

We hence want to constrain the maximum reactive power generation at bus 1 and the maximum active power generation at bus 3. Motivated by the relation  $\lambda(\varepsilon) = \Phi^{-1}(1 - \varepsilon)$  that holds for Gaussian random variables, see Remark 4.2 we choose  $\lambda(\varepsilon_q) = 2.6$  and  $\lambda(\varepsilon_p) = 1.6$ . For (6.2) this amounts to  $\varepsilon_q = 0.99$  and  $\varepsilon_p = 0.95$ . The numerical results of this now chance-constrained AC-OPF are shown in Table 6.5 for the PCE coefficients of the active and reactive power of the generators. Table 6.5 also shows the empirical constraint satisfaction: the required constraint satisfaction is achieved. Hence, the Gaussian heuristic is good enough for this specific example. The total expected cost becomes 14,052.92, an increase of about 3.42% compared to the unconstrained solution. Figure 6.9 shows the PDFs of the injected powers compared to the PDFs from the previous OPF solution. As intended, the reactive power at bus 1 is decreased, so is the active power at bus 3. The small effect on the voltage magnitudes is shown in Figure 6.10. Again, the active power line flows are more sensitive, especially the flows across lines 2 and 5, as shown in Figure 6.11.

With the inequality constraints the time to solve the PCE-overloaded NLP becomes 55 ms.

## 6.4 Chance-constrained DC optimal power flow

Let us study the chance-constrained AC-OPF problem under DC conditions, see Assumption 4.3. DC conditions sacrifice physical exactness for computational

Table 6.6: PCE-based solution for generators for chance-constrained DC-OPF.

$x$	$k^{\text{th}}$ PCE coefficient					Moments		Constraint satisfaction	
	0	1	2	3	4	$\mathbb{E}(x)$	$\sqrt{\mathbb{V}(x)}$	Required	Achieved
$p_1$	1.334	0.687	0.0	0.0	0.0	1.334	0.431	-	-
$p_3$	2.736	0.313	0.0	0.0	0.0	2.736	0.196	95.0 %	96.8 %

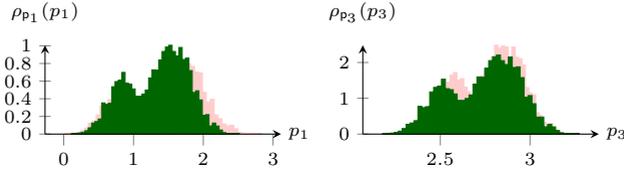


Figure 6.12: (CC-DC-OPF) PDFs of active power at buses  $\{1, 4\}$ . CC-OPF solution in light red.

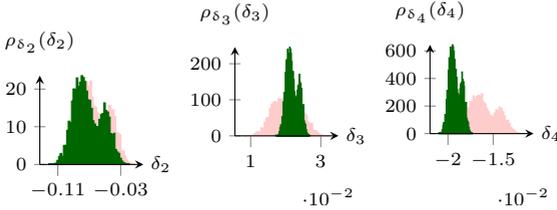


Figure 6.13: (CC-DC-OPF) PDFs of voltage angles at buses  $\{2, 3, 4\}$ . CC-OPF solution in light red.

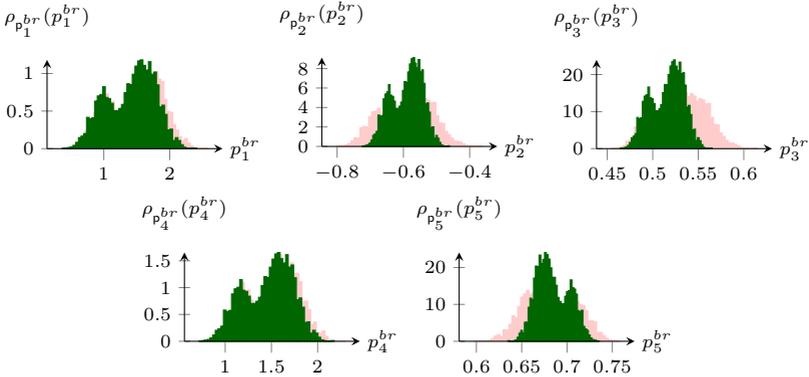


Figure 6.14: (CC-DC-OPF) PDFs of active power for all lines. CC-OPF solution in light red.

tractability. Recall that for DC power flow the grid state is given by the active

power and the voltage angles, shrinking the number of decision variables to one half.

Before we can solve Problem 4.4 we need to satisfy Assumption 4.2: we have to construct the orthogonal basis, and to parameterize the buses in terms of given and exact PCE coefficients.

### Construction of orthogonal basis

We leave the uncertainty model untouched; the basis from PPF from Section 6.2 is used.

### PCE coefficients for bus specifications

Borrowing from the previous PPF problem from Section 6.2 we find

$$\begin{bmatrix} \hat{p}_{2,0} & \hat{p}_{2,1} & \hat{p}_{2,k} \\ \hat{p}_{4,0} & \hat{p}_{4,1} & \hat{p}_{4,k} \end{bmatrix} = \begin{bmatrix} -2.87 & -1.00 & 0.00 \\ -1.20 & 0.00 & 0.00 \end{bmatrix},$$

where  $k \in \{2, \dots, \hat{k}\}$  are the indices relative to degrees greater than one. Note that no reactive power bus specifications are required due to the DC conditions.

### Numerical solution

We use the cost coefficients from (6.1). Mosek is able to solve the PCE-overloaded second-order cone program (SOCP) from Problem 4.4 in 1.6 ms. Table 6.6 shows the numerical values of the PCE coefficients for the active power at the generators, yielding an objective value of 13,117.51. Comparing these figures, especially the moments, to their AC counterparts from Table 6.5 we see that the values for order zero and order one are close; the PCE coefficients of order  $k \geq 2$  are zero in the DC case, because the power flow satisfaction is achieved with just an affine basis. Figure 6.12 shows the histograms for the injected power at buses 1 and 3 in comparison to their AC counterparts: there is a significant overlap. Differences for the DC solution become

more apparent for the voltage angles, shown in Figure 6.13, and especially the power flows, shown in Figure 6.14.

## 6.5 Comparison

Table 6.7 provides a comparison of the solution characteristics for all considered problems. Note that the number of variables for PPF, OPF, and CC-OPF should read  $4N_b(\hat{k} + 1) = 80$ , but we need not define decision variables for the active and reactive power PCE coefficients at buses 2 and 4, reducing the number to 60 unknowns. The same argument holds for the DC case. From Table 6.7 we clearly see how an increasing problem complexity corresponds to an increasing number of iterations and increasing solution times. The CC-OPF problem under DC conditions exhibits a specific problem structure (SOCP) for which there exist tailored solution algorithms, thus reducing the computation time. The computational advantages of DC conditions are evident. However, it is highly case-dependent whether or not the deviations from the “true” AC solution are acceptable.

Table 6.7: Comparison of solution characteristics.

Case	Problem type	Variables	Solver	Objective	Iterations	Time
PPF	Feasibility	60	Ipopt	30,664.30	4	20 ms
OPF	NLP	60	Ipopt	13,588.37	7	25 ms
CC-AC-OPF	NLP	60	Ipopt	14,052.92	18	55 ms
CC-DC-OPF	SOCP	10	Mosek	13,117.51	9	1.6 ms

## 7 Summary and outlook

While societies are continuously increasing the share of renewable energy sources, they are penetrating the power system with considerable uncertainties: no longer is it certain where and when and how much electrical energy is being supplied. This leads to intriguing research questions—which the present thesis sheds light upon. Most importantly we ask: How to formulate (optimal) power flow problems rigorously in the presence of uncertainties? How to solve these problems coherently? Figure 1.1 from page 5 summarizes our approach: first, study generic mappings under uncertainty using the Hilbert space technique polynomial chaos expansion (PCE), and then apply these ideas to specific (optimal) power flow problems. To deal with the computational aspects of (intrusive) PCE we created the Julia package *PolyChaos.jl*. It allows to compute orthogonal polynomials, quadrature rules, and polynomial chaos expansions of random variables. As with anything in research and science: a single thesis can cover only so much. There are numerous open threads worth pursuing in the future. We examine the previous chapters in light of scientific contributions and open topics.

### Chapter 2 – Mappings under uncertainty

Given a mapping, and given random variables, we show how to propagate these random variables through the mapping to obtain image random variables. Our method of choice is the Hilbert space technique called PCE. The theoretical foundations of PCE lie with measure theory, Hilbert space theory, and orthogonal polynomials. We discuss the core concepts of each topic and provide a self-contained introduction to PCE. Our contributions are twofold: pedagogically, we provide a thorough problem formulation of mappings under uncertainty; scientifically, we study PCE truncation errors that result from using not infinitely many but finitely many coefficients to represent random

variables. For explicit polynomial or non-polynomial mappings we derive error bounds. Possible extensions for future work include:

- Quantify PCE truncation errors for implicit mappings.
- Quantify PCE truncation errors for non-intrusive methods.
- Assess whether and how a decaying truncation error can be used for stability in systems and control.

### **Chapter 3 – Power flow under uncertainty**

This chapter draws the attention from abstract mappings to concrete examples from power systems: how to solve power flow problems in the presence of uncertainties. We study the generic AC power flow problem as well as power flow for radial grids and so-called optimal adaptive linearizations. These linearizations provide easy-to-evaluate proxies to the full AC power flow solution whilst accounting for operating ranges modeled by random variables. We show how PCE helps with all these intrinsically infinite-dimensional problems: it renders them finite-dimensional by reformulating the problems in the PCE coefficients. The methods and contributions from this chapter are among the first PCE-based formulations of power flow under uncertainty. The upper half of Table 7.1 provides an in-depth overview on how we reformulate infinite-dimensional power flow problems under uncertainty as finite-dimensional counterparts, with the help of PCE. Besides the names and spaces of the decision variables, Table 7.1 also provides a short discussion about the kind of the problem, the number of decision variables, and whether zero truncation errors are attainable.

It is especially the numerics of PCE-overloaded problem formulations that deserve future attention. We envision improvements as follows:

- Tailor algorithms to solving large-scale PCE-overloaded PPF problems.
- Exploit the decay of the PCE coefficients in numerical routines.
- Use pre-conditioning techniques for PCE-overloaded PPF problems.
- Consider other error metrics for optimal adaptive linearizations.

Table 7.1: Overview of problem formulations covered. For all  $i \in \mathcal{N}$ , for all  $j \in \mathcal{N} \setminus \mathcal{N}_{\text{sl}}$ , for all  $k \in \mathcal{K}$ .

Name	Infinite-dimensional		Finite-dimensional		Discussion
	Problem Variables	Problem Variables	Problem Variables	Variables	
PPF	3.1	$\mathbf{p}_i, \mathbf{q}_i, \mathbf{e}_i, \mathbf{f}_i \in L^2(\Omega, \mu; \mathbb{R})$	3.2	$p_{i,k}, q_{i,k}, e_{i,k}, f_{i,k} \in \mathbb{R}$	<ul style="list-style-type: none"> <li>- Feasibility problem</li> <li>- <math>4N_b(\hat{k}+1)</math> decision variables</li> <li>- Non-zero truncation error in general</li> </ul>
BFS	3.3	$\mathbf{p}_i, \mathbf{q}_i, \mathbf{e}_i, \mathbf{f}_i \in L^2(\Omega, \mu; \mathbb{R})$	3.4	$p_{i,k}, q_{i,k}, e_{i,k}, f_{i,k} \in \mathbb{R}$	<ul style="list-style-type: none"> <li>- Iterative problem</li> <li>- <math>4N_b(\hat{k}+1)</math> decision variables</li> <li>- Truncation errors depend on load model</li> </ul>
Optimal adaptive linearization	3.5	$\ell_0, \ell_{p,j}, \ell_{q,j} \in \mathbb{R}, \mathbf{p}_i, \mathbf{q}_i, \mathbf{e}_i, \mathbf{f}_i \in L^2(\Omega, \mu; \mathbb{R})$	3.6	$\ell_0, \ell_{p,j}, \ell_{q,j}, p_{i,k}, q_{i,k}, e_{i,k}, f_{i,k} \in \mathbb{R}$	<ul style="list-style-type: none"> <li>- Feasibility problem + unconstrained QP</li> <li>- <math>4N_b(\hat{k}+1)</math> decision variables for PPF, <math>1 + 2 \mathcal{N} \setminus \mathcal{N}_{\text{sl}} </math> for unconstrained QP</li> <li>- Non-zero truncation error in general for PPF</li> </ul>
CC-AC-OPF	4.1	$\mathbf{p}_i^c, \mathbf{q}_i^c, \mathbf{p}_i, \mathbf{q}_i, \mathbf{e}_i, \mathbf{f}_i \in L^2(\Omega, \mu; \mathbb{R})$	4.2	$p_{i,k}^c, q_{i,k}^c, p_{i,k}, q_{i,k}, e_{i,k}, f_{i,k} \in \mathbb{R}$	<ul style="list-style-type: none"> <li>- NLP</li> <li>- <math>4N_b(\hat{k}+1)</math> decision variables</li> <li>- Non-zero truncation error in general</li> </ul>
CC-DC-OPF	4.3	$\mathbf{p}_i^c, \mathbf{p}_i \in L^2(\Omega, \mu; \mathbb{R})$	4.4	$p_{i,k}^c, p_{i,k} \in \mathbb{R}$	<ul style="list-style-type: none"> <li>- SOCP</li> <li>- <math>N_b(\hat{k} + 1)</math> decision variables</li> <li>- Zero truncation error</li> </ul>
Multi-stage CC-DC-OPF	4.5	$\mathbf{p}_i^c, \mathbf{p}_i^s, \mathbf{e}_i, \mathbf{p}_i \in L^2(\Omega, \mu; \mathbb{R})$	4.6	$\hat{p}_i^c, \hat{p}_i^s \in \mathbb{R}^T, P^c, P^s \in \mathbb{R}^T \times T$	<ul style="list-style-type: none"> <li>- SOCP</li> <li>- <math>N_b(\hat{k} + 1)</math> decision variables</li> <li>- Zero truncation error</li> </ul>

## Chapter 4 – Optimal power flow under uncertainty

Building on the problem formulations from Chapter 3, Chapter 4 turns from power flow problems to *optimal* power flow problems. This leads to the question: how to operate electrical grids optimally despite uncertainties *and* whilst respecting engineering limits? We apply PCE to three specific problems: AC-OPF under uncertainty, DC-OPF under uncertainty, DC-OPF under uncertainty in the multi time-step setting. To reformulate the optimization problems in the presence of uncertainty, we introduce chance constraints and their deterministic moment-based reformulations. To the best of the author's knowledge, PCE is currently the only method that allows to tackle the full nonlinear AC power flow equations together with moment-based reformulations of individual chance constraints. In the DC setting PCE allows to formulate the OPF problem under uncertainty as a second-order cone program (SOCP). The lower half of Table 7.1 provides an in-depth overview on how we reformulate infinite-dimensional optimal power flow problems under uncertainty as finite-dimensional counterparts, with the help of PCE. Besides the names and spaces of the decision variables, Table 7.1 also provides a short discussion about the kind of the problem, the number of decision variables, and whether zero truncation errors are attainable.

Similar to the open issues for Chapter 3, we sense future improvements especially with the numerics of PCE-overloaded OPF problems:

- Derive tailored constraint generation algorithms for the PCE-overloaded AC-OPF problem.
- Exploit the decay of the PCE coefficients in numerical routines.
- Use pre-conditioning techniques for PCE-overloaded OPF problems.
- Derive non-intrusive schemes to solving PCE-overloaded OPF problems.
- Derive tailored constraint generation algorithms for the SOCPs for OPF under DC conditions.
- Explore the connection between PCE and distributionally robust chance constraint reformulations.

- Exploit PCE with AC-OPF problems that use linearized AC power flow equations and/or convex relaxations.
- Incorporate explicit dependencies between policies for generating units, e.g. between active and reactive power.
- Employ different cost functions, e.g. accounting for reserve costs.
- Introduce mixed-integer variables to PCE-overloaded problems to account for security-constrained OPF.

## Chapter 5 – PolyChaos.jl

Polynomial chaos requires efficient software to compute orthogonal bases and (tensorized) scalar products. We created *PolyChaos.jl*, a software package written in the Julia programming language that allows to compute orthogonal polynomials, quadrature rules, and polynomial chaos expansions. The package is open source and available free of charge. The founding principle of *PolyChaos.jl* is to allow to compute orthogonal polynomials for arbitrary, user-specified probability density functions (or generally absolutely continuous, non-negative measures). The package comes with a straightforward syntax and a comprehensible online documentation. As with any (numerical) software, possible future improvements or additional features are multifold:

- Add support for non-intrusive PCE such as collocation.
- Add automated plotting functionality.
- Provide functionality for code generation for specific classes of functions such as polynomials or linear ordinary differential equations.
- Add basis-adaptive and sparse PCE.
- Add sparse quadrature rules.
- Add support for advanced sampling techniques.
- Extend the documentation and collect more use cases.

## Chapter 6 – Case studies

This chapter provides case studies for three selected settings: power flow under uncertainty, AC-OPF, and DC-OPF. For all settings we study the same grid under the same uncertainty. This allows to compare the results straightforwardly, for instance comparing the accuracy of OPF for the AC and the DC setting. Although we provide references to larger test cases in the chapter, we are aware that test cases with many buses *and* many uncertainties are desirable. Assuming that some of the aforementioned numerics-related improvements for PCE were successful, future case studies can be extended to account for more general problems. However, these problems may require to revisit problem formulations too:

- Combine AC and HVDC lines.
- Apply distributed optimization to solve problems region-wise.
- Add contingencies and N-1 constraints.
- Compare chance constraint reformulations (individual vs. joint).
- Investigate coupled transmission systems and distribution systems.

The modern world wants to uphold the luxury and comfort of omniscient electrical energy whilst reducing the dependency on naturally limited fossil fuels. If this thesis contributed in the faintest way, then it has paid its dues.

# A Appendix

## A.1 Orthogonal polynomials

Orthogonal polynomials possess remarkable properties:

- If the absolutely continuous measure  $\mu$  is symmetric with respect to the origin, that is its density satisfies  $\rho(t) = \rho(-t)$ , then  $\phi_k$  is an even/odd function in case  $k$  is even/odd [64, Thm 1.17].
- All zeros of  $\phi_k$  for  $k \in \mathbb{N}$  are real, simple, and located in the interior of the support of  $\mu$  [64, Thm 1.19]; they can be computed as the eigenvalues of a symmetric tridiagonal matrix [64, 71]. In case the measure is symmetric with respect to a point that is not the origin, the coordinates can be shifted such that one attains symmetry with respect to the origin in the shifted coordinates.
- The zeros of  $\phi_{k+1}$  alternate with those of  $\phi_k$  [64, Thm 1.20].
- Let  $\mathcal{P}_\mu = \{\phi_k\}_{k \in \mathcal{K}}$  satisfy Definition 2.13, then

$$\begin{aligned}\phi_{k+1}(\tau) &= (\tau - \alpha_k)\phi_k(\tau) - \beta_k\phi_{k-1}(\tau) \quad \forall k \in \mathcal{K} \\ \phi_0(\tau) &= 1, \\ \phi_{-1}(\tau) &= 0,\end{aligned}\tag{A.1a}$$

where

$$\alpha_k = \frac{\langle \tau \phi_k, \phi_k \rangle}{\langle \phi_k, \phi_k \rangle}, \quad \forall k \in \mathcal{K}, \tag{A.1b}$$

$$\beta_k = \frac{\langle \phi_k, \phi_k \rangle}{\langle \phi_{k-1}, \phi_{k-1} \rangle}, \quad \forall k \in \mathcal{K} \setminus \{0\}, \tag{A.1c}$$

are called the recurrence coefficients. It is convenient to define<sup>1</sup>

$$\beta_0 = \langle \phi_0, \phi_0 \rangle. \quad (\text{A.1d})$$

For many well-known densities the recurrence coefficients of the respective orthogonal polynomials are known in closed form. Users can get a lot of mileage from these textbook definitions, see e.g. [64]. In case the recurrence coefficients are not known, there are numerical procedures to compute the recurrence coefficients for absolutely continuous measures such as the Stieltjes procedure, the Lanczos procedure, or the (modified) Chebyshev algorithm [64].

According to [64] “[the] three-term recurrence relation [(A.1)] [...] is arguably the single most important piece of information for the constructive and computational use of orthogonal polynomials.” The relation (A.1) can be used for evaluating the polynomials, but more importantly to compute quadrature rules, see Appendix A.2.

It is fair to ask: given an absolutely continuous measure  $d\mu(\tau) = \rho d\tau$ , what are the respective orthogonal polynomials? The Stieltjes procedure and the Lanczos procedure are two numerical methods that answer this question. As usual with numerical methods we need to discretize, hence we identify with an absolutely continuous measure its discretized counterpart  $\mu_N$  in terms of  $N$  pairs of nodes and weights  $\{(\lambda_i, t_i)\}_{i \in \{1, \dots, N\}}$ .

## Stieltjes procedure

The Stieltjes procedure—dating back to ideas from 1884—is an iterative method to compute the three-term recurrence coefficients of orthogonal polynomials. The idea is to evaluate the expressions from (A.1) successively based on the insight: to compute  $\alpha_k$  we need to know the  $k^{\text{th}}$  basis polynomial, and to compute  $\beta_k$  we need to know the  $k^{\text{th}}$  and the  $(k - 1)^{\text{th}}$  basis polynomials.

---

<sup>1</sup> For probability measures we have  $\beta_0 = 1$  by definition.

By definition we know  $\beta_0$ , see (A.1d). Starting from  $\phi_0 = 1$ , the coefficient  $\alpha_0$  can be computed from (A.1b),

$$\alpha_0 = \frac{\langle \tau \phi_0, \phi_0 \rangle}{\langle \phi_0, \phi_0 \rangle} = \frac{\langle \tau, 1 \rangle}{\beta_0}.$$

For probability density functions we have  $\alpha_0 = \mathbb{E}(\tau)$ . Knowing  $\alpha_0$  and  $\beta_0$  we can compute  $\phi_1$  from (A.1a). With this we can go back to (A.1b) and compute  $\alpha_1$ , and from (A.1c) we obtain  $\beta_1$ . Knowing  $\alpha_1$  and  $\beta_1$  we obtain  $\phi_2$  from (A.1a). We continue to cycle through the definition of the recurrence coefficients (A.1b) and (A.1a), and the computation of the basis polynomial (A.1a) until a desired degree is reached.

Usually, the Stieltjes procedure relies on a quadrature rule to solve all occurring integrals—which is but a discretization of the original absolutely continuous measure. As pointed out in [64], it is advisable to choose the order of the quadrature rule much larger than the number of desired recursion coefficients. Potential over- or underflows can be avoided by scaling the quadrature weights or the basis polynomials.

## Lanczos procedure

The Lanczos algorithm allows to tri-diagonalize a given symmetric matrix  $A$  [70, Ch. 9]. More specifically, a real symmetric matrix  $A$  allows the transformation  $Q^\top A Q = T$ , where  $Q$  is orthogonal and  $T$  is symmetric and tridiagonal. Lanczos' algorithm produces the matrices  $Q$  and  $T$ , given  $A$ . In light of orthogonal polynomials, the Lanczos *procedure* means to construct  $A$  such that the output of the Lanczos algorithm is the Jacobi matrix (A.6), from which the recurrence coefficients can be read off. As shown in [64, Section 2.2.3.2], the matrix  $A$  contains the nodes and the weights of the discretized measure  $\mu_N$ :

$$A = \begin{bmatrix} 1 & \sqrt{t_1} & \sqrt{\lambda_2} & \dots & \sqrt{\lambda_N} \\ \sqrt{\lambda_1} & t_1 & 0 & & 0 \\ \sqrt{\lambda_2} & 0 & t_2 & & \\ \vdots & & & \ddots & 0 \\ \sqrt{\lambda_N} & 0 & & 0 & t_N \end{bmatrix}, \quad (\text{A.2})$$

yielding the tridiagonal matrix

$$T = \begin{bmatrix} 1 & \sqrt{\beta_0} e_1^\top \\ \sqrt{\beta_0} e_1 & J_N(\mu_N) \end{bmatrix}, \quad (\text{A.3})$$

with  $e = [1 \ 0 \ \dots \ 0]^\top \in \mathbb{R}^N$ , and  $J_N(\mu_N)$  is the  $N$ -point Jacobi matrix for the discretized measure  $\mu_N$ , see (A.6).

For details on the Lanczos algorithm itself we refer to the dedicated literature, for instance [70, Ch. 9] or [77].

## A.2 Gauss quadrature

This section is based on [64, 65, 69, 71].

Consider an absolutely continuous positive measure  $\mu$  on the real axis whose moments

$$m_r = \int_{\mathbb{R}} \tau^r d\mu(\tau) \quad (\text{A.4})$$

exist for all  $r \in \mathbb{N}_0$ . An  $n$ -point Gauss quadrature rule for the measure  $\mu$  is given by the pairs  $\{(\lambda_i, t_i)\}_{i=1}^n$  such that

$$\int_{\mathbb{R}} f(\tau) d\mu(\tau) = \sum_{i=1}^n \lambda_i f(t_i) + R_n(f) \quad (\text{A.5})$$

with—and that is key to making a quadrature rule a *Gauss* quadrature rule— $R_n(f) = 0$  whenever  $f$  is a polynomial of degree  $\leq 2n - 1$ . That means an integral involving a polynomial of degree up to  $2n - 1 = 5$  can be computed exactly by a weighted sum of just  $n = 3$  terms. And how to find the Gauss quadrature rule? It turns out that the nodes  $\{t_i\}_{i=1}^n$  are the zeros of the univariate monic polynomial  $\phi_n$  of degree  $n$  that is orthogonal relative to the measure  $\mu$  [64, 154]. Recall from Appendix A.1 that these zeros are real, simple and located in the interior of the support of the measure. Both the nodes  $t_i$  and the corresponding weights  $\lambda_i$  can be obtained elegantly from

the solution of an eigenvalue problem that involves the three-term recurrence relation. We define the  $n$ -point symmetric tridiagonal Jacobi matrix

$$J_n(\mu) = \begin{bmatrix} \alpha_0 & \sqrt{\beta_1} & & & 0 \\ \sqrt{\beta_1} & \alpha_1 & \sqrt{\beta_2} & & \\ & \sqrt{\beta_2} & \alpha_2 & \sqrt{\beta_3} & \\ & & & \ddots & \\ 0 & & & & \sqrt{\beta_{n-1}} \\ & & & & \sqrt{\beta_{n-1}} & \alpha_{n-1} \end{bmatrix}, \quad (\text{A.6})$$

where  $(\alpha_i, \beta_i)$  for  $i \in \{0, \dots, n-1\}$  are the recurrence coefficients of the system of orthogonal polynomials, see Appendix A.1. Then, the nodes  $\{t_i\}_{i=1}^n$  are the eigenvalues of  $J_n(\mu)$ , and the weights  $\{\lambda_i\}_{i=1}^n$  satisfy

$$\lambda_i = m_0 v_{i,1}^2 \quad (\text{A.7})$$

with  $v_{i,1}$  being the first component of the normalized eigenvector  $v_i$  corresponding to the eigenvalue  $t_i$ , hence  $J_n(\mu)v_i = t_i v_i$  for all  $i \in \{1, \dots, n\}$  [64, 65, 71, 154]. In case the integrand is not polynomial, but a smooth function, then there exist bounds on the remainder term  $R_n(f)$  from (A.5) [65].

There exist variants of the classic Gauss quadrature that consider specific points in the quadrature rule in case the support is finite. For example, a Gauss-Radau quadrature rule is an  $(n+1)$ -point quadrature rule of the form

$$\int_a^b f(\tau) d\mu(\tau) = \lambda_0 f(a) + \sum_{i=1}^n \lambda_i f(t_i) + R_n^a(f) \quad (\text{A.8})$$

where  $R_n^a(f) = 0$  holds for all functions  $f$  that are polynomials of degree at most  $2n$ . Interestingly enough, the nodes and weights  $(\lambda_i, t_i)_{i=0}^n$  with  $t_0 = a$  can again be computed by solving an eigenvalue problem. Consider a modified Jacobi matrix

$$J_{n+1}^a(\mu) = \begin{bmatrix} J_n(\mu) & \sqrt{\beta_n} e_n \\ \sqrt{\beta_n} e_n^\top & a - \beta_n \frac{\phi_{n-1}(a)}{\phi_n(a)} \end{bmatrix}, \quad (\text{A.9})$$

where  $e_n^\top = [0 \ 0 \ \dots \ 1] \in \mathbb{R}^n$  is the  $n^{\text{th}}$  unit vector of  $\mathbb{R}^n$ . Then, the nodes  $\{t_i\}_{i=0}^n$  are the eigenvalues of the modified Jacobi matrix  $J_{n+1}^a(\mu)$ , and the weights  $\{\lambda_i\}_{i=0}^n$  are again characterized by the relation (A.7) (in terms of the eigenvectors of  $J_{n+1}^a(\mu)$ ). The Gauss-Radau quadrature can similarly be defined to include the end point  $b$  of the support rather than the beginning point  $a$  [65, 69]. If we are to include both end points  $a$  and  $b$  of the support, then we obtain Gauss-Lobatto quadrature rules. These are  $(n + 2)$ -point quadrature rules

$$\int_a^b f(\tau) d\mu(\tau) = \lambda_0 f(a) + \sum_{i=1}^n \lambda_i f(t_i) + \lambda_{n+1} f(b) + R_n^{a,b}(f) \quad (\text{A.10})$$

for which  $R_n^{a,b}(f) = 0$  for integrands  $f$  that are polynomials of degree at most  $2n + 1$ . Once again, the nodes and weights  $(\lambda_i, t_i)_{i=0}^{n+1}$  with  $t_0 = a$  and  $t_{n+1} = b$  can be computed by solving an eigenvalue problem. Consider the modified Jacobi matrix

$$J_{n+2}^{a,b}(\mu) = \begin{bmatrix} J_{n+1}(\mu) & \sqrt{\beta_{n+1}^L} e_{n+1} \\ \sqrt{\beta_{n+1}^L} e_{n+1}^\top & \alpha_{n+1}^L \end{bmatrix}, \quad (\text{A.11})$$

where  $\alpha_{n+1}^L$  and  $\beta_{n+1}^L$  satisfy

$$\begin{bmatrix} \phi_{n+1}(a) & \phi_n(a) \\ \phi_{n+1}(b) & \phi_n(b) \end{bmatrix} \begin{bmatrix} \alpha_{n+1}^L \\ \beta_{n+1}^L \end{bmatrix} = \begin{bmatrix} a\phi_{n+1}(a) \\ b\phi_{n+1}(b) \end{bmatrix}. \quad (\text{A.12})$$

Then, the nodes  $\{t_i\}_{i=0}^{n+1}$  are the eigenvalues of the modified Jacobi matrix  $J_{n+2}^{a,b}(\mu)$ , and the weights  $\{\lambda_i\}_{i=0}^{n+1}$  are again characterized by the relation (A.7) (in terms of the eigenvectors of  $J_{n+2}^{a,b}(\mu)$ ).

The computation of quadrature rules by means of an eigenvalue problem of a (modified) Jacobi matrix can be extended even to the nested Gauss-Kronrod quadrature rule [65].

### A.3 Number of basis polynomials

Assume we have  $i \in \{1, \dots, m\}$  univariate sets of orthogonal polynomials  $\{\phi_k^{(i)}\}_{k=0}^{\hat{k}}$  of degree at most  $\hat{k}$ . What is the dimension of the corresponding  $m$ -variate basis of degree at most  $\hat{k}$ ? This is equivalent to asking for the number of non-negative integer solutions to the system of equations

$$z_1 + z_2 + \dots + z_m = k, \quad \forall k \in \{0, 1, \dots, \hat{k}\}. \quad (\text{A.13})$$

For each fixed  $k$  the number of non-negative integer solutions is known to be (so-called stars and bars method)

$$s_k = \binom{m}{k} = \binom{m+k-1}{k}. \quad (\text{A.14})$$

Hence, the total number  $s$  is given by

$$s = \sum_{k=0}^{\hat{k}} \binom{m}{k} = \binom{\hat{k}+1}{m} = \binom{m+\hat{k}}{m} = \frac{(m+\hat{k})!}{m!\hat{k}!}. \quad (\text{A.15})$$

### A.4 Polynomial chaos and stochastic processes

Random vectors can be interpreted as stochastic processes with a finite index set, see Section 2.2.2.<sup>2</sup> Hence, PCE is applicable to stochastic processes. Specifically, let us consider stochastic processes that are  $L^2$ -functions relative to both domains.

**Definition A.1** (Square-integrable real-valued stochastic process). *Let  $\mathbf{x}: \mathcal{T} \times \Omega$  be a stochastic process according to Definition 2.8. For every  $t \in \mathcal{T}$  let  $\mathbf{x}(t, \cdot) = \mathbf{x}(t) \in L^2(\Omega, \mu; \mathbb{R})$  be a square-integrable real-valued random variable, and for every  $\omega \in \Omega$  let  $\mathbf{x}(\cdot, \omega) \in L^2(\mathcal{T}, dt; \mathbb{R})$  be a square-integrable real-valued function. We call  $\mathbf{x} \in L^2(\Omega, \mu; \mathbb{R}) \otimes L^2(\mathcal{T}, dt; \mathbb{R})$  a square-integrable real-valued process. In case  $\mathcal{T}$  has finitely many elements  $dt$  is the discrete Lebesgue*

<sup>2</sup> Similarly, we can view continuous stochastic processes as infinite-dimensional random vectors.

measure defined at points  $t \in \mathcal{T}$ . Then, we call  $x \in L^2(\Omega, \mu; \mathbb{R}) \otimes L^2(\mathcal{T}, dt; \mathbb{R})$  a square-integrable discrete stochastic process.<sup>3</sup>

With this definition we can apply PCE to every random variable  $x(t, \cdot)$  of the stochastic process, yielding PCE functions that are functions with domain  $\mathcal{T}$ .

**Definition A.2** (PCE of square-integrable real-valued stochastic process). *Let  $x$  be a square-integrable real-valued stochastic process according to Definition A.1. The PCE of the stochastic process  $x$  is then given by the PCE of every random variable  $x(t, \cdot)$  according to Definition 2.15, yielding*

$$x(t, \cdot) = \sum_{k \in \mathbb{N}_0} x_k(t) \phi_k(\cdot), \quad (\text{A.16})$$

for which we write in short-hand

$$x(t) = \sum_{k \in \mathbb{N}_0} x_k(t) \phi_k. \quad (\text{A.17})$$

The functions  $x_k : \mathcal{T} \rightarrow \mathbb{R}$  are called the stochastic modes.

Hence, knowing the stochastic modes *and* knowing the respective orthogonal basis, we can model a given stochastic process. Although the decomposition from Definition A.2 is mathematically appealing, it is fair to ask whether and how the stochastic modes can be obtained from data of a given stochastic process. More often than not we are able to model a stochastic process in terms of its mean and covariance function. In that case there exists another elegant decomposition: the Karhunen-Loève (KL) decomposition, which is orthogonal not just in the probability space but also in the index set  $\mathcal{T}$ .

**Definition A.3** (Karhunen-Loève (KL) decomposition). *Let  $x : \mathcal{T} \times \Omega$  be a square-integrable real-valued process according to Definition A.1 with mean zero, i.e. for all  $t \in \mathcal{T} : \mathbb{E}(x) = 0$ , and a continuous and square-integrable covariance function  $c : \mathcal{T} \times \mathcal{T}$  with  $c(s, t) = \mathbb{E}(x(s, \cdot)x(t, \cdot))$ . Then*

$$x(t) = x(t, \cdot) = \sum_{n \in \mathbb{N}} \sqrt{\lambda_n} z_n(\cdot) \varphi_n(t) = \sum_{n \in \mathbb{N}} \sqrt{\lambda_n} z_n \varphi_n(t), \quad (\text{A.18})$$

---

<sup>3</sup> In other words, for each realization  $\omega \in \Omega$  we obtain a square-summable sequence  $x(\cdot, \omega)$ .

where  $(\lambda_n, \varphi_n)$  are the eigen-pairs of the covariance operator

$$\int_{\mathcal{T}} c(s, t) \varphi_n(s) ds = \lambda_n \varphi_n(t) \quad (\text{A.19})$$

for all  $n \in \mathbb{N}$  such that

$$\int_{\mathcal{T}} \varphi_n(t) \varphi_k(t) dt = \delta_{nk} \quad (\text{A.20})$$

for all  $n, k \in \mathbb{N}$ . The random variables  $z_n$  are defined by

$$z_n = z_n(\cdot) = \frac{1}{\sqrt{\lambda_n}} \int_{\mathcal{T}} x(t, \cdot) \varphi_n(t) dt \quad (\text{A.21})$$

for all  $n \in \mathbb{N}$ . They are centered, uncorrelated, and have unit variance

$$\mathbb{E}(z_n) = 0, \quad \mathbb{E}(z_n z_k) = \delta_{nk}. \quad (\text{A.22})$$

Similar to PCE, KL decompositions are truncated after finitely many terms in practice. The main computational challenge with KL is to find the eigen-pairs by solving the Fredholm integral equation of the second kind. It is fair to ask for the connection between PCE and KL. Given that the random variables  $z_n$  in the KL decomposition are themselves square-integrable real-valued random variables we can write their PCE

$$z_n = \sum_{k \in \mathbb{N}_0} z_k^{(n)} \phi_k. \quad (\text{A.23})$$

We can substitute this PCE in Definition A.3 and formally compare with the PCE of a stochastic process according to Definition A.2, which yields [95]

$$x_k(t) = \sum_{n \in \mathbb{N}} \sqrt{\lambda_n} \varphi_n(t) z_k^{(n)} \quad (\text{A.24})$$

for all  $k \in \mathbb{N}_0$ . The difficult-to-compute expression (A.24) simplifies tremendously in case the considered stochastic process is a Gaussian process.

**Definition A.4** (Gaussian process). Let  $F: \mathcal{T} \times \Omega$  be a stochastic process according to Definition 2.8. We call  $F$  a Gaussian process if  $(F(t_1), \dots, F(t_n))$  follows an  $n$ -dimensional normal distribution for all  $t_1, \dots, t_n \in \mathcal{T}$  with  $n \in \mathbb{N}$ .

**Proposition A.1** (PCE for Gaussian process). Let  $x: \mathcal{T} \times \Omega$  be a continuous Gaussian process according to Definition A.4 with mean zero and a continuous and square-integrable covariance function  $c: \mathcal{T} \times \mathcal{T}$  with  $c(s, t) = \mathbb{E}(x(s, \cdot)x(t, \cdot))$ . Then, the PCE of  $x$  is given by

$$x(t) = \sum_{k \in \mathbb{N}_0} x_k(t) \phi_k, \quad (\text{A.25})$$

where

$$\{\phi_k\}_{k \in \mathbb{N}_0} = \{1, \xi_1, \xi_2, \dots, \xi_n, \dots\}, \quad (\text{A.26})$$

$$x_k(t) = \begin{cases} 0, & k = 0, \\ \sqrt{\lambda_k} \varphi_k(t), & k \geq 1, \end{cases} \quad (\text{A.27})$$

and  $(\lambda_k, \varphi_k)$  are the eigen-pairs of the covariance-operator from Definition A.3 for all  $k \in \mathbb{N}$ .

*Proof.* The KL decomposition of the Gaussian process  $x$  is [95, 154, 172]

$$x(t) = \sum_{n \in \mathbb{N}} \sqrt{\lambda_n} z_n \varphi_n(t) = \sum_{n \in \mathbb{N}} \sqrt{\lambda_n} \xi_n \varphi_n(t), \quad (\text{A.28})$$

where  $\xi_n$  are standard Gaussian random variables. The collection of all  $\xi_n$  augmented with the unity function

$$\{\phi_n\}_{n \in \mathbb{N}_0} = \{1, \xi_1, \xi_2, \dots\} \quad (\text{A.29})$$

forms an infinite-dimensional orthogonal basis of total degree less than or equal to one relative to the “formal” probability density proportional to [154]

$$\prod_{n \in \mathbb{N}} \exp\left(-\frac{\tau_n^2}{2}\right). \quad (\text{A.30})$$

Re-arranging the  $\kappa\text{L}$  decomposition yields

$$x(t) = \underbrace{0}_{=: x_0(t)} + \sum_{n \in \mathbb{N}} \underbrace{\sqrt{\lambda_n} \varphi_n(t)}_{=: x_n(t)} \underbrace{\xi_n}_{=: \phi_n}, \quad (\text{A.31})$$

which completes the proof.  $\square$

## A.5 Bus admittance matrix

We provide a derivation of the bus admittance matrix by means of graph theory, inspired by [78, Ch. 7] and [152, Ch. 8].

We characterize the steady state of an electrical network by

1. its digraph  $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ , where  $\mathcal{V}$  comprises the set of vertices or nodes, and  $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$  comprises the set of edges which we numerate increasingly;
2. vertex voltages  $v \in \mathbb{C}^{|\mathcal{V}|}$  and vertex currents  $i \in \mathbb{C}^{|\mathcal{V}|}$ , and branch voltages  $v_{\text{br}} \in \mathbb{C}^{|\mathcal{E}|}$  and branch currents  $i_{\text{br}} \in \mathbb{C}^{|\mathcal{E}|}$ ;
3. a constitutive relation that maps branch voltages to branch currents.

Let  $|\mathcal{V}| = n$  be the number of nodes and  $|\mathcal{E}| = m$  be the number of edges of the digraph  $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ . The incidence matrix  $A \in \mathbb{Z}^{m \times n}$  is defined as

$$A_{ij} = \begin{cases} -1, & \text{if edge } i \text{ starts at node } j, \\ 1, & \text{if edge } i \text{ ends at node } j, \\ 0, & \text{otherwise,} \end{cases} \quad (\text{A.32})$$

for all  $i \in \{1, \dots, m\}$  and  $j \in \{1, \dots, n\}$ . The incidence matrix allows to convert nodal quantities to branch quantities. For example, let  $v \in \mathbb{C}^n$  be the vector of node voltages, then the vectors in the column space of  $A$

$$v_{\text{br}} = Av \in \mathbb{C}^m, \quad (\text{A.33})$$

correspond to the branch voltages, or potential differences. According to Kirchhoff's voltage law, the components of  $Av$  add to zero around every closed loop. This follows mathematically from the nullspace of  $A$  being

$$\text{Null } A = \{x \in \mathbb{R}^n : x = \alpha \mathbf{1}_n, \alpha \in \mathbb{R}\}, \quad (\text{A.34})$$

hence  $\text{rank } A = n - 1$ . Kirchhoff's current law is given by the left nullspace of  $A$ , namely

$$A^\top i_{\text{br}} = 0_n, \quad (\text{A.35})$$

where  $i_{\text{br}} \in \mathbb{C}^m$  denotes the branch currents. In other words: the current flowing in equals the current flowing out at each node, because Kirchhoff's current law is equivalent to a steady-state charge balance. If we draw/inject currents  $i$  at the nodes, then the steady-state charge balance has a sink/source term, thus reading

$$\underbrace{A^\top i_{\text{br}}}_{\text{inflow/outflow}} + \underbrace{i}_{\text{sink/source}} = 0_n, \quad (\text{A.36})$$

with the convention

$$i_j \begin{cases} < 0, & \text{for sinks,} \\ > 0, & \text{for sources,} \end{cases} \quad (\text{A.37})$$

for all nodes  $j \in \{1, \dots, n\}$ .

**Remark A.1.** *The four fundamental subspaces of  $A$ , i.e. column space, null space, row space, and left nullspace, characterize the graph in terms of branch and nodal quantities and represent Kirchhoff's laws. This is true for the concept of flows and potentials and not specific to electrical networks [152].*

It remains to relate the voltages to currents, which requires a constitutive law. The simplest possible constitutive equation relates the flow and its driving force linearly with a proportionality constant specific to the material. For electrical networks this is Ohm's law<sup>4</sup>

$$i_{\text{br}} = -Y_{\text{br}} v_{\text{br}}, \quad (\text{A.38})$$

where  $Y_{\text{br}} \in \mathbb{C}^{m \times m}$  is a complex regular and usually symmetric matrix, the so-called primitive or branch admittance matrix [78]. The minus sign is nec-

---

<sup>4</sup> Other such relations are Hooke's law, Fick's law, Newton's/Fourier's law of cooling.

essary to ensure that current flows from higher to lower potentials. Substituting Ohm's law (A.38) in Kirchhoff's current law (A.35), and using the relation (A.33) yields

$$-A^\top Y_{\text{br}} v_{\text{br}} + i = -A^\top Y_{\text{br}} A v + i = 0_n \iff i = A^\top Y_{\text{br}} A v =: Y v. \quad (\text{A.39})$$

The matrix  $Y$  is the bus admittance matrix; it combines topological information of the power system with physical values of the electrical lines. Recall that  $\text{rank } A = n - 1$  and consequently  $\text{rank } Y = n - 1$ . Also note that  $Y$  is symmetric whenever  $Y_{\text{br}}$  is symmetric, and that  $Y$  corresponds to the weighted Laplacian of the graph in case it is symmetric.

Thus, to find the steady state of an electrical network, the incidence matrix  $A$  and the branch admittance matrix  $Y_{\text{br}}$  have to be specified. For example, in case we model a transmission line by its  $\pi$ -line equivalent, then two branches have to account for the ground connection, i.e.

$$Y_{\text{br}} = \begin{bmatrix} y & 0 & 0 \\ 0 & \frac{y_l}{2} & 0 \\ 0 & 0 & \frac{y_l}{2} \end{bmatrix}, \quad A = \begin{bmatrix} 0 & -1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix}, \quad Y = \begin{bmatrix} y_l & -\frac{y_l}{2} & -\frac{y_l}{2} \\ -\frac{y_l}{2} y + \frac{y_l}{2} & -y & -y \\ -\frac{y_l}{2} & -y & y + \frac{y_l}{2} \end{bmatrix}.$$

For a transformer with a specified (real) tap setting  $t \in \mathbb{R}$ , the matrices are

$$Y_{\text{br}} = \begin{bmatrix} ty & 0 & 0 \\ 0 & t(t-1)y & 0 \\ 0 & 0 & (1-t)y \end{bmatrix}, \quad A = \begin{bmatrix} 0 & -1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix},$$

$$Y = \begin{bmatrix} (t-1)^2 y & t(1-t)y & (t-1)y \\ t(1-t)y & t^2 y & -ty \\ (t-1)y & -ty & y \end{bmatrix}.$$

To overcome the rank deficiency of the bus admittance matrix, a reference bus needs to be specified.

There exist explicit formulæ for constructing the bus admittance matrix, including all nodal shunts and line shunts as well as transformers. We refer to [8, 61].

## A.6 AC and DC power flow equations

In Section 3.1 we derived the AC power flow equations based on the branch flows. An alternative is to use the bus admittance matrix from Appendix A.5. Recall from Section 3.1 that the state of every bus  $i \in \mathcal{N}$  is described by its complex power and its complex voltage. Using rectangular coordinates for the complex power and polar coordinates for the complex voltage we define

$$z_i^{\text{AC}} = \left[ p_i \ q_i \ v_i \ \delta_i \right]^\top \in \mathbb{R}^4 \quad (\text{A.40})$$

to be the state of bus  $i \in \mathcal{N}$ ; the voltage magnitude at bus  $i$  is  $v_i$ , the corresponding voltage angle is  $\delta_i$ . With the bus admittance matrix  $Y = G + jB \in \mathbb{C}^{N_b \times N_b}$ , see Appendix A.5, the power injection at every bus is<sup>5</sup>

$$p_i - jq_i = v_i e^{-j\delta_i} \sum_{k \in \mathcal{N}} Y_{ik} v_k e^{j\delta_k}, \quad (\text{A.41})$$

see Figure 3.2. We separate the real and imaginary parts to obtain

$$p_i = \sum_{j \in \mathcal{N}} v_i v_j (G_{ij} \cos(\delta_i - \delta_j) + B_{ij} \sin(\delta_i - \delta_j)), \quad (\text{A.42a})$$

$$q_i = \sum_{j \in \mathcal{N}} v_i v_j (G_{ij} \sin(\delta_i - \delta_j) - B_{ij} \cos(\delta_i - \delta_j)). \quad (\text{A.42b})$$

To highlight the difference to the branch flow model from Section 3.1 the formulation (A.42) is called the bus injection model. Regardless of the formulation, the AC power flow equations constitute a system of nonlinear algebraic equations.

Let us make the following simplifying assumption.

**Assumption A.1** (DC power flow). *We study a power system under Assumption 3.1. Let the state of every bus  $i \in \mathcal{N}$  be given by (A.40), and let the following conditions hold:*

---

<sup>5</sup> The bus admittance matrix can be constructed to account for nodal and line shunts [8, 61].

AA.1 the Ohmic losses across each line  $(j, k) \in \mathcal{L}$  are negligible such that  $g_{jk} = 0$  holds for the  $\Pi$ -line model from Figure 3.1;

AA.2 the voltage angle differences  $(\delta_j - \delta_k)$  are small across all lines  $(j, k) \in \mathcal{L}$ ;

AA.3 the voltage magnitudes  $v_i$  are constant at one per unit for all buses  $i \in \mathcal{N}$ .

A word of caution: the wording “DC” is unfortunate. We are still in an AC setting, however the equations mathematically resemble Ohm’s law from the “true” DC conditions, hence the name. The consequences of Assumption A.1 for the AC power flow equations according to (A.42) are immediate:

$$p_i = \sum_{i \in \mathcal{N}} B_{ij}(\delta_i - \delta_j), \quad (\text{A.43})$$

$$q_i = \sum_{i \in \mathcal{N}} -B_{ij} = \text{const}. \quad (\text{A.44})$$

Hence, for every bus  $i \in \mathcal{N}$  both the voltage magnitude and the reactive power become constants, and

$$z_i^{\text{DC}} = \begin{bmatrix} p_i & \delta_i \end{bmatrix}^{\top} \in \mathbb{R}^2 \quad (\text{A.45})$$

fully characterizes the state of bus  $i$  under DC conditions. From (A.43) we obtain a relation between the vector of net active powers  $p = [p_1 \dots p_{N_b}]^{\top} \in \mathbb{R}^{N_b}$  and the vector of phase angles  $\delta = [\delta_1 \dots \delta_{N_b}]^{\top} \in \mathbb{R}^{N_b}$ , namely

$$p = -B\delta, \quad (\text{A.46})$$

where  $B = \text{Im}(Y) = A^{\top} B^{\text{br}} A \in \mathbb{R}^{N_b \times N_b}$  is the imaginary part of the bus admittance matrix, and  $B^{\text{br}} = \text{Im}(Y_{\text{br}}) \in \mathbb{R}^{N_{\text{br}} \times N_{\text{br}}}$  is the imaginary part of the branch admittance matrix from (A.38), where  $N_{\text{br}} = |\mathcal{L}|$  is the number of lines. We see: in the DC setting the power flow equations constitute  $N_b$  linear equations. If we know the voltage angles  $\delta$  we can compute the net bus injections  $p$ . Can we also compute the voltage angles from the net bus injections given that  $B$  is rank deficient? We can answer in the affirmative upon introducing a slack bus: we assume the system has a single slack bus  $i \in$

$\mathcal{N}_{\text{sl}} = \{1\}$  at bus number one for which  $\delta_1 = 0$  must hold. In that case we can write

$$\hat{\delta} = -\hat{B}^{-1}\hat{p}, \quad (\text{A.47})$$

where  $\hat{\delta} \in \mathbb{R}^{N_b-1}$  is  $\delta$  with the first entry—corresponding to the slack-removed,  $\hat{B} \in \mathbb{R}^{(N_b-1) \times (N_b-1)}$  is  $B$  with the first row and the first column removed, and  $\hat{p} \in \mathbb{R}^{N_b-1}$  corresponds to  $p$  with the first entry removed.

According to AA.1 from Assumption A.1 there are no Ohmic losses. Intuitively, this means that the sum of the net power injections must add to zero. This is true indeed

$$\sum_{i \in \mathcal{N}} p_i = \mathbf{1}_{N_b}^\top p_i = -\mathbf{1}_{N_b}^\top B \delta = 0 \cdot \delta, \quad (\text{A.48})$$

because  $\mathbf{1}_{N_b}$  is a left-eigenvector to the zero eigenvalue of the graph Laplacian  $B$ ; this follows from (A.39) in combination with (A.34).

From Assumption A.1 we can draw another conclusion: the power flows across the lines are linearly related to the net bus power injections. To see this we first write the steady state energy balance using the incidence matrix  $A$

$$0_{N_b} = A^\top p^{\text{br}} + p = A(p^{\text{br}} - B^{\text{br}} A \delta), \quad (\text{A.49})$$

where  $p^{\text{br}} \in \mathbb{R}^{N_{\text{br}}}$  is the vector of active power flows.<sup>6</sup> From (A.49) we get

$$p^{\text{br}} = B^{\text{br}} A \delta, \quad (\text{A.50})$$

which relates the voltage angles linearly to the branch flows. We assume the system has a single slack bus  $i \in \mathcal{N}_{\text{sl}} = 1$  at bus number one for which  $\delta_1 = 0$  must hold. In that case we can write

$$p^{\text{br}} = B^{\text{br}} A \delta = B^{\text{br}} \hat{A} \hat{\delta}, \quad (\text{A.51})$$

---

<sup>6</sup> The entries of  $p^{\text{br}}$  correspond to the row-wise entries of the incidence matrix  $A$ .

where  $\hat{A} \in \mathbb{Z}^{|\mathcal{L}| \times (N_b - 1)}$  is the incidence matrix  $A$  with the first column removed, and  $\hat{\delta} \in \mathbb{R}^{N_b - 1}$  is  $\delta$  with the first entry—corresponding to the slack—removed. We can use the DC power flow equation (A.46) to write

$$\hat{\delta} = -(\hat{A}^\top B^{\text{br}} \hat{A})^{-1} \hat{p}, \quad (\text{A.52})$$

where  $\hat{p} \in \mathbb{R}^{N_b - 1}$  corresponds to  $p$  with the first entry removed. We can combine (A.52) with (A.51) and obtain

$$p^{\text{br}} = -B^{\text{br}} \hat{A} (\hat{A}^\top B^{\text{br}} \hat{A})^{-1} \hat{p} = \underbrace{\begin{bmatrix} 0_{N_{\text{br}}} & -B^{\text{br}} \hat{A} (\hat{A}^\top B^{\text{br}} \hat{A})^{-1} \end{bmatrix}}_{=: \Psi} p. \quad (\text{A.53})$$

The matrix  $\Psi \in \mathbb{R}^{N_{\text{br}} \times N_b}$  is the so-called power transfer distribution factor (PTDF) matrix which maps the net power injections linearly to the line flows.

## A.7 Backward-forward sweep method

The setting for the backward-forward sweep method is described in the beginning of Section 3.4. We derive the BFS method with the help of the following ingredients:

- Governing equations

- ▷ Kirchhoff's current law is given by, see A.5,

$$0 = A_r^\top i^{\text{br}} + i - i^{\text{sh}}, \quad (\text{A.54})$$

where  $i^{\text{br}} \in \mathbb{C}^N$  is the vector of branch currents and  $i \in \mathbb{C}^N$  is the vector of nodal currents, and  $i^{\text{sh}} \in \mathbb{C}^N$  is the vector of branch shunt currents.

- ▷ Kirchhoff's voltage law in terms of the reduced incidence matrix is given by

$$v^{\text{br}} = A_r v - v_{\text{sl}} a_0, \quad (\text{A.55})$$

where  $v^{\text{br}} \in \mathbb{C}^N$  is the vector of branch voltages,  $v \in \mathbb{C}^N$  is the vector of nodal voltages,  $v_{\text{sl}} \in \mathbb{C}$  is the voltage at the root node, and  $e_1 \in \mathbb{R}^N$  is the first unit vector of  $\mathbb{R}^N$ .

– Constitutive laws and load/shunt modeling:

- ▷ The loads are modeled in terms of some function  $f : \mathbb{C}^N \rightarrow \mathbb{C}^N$  with  $i = f(v)$ .
- ▷ The branch shunt currents are modeled in terms of some function  $h : \mathbb{C}^N \rightarrow \mathbb{C}$  with  $i^{\text{sh}} = h(v)$ .
- ▷ The constitutive equation  $g : \mathbb{C}^N \rightarrow \mathbb{C}^N$  relates potential differences to flows with  $i^{\text{br}} = g(v^{\text{br}})$ .

The overall scheme of the BFS method is shown in Figure 3.5: it shows how Kirchhoff’s laws are employed together with the additional modeling equations. Let us now derive the relevant equations for the BFS method.

### Backward sweep

In the backward sweep, Kirchhoff’s current law and the load model are used to compute the branch currents from given nodal voltages, i.e.

$$i^{\text{br}} = (A_r^\top)^{-1}(i^{\text{sh}} - i) = (A_r^\top)^{-1}(h(v) - f(v)). \quad (\text{A.56})$$

Using constant impedance loads  $f(v) = Z_{\text{const}}^{-1}v$  we have

$$i^{\text{br}} = (A_r^\top)^{-1}(h(v) - Z_{\text{const}}^{-1}v). \quad (\text{A.57})$$

If instead we use a constant power model and account for bus shunt currents, we have (the superscript  $*$  denotes the complex conjugate)

$$\begin{aligned} s_k &= v_k \hat{i}_k^* \\ i_k &= \hat{i}_k + y_k^{\text{sh}} v_k \end{aligned} \quad (\text{A.58})$$

and thus

$$i_k = \frac{s_k^*}{v_k^*} + y_k^{\text{sh}} v_k \quad (\text{A.59})$$

for every node  $k \in \{1, \dots, N_b\}$ . Then, the load model function  $f$  follows from

$$\begin{bmatrix} i_1 \\ i_2 \\ \vdots \\ i_{N_b} \end{bmatrix} = \begin{bmatrix} 1/v_1^* & 0 & \dots \\ 0 & 1/v_2^* & \\ \vdots & & \ddots \\ & & & 1/v_{N_b}^* \end{bmatrix} \begin{bmatrix} s_1^* \\ s_2^* \\ \vdots \\ s_{N_b}^* \end{bmatrix} + \begin{bmatrix} y_1^{\text{sh}} & 0 & \dots \\ 0 & y_2^{\text{sh}} & \\ \vdots & & \ddots \\ & & & y_{N_b}^{\text{sh}} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_{N_b} \end{bmatrix}. \quad (\text{A.60})$$

The total branch shunt current at node  $k$  is given by

$$i_k^{\text{sh}} = \sum_{j \in \text{adj}(k)} y_{jk}^{\text{br,sh}} v_k = \sum_{j \in \text{adj}(k)} j \frac{|b_{jk}^{\text{br,sh}}|}{2} v_k, \quad (\text{A.61})$$

where the set  $\text{adj}(k)$  contains all buses adjacent to bus  $k$ . From this, the function  $i^{\text{sh}} = h(v)$  becomes

$$i^{\text{sh}} = (Q \mathbf{1}_{N+1})_r * v, \quad (\text{A.62})$$

where  $Q \in \mathbb{C}^{(N+1) \times (N+1)}$  is the branch shunt adjacency matrix,<sup>7</sup> and  $(\cdot)_r$  refers to reducing, i.e. dropping the root node index. Note that  $*$  means element-wise multiplication (Hadamard product).

The term *backwards* stems from upper-triangularity of the matrix  $(A_r^\top)^{-1}$ : the leaf node branch currents are computed first, which are then used to find the branch currents of the respective parents. Just like when solving systems of linear equations using LU factorization the solutions are backward-substituted.

---

<sup>7</sup> I.e.  $A^\top Y_{\text{sh}} A = D - Q$ , where  $D$  is the degree matrix and  $Y_{\text{sh}} \in \mathbb{C}^{(N+1) \times (N+1)}$  is the branch shunt admittance matrix, usually  $Y_{\text{sh}} = j/2 \text{diag}(|b_{jk}^{\text{br,sh}}|)$  for all lines  $(j, k) \in \mathcal{L}$ .

**Forward sweep**

In the forward sweep, Kirchhoff's voltage law and the constitutive equation are used to compute the nodal voltages given the branch currents, i.e.

$$v = A_r^{-1}(v^{\text{br}} + v_{\text{SL}}a_0) = A_r^{-1}(g^{-1}(i^{\text{br}}) + v_{\text{SL}}a_0) = A_r^{-1}g^{-1}(i^{\text{br}}) + v_{\text{SL}}\mathbf{1}_{N_b}. \quad (\text{A.63})$$

Using Ohm's law  $g(v^{\text{br}}) = -Y_{\text{br}}v^{\text{br}} = -(Z^{\text{br}})^{-1}v^{\text{br}}$ , see Appendix A.5, we have

$$v = -A_r^{-1}Z^{\text{br}}i^{\text{br}} + v_{\text{SL}}\mathbf{1}_{N_b}. \quad (\text{A.64})$$

The term *forward* stems from lower-triangularity of the matrix  $A_r^{-1}$ : the nodal voltage of the first node can be readily computed, which is then used to compute the nodal voltages of its leaves. Just like when solving systems of linear equations using LU factorization the solutions are forward-substituted.

---

**Algorithm 3:** BFS method.

---

**Data:** load model  $f(\cdot)$ ; shunt model  $h(\cdot)$ ; constitutive law  $g(\cdot)$ ;  $v_{\text{SL}}$ ;  $A_r$ ;  $v_{\text{init}}$ ;  $\varepsilon > 0$ .

**Result:**  $v, i^{\text{br}}$ .

Set  $\tau \leftarrow 0$ ;

Initialize  $v(\tau) = v_{\text{init}}$ ;

**do**

    # Backward sweep

    Evaluate

$$i^{\text{br}}(\tau) = (A_r^{-1})^\top (h(v(\tau)) - f(v(\tau)));$$

    # Forward sweep

    Evaluate

$$v(\tau + 1) = A_r^{-1}g^{-1}(i^{\text{br}}(\tau)) + v_{\text{SL}}\mathbf{1}_{N_b};$$

    Set  $\tau \leftarrow \tau + 1$ ;

**while**  $\|v(\tau + 1) - v(\tau)\| > \varepsilon$ ;

**Return**  $v \leftarrow v(\tau + 1), i^{\text{br}} \leftarrow i^{\text{br}}(\tau)$ ;

---

### Convergence analysis

We usually initiate any BFS scheme with a flat start by setting the bus voltages equal to the voltage at the root node

$$v_{\text{sl}} = v_1(0) = \dots = v_{N_b}(0), \quad (\text{A.65})$$

where the superscript denotes the iteration index. Then, we repeatedly apply the backward sweep and the forward sweep until a convergence criterion is satisfied, e.g.  $\|v(k+1) - v(k)\|_\infty < \varepsilon$ .

We can study convergence of the BFS method easily in case the load models are constant impedance loads (A.57), and in the absence of shunts. In that case, we can merge the backward and the forward sweep to obtain a discrete-time linear time-invariant system in terms of the bus voltages

$$v(k+1) = (A_r^\top Y_{\text{br}} A_r)^{-1} Z_{\text{const}}^{-1} v(k) + v_{\text{sl}} \mathbf{1}_{N_b}, \quad v(0) = v_{\text{sl}} \mathbf{1}_{N_b}. \quad (\text{A.66})$$

We see that BFS converges in case the magnitude of every eigenvalue of  $(A_r^\top Y_{\text{br}} A_r)^{-1} Z_{\text{const}}^{-1}$  is inside the unit circle. The steady state bus voltage  $\bar{v}$  is then

$$\bar{v} = v_{\text{sl}} (I_{N_b} - (A_r^\top Y_{\text{br}} A_r)^{-1} Z_{\text{const}}^{-1})^{-1} \mathbf{1}_{N_b}. \quad (\text{A.67})$$



# References

- [1] „20% wind energy by 2030: Increasing wind energy’s contribution to U.S. electricity supply“, Department of Energy (DOE), Tech. Rep. DOE/GO-102008-2567, 2008.
- [2] B. Adams, L. Bauman, W. Bohnhoff, K. Dalbey, M. Ebeida, J. Eddy, M. Eldred, P. Hough, K. Hu, J. Jakeman, J. Stephens, L. Swiler, D. Vigil, and T. Wildey, „Dakota, a multilevel parallel object-oriented framework for design optimization, parameter estimation, uncertainty quantification, and sensitivity analysis: version 6.0 user’s manual“, Sandia National Lab SAND2014-4633, Tech. Rep., 2014.
- [3] H. Ahmadi, J. R. Martí, and A. von Meier, „A linear power flow formulation for three-phase distribution systems“, *IEEE Transactions on Power Systems*, vol. 31, no. 6, pp. 5012–5021, Nov. 2016.
- [4] 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, no. 4, pp. 2233–2241, Nov. 2012.
- [5] M. Alipour, B. Mohammadi-Ivatloo, and K. Zare, „Stochastic scheduling of renewable and CHP-based microgrids“, *IEEE Transactions on Industrial Informatics*, vol. 11, no. 5, pp. 1049–1058, Oct. 2015.
- [6] R. Allan, A. Leite Da Silva, and R. Burchett, „Evaluation methods and accuracy in probabilistic load flow solutions“, *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-100, no. 5, pp. 2539–2546, May 1981.
- [7] J. An and A. Owen, „Quasi-regression“, *Journal of Complexity*, vol. 17, no. 4, pp. 588 –607, 2001.
- [8] G. Andersson, *Lecture notes on power system analysis, power flow analysis, fault analysis, power system dynamics and stability*, Sep. 2015.
- [9] R. Appino, T. Mühlfordt, T. Faulwasser, and V. Hagenmeyer, „On solving probabilistic load flow for radial grids using polynomial chaos“, in *2017 IEEE Manchester PowerTech*, Jun. 2017, pp. 1–6.
- [10] G. Arfken, *Mathematical Methods for Physicists*, 3rd ed. Orlando, Florida: Academic Press, 1985.
- [11] F. Augustin, A. Gilg, M. Paffrath, P. Rentrop, and U. Wever, „Polynomial chaos for the approximation of uncertainties: chances and limits“, *European Journal of Applied Mathematics*, vol. 19, no. 2, pp. 149–190, Apr. 2008.
- [12] C. Auxenfans, R. Baumann, T. Kapetanovic, P. De Leener, R. Paprocki, A. Wirth, and D. Klaar, „European transmission system operators’ cooperation for enhanced system security, integration of renewables and market support“, Cigré, Paris, Tech. Rep. C2-113, 2014.
- [13] M. Avila Rosales, „Wind power grid integration and cross-border export in Mexico“, Cigré, Paris, Tech. Rep. C5-310, 2014.
- [14] K. Baker, E. Dall’Anese, and T. Summers, „Distribution-agnostic stochastic optimal power flow for distribution grids“, in *2016 North American Power Symposium (NAPS)*, Sep. 2016, pp. 1–6.
- [15] M. Baran and F. Wu, „Optimal sizing of capacitors placed on a radial distribution system“, *IEEE Transactions on Power Delivery*, vol. 4, no. 1, pp. 735–743, 1989.
- [16] M. Baudin, A. Duffoy, B. Iooss, and A.-L. Popelin, „OpenTURNS: An industrial software for uncertainty quantification in simulation“, in *Handbook of Uncertainty Quantification*, R. Ghanem, D. Higdon, and H. Owahdi, Eds. Cham: Springer International Publishing, 2017.

- [17] F. Becker, „A Julia framework for uncertainty quantification using polynomial chaos expansion“, Master’s thesis, Karlsruhe Institute of Technology, 2017.
- [18] A. Bemporad, M. Morari, V. Dua, and E. Pistikopoulos, „The explicit solution of model predictive control via multiparametric quadratic programming“, in *Proceedings of the 2000 American Control Conference (ACC)*, vol. 2, 2000, pp. 872–876.
- [19] J. Bezanson, A. Edelman, S. Karpinski, and V. Shah, „Julia: A fresh approach to numerical computing“, *SIAM Review*, vol. 59, no. 1, pp. 65–98, 2017.
- [20] D. Bienstock, M. Chertkov, and S. Harnett, „Chance-constrained optimal power flow: Risk-aware network control under uncertainty“, *SIAM Review*, vol. 56, no. 3, pp. 461–495, 2014.
- [21] D. Bienstock and A. Shukla, „Variance-aware optimal power flow“, in *2018 Power Systems Computation Conference (PSCC)*, Jun. 2018, pp. 1–8.
- [22] —, „Variance-aware optimal power flow: Addressing the trade-off between cost, security and variability“, *IEEE Transactions on Control of Network Systems*, pp. 1–1, 2019.
- [23] G. Blatman, „Adaptive sparse polynomial chaos expansions for uncertainty propagation and sensitivity analysis“, Dissertation, Université Blaise Pascal, 2009.
- [24] G. Blatman and B. Sudret, „Adaptive sparse polynomial chaos expansion based on least angle regression“, *Journal of Computational Physics*, vol. 230, no. 6, pp. 2345–2367, 2011.
- [25] —, „An adaptive algorithm to build up sparse polynomial chaos expansions for stochastic finite element analysis“, *Probabilistic Engineering Mechanics*, vol. 25, no. 2, pp. 183–197, 2010.
- [26] E. Bompard, E. Carpaneto, G. Chicco, and R. Napoli, „Convergence of the backward/forward sweep method for the load-flow analysis of radial distribution systems“, *International Journal of Electrical Power & Energy Systems*, vol. 22, no. 7, pp. 521–530, 2000.
- [27] B. Borkowska, „Probabilistic load flow“, *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-93, no. 3, pp. 752–759, May 1974.
- [28] I. Bronstein, H. Mühlig, G. Musiol, and K. Semendjajew, *Taschenbuch der Mathematik*. Europa Lehrmittel Verlag, 2013.
- [29] „Bundesgesetzblatt“, Bundesanzeiger Verlag GmbH, Tech. Rep. G 5702 Teil I, Nr. 23, 1998.
- [30] „Bundesgesetzblatt“, Bundesanzeiger Verlag GmbH, Tech. Rep. G 5702 Teil I, Nr. 13, 2000.
- [31] „Bundesgesetzblatt“, Bundesanzeiger Verlag GmbH, Tech. Rep. G 5702 Teil I, Nr. 43, 2011.
- [32] G. Calafiore and L. El Ghaoui, „On distributionally robust chance-constrained linear programs“, *Journal of Optimization Theory and Applications*, vol. 130, no. 1, pp. 1–22, 2006.
- [33] M. Campi and G. Calafiore, „Decision making in an uncertain environment: The scenario-based optimization approach“, *Working paper, Università di Brescia, Brescia, Italy*, 2004.
- [34] M. Capiński and P. Kopp, *Measure, Integral and Probability*, 2nd, ser. Springer undergraduate mathematics series. London, Berlin, Heidelberg: Springer-Verlag, 2004.
- [35] J. Carpentier, „Contribution to the economic dispatch problem“, *Bulletin de la Société Française des Electriciens*, vol. 3, no. 8, pp. 431–447, 1962.
- [36] P. Chen, Z. Chen, and B. Bak-Jensen, „Probabilistic load flow: A review“, in *2008 Third International Conference on Electric Utility Deregulation and Restructuring and Power Technologies*, Apr. 2008, pp. 1586–1591.
- [37] Y. Chen, J. Wen, and S. Cheng, „Probabilistic load flow method based on nataf transformation and latin hypercube sampling“, *IEEE Transactions on Sustainable Energy*, vol. 4, no. 2, pp. 294–301, Apr. 2013.
- [38] C. Cheng and D. Shirmohammadi, „A three-phase power flow method for real-time distribution system analysis“, *IEEE Transactions on Power Systems*, vol. 10, no. 2, pp. 671–679, May 1995.

- [39] P. Conrad and Y. Marzouk, „Adaptive Smolyak pseudospectral approximations“, *SIAM Journal on Scientific Computing*, vol. 35, no. 6, A2643–A2670, 2013.
- [40] S. Conti and S. Raiti, „Probabilistic load flow using Monte Carlo techniques for distribution networks with photovoltaic generators“, *Solar Energy*, vol. 81, no. 12, pp. 1473–1481, 2007.
- [41] D. Cox, J. Little, and D. O’Shea, *Ideals, Varieties, and Algorithms, An Introduction to Computational Algebraic Geometry and Commutative Algebra*, 12th ed. New York: Springer-Verlag, 1997.
- [42] —, *Using Algebraic Geometry*. New York: Springer Science+Business Media, LLC, 1998.
- [43] E. Dall’Anese, K. Baker, and T. Summers, „Chance-constrained AC optimal power flow for distribution systems with renewables“, *IEEE Transactions on Power Systems*, vol. 32, no. 5, pp. 3427–3438, Sep. 2017.
- [44] E. Dall’Anese, K. Baker, and T. Summers, „Optimal power flow for distribution systems under uncertain forecasts“, in *2016 IEEE 55th Conference on Decision and Control (CDC)*, Dec. 2016, pp. 7502–7507.
- [45] S. Dasgupta, D. Hsu, and N. Verma, „A concentration theorem for projections“, in *22nd Conference on Uncertainty in Artificial Intelligence (UAI), 2006*.
- [46] B. Debusschere, K. Sargsyan, C. Safta, and K. Chowdhary, „Uncertainty quantification toolkit (UQTK)“, in *Handbook of Uncertainty Quantification*, R. Ghanem, D. Higdon, and H. Owaldi, Eds. Cham: Springer International Publishing, 2016, pp. 1–21.
- [47] B. Debusschere, H. Najm, P. Pébay, O. Knio, R. Ghanem, and O. Le Maître, „Numerical challenges in the use of polynomial chaos representations for stochastic processes“, *SIAM Journal on Scientific Computing*, vol. 26, no. 2, pp. 698–719, 2004.
- [48] H. Ding, P. Pinson, Z. Hu, and Y. Song, „Optimal offering and operating strategies for wind-storage systems with linear decision rules“, *IEEE Transactions on Power Systems*, vol. 31, no. 6, pp. 4755–4764, Nov. 2016.
- [49] J. Dopazo, O. Klitin, and A. Sasson, „Stochastic load flows“, *IEEE Transactions on Power Apparatus and Systems*, vol. 94, no. 2, pp. 299–309, Mar. 1975.
- [50] C. Duan, W. Fang, L. Jiang, L. Yao, and J. Liu, „Distributionally robust chance-constrained approximate AC-OPF with Wasserstein metric“, *IEEE Transactions on Power Systems*, vol. 33, no. 5, pp. 4924–4936, Sep. 2018.
- [51] I. Dunning, J. Huchette, and M. Lubin, „JuMP: A modeling language for mathematical optimization“, *SIAM Review*, vol. 59, no. 2, pp. 295–320, 2017.
- [52] U. Eminoglu and M. Hocaoglu, „Distribution systems forward/backward sweep-based power flow algorithms: A review and comparison study“, *Electric Power Components and Systems*, vol. 37, no. 1, pp. 91–110, 2008.
- [53] A. Engelmann, Y. Jiang, T. Mühlpfordt, B. Houska, and T. Faulwasser, „Toward distributed OPF using ALADIN“, *IEEE Transactions on Power Systems*, vol. 34, no. 1, pp. 584–594, Jan. 2019.
- [54] A. Engelmann, T. Mühlpfordt, Y. Jiang, B. Houska, and T. Faulwasser, „Distributed AC optimal power flow using ALADIN“, *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 5536–5541, 2017, 20th IFAC World Congress.
- [55] —, „Distributed stochastic AC optimal power flow based on polynomial chaos expansion“, in *2018 Annual American Control Conference (ACC)*, Jun. 2018, pp. 6188–6193.
- [56] P. Esfahani and D. Kuhn, „Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations“, *Mathematical Programming*, vol. 171, no. 1, pp. 115–166, Sep. 2018.
- [57] L. Fagiano and M. Khammash, „Nonlinear stochastic model predictive control via regularized polynomial chaos expansions“, in *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*, Dec. 2012, pp. 142–147.

- [58] T. Faulwasser, A. Engelmann, T. Mühlpfordt, and V. Hagenmeyer, „Optimal power flow: An introduction to predictive, distributed and stochastic control challenges“, *at – Automatisierungstechnik*, vol. 66, no. 7, pp. 573–589, Jan. 2018.
- [59] J. Feinberg and H. Langtangen, „Chaospy: An open source tool for designing methods of uncertainty quantification“, *Journal of Computational Science*, vol. 11, pp. 46–57, 2015.
- [60] R. Field Jr. and M. Grigoriu, „On the accuracy of the polynomial chaos approximation“, *Probabilistic Engineering Mechanics*, vol. 19, no. 1–2, pp. 65–80, 2004, Fourth International Conference on Computational Stochastic Mechanics.
- [61] S. Frank and S. Rebennack, „An introduction to optimal power flow: Theory, formulation, and examples“, *IEEE Transactions*, vol. 48, no. 12, pp. 1172–1197, 2016.
- [62] S. Frank, I. Steponavice, and S. Rebennack, „Optimal power flow: A bibliographic survey I“, *Energy Systems*, vol. 3, no. 3, pp. 221–258, 2012.
- [63] —, „Optimal power flow: A bibliographic survey II“, *Energy Systems*, vol. 3, no. 3, pp. 259–289, Sep. 2012.
- [64] W. Gautschi, *Orthogonal Polynomials: Computation and Approximation*. Oxford: Oxford University Press, 2004.
- [65] —, „The interplay between classical analysis and (numerical) linear algebra – A tribute to gene h. golub“, *Electronic Transactions on Numerical Analysis*, vol. 13, pp. 119–147, 2002.
- [66] A. Georghiou, D. Kuhn, and W. Wiesemann, „The decision rule approach to optimization under uncertainty: Methodology and applications“, *Computational Management Science*, Nov. 2018.
- [67] R. Ghanem and P. Spanos, *Stochastic Finite Elements*, Rev. ed. Mineola, New York: Dover Publications, Inc., 1991.
- [68] C. Goebel, H.-A. Jacobsen, V. del Razo, C. Doblander, J. Rivera, J. Ilg, C. Flath, H. Schmeck, C. Weinhardt, D. Pathmaperuma, H.-J. Appelrath, M. Sonnenschein, S. Lehnhoff, O. Kramer, T. Staake, E. Fleisch, D. Neumann, J. Strüker, K. Ereke, R. Zarnekow, H. Ziekow, and J. Lässig, „Energieinformatik“, *Wirtschaftsinformatik*, vol. 56, no. 1, pp. 31–39, 2014.
- [69] G. Golub, „Some modified matrix eigenvalue problems“, *SIAM Review*, vol. 15, no. 2, pp. 318–334, 1973.
- [70] G. Golub and C. Van Loan, *Matrix Computations*, Fourth. Baltimore, Maryland: The Johns Hopkins University Press, 1983.
- [71] G. Golub and J. Welsch, „Calculation of Gauss quadrature rules“, *Mathematics of Computation*, pp. 221–230, 23 1969.
- [72] V. Goncharov, „The theory of best approximation of functions“, *Journal of Approximation Theory*, vol. 106, no. 1, pp. 2–57, 2000.
- [73] J. González Ordiano, W. Doneit, S. Waczowicz, L. Gröll, R. Mikut, and V. Hagenmeyer, „Nearest-neighbor based non-parametric probabilistic forecasting with applications in photovoltaic systems“, *arXiv e-prints*, arXiv:1701.06463, arXiv:1701.06463, 2017.
- [74] J. González Ordiano, S. Waczowicz, V. Hagenmeyer, and R. Mikut, „Energy forecasting tools and services“, *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 8, no. 2, e1235, 2018.
- [75] J. González Ordiano, S. Waczowicz, M. Reischl, R. Mikut, and V. Hagenmeyer, „Photovoltaic power forecasting using simple data-driven models without weather data“, *Computer Science – Research and Development*, vol. 32, no. 1, pp. 237–246, 2017.
- [76] N. Good and P. Mancarella, „Flexibility in multi-energy communities with electrical and thermal storage: A stochastic, robust approach for multi-service demand response“, *IEEE Transactions on Smart Grid*, vol. 10, no. 1, pp. 503–513, Jan. 2019.

- [77] W. Gragg and W. Harrod, „The numerically stable reconstruction of Jacobi matrices from spectral data“, *Numerische Mathematik*, vol. 44, no. 3, pp. 317–335, 1984.
- [78] J. Grainger and W. Stevenson, *Power System Analysis*. McGraw-Hill Education, 1994.
- [79] Y. Guo, K. Baker, E. Dall’Anese, Z. Hu, and T. Summers, „Stochastic optimal power flow based on data-driven distributionally robust optimization“, in *2018 Annual American Control Conference (ACC)*, Jun. 2018, pp. 3840–3846.
- [80] Y. Guo, K. Baker, E. Dall’Anese, Z. Hu, and T. Summers, „Data-based distributionally robust stochastic optimal power flow – Part I: Methodologies“, *IEEE Transactions on Power Systems*, vol. 34, no. 2, pp. 1483–1492, Mar. 2019.
- [81] —, „Data-based distributionally robust stochastic optimal power flow – Part II: Case studies“, *IEEE Transactions on Power Systems*, vol. 34, no. 2, pp. 1493–1503, Mar. 2019.
- [82] R. Hable, *Einführung in die Stochastik. Ein Begleitbuch zur Vorlesung*. Berlin, Heidelberg: Springer Spektrum, 2015.
- [83] J. Hoffmann-Jørgensen, *Probability with a View Toward Statistics*. New York, London: Chapman & Hall, 1994, vol. I.
- [84] —, *Probability with a View Toward Statistics*. New York, London: Chapman & Hall, 1994, vol. II.
- [85] M. Hohmann, J. Warrington, and J. Lygeros, „Optimal linearizations of power systems with uncertain supply and demand“, *IEEE Transactions on Power Systems*, vol. 34, no. 2, pp. 1504–1512, Mar. 2019.
- [86] H. Hong, „An efficient point estimate method for probabilistic analysis“, *Reliability Engineering & System Safety*, vol. 59, no. 3, pp. 261–267, 1998.
- [87] A. Hunt and D. Thomas, *The Pragmatic Programmer: From Journeyman to Master*. Reading, MA, USA: Addison-Wesley Longman Inc., 2000.
- [88] E. Janecek and D. Georgiev, „Probabilistic extension of the backward/forward load flow analysis method“, *IEEE Transactions on Power Systems*, vol. 27, no. 2, pp. 695–704, May 2012.
- [89] I. Jha, Y. Sehgal, S. Sen, and K. Bhambhani, „Grid integration of large scale renewable generation – Initiatives in Indian power system“, Cigré, Paris, Tech. Rep. C1-205, 2014.
- [90] P. Kall, „Stochastic programming“, *European Journal of Operational Research*, vol. 10, no. 2, pp. 125–130, 1982.
- [91] P. Kall and S. Wallace, *Stochastic Programming*, Second. Chichester: John Wiley & Sons, Inc., 1994.
- [92] K. Kim and R. Braatz, „Generalised polynomial chaos expansion approaches to approximate stochastic model predictive control“, *International Journal of Control*, vol. 86, no. 8, pp. 1324–1337, 2013.
- [93] „Kommission Wachstum, Strukturwandel und Beschäftigung“, Bundesministerium für Wirtschaft und Energie (BMWi), Tech. Rep., 2019.
- [94] C. Kuehn, „Moment Closure – A Brief Review“, *arXiv e-prints*, arXiv:1505.02190, arXiv:1505.02190, 2015.
- [95] O. Le Maître and O. Knio, *Spectral Methods for Uncertainty Quantification with Applications to Computational Fluid Dynamics*, ser. Scientific Computation. Dordrecht Heidelberg London New York: Springer Science+Business Media, 2010.
- [96] B. Li, R. Jiang, and J. Mathieu, „Distributionally robust risk-constrained optimal power flow using moment and unimodality information“, in *2016 IEEE 55th Conference on Decision and Control (CDC)*, Dec. 2016, pp. 2425–2430.
- [97] B. Li, J. Mathieu, and R. Jiang, „Distributionally robust chance constrained optimal power flow assuming log-concave distributions“, in *2018 Power Systems Computation Conference (PSCC)*, Jun. 2018, pp. 1–7.
- [98] X. Li and K. Hedman, *Data driven linearized AC power flow model with regression analysis*, 2018.

- [99] M. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret, „Applications of second-order cone programming“, *Linear Algebra and its Applications*, vol. 284, no. 1, pp. 193–228, 1998.
- [100] A. Lorca and X. A. Sun, „The adaptive robust multi-period alternating current optimal power flow problem“, *IEEE Transactions on Power Systems*, vol. 33, no. 2, pp. 1993–2003, 2018.
- [101] R. Louca and E. Bitar, „Robust AC optimal power flow“, *arXiv:1706.09019*, Jun. 2017.
- [102] M. Lubin, Y. Dvorkin, and S. Backhaus, „A robust approach to chance constrained optimal power flow with renewable generation“, *IEEE Transactions on Power Systems*, vol. 31, no. 5, pp. 3840–3849, Sep. 2016.
- [103] S. Lucia, P. Zometa, M. Kögel, and R. Findeisen, „Efficient stochastic model predictive control based on polynomial chaos expansions for embedded applications“, in *54th Annual Conference on Decision and Control*, Dec. 2015, pp. 3006–3012.
- [104] D. Luenberger, *Optimization by Vector Space Methods*, ser. Series in Decision and Control. New York: John Wiley & Sons, Inc., 1969.
- [105] J. Maciejowski, *Predictive Control with Constraints*. Prentice Hall, 2002.
- [106] S. Marelli and B. Sudret, „UQLab: a framework for uncertainty quantification in Matlab“, in *Vulnerability, Uncertainty, and Risk*. 2014, pp. 2554–2563.
- [107] J. Marti, H. Ahmadi, and L. Bashualdo, „Linear power-flow formulation based on a voltage-dependent load model“, *IEEE Transactions on Power Delivery*, vol. 28, no. 3, pp. 1682–1690, Jul. 2013.
- [108] A. Mesbah, „Stochastic model predictive control: An overview and perspectives for future research“, *IEEE Control Systems*, vol. 36, no. 6, pp. 30–44, Dec. 2016.
- [109] A. Mesbah, S. Streif, R. Findeisen, and R. Braatz, „Stochastic nonlinear model predictive control with probabilistic constraints“, in *American Control Conference*, Jun. 2014, pp. 2413–2419.
- [110] R. Mieth and Y. Dvorkin, „Data-driven distributionally robust optimal power flow for distribution systems“, *IEEE Control Systems Letters*, vol. 2, no. 3, pp. 363–368, Jul. 2018.
- [111] S. Misra, D. Molzahn, and K. Dvijotham, „Optimal adaptive linearizations of the AC power flow equations“, in *2018 Power Systems Computation Conference (PSCC)*, Jun. 2018, pp. 1–7.
- [112] P. Mogensen and A. Riseth, „Optim: A mathematical optimization package for Julia“, *Journal of Open Source Software*, vol. 3, no. 24, pp. 1–3, 2017.
- [113] D. Molzahn and L. Roald, „Towards an AC optimal power flow algorithm with robust feasibility guarantees“, in *20th Power Systems Computation Conference*, Dublin, 2018.
- [114] D. Molzahn and I. Hiskens, „A survey of relaxations and approximations of the power flow equations“, *Foundations and Trends in Electric Energy Systems*, vol. 4, no. 1-2, pp. 1–221, 2019.
- [115] D. Montgomery, *Design and Analysis of Experiments*. New York: John Wiley and Sons, 2004.
- [116] J. Morales and J. Perez-Ruiz, „Point estimate schemes to solve the probabilistic power flow“, *IEEE Transactions on Power Systems*, vol. 22, no. 4, pp. 1594–1601, Nov. 2007.
- [117] T. Mühlpfordt, T. Faulwasser, and V. Hagenmeyer, „A generalized framework for chance-constrained optimal power flow“, *Sustainable Energy, Grids and Networks*, vol. 16, pp. 231–242, 2018.
- [118] —, „Optimal power flow subject to non-Gaussian stochastic uncertainties: an L2-based approach“, *Proceedings in Applied Mathematics and Mechanics (PAMM)*, vol. 18, no. 1, 2018.
- [119] —, „Solving stochastic AC power flow via polynomial chaos expansion“, in *IEEE International Conference on Control Applications*, 2016, pp. 70–76.
- [120] —, „Uncertainty quantification for optimal power flow problems“, *Proceedings in Applied Mathematics and Mechanics (PAMM)*, vol. 19, no. 1, 2019.
- [121] T. Mühlpfordt, T. Faulwasser, L. Roald, and V. Hagenmeyer, „Solving optimal power flow with non-Gaussian uncertainties via polynomial chaos expansion“, in *56th IEEE Conference on Decision and Control (CDC)*, Dec. 2017, pp. 4490–4496.

- [122] T. Mühlpfordt, R. Findeisen, V. Hagenmeyer, and T. Faulwasser, „Comments on truncation errors for polynomial chaos expansions“, *IEEE Control Systems Letters*, vol. 2, no. 1, pp. 169–174, Jan. 2018.
- [123] —, „Errata for “comments on truncation errors for polynomial chaos expansions”“, *IEEE Control Systems Letters*, pp. 1–1, 2019.
- [124] T. Mühlpfordt, V. Hagenmeyer, and T. Faulwasser, „The price of uncertainty: Chance-constrained OPF vs. in-hindsight OPF“, in *2018 Power Systems Computation Conference (PSCC)*, Jun. 2018, pp. 1–7.
- [125] T. Mühlpfordt, V. Hagenmeyer, D. Molzahn, and S. Misra, „Optimal adaptive power flow linearizations: Expected error minimization using polynomial chaos expansion“, in *2019 IEEE Milan PowerTech*, 2019, pp. 1–6.
- [126] 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.
- [127] T. Mühlpfordt, F. Zahn, F. Becker, T. Faulwasser, and V. Hagenmeyer, *timueh/PolyChaos.jl: v0.2.1*, Jun. 2019.
- [128] D. Muñoz-Alvarez, E. Bitar, L. Tong, and J. Wang, „Piecewise affine dispatch policies for economic dispatch under uncertainty“, in *2014 IEEE PES General Meeting | Conference & Exposition*, IEEE, Jul. 2014.
- [129] A. Nasri, S. Kazempour, A. Conejo, and M. Ghandhari, „Network-constrained AC unit commitment under uncertainty: A Benders’ decomposition approach“, *IEEE Transactions on Power Systems*, vol. 31, no. 1, pp. 412–422, Jan. 2016.
- [130] Y. Ng, S. Misra, L. Roald, and S. Backhaus, „Statistical learning for DC optimal power flow“, in *2018 Power Systems Computation Conference (PSCC)*, Jun. 2018, pp. 1–7.
- [131] F. Ni, P. Nguyen, and J. Cobben, „Basis-adaptive sparse polynomial chaos expansion for probabilistic power flow“, *IEEE Transactions on Power Systems*, vol. 32, no. 1, pp. 694–704, Jan. 2017.
- [132] K. Ogimoto, Y. Udagawa, and T. Ikegami, „Impacts of variable renewable energy source integration into power system operation and implications for Japan’s future power market“, Cigré, Paris, Tech. Rep. C5-309-2014, 2014.
- [133] F. Oldewurtel, L. Roald, G. Andersson, and C. Tomlin, „Adaptively constrained stochastic model predictive control applied to security constrained optimal power flow“, in *IEEE American Control Conference (ACC)*, 2015, pp. 931–936.
- [134] B. Oswald, *Berechnung von Drehstromnetzen. Berechnung stationärer und nichtstationärer Vorgänge mit symmetrischen Komponenten und Raumzeigern*, 2nd ed. Wiesbaden: Springer Vieweg, 2013.
- [135] J. Qin and R. Rajagopal, „Price of uncertainty in multistage stochastic power dispatch“, in *53rd IEEE Conference on Decision and Control*, Dec. 2014, pp. 4065–4070.
- [136] C. Rackauckas and Q. Nie, „DifferentialEquations.jl – A performant and feature-rich ecosystem for solving differential equations in Julia“, *Journal of Open Research Software*, vol. 5, no. 15, pp. 1–10, 2017.
- [137] C. Rasmussen and C. Williams, *Gaussian Processes for Machine Learning*. Cambridge, MA, USA: The MIT Press, 2006.
- [138] J. Rawlings and D. Mayne, *Model Predictive Control: Theory and Design*. Nob Hill Publishing, 2009.
- [139] Z. Ren, W. Li, R. Billinton, and W. Yan, „Probabilistic power flow analysis based on the stochastic response surface method“, *IEEE Transactions on Power Systems*, vol. 31, no. 3, pp. 2307–2315, May 2016.
- [140] L. Roald and G. Andersson, „Chance-constrained AC optimal power flow: Reformulations and efficient algorithms“, *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 2906–2918, May 2018.

- [141] L. Roald, T. Krause, and G. Andersson, „Integrated balancing and congestion management under forecast uncertainty“, in *2016 IEEE International Energy Conference (ENERGYCON)*, Apr. 2016, pp. 1–6.
- [142] L. Roald, S. Misra, M. Chertkov, and G. Andersson, „Optimal power flow with weighted chance constraints and general policies for generation control“, in *2015 54th IEEE Conference on Decision and Control (CDC)*, Dec. 2015, pp. 6927–6933.
- [143] L. Roald and D. Molzahn, „Implied constraint satisfaction in power system optimization: The impacts of load variations“, *ArXiv e-prints*, Apr. 2019, 1904.01757.
- [144] L. Roald, F. Oldewurtel, T. Krause, and G. Andersson, „Analytical reformulation of security constrained optimal power flow with probabilistic constraints“, in *2013 IEEE Grenoble Conference*, Jun. 2013, pp. 1–6.
- [145] L. Roald, F. Oldewurtel, B. Van Parys, and G. Andersson, „Security constrained optimal power flow with distributionally robust chance constraints“, *ArXiv e-prints*, Aug. 2015, 1508.06061.
- [146] L. Roald, „Optimization methods to manage uncertainty and risk in power systems operation“, PhD thesis, Eidgenössische Technische Hochschule Zürich (ETH), 2016.
- [147] M. Schilling, A. Leite da Silva, R. Billinton, and M. El-Kady, „Bibliography on power system probabilistic analysis (1962-88)“, *IEEE Transactions on Power Systems*, vol. 5, no. 1, pp. 1–11, Feb. 1990.
- [148] J. Schmidli, L. Roald, S. Chatzivasileiadis, and G. Andersson, „Stochastic AC optimal power flow with approximate chance-constraints“, in *2016 IEEE Power and Energy Society General Meeting (PESGM)*, Jul. 2016, pp. 1–5.
- [149] A. Schwab, *Elektroenergiesysteme*, 3rd ed. Berlin, Heidelberg: Springer-Verlag, 2012.
- [150] D. Shirmohammadi, H. Hong, A. Semlyen, and G. Luo, „A compensation-based power flow method for weakly meshed distribution and transmission networks“, *IEEE Transactions on Power Systems*, vol. 3, no. 2, pp. 753–762, May 1988.
- [151] E. Stein and R. Shakarchi, *Real Analysis. Measure Theory, Integration, and Hilbert Spaces*, ser. Princeton Lectures in Analysis. Princeton: Princeton University Press, 2005, vol. III.
- [152] G. Strang, *Introduction to Linear Algebra*, 3rd. Wellesley: Wellesley-Cambridge Press, 2003.
- [153] C.-L. Su, „Probabilistic load-flow computation using point estimate method“, *IEEE Transactions on Power Systems*, vol. 20, no. 4, pp. 1843–1851, Nov. 2005.
- [154] T. Sullivan, *Introduction to Uncertainty Quantification*, 1st ed. Switzerland: Springer International Publishing, 2015, vol. 63.
- [155] A. Venzke and S. Chatzivasileiadis, „Convex relaxations of probabilistic AC optimal power flow for interconnected AC and HVDC grids“, *IEEE Transactions on Power Systems*, vol. 34, no. 4, pp. 2706–2718, Jul. 2019.
- [156] —, „Convex relaxations of security constrained AC optimal power flow under uncertainty“, in *2018 Power Systems Computation Conference (PSCC)*, Jun. 2018, pp. 1–7.
- [157] A. Venzke, L. Halilbasic, U. Markovic, G. Hug, and S. Chatzivasileiadis, „Convex relaxations of chance constrained AC optimal power flow“, *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 2829–2841, May 2018.
- [158] G. Verbić and C. Cañizares, „Probabilistic optimal power flow in electricity markets based on a two-point estimate method“, *IEEE Transactions on Power Systems*, vol. 21, no. 4, pp. 1883–1893, Nov. 2006.
- [159] M. Vrakopoulou and I. Hiskens, „Optimal control policies for reserve deployment with probabilistic performance guarantees“, in *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*, Dec. 2017, pp. 4470–4475.
- [160] M. Vrakopoulou and I. Hiskens, „Optimal policy-based control of generation and HVDC lines in power systems under uncertainty“, in *2017 IEEE Manchester PowerTech*, IEEE, Jun. 2017.

- [161] M. Vrakopoulou, M. Katsampani, K. Margellos, J. Lygeros, and G. Andersson, „Probabilistic security-constrained AC optimal power flow“, in *PowerTech*, Grenoble, France, Jun. 2013.
- [162] —, „Probabilistic security-constrained AC optimal power flow“, in *IEEE PowerTech Grenoble*, Jun. 2013, pp. 12–6.
- [163] M. Vrakopoulou, K. Margellos, J. Lygeros, and G. Andersson, „Probabilistic guarantees for the N-1 security of systems with wind power generation“, in *Proceedings of PMAPS*, 2012, pp. 858–863.
- [164] M. Vrakopoulou, J. Mathieu, and G. Andersson, „Stochastic optimal power flow with uncertain reserves from demand response“, in *2014 47th Hawaii International Conference on System Sciences*, IEEE, 2014, pp. 2353–2362.
- [165] J. Waldvogel, „Fast construction of the Fejér and Clenshaw-Curtis quadrature rules“, *BIT Numerical Mathematics*, vol. 46, no. 1, pp. 195–202, Mar. 2006.
- [166] J. Warrington, P. Goulart, S. Mariéthoz, and M. Morari, „Policy-based reserves for power systems“, *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 4427–4437, Nov. 2013.
- [167] N. Wiener, „The homogeneous chaos“, *American Journal of Mathematics*, pp. 897–936, 1938.
- [168] A. Wood and B. Wollenberg, *Power Generation, Operation, and Control*. John Wiley & Sons, 2012.
- [169] H. Wu, Y. Zhou, S. Dong, and Y. Song, „Probabilistic load flow based on generalized polynomial chaos“, *IEEE Transactions on Power Systems*, vol. 32, no. 1, pp. 820–821, Jan. 2017.
- [170] Y. Xiang, J. Liu, and Y. Liu, „Robust energy management of microgrid with uncertain renewable generation and load“, *IEEE Transactions on Smart Grid*, vol. 7, no. 2, pp. 1034–1043, Mar. 2016.
- [171] W. Xie and S. Ahmed, „Distributionally robust chance constrained optimal power flow with renewables: A conic reformulation“, *IEEE Transactions on Power Systems*, vol. 33, no. 2, pp. 1860–1867, Mar. 2018.
- [172] D. Xiu, *Numerical Methods for Stochastic Computations*. Princeton, New Jersey: Princeton University Press, 2010.
- [173] D. Xiu and G. E. Karniadakis, „The Wiener-Askey polynomial chaos for stochastic differential equations“, *SIAM Journal on Scientific Computing*, vol. 24, no. 2, pp. 619–644, 2002.
- [174] H. Yu, C. Chung, K. Wong, H. Lee, and J. Zhang, „Probabilistic load flow evaluation with hybrid latin hypercube sampling and Cholesky decomposition“, *IEEE Transactions on Power Systems*, vol. 24, no. 2, pp. 661–667, May 2009.
- [175] H. Zhang and P. Li, „Application of sparse-grid technique to chance constrained optimal power flow“, *IET Generation, Transmission Distribution*, vol. 7, no. 5, pp. 491–499, May 2013.
- [176] —, „Chance constrained programming for optimal power flow under uncertainty“, *IEEE Transactions on Power Systems*, vol. 26, no. 4, pp. 2417–2424, Nov. 2011.
- [177] —, „Probabilistic analysis for optimal power flow under uncertainty“, English, *IET Generation, Transmission & Distribution*, vol. 4, 553–561(8), 5 May 2010.
- [178] P. Zhang and S. Lee, „Probabilistic load flow computation using the method of combined cumulants and Gram-Charlier expansion“, *IEEE Transactions on Power Systems*, vol. 19, no. 1, pp. 676–682, 2004.
- [179] Y. Zhang, S. Shen, B. Li, and J. Mathieu, „Two-stage distributionally robust optimal power flow with flexible loads“, in *2017 IEEE Manchester PowerTech*, Jun. 2017, pp. 1–6.
- [180] Y. Zhang, S. Shen, and J. Mathieu, „Distributionally robust chance-constrained optimal power flow with uncertain renewables and uncertain reserves provided by loads“, *IEEE Transactions on Power Systems*, vol. 32, no. 2, pp. 1378–1388, Mar. 2017.



## Conference publications

- [9] 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*, Jun. 2017, pp. 1–6.
- [54] A. Engelmann, T. Mühlpfordt, Y. Jiang, B. Houska, and T. Faulwasser, „Distributed AC optimal power flow using ALADIN“, *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 5536–5541, 2017, 20th IFAC World Congress.
- [55] —, „Distributed stochastic AC optimal power flow based on polynomial chaos expansion“, in *2018 Annual American Control Conference (ACC)*, Jun. 2018, pp. 6188–6193.
- [118] T. Mühlpfordt, T. Faulwasser, and V. Hagenmeyer, „Optimal power flow subject to non-Gaussian stochastic uncertainties: an L2-based approach“, *Proceedings in Applied Mathematics and Mechanics (PAMM)*, vol. 18, no. 1, 2018.
- [119] —, „Solving stochastic AC power flow via polynomial chaos expansion“, in *IEEE International Conference on Control Applications*, 2016, pp. 70–76.
- [120] —, „Uncertainty quantification for optimal power flow problems“, *Proceedings in Applied Mathematics and Mechanics (PAMM)*, vol. 19, no. 1, 2019.
- [121] T. Mühlpfordt, T. Faulwasser, L. Roald, and V. Hagenmeyer, „Solving optimal power flow with non-Gaussian uncertainties via polynomial chaos expansion“, in *56th IEEE Conference on Decision and Control (CDC)*, Dec. 2017, pp. 4490–4496.
- [124] T. Mühlpfordt, V. Hagenmeyer, and T. Faulwasser, „The price of uncertainty: Chance-constrained OPF vs. in-hindsight OPF“, in *2018 Power Systems Computation Conference (PSCC)*, Jun. 2018, pp. 1–7.
- [125] T. Mühlpfordt, V. Hagenmeyer, D. Molzahn, and S. Misra, „Optimal adaptive power flow linearizations: Expected error minimization using polynomial chaos expansion“, in *2019 IEEE Milan PowerTech*, 2019, pp. 1–6.

## Journal publications

- [53] A. Engelmann, Y. Jiang, T. Mühlpfordt, B. Houska, and T. Faulwasser, „Toward distributed OPF using ALADIN“, *IEEE Transactions on Power Systems*, vol. 34, no. 1, pp. 584–594, Jan. 2019.
- [58] T. Faulwasser, A. Engelmann, T. Mühlpfordt, and V. Hagenmeyer, „Optimal power flow: An introduction to predictive, distributed and stochastic control challenges“, *at – Automatisierungstechnik*, vol. 66, no. 7, pp. 573–589, Jan. 2018.
- [117] T. Mühlpfordt, T. Faulwasser, and V. Hagenmeyer, „A generalized framework for chance-constrained optimal power flow“, *Sustainable Energy, Grids and Networks*, vol. 16, pp. 231–242, 2018.
- [122] T. Mühlpfordt, R. Findeisen, V. Hagenmeyer, and T. Faulwasser, „Comments on truncation errors for polynomial chaos expansions“, *IEEE Control Systems Letters*, vol. 2, no. 1, pp. 169–174, Jan. 2018.

- [123] —, „Errata for “comments on truncation errors for polynomial chaos expansions”“, *IEEE Control Systems Letters*, pp. 1–1, 2019.
- [126] 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.

## Curriculum vitæ

- 07/18 - 10/18    Research stay  
École Polytechnique Fédérale de Lausanne, Lausanne, CH
- 07/17 - 08/17    Summer student  
Los Alamos National Laboratory, Los Alamos, NM, USA
- since 07/15      Research associate and PhD candidate  
Karlsruhe Institute of Technology, Karlsruhe, Germany
- 10/14 - 05/15    Non-degree visiting student  
Massachusetts Institute of Technology, Cambridge, MA, USA
- 10/13 - 05/15    M. Sc. in systems engineering and engineering cybernetics  
Otto-von-Guericke-University, Magdeburg, Germany
- 04/13 - 08/13    Intern in power plant division  
BASF SE, Ludwigshafen, Germany
- 08/12 - 02/13    Occupational trainee  
University of Melbourne, Melbourne, Australia
- 10/09 - 08/13    B. Sc. in systems engineering and engineering cybernetics  
Otto-von-Guericke-University, Magdeburg, Germany
- 06/08            Abitur  
Staatliches Gymnasium, Arnstadt, Germany
- 08/05 - 06/06    Junior high school year  
Waterford Mott High School, Waterford, MI, USA