UNIVERSIDADE ESTADUAL DE CAMPINAS
Faculdade de Engenharia Elétrica e de Computação

Iuri Abrahão Monteiro

An Investigation on Methods of Determining Salient-Pole Synchronous Machines States and Parameters

Um Estudo sobre Métodos de Determinação de Estados e Parâmetros de Máquinas Síncronas de Polos Salientes

Campinas
2020

Iuri Abrahão Monteiro

An Investigation on Methods of Determining Salient-Pole Synchronous Machines States and Parameters

Um Estudo sobre Métodos de Determinação de Estados e Parâmetros de Máquinas Síncronas de Polos Salientes

Dissertation presented to the School of Electrical and Computer Engineering of the University of Campinas in partial fulfillment of the requirements for the degree of Master in Electrical Engineering, in the area of Automation

Dissertação apresentada à Faculdade de Engenharia Elétrica e de Computação da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Mestre em Engenharia Elétrica, na área de Automação

Supervisor: Prof. Dr. Mateus Giesbrecht

This copy corresponds to the final version of the dissertation defended by student Iuri Abrahão Monteiro and supervised by Prof. Dr. Mateus Giesbrecht

Campinas
2020

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca da Área de Engenharia e Arquitetura
Elizangela Aparecida dos Santos Souza - CRB 8/8098

M764i

Monteiro, Iuri Abrahão, 1993-
An investigation on methods of determining salient-pole synchronous machines states and parameters / Iuri Abrahão Monteiro. – Campinas, SP : [s.n.], 2020.

Orientador: Mateus Giesbrecht.
Dissertação (mestrado) – Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação.

1. Máquinas elétricas síncronas - Mpodelos matemáticos. 2. Sistemas dinâmicos. 3. Sistemas de energia elétrica - Estimação de estado. 4. Identificação de sistemas. I. Giesbrecht, Mateus, 1984-. II. Universidade Estadual de Campinas. Faculdade de Engenharia Elétrica e de Computação. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Um estudo sobre métodos de determinação de estados e parâmetros de máquinas síncronas de polos salientes
Palavras-chave em inglês:
Electric machinery, Synchronous - Mathematical models
Dynamical systems
Electric power systems - State estimation
System identification
Área de concentração: Automação
Titulação: Mestre em Engenharia Elétrica
Banca examinadora:
Mateus Giesbrecht [Orientador]
José Roberto Boffino de Almeida Monteiro
Gilmar Barreto
Data de defesa: 08-05-2020
Programa de Pós-Graduação: Engenharia Elétrica

Identificação e informações acadêmicas do(a) aluno(a)
- ORCID do autor: https://orcid.org/0000-0002-0380-8262
- Currículo Lattes do autor: http://lattes.cnpq.br/3258592297949316

COMISSÃO JULGADORA – DISSERTAÇÃO DE MESTRADO

Candidato: Iuri Abrahão Monteiro, RA: 211508

Data da defesa: 08 de maio de 2020

Título da dissertação: An Investigation on Methods of Determining Salient-Pole Synchronous Machines States and Parameters / Um Estudo Sobre Métodos de Determinação de Estados e Parâmetros de Máquinas Síncronas de Polos Salientes

Prof. Dr. Mateus Giesbrecht (FEEC/UNICAMP – Presidente)

Prof. Dr. José Roberto Boffino de Almeida Monteiro (EESC/USP – Membro Externo)

Prof. Dr. Gilmar Barreto (FEEC/UNICAMP – Membro Interno)

A ata de defesa, com as respectivas assinaturas dos membros da Comissão Julgadora, encontra-se no SIGA (Sistema de Fluxo de Dissertação/Tese) e na Secretaria de Pós-Graduação da Faculdade de Engenharia Elétrica e de Computação.

Dedicated to the memory of my grandmother Laura, my true guardian angel; without her, this work and my many other dreams would not have become reality.

Acknowledgments

At first, I would like to thank God, the center and foundation of everything in my life, and the Orixás, for renewing, at every moment, my strength and disposition and for the discernment given to me throughout this tough journey.

My parents, Alexandro and Denise, for their unconditional love, unrestricted support, affection, for believing in me, and for always being my official sponsors, without whom this work could not have been completed to its fullest.

My brother, Pedro Henrique, for his companionship, for always supporting me in times of trouble, and for taking care of me.

My grand-aunt Raquel and my grandmother Maria, for cheering with every single victory I have and for supporting me in every possible way.

My advisor, Prof. Dr. Mateus Giesbrecht, for his tireless supervision, for his interest and trust, for his advice and vast intelligence, and especially for his friendship. The door to his office was always open whenever I ran into a trouble spot or had a question about my research or writing. He consistently allowed this paper to be my own work, but steered me in the right the direction whenever he thought I needed it.

My special friend and coworker, Yara Quilles, for helping me with my afflictions, my presentation training, and for being an always-available ear.

Prof. Dr. Paulo Valente, for his excellent course on *Mathematical Optimization*. Prof. Dr. Pedro Peres, for teaching me the basis of *Linear Analysis of Signals and Systems*. Prof. Dr. Mateus Giesbrecht, for *Methods for Subspace System Identification*. And Prof. Dr. Edson Bim, for *Multiphase Electrical Machines*.

The members of the examining board, for accepting the invitation and for willing to enhance this work. I also thank you in advance for your comments and recommendations.

Luís Alfredo Esteves Meneses and Empresa de Energía del Pacífico for performing tests and acquiring the data necessary to verify some of this work methodologies.

The Brazilian people, for the portion of the numerous taxes that became my major financial support.

This study was financed in part by CNPq, Conselho Nacional de Desenvolvimento Científico e Tecnológico – Brasil.

And everyone who, somehow, got involved in this campaign.

"Não me misturo, não me dobro. A rainha do mar anda de mãos dadas comigo, me ensina o baile das ondas e canta, canta, canta pra mim. É do ouro de Oxum que é feita a armadura que cobre o meu corpo, garante o meu sangue e minha garganta: o veneno do mal não acha passagem. Em meu coração, Maria, acende a sua luz e me aponta o caminho. Me sumo no vento, cavalgo no raio de Iansã, giro o mundo, viro e reviro. Tô no recôncavo, tô em Fez. Voo entre as estrelas, brinco de ser uma. Traço o Cruzeiro do Sul com a tocha da fogueira de João menino, rezo com as Três Marias. Vou além: me recolho no esplendor das nebulosas, descanso nos vales e montanhas. Durmo na forja de Ogum e me mergulho no calor da lava dos vulcões – corpo vivo de Xangô."

Maria Bethânia, *Carta de Amor*

Abstract

Salient-pole synchronous machines play a fundamental role in the stability analysis of electrical power systems, especially in countries where most of the generated energy comes from hydraulic sources. The electrical equivalent models that describe the behavior of these machines are composed of several electrical parameters, which are used in a wide range of studies. In the present work, techniques for estimating states and parameters of salient-pole synchronous machines are studied and proposed.

A priori, the voltage, flux linkage, power, and motion equations are developed with the appropriate units included, both in machine variables and in variables projected on an orthogonal plane rotating in the rotor's electrical speed. In most of the literature, these units are not explained in the equation process.

Among the electrical parameters, the magnetizing reactances are the ones that most influence the machine behavior under transient and steady-state conditions. In this way, a new approach to estimate the load angle of these machines and the subsequent calculation of the magnetizing reactances from specific load conditions are presented – the performance of the proposed method is evaluated by means of simulation data and by operating data of a large synchronous generator.

Some approaches to determine parameters require the machine to be taken out of operation, so that specific tests may be performed. Among them, one of the most used to determine transient and steady-state parameters is the load rejection test; thus, this test is also analyzed and refined by an automated method based on variable projection for separating the resulting sum-of-exponentials.

Since the machines are highly nonlinear, multivariate, dynamic systems, different state observers seek to solve the state estimation problem in a timely manner and with satisfactory accuracy. This work presents a nonlinear and recursive approach for the estimation of flux linkages per second, amortisseur winding currents, load angle, and magnetizing reactances of salient-pole synchronous machines by means of the particle filtering. An eighth-order nonlinear model is considered, and only measurements taken at the machine terminals are necessary to estimate these quantities.

Resumo

As máquinas síncronas de polos salientes desempenham um papel fundamental na análise de estabilidade de sistemas elétricos de potência, especialmente em países cuja maior parte da energia gerada provém de fontes hidráulicas. Os modelos elétricos equivalentes que descrevem o comportamento dessas máquinas são compostos por diversos parâmetros, os quais são utilizados em uma ampla gama de estudos.

No presente trabalho, estudam-se e propõem-se técnicas de estimação de estados e parâmetros de máquinas síncronas de polos salientes. A princípio, as equações de tensão, de fluxos concatenados, de potência e de movimento são desenvolvidas com as devidas unidades de medida, tanto em variáveis de máquina quanto em variáveis projetadas sobre um plano ortogonal que gira na velocidade elétrica do rotor. Na maior parte da literatura, essas unidades não são explicitadas no equacionamento.

Dentre os parâmetros elétricos dos modelos das máquinas síncronas de polos salientes, as reatâncias de magnetização são os que mais influenciam o comportamento da máquina em condições de regime permanente senoidal. Desta forma, apresenta-se uma nova abordagem à estimação do ângulo de carga dessas máquinas e o subsequente cálculo das reatâncias de magnetização a partir de condições de carga específicas – o desempenho do método proposto é avaliado em dados de simulação e em dados reais de operação de um gerador síncrono de grande porte.

Algumas abordagens à determinação de parâmetros requerem que a máquina seja posta fora de operação para que ensaios específicos possam ser realizados. Dentre eles, um dos mais empregados na determinação de parâmetros transitórios e de regime permanente é o ensaio de rejeição de carga; assim, este ensaio também é analisado e aperfeiçoado por um método automatizado de separação de soma de exponenciais baseado em projeção de variáveis.

Por tratar-se de um sistema multivariável e altamente não linear, diferentes observadores de estado também são utilizados para se determinarem estados e parâmetros de máquinas síncronas em tempo hábil e com precisão satisfatória. Este trabalho apresenta uma abordagem não linear recursivamente aplicável à estimação de fluxos concatenados, correntes de enrolamentos amortecedores, ângulo de carga e reatâncias de magnetização de máquinas síncronas de polos salientes por meio da filtragem de partículas. Um modelo não linear de oitava ordem é considerado e apenas as medições realizadas nos terminais da armadura e do campo durante regime permanente se fazem necessárias para estimar as referidas grandezas.

List of Figures

List of Tables

List of Abbreviations and Acronyms

Abbreviation/Acronym	Meaning
AC	Alternating Current
BLDC	Brushless Direct Current
BP	Bayesian Processor
cdf	Cumulative distribution function
DC	Direct Current
GA	Genetic Algorithm
emf	Electromotive force
EKF	Extended Kalman Filter
IEEE	Institute of Electrical and Electronics Engineers
KF	Kalman Filter
MAP	Maximum *a posteriori*
MC	Monte Carlo
ML	Maximum likelihood
mmf	Magnetomotive force
MSE	Mean-squared error
MV	Minimum variance
N4SID	Numerical Algorithms for Subspace State Space System Identification
OEL	Over-excitation limiter
pdf	Probability density function
PF	Particle Filter
PID	Proportional-integral-derivative
PMU	Phasor Measurement Unit
PRBS	Pseudorandom Binary Sequence
PSS	Power system stabilizer
PT	Park's Transformation
pu	Per unit
rms	Root mean square
rpm	Revolutions per minute
SMC	Sequential Monte Carlo
SPT	Sigma-Point Transformation

(continue on the next page)

(continued)

Abbreviation	Meaning
UEL	Under-excitation limiter
UKF	Unscented Kalman Filter

List of Notation

"A good notation should be unambiguous, pregnant, easy to remember; it should avoid harmful second meanings and take advantage of useful second meanings; the order and connection of signs should suggest the order and connection of things."

— George Polya, *How to Solve It (1957)*

Symbol	Meaning
$\mathbf{a}$	Lowercase bold letter – vector
$\mathbf{A}$	Uppercase bold letter – matrix
$\mathbf{B}$	Magnetic field intensity
$\mathcal{B}$	Borel field or binomial distribution function
$\mathbb{C}^n$	Complex vector space of dimension n
$\mathbf{E}$	Electric field intensity
f'	A quantity with appropriate turns ratios included
$\dot{f}$	The derivative of a quantity f with respect to time
$\bar{f}$	A quantity in per unit (pu)
$f_{\mathcal{X}}$	Probability density function
F	An event – a subset of the sample space
$F_{\mathcal{X}}$	Probability distribution function
$\mathcal{F}$	$\sigma-$algebra
i	Electric current
J	Inertia
$\mathbf{K}$	Park's Transformation matrix
ℓ	Inductive reactance
L_x	Inductance; $x = l$ represents the leakage component, and $x = m$, the magnetizing component
$\mathcal{M}$	Multinomial distribution function
$\mathcal{N}(\text{arg}_1, \text{arg}_2)$	Gaussian distribution function of mean arg_1 and variance arg_2
N_x	Number of coils of winding x
P	Active power
$\mathcal{P}$	Instantaneous power
p	An operator for the derivative of a function with respect to time
Q	Reactive power
r	Electric resistance
$\mathbb{R}^n$	Real vector space of dimension n

(continue on the next page)

(continued)

Symbol	Meaning
$\mathfrak{R}$	Magnetic reluctance
$\mathcal{T}_e$	Electrical torque
$\mathcal{T}_\ell$	Net mechanical shaft torque
$\boldsymbol{u}$	Input vector
v	Terminal voltage
$\mathcal{W}_f$	Energy stored in the coupling field
x	Inductive reactance or state variable
$\mathcal{X}, \mathcal{Y}, \mathcal{Z}$	Scalar random variables (or, simply, random variables)
$\boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{Y}}, \boldsymbol{\mathcal{Z}}$	Vector random variable (or, simply, random vectors)
$\boldsymbol{\mathfrak{X}}$	Set of sigma-points
$\boldsymbol{y}$	Measurement vector
$\boldsymbol{\mathfrak{Y}}$	An arbitrary nonlinear transformation of $\boldsymbol{\mathfrak{X}}$
$\emptyset$	Empty set
Greek Letter	
δ	Load/rotor angle, in electrical radians
δ_m	Load/rotor angle, in mechanical radians
θ_r	Mechanical rotor angle – a measure of q-axis leading the as-axis in the direction of rotation
κ_d	Damping-torque coefficient, in newton-meter-second
ρ	Number of pole pairs
Φ	Instantaneous value of a time-varying magnetic flux
Φ_l	Leakage component of flux
Φ_m	Magnetizing component of flux
Ψ	Flux linkage per second
ψ	Total flux linkage
ω	A single elementary outcome of an experiment
ω_e	Angular frequency of the generated voltage, in electrical radians per second
ω_r	Angular frequency of the mechanical shaft, in electrical radians per second
ω_s	Synchronous angular frequency, in electrical radians per second
ω_{sm}	Synchronous angular frequency, in mechanical radians per second
Ω	Sample space

Contents

Works Developed by the Author

[1] MONTEIRO, I. A.; VIANNA, L. M. S.; GIESBRECHT, M. Nonlinear estimation of salient-pole synchronous machines parameters via Particle Filter. In: 2019 IEEE PES Innovative Smart Grid Technologies Conference – Latin America (ISGT Latin America). Gramado, RS, BR: IEEE, Sept. 2019. P. 1–6. DOI: 10.1109/ISGT-LA.2019.8895417

[2] MONTEIRO, I. A.; VIANNA, L. M. S.; GIESBRECHT, M. Observador de fluxos, correntes e ângulo de carga de máquinas síncronas por meio da filtragem de partículas. In: ANAIS do XIV Simpósio Brasileiro de Automação Inteligente. Ouro Preto, MG, BR: Galoá, Oct. 2019. v. 1. DOI: 10.17648/sbai-2019-111220

[3] MONTEIRO, I. A.; MENESES, L.; GIESBRECHT, M. A novel approach on the determination of salient-pole synchronous machine magnetizing reactances from on-line measurements. In: 2020 IEEE 29th International Symposium on Industrial Electronics (ISIE). In Press: [s.n.], 2020

[4] VIANNA, L. et al. Detecção de falhas de alimentação de um motor CC sem escovas via Filtro de Partículas. In: ANAIS do XIV Simpósio Brasileiro de Automação Inteligente. Ouro Preto, MG, Brazil: Galoá, Oct. 2019. v. 1. DOI: 10.17648/sbai-2019-111202

Chapter 1

Introduction

"There is another world, but it is in this one."
— William Butler Yeats (1865—1939),
The Secret Rose

Ever since Thomas A. Edison started to work with the electric light and formulated the concept of centrally located power stations in 1878, the power system has undergone many changes. From distributed lighting systems capable of supplying 30 kW [5], the electric grid evolved into a complex system divided into several subsystems: generation, transmission, substation, distribution, and consumption [6]. A typical electric system is composed of a few hundreds of generators interconnected by a transmission network.

In recent years, there has been a notable increase of distributed energy resources on distribution grids, either at medium- or low-voltage levels [5]. Renewable energy sources like wind and sun are reliable alternatives to traditional energy sources, such as oil, natural gas, or coal. Distributed power generation systems based on renewable energy sources experience large development worldwide, with Germany, Denmark, Japan, and the United States as leaders in this field [7]. By the end of 2013, there were 12.1 GW installed in solar photovoltaic systems in the United States alone [8]. This shift alters the way electricity is being generated, transmitted, and managed, thus necessitating a change in how utilities plan and integrate those resources [9].

Even in that context, one of the most important components in a power system is the synchronous generator. Specially in countries where the electric power generation is based on hydraulic sources, salient-pole synchronous machines generate most of the electric power and are capable of considerably influencing the behavior of these systems during transient- and steady-state conditions [5]. Almost one century after the first publications in this area [10, 11], modeling synchronous machines is still a challenging and attractive research topic: today's most mature science of power generation is still based on synchronous-generator technologies [12].

Models of power system components are crucial for power systems stability studies. Generally, these models have a known parametric structure, whose parameters must be determined (by means of well-established tests [13]) or estimated (by means of states observers, for example) to represent a given component. In a first approach, the parameters from each component may be obtained from manufacturers' data. This approach is not recommended since some design data may be inaccurate [14]. Furthermore, within the state space framework, the dynamic states of synchronous machines are the minimum set of variables (including rotor angles and speeds) that may uniquely determine the machine's dynamic status [15] and may be used in various advanced control methods [16].

In fact, two significant power system outages happened in the Western North American Power System during 1996, where the power system simulations were unable to reflect the real extension of those outages due to inaccurate model parameters [17, 18]. Therefore, an accurate estimation of synchronous generators states and parameters is fundamental to the determination of accurate and adequate power system models, since both electric and electromechanical behaviors of synchronous machines can be predicted by means of equations that describe them [19]. Estimation of dynamic states becomes increasingly challenging and important with the transition from a traditional power system to the smart grid, where faster and system-wide control is desired [20].

By considering this perspective, it is important to realize that the electrical parameters of synchronous machine are used in a variety of power system studies, including short-circuit computation [5], power system stability [21], and sub-synchronous resonances [22]. In steady-state conditions, the knowledge of quadrature- and direct-axis synchronous reactances is necessary to determine, after appropriate saturation adjustments, the maximum value of the reactive output power – which is a function of the field excitation [19].

In short-circuit analyses, the resulting fault current is determined by means of the internal voltage of synchronous generators and the system impedances between the machine voltages and the fault [5]. Furthermore, for transmission lines longer than 300 km, steady-state stability is a factor that imposes limitations on the system operation. Stability refers to the ability of synchronous machines on either end of a line to remain in synchronism [23], after moving from one steady-state operating point to another after a disturbance [24].

Stability programs combine power-flow equations and machine-dynamic equations to compute the angular swings of machines during disturbances. System disturbances can be caused by sudden loss of a generator or a transmission line, sudden load increases or decreases, short-circuits, and line-switching operations [5].

Real-time and accurate data must flow all the way to and from the large central generators, substations, customer loads, and distributed generators, and are necessary for near real-time decision-making and automated actions [25]. On-line monitoring and analysis of power system dynamics using real-time data several times a cycle will make it possible for appropriate control actions to mitigate transient stability problems in a more effective and efficient fashion [26].

Moreover, it is known that the parameters of synchronous machines may drift due to a variety of factors such as: machine-internal temperature, machine aging, magnetic saturation, the coupling effect between the system and the external systems, and so forth [27]. The need for accurate states and parameters estimation arises particularly in on-line stability analysis in which the operational-model parameters may deviate substantially from their rated values.

1.1 Objectives

The general objective of this work is to propose and analyze methods for estimating states and physical parameters of salient-pole synchronous machines. The specific objectives are: (i) to evaluate the load rejection test and to propose an automated analytical approach to it; (ii) to apply the particle filtering on states and parameters estimation and evaluate its performance; and (iii) to propose a simplified approach on the calculation of quadrature- and direct-axis magnetizing reactances from certain load conditions.

1.2 Dissertation structure

Chapter 2 presents essential concepts in the study of salient-pole synchronous machines: such as voltages equations, Park's Transformation, transient- and steady-state operation, and a widely applied off-line method for parameters estimation, which is the load rejection test.

Chapter 3 aims at adapting the machine equations into the state-space representation, which is a very useful tool for states and parameters estimation. In order to do so, Chapter 3 deals with elementary dynamical system analysis concepts.

Since the approach developed in this work to estimate states and parameters of salient-pole synchronous machines is based on the Particle Filter (PF), which is a probability-based, Sequential Monte Carlo (SMC) processor, Chapter 4 presents the Bayesian approach to states estimation.

Chapter 5 brings a literature review, the state-of-the-art, on the different approaches to the estimation of salient-pole synchronous machines physical parameters.

Chapter 6 presents the proposed methodology to determine the machine parameters from certain loading conditions. When this loading condition is met, it becomes possible to estimate the load angle and, from it, calculate the referred parameters. Moreover, Chapter 6 discusses the methodology used for particle filtering and for an automated load-rejection test.

Chapter 7 illustrates the results obtained with the developed methodology, both on simulated and real machine data, as well as observability analyses of different machine models.

Chapter 8 presents final considerations and proposals for future work.

Chapter 2

Salient-Pole Synchronous Generator

"Synchronous machines, when compared to other alternating-current machines, have a great advantage: they operate under the three possible power factors – inductive, capacitive, and resistive – with greater efficiency by simply adjusting their field current."
— Edson Bim, *Máquinas Elétricas e Acionamento*[1]

In this chapter, the fundamental concepts involved in the study of synchronous generators are described. Given the focus of this dissertation, the concepts and models presented throughout this chapter mainly refer to salient-pole synchronous generators.

In practical configurations, such as in a polyphase synchronous machine, the number of terminal pairs is great enough to make the mathematical description seems lengthy. Although it is mathematically complex, the analysis of rotating machines is conceptually simple. As its treatment unfolds, it will become clear that there are geometrical and mathematical symmetries that imply simplification techniques. These techniques have been developed to a high degree of sophistication and are essential in the analysis of machine systems – which may be found in other texts such as the work of White and Woodson [29].

The majority of concepts involved in this chapter are based on the works of Krause et al. [19], Anderson and Fouad [22], Adkins [30], Concordia [31], Elgerd [32], Kundur [33], Kostenko and Piotrovsky [34], Padiyar [35], and Lipo [36]. One of the major contributions of this master's dissertation is the inclusion of appropriate units[2] in every single equation.

2.1 Introduction

Synchronous machines are electromechanical rotating converters that operate at constant speed when in steady state and are mainly used to convert certain sources of mechanical energy into electrical energy [34].

---

[1]Freely translated quotation of "*As máquinas síncronas, quando comparadas com as demais máquinas de corrente alternada [...], têm uma grande vantagem, que é a de funcionar com os três possíveis fatores de potência – indutivo, capacitivo e resistivo – pelo ajuste da corrente de campo e com eficiência maior*" [28].

[2]An important, although brief, compiled of the International System of Units may be found in Annex C.

The main characteristics of these machines consist in:

i) their operating speed, in a steady-state condition, be proportional to the frequency of their armature current, that is,

$$\omega_{sm} = \frac{\omega_e}{\rho} \qquad [\text{mechanical rad/s}] \ , \tag{2.1}$$

where ω_{sm} is the angular frequency of the mechanical shaft, in mechanical radians per second; ω_e is the angular frequency of the generated voltage, in electrical radians per second; and ρ is the number of pole pairs;

ii) their rotor, as well as the magnetic field created by the Direct Current (DC) through the field winding, rotate in synchronism with the rotating magnetic field produced by the armature currents, resulting in a constant torque.

2.2 Physical description

A synchronous generator is essentially composed of two elements: the first element, which is stationary, to produce a rotating magnetic field and the second to couple with the field and to rotate relative to the stationary element, and, thereby, produce electromechanical energy conversion [36]. Voltages are produced in the first element (a set of armature coils) by the relative motion between those two elements. In usual modern machines, the field structure rotates within a stator that supports and provides a magnetic-flux path for the armature winding. The exciting magnetic field is ordinarily produced by a set of coils (the field windings) on the moving element, the so-called rotor [31].

Such synchronous machines configuration is due to the fact that the great majority of them are built to operate under voltage levels above 20 kV and under currents of thousands of amperes[3]; under these conditions, the operation with collector rings, as in DC machines, becomes impractical [34].

By an appropriate excitation of the windings, the field distribution of magnetic flux density in the space that separates the aforementioned elements (the air gap) can be made to rotate relative to the stationary element (synchronous machines), relative to the rotatory element (DC machines), or relative to both elements (induction machines). The interaction of the flux components produced by the stationary and the rotatory elements results in the production of torque.

The construction of a synchronous machine, more specifically of its rotor, depends, fundamentally, on the desired speed of operation. Considering an operating frequency of 60 Hz

[3]The ampere (symbol A) is the base unit of electric current in the International System of Units. It is named after André-Marie Ampère (1775-1836), French mathematician and physicist, considered the father of electrodynamics. He is also the inventor of numerous applications, such as the solenoid – a term coined by him – and the electrical telegraph.

and the velocity-frequency relationship expressed in (2.1), machines of one or two pole pairs rotate at 3600 revolutions per minute (rpm) and 1800 rpm, respectively; while those of 39 pole pairs, such as the ones of Itaipu[4], operate at 92 rpm approximately.

For machines operating at high speeds, the excitation winding is required to be distributed over the entire rotor surface for greater mechanical stiffness, for better resistance to high-intensity centrifugal forces, and for better accommodation to it. These requirements are met by cylindrical rotors of non-salient poles [34].

On the other hand, for the same operating frequency, as the number of pole pairs increases, the operating speed decreases proportionally – accordingly to (2.1). Kostenko and Piotrovsky [34] state that synchronous machines of more than three pole pairs may be constructed with rotors of salient poles aiming at a more simplified construction and, consequently, cost reduction.

The salient-pole rotor consists of a uniform array of magnetic poles projected radially outwards its mechanical axis. The field windings, operated in DC, are concentrated and wrapped around each pole, which must alternate in polarity. Each pole may be dovetailed so that it fits into a wedge-shaped recess or be bolted onto a magnetic wheel called spider[5] [38], which is itself keyed to the shaft [39]. A schematic diagram of such dovetailed configuration is shown in Figure 2.1.

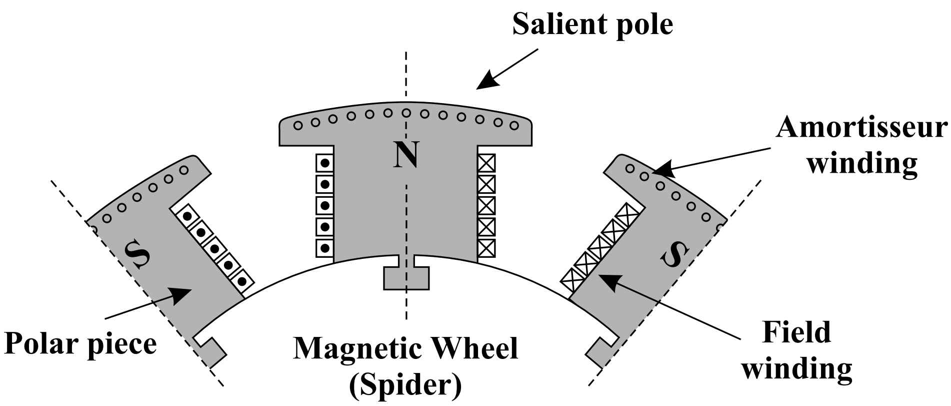

Figure 2.1: Schematic diagram of a salient-pole rotor.

In addition, amortisseur (also known as damper) windings, usually consisting of a set of copper or brass bars, may be attached to the pole-face slots and connected at the ends of the machine, as shown in Figure 2.2. This amortisseur winding has several useful functions, including: to permit the starting of synchronous motors as induction motors using the amortisseur as equivalent to the squirrel cage of an induction-motor rotor; to assist in damping rotor

[4]The Itaipu Hydroelectric Power Plant (launch in 1984) is a bi-national hydroelectric power plant located on the Paraná River, on the border between Brazil and Paraguay, whose generating units have 39 pole pairs [28].

[5]A structure supporting the core or poles of a rotor from the shaft, and typically consisting of a hub, spokes, and rim, or some modified arrangement of these [37, p. 1086].

oscillations; to reduce overvoltages under certain short-circuit conditions; and to aid at the machine synchronization [31]. The space harmonics of the armature magnetomotive force (mmf) contribute to surface Foucault current[6] losses [40]; therefore, the pole faces of salient-pole machines are usually laminated [33].

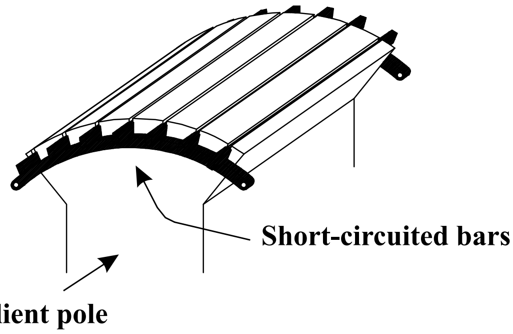

Figure 2.2: Schematic diagram of amortisseur windings.
Adapted from Bim [28, p. 191].

The stator of synchronous machines much resembles that of asynchronous machines, being composed of thin sheets of highly permeable steel to reduce core losses. These sheets are held superimposed by the action of the fingers and pressing plates, creating the stator core. The fingers are manufactured to avoid conducting magnetic flux and the pressing plates are in the back of the core, and can be manufactured with regular steel. The stator core is keyed to the stator frame, which provides mechanical support to the machine. Inside the stator core, there are several slots, whose function is to accommodate the thick armature conductors [38]. In a conventional three-phase synchronous machine, the armature conductors are symmetrically spaced to form a balanced three-phase winding. For large machines, although it is more common to adopt a fractional number of slots per pole per phase, another possible winding pattern is shown in Figure 2.3 for a three-phase, two-pole-pair, 36-slot machine – as it can be verified, there are three slots per pole per phase.

The armature of most synchronous machines is coiled with three separated independent windings to generate three-phase power. Each of these windings represents one of the three phases of a three-phase machine. To ensure that the generated electromotive forces (emfs) are periodic waves, close to sinusoids, and lagged at $2\pi/3$ radians in time, the windings are identical in shape and are spaced apart from each other by $2\pi/3$ electrical radians in space.

The steady-state voltages produced, under balanced load conditions, are always $2\pi/3$ radians apart in phase regardless of the speed of rotation of the field. That is:

1. because $1/\rho$ revolution (a displacement equal to the space occupied by one pole pair) will always correspond to one cycle of the generated voltage (i.e., the fundamental frequency will always be exactly ρ times the speed of rotation);

[6]Foucault current is the name given to induced currents in a relatively large conductive material when subjected to a variable magnetic flux. The name was given in acknowledgment to the French physicist Jean Bernard Léon Foucault (1819-1868), who studied that effect in 1855.

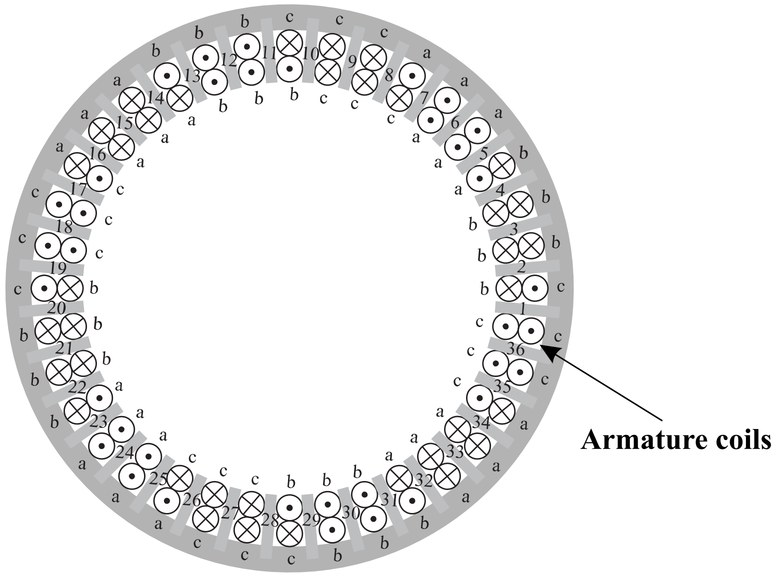

Figure 2.3: Schematic diagram of a stator double-layer winding for a three-phase, two-pole-pair, 36-slot machine. Adapted from Krause et al. [19, p. 62].

2. and, because with constant rate of rotation, the time required for the rotor to move any given distance is proportional to the distance moved,

the time required for the field to move from any given position with respect to one coil to the corresponding position with respect to the equivalent coil of the following phase is just one third of a cycle, or $2\pi/3$ electrical radians [31].

When carrying balanced three-phase currents, the armature will produce a magnetic field in the air gap rotating at synchronous speed. The magnetic field produced by the direct current in the rotor winding, on the other hand, revolves with the rotor. For a constant torque production, the stator and rotor magnetic fields must rotate at the same speed. Therefore, the rotor must precisely run at the electrical synchronous speed [33].

2.3 Direct and quadrature axes

In the analysis of electric machines, two important concepts are commonly used: the direct and quadrature axes. A precise definition for them is found in the *Authoritative Dictionary of Institute of Electrical and Electronics Engineers (IEEE) Standards Terms*:

> **direct-axis (synchronous machines)**: the axis that represents the direction of the plane of symmetry of the no-load magnetic-flux density, produced by the main field winding current, normally coinciding with the radial plane of symmetry of a field pole [37, p. 310];
>
> **quadrature-axis (synchronous machines)**: the axis that represents the direction of the radial plane along which the main field winding produces no magnetization, normally coinciding with the radial plane midway between adjacent poles. The positive direction of the quadrature-axis is 90 [electrical] degrees ahead of the positive direction of the direct-axis, in the direction of rotation of the field relative to the armature [37, p. 899].

Therefore, one important assumption to derive the salient-pole synchronous machine equations is that the magnetic circuits and all rotor windings are symmetrical with respect to both polar and inter-polar axes.

Although the selection of the quadrature-axis as leading the direct-axis may be purely arbitrary [33], this work bases itself on the widely used [19, 28, 34, 36] IEEE convention shown above. Alternatively, some works [22, 41, 42] choose the quadrature-axis to lag the direct-axis by $\pi/2$ electrical radians.

In some works, the rotor's position relative to the stator is measured by the angle between the direct-axis and the magnetic axis of phase-*a* winding [31, 33, 36, 43, 44]. This work, on the contrary, follows the notation used by Krause et al. [19], measuring the aforementioned position by the angle from the magnetic axis of phase-*a* winding to the quadrature-axis.

The concept of resolving synchronous-machine armature quantities into two rotating components – as will be demonstrated – was introduced as a means of facilitating the analyses of salient-pole machines.

2.4 Mathematical description

In order to achieve a complete understanding of the behavior of a synchronous machine in transient and steady-state operating conditions, it becomes mandatory to develop its equations. Some hypotheses are made to simplify and ease the following development and will be presented as necessary.

Elgerd [32] corroborates rather brilliantly why the method used in this work should be applied:

> Classically, the theory of synchronous machine was presented in terms of traveling air-gap flux, current, and emf waves. This theory has the advantage of close adherence to the physical realities within the machine and serves the limited purpose of explaining its elementary steady-state operating characteristics. This approach becomes extremely impractical when it becomes necessary to expose the behavior of the machine under transient conditions and its interplay with the external network. [...] the central feature of the method to be used is the exclusive use of the circuit concept; the machine is considered as a set of magnetically coupled circuits, the main parameters of which are time-variant. [32, p. 77]

The following development is based on the works of Krause et al. [19], Adkins [30], Concordia [31], Elgerd [32], Kundur [33], Lipo [36], and Kron [45], to which one should refer for further details.

A brief note on the notation to be used:

The terminology and notation used in developing the general theory follow, in most respects, those used in the papers and books listed in the bibliography. The symbols and names used for the constants of the synchronous machine, for example, are very well established.

In the differential equations, the Heaviside[7] notation is used. It is used by Adkins [30] and Kron [45] and many other writers on electrical machine theory. According to Adkins [30], the Heaviside notation is advantageous for expressing the general equations of machines because they are non-linear. The Laplace transform notation, on the other hand, is suitable for the study of circuits and control systems because, for these subjects, the equations used in developing the basic theory are linear. Furthermore, the Heaviside method can be used for manipulating the equations under certain conditions, for example, when some are linear, and some are non-linear. Laplace transforms cannot be used for this purpose [30].

The Heaviside operational method [46], introduced by Heaviside in the early days of circuit analysis, replaces d/dt by p in the equations, and threats the operator p as an algebraic quantity. Operational calculus is of great assistance in handling differential equations arising in the analysis of electrical machines. It is valuable for stating the equations in an abbreviated form, for manipulating them, and, in certain types of problem, for obtaining the solution.

2.4.1 Flux linkage and inductance

When a magnetic field varies with time, an electric field is produced in space as determined by Faraday[8]'s law:

$$\oint_C \mathbf{E} \cdot \mathbf{ds} = -p \int_S \mathbf{B} \cdot \mathbf{da}\,, \tag{2.2}$$

which states that the line integral of the electric field intensity **E** around a closed contour C is equal to the time rate of change of the magnetic flux passing through that contour. In magnetic structures with windings of high-electrical conductivity, it can be shown that the electric field in the wire is extremely small and can be neglected, so that the left-hand side of Faraday's Law reduces to the negative of the induced voltage e at the winding terminals. In addition, the flux on the right-hand side is dominated by the core flux [44]. Since the winding links the core flux N times, Faraday's law reduces to:

$$e = -Np\Phi \tag{2.3a}$$

$$= -p\psi \qquad [\mathrm{V}]\,, \tag{2.3b}$$

[7]Oliver Heaviside (1850–1925), Fellow of the Royal Society, was an English self-taught electrical engineer, mathematician, and physicist who adapted complex numbers to the study of electrical circuits, invented mathematical techniques for the solution of differential equations (equivalent to Laplace transforms), reformulated Maxwell's field equations in terms of electric and magnetic forces and energy flux, and independently co-formulated vector analysis. Although at odds with the scientific establishment for most of his life, Heaviside changed the face of telecommunications, mathematics, and science for years to come.

[8]Michael Faraday (1791–1867) was a British scientist who contributed to the study of electromagnetism and electrochemistry. His main discoveries include the principles underlying electromagnetic induction, diamagnetism, and electrolysis.

where ψ[9] is the total flux linkage[10] of the winding; and Φ is the instantaneous value of a time-varying flux.

In an idealization of an actual-magnetic system, the flux produced by a coil can be separated into two components: a leakage component and a magnetizing component. The distinction between them is not always precise. However, leakage flux is associated with flux that does not travel across the air gap or couple both the rotor and the stator windings. Magnetizing flux linkage, on the other hand, is associated with radial-flux flow across the air gap and links both the stator and rotor windings [19].

As an example, let the magnetic circuit shown in Figure 2.4. It shows two stationary electric circuits that are magnetically coupled. The two coils consist of turns N_1 and N_2, respectively, and they are wound on a common core with a large permeability[11] if compared to that of the air.

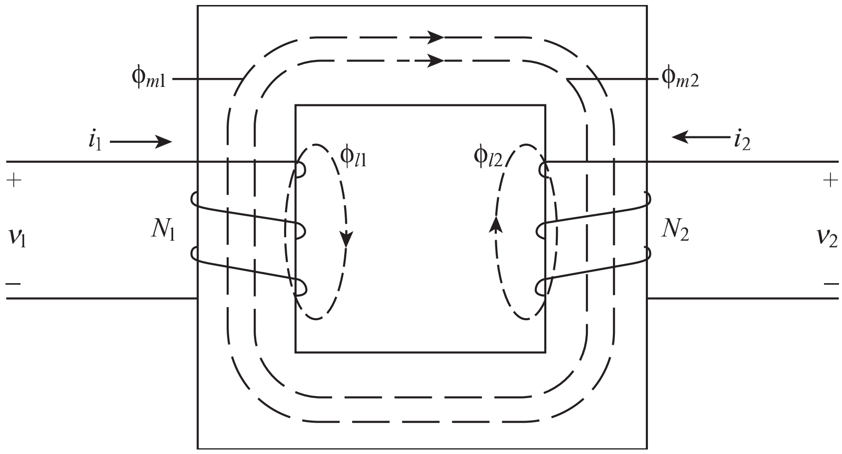

Figure 2.4: Two magnetically coupled stationary circuits.
Adapted from Krause et al. [19, p. 2].

The flux linking each coil may be expressed as

$$\Phi_1 = \Phi_{l1} + \Phi_{m1} + \Phi_{m2} \qquad [\mathrm{Wb}] \ , \tag{2.4}$$

$$\Phi_2 = \Phi_{l2} + \Phi_{m2} + \Phi_{m1} \qquad [\mathrm{Wb}] \ . \tag{2.5}$$

The leakage flux Φ_{l1} is produced by current flowing in coil 1, and it links only the turns of coil 1. The magnetizing flux Φ_{m1} is produced by current flowing in coil 1, and it links all turns of coils 1 and 2. The same analysis follows to coil 2.

---

[9]In circuit analysis, the symbol λ is commonly used to denote flux linkage, whereas in the most of the literature on synchronous machines and power system stability the symbol ψ is used. Here, the latter practice is followed to correspond with the published literature and to avoid confusion to the common use of λ to denote eigenvalues.

[10]Flux linkage is measured in units of webers (or equivalently weber-turns). The weber is named after the German physicist Willheim Eduard Weber (1804-1891) who, together with Carl Friedrich Gauss, invented the first electromagnetic telegraph.

[11]The magnetic permeability of free space, μ_0, is $4\pi \times 10^{-7}$ H/m. The permeability of other materials is expressed as $\mu = \mu_r \mu_0$, where μ_r is the relative permeability. In the case of transformer steel, the relative permeability may be as high as 2000-4000 [19].

If saturation is neglected, the magnetic system is magnetically linear and there is a proportional relation between currents and fluxes. This first hypothesis is important and make it possible to consider the concept of inductance[12]: when the magnetic system is linear, the flux linkages are generally expressed in terms of inductances and currents.

In terms of flux linkages, (2.4) becomes

$$\psi_1 = \frac{N_1^2}{\mathfrak{R}_{l1}} i_1 + \frac{N_1^2}{\mathfrak{R}_m} i_1 + \frac{N_1 N_2}{\mathfrak{R}_m} i_2 \qquad \text{[Wb-t]} , \tag{2.6}$$

where $\mathfrak{R}_{l1}$ is the reluctance[13] of the leakage path; $\mathfrak{R}_m$ is the reluctance of the magnetizing flux path; and i_1 and i_2 are the currents flowing through coils 1 and 2, respectively.

The coefficients of the first two terms on the right-hand side of (2.6) depend upon the turns of coil 1 and the reluctance of the magnetic system, independent of the existence of coil 2. The last term relates both coils 1 and 2.

Hence, the self-inductance L_1 of coil 1 is defined by the coefficients of the first-two terms on the right-hand side of (2.6) as

$$L_1 = \frac{N_1^2}{\mathfrak{R}_{l1}} + \frac{N_1^2}{\mathfrak{R}_m} \tag{2.7a}$$

$$= L_{l1} + L_{m1} \qquad \text{[H]} , \tag{2.7b}$$

and the mutual inductances by the coefficient of the third term on the right-hand side:

$$L_{12} = \frac{N_1 N_2}{\mathfrak{R}_m} \qquad \text{[H]} . \tag{2.8}$$

An analogous statement may be made regarding coil 2.

The flux linkages may now be written in matrix form as

$$\boldsymbol{\psi} = \mathbf{L}\mathbf{i} \qquad \text{[Wb-t]} , \tag{2.9}$$

where

$$\mathbf{L} = \begin{bmatrix} L_{11} & L_{12} \\ L_{21} & L_{22} \end{bmatrix} = \begin{bmatrix} L_{l1} + L_{m1} & \frac{N_2}{N_1} L_{m1} \\ \frac{N_1}{N_2} L_{m2} & L_{l2} + L_{m2} \end{bmatrix} \qquad \text{[H]} , \tag{2.10}$$

$\mathbf{L} \in \mathbb{R}^{l \times l}$, $\mathbf{i} \in \mathbb{R}^l$, and $\boldsymbol{\psi} \in \mathbb{R}^l$, where l is the number of coils in the magnetic circuit.

The expansion of (2.9) results in

$$\psi_1 = L_{l1} i_1 + L_{m1} \left(i_1 + \frac{N_2}{N_1} i_2 \right) \qquad \text{[Wb-t]} , \tag{2.11}$$

$$\psi_2 = L_{l2} i_2 + L_{m2} \left(i_2 + \frac{N_1}{N_2} i_1 \right) \qquad \text{[Wb-t]} . \tag{2.12}$$

[12]Inductance is measured in henrys (H) or weber-turns per ampere. The unit is named after Joseph Henry (1797-1878), the American scientist who discovered electromagnetic induction independently of and at about the same time as Michael Faraday in England.

[13]Magnetic reluctance is a concept used in the analysis of magnetic circuits. It is defined as the ratio of mmf to magnetic flux. It represents the opposition to magnetic flux and depends on the geometry and composition of an object. The term was coined in 1888 by Oliver Heaviside, and first mentioned as a "magnetic resistance" by James Joule in 1840.

2.4.2 The elementary parameters of a synchronous machine

For the purpose of energy conversion, all conventional machines rely upon magnetic fields. A valid approach to the study of electric machines is to deal directly with these electromagnetic fields. The complete knowledge of the field distribution leads to a deeper understanding of where the fluxes are concentrated, where the electric currents flow, where the forces appear, and where heat is generated within the machine. Such detailed information is very important, since relatively small alterations in the design can often lead to substantial improvements in efficiency, cost, or reliability. Unfortunately, the analysis of machines as a fields problem involves the solution of Laplace[14]'s or Poisson[15]'s equation. The machines geometry leads to complicated boundary conditions even for simplified cases.

The approach adopted in this work aims at characterizing the machine in terms of coupled magnetic circuits rather than magnetic fields. The primary interest is restricted to the terminal rather than internal characteristics of machines. Although the exact spatial distribution of currents and fluxes is lost, the problem becomes immensely simplified. Furthermore, the significant effects of the rotating fields must be properly expressed in terms of flux linkages in rotating coupled circuits. Since flux linkage is proportional to inductance, the ability to characterize winding distributions and utilize this characterization in the calculation of winding inductances is of central importance for determining the machines parameters.

All the elementary parameters of a synchronous machine and their related equations are derived considering the one-pole-pair, three-phase, wye-connected, salient-pole synchronous machine shown in Figure 2.5. For the sake of simplicity, only one damper winding is explicitly assumed in each axis. However, an arbitrary number of such circuits is implicitly considered; a subscript k is used to denote this.

Concerning this matter, Krause et al. [19] affirm:

> The behavior of low-speed hydro turbine generators, which are always salient-pole [synchronous] machines, is generally predicted sufficiently by one equivalent damper winding in the quadrature-axis. [On the other hand,] it is necessary, in most cases, to include three damper windings in order to portray adequately the transient characteristics of the stator variables and the electromagnetic torque of solid iron rotor machines [19, p.145].

The statement above justifies the use of only one damper winding in the quadrature-axis in this work, as it concerns the study of salient-pole synchronous machines.

In Figure 2.5, the stator windings are identical, displaced $2\pi/3$ electrical radians apart from one another. The rotor is equipped with a field winding and two damper windings. The

[14]Pierre-Simon, Marquis de Laplace (1749–1827) was a French scholar whose work was important to the development of engineering, mathematics, statistics, physics, and astronomy. His work translated the geometric study of classical mechanics to one based on calculus, opening up a broader range of problems. In statistics, the Bayesian interpretation of probability was developed mainly by Laplace.

[15]Baron Siméon Denis Poisson (1781–1840) was a French mathematician, engineer, and physicist who made important contributions to potential theory, optics, pure mathematics, mechanics, and others.

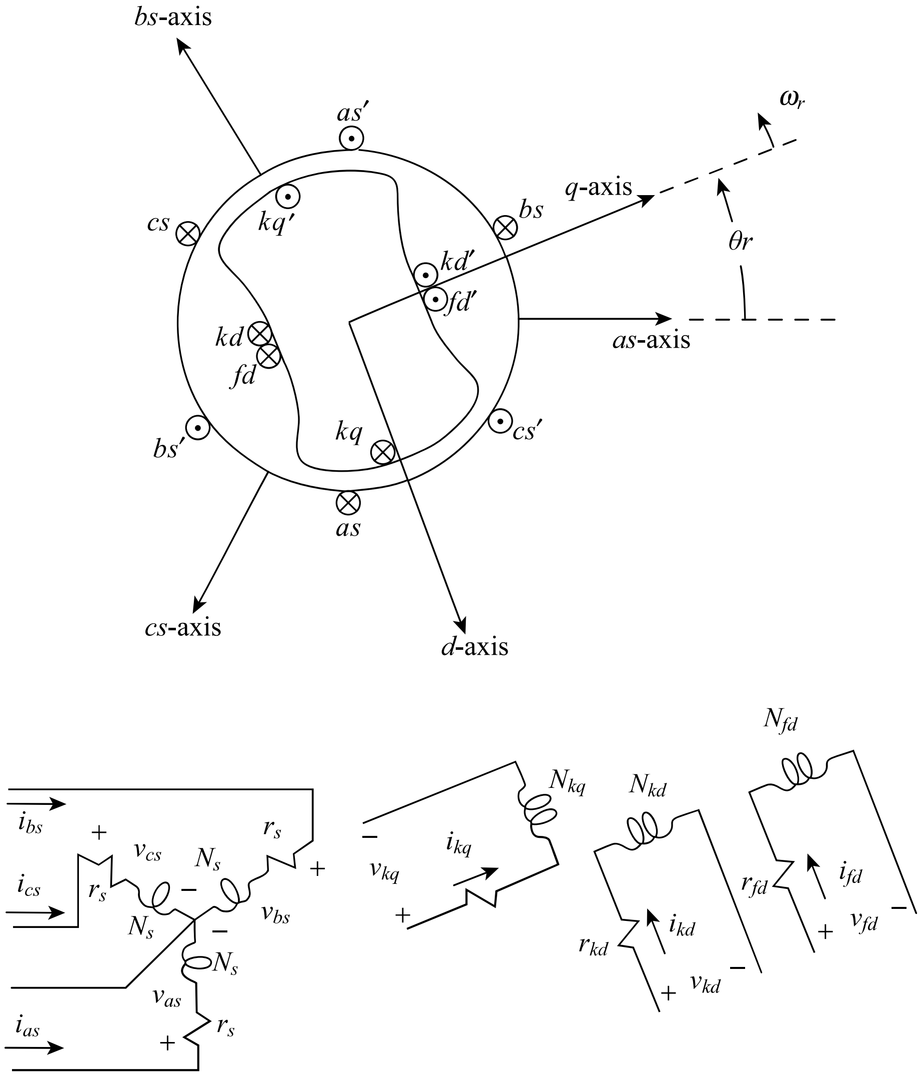

Figure 2.5: A one-pole-pair, three-phase, wye-connected, salient-pole synchronous machine. Adapted from Krause et al. [19, p. 144].

field winding, fd, has N_{fd} equivalent turns with resistance r_{fd}. The direct-axis damper winding, the kd winding, has the same magnetic axis as the field winding. It has N_{kd} equivalent turns with resistance r_{kd}. The magnetic axis of the second winding, the kq winding, is displaced $\pi/2$ electrical radians ahead of the magnetic axis of the fd winding. The kq winding has N_{kq} equivalent turns with resistance r_{kq}.

Furthermore, the magnetic axes of the stator windings are denoted by the as, bs, and cs. The quadrature-axis (q-axis) and direct-axis (d-axis) are also shown. The q-axis is the magnetic axis of the kq winding, while the d-axis is the magnetic axis of the fd and kd windings.

The mechanical rotor angle, θ_r, is defined as the angle by which the q-axis leads the as-axis in the direction of rotation. Since the rotor is rotating with respect to the stator, the angle θ_r is continuously increasing and is related to the rotor angular speed, ω_r, and time, t, by

$$\theta_r = \omega_r t \qquad [\text{electrical rad}] \ , \tag{2.13}$$

where the angle θ_r is measured in electrical radians; and the velocity ω_r, in electrical radians per second.

Moreover, it is important to notice that although the damper windings are shown with provisions to apply a voltage, they are, in fact, short-circuited windings that represent the paths for induced rotor currents [19]. As the rotor of salient-pole synchronous machines is laminated, the damper winding currents are confined, for the most part, to the cage windings embedded in the rotor.

All presented elements are briefly summarized in Table 2.1.

Table 2.1: Summary of the elements of Figure 2.5.

Element	Meaning
as, bs, cs	stator phase windings
fd	field winding
kq	q-axis amortisseur winding
kd	d-axis amortisseur winding
θ_r	angle by which the q-axis leads the magnetic axis of phase *as*
ω_r	rotor angular velocity

■

To derive the armature and rotor self- and mutual inductances, as well as the mutual inductances between stator and rotor, the following assumptions are initially made:

i) the rotor-magnetic paths and all its electric circuits are symmetrical about both the pole and interpole axes for a salient-pole machine. This assumption has the virtue of making all mutual inductances and resistances between direct- and quadrature-axis rotor circuits equal to zero;

ii) the field winding is separate from the others and has its axis in line with the pole axis. Although this winding is generally concentrated, its effects are represented by an equivalent sinusoidally distributed winding which produces the same fundamental component of mmf in the air gap;

iii) the amortisseur bars are all connected in a more or less continuous mesh;

iv) the quadrature-axis is taken as $\pi/2$ electrical radians ahead of the direct-axis in the direction of normal-rotor rotation;

v) all mutual inductances between stator and rotor circuits are periodic functions of rotor angular position;

vi) because of the rotor salience, the mutual inductances between any two stator phases are also periodic functions of rotor-angular position;

vii) the stator windings are sinusoidally distributed along the air gap as far as all mutual effects with the rotor are concerned;

viii) the stator slots cause no appreciable variation of any of the rotor inductances with rotor angle;

ix) all electrical parameters are assumed constant, independent of temperature and frequency.

Assumptions (i)-(vi) lead to a set of differential equations most of whose coefficients are periodic functions of rotor angle, so that even in the case of constant rotor speed – when the equations are linear if saturation is neglected – they are awkward to handle and difficult to solve. However, if certain reasonable assumptions are made, a relatively simple transformation of variable will eliminate all these troublesome functions of angle from the equations.

2.4.2.1 Armature self-inductances

The self-inductance of any armature winding varies periodically from a maximum, when the pole axis is aligned with the phase axis, to a minimum, when the interpole axis is aligned with the phase axis. Because of the symmetry of the rotor, the armature self-inductance must have a period of π electrical radians and must be expressed by a series of cosines of even harmonics of angle [31]. Under assumption (vii), only the first two terms of the series are significant.

Therefore, the inductance variation is considered harmonic, i.e.,

$$\ell_{aa} = \ell_{aa0} + \ell_{aa2} \cos 2\theta_r \qquad [\mathrm{H}] \ , \tag{2.14a}$$

where θ_r is the angle of the quadrature-axis from the axis of phase-a, measured in the direction of rotor rotation; and the ℓ's are inductances to be defined later, whose subscripts refer to the circuits under analysis. Similarly,

$$\ell_{bb} = \ell_{aa0} + \ell_{aa2} \cos\left[2\left(\theta_r - 2\pi/3\right)\right] \qquad [\mathrm{H}] \ , \tag{2.14b}$$

$$\ell_{cc} = \ell_{aa0} + \ell_{aa2} \cos\left[2\left(\theta_r - 4\pi/3\right)\right] \qquad [\mathrm{H}] \ . \tag{2.14c}$$

When it comes to magnetic fluxes, because of assumption (vii) of sinusoidal distribution of stator windings along the air gap, the electric current in phase-a produces a mmf space wave in the air gap which is only of fundamental span frequency as far as the rotor is concerned. This may be conveniently broken up into two components proportional to $(\sin\theta_r)$ and $(-\cos\theta_r)$ acting in direct- and quadrature-axis, respectively [31].

These components of mmf in phase-a produce corresponding components of flux, having space fundamental components of magnitude

$$\Phi_d = \mathcal{P}_d \sin\theta_r \qquad [\mathrm{Wb}] \ , \tag{2.15a}$$

$$\Phi_q = -\mathcal{P}_q \cos\theta_r \qquad [\mathrm{Wb}] \ , \tag{2.15b}$$

where $\mathcal{P}_d$ and $\mathcal{P}_q$ are proportional to effective permeance coefficients in the direct and quadrature axes, respectively, and to the mmf. The linkage with phase-a caused by this flux is then proportional to:

$$\Phi_d \sin\theta_r - \Phi_q \cos\theta_r = \mathcal{P}_d \sin^2\theta_r + \mathcal{P}_q \cos^2\theta_r \tag{2.16a}$$

$$= \frac{\mathcal{P}_d + \mathcal{P}_q}{2} + \frac{\mathcal{P}_q - \mathcal{P}_d}{2}\cos 2\theta_r \tag{2.16b}$$

$$= \mathcal{K}_1 + \mathcal{K}_2 \cos 2\theta_r \qquad [\mathrm{Wb}]\ . \tag{2.16c}$$

There is also some flux linking phase-a that does not link the rotor. This flux has no relation with the rotor position and, thus, adds only to the $\mathcal{K}_1$ constant in (2.16c) [31].

In summary, due to the salience of the rotor, the stator windings experience a change in self-inductance as the rotor rotates, which may be approximated as a double-angle variation about an average value [19, 31, 32].

2.4.2.2 Armature mutual inductances

To determine the form of the mutual inductance between, e.g., phases a and b, it is important to recognize that there may be a component of mutual flux that does not link the rotor and is thus independent of angle. Then, considering the mmf generated in phase-a, the components of air gap flux are, as before, those shown in (2.15), and the linkage with phase b due to these components is proportional to

$$\Phi_d \sin\theta_b - \Phi_q \cos\theta_b = \mathcal{P}_d \sin\theta_r \sin\theta_b + \mathcal{P}_q \cos\theta_r \cos\theta_b \tag{2.17a}$$

$$= \mathcal{P}_d \sin\theta_r \sin(\theta_r - \pi/3) + \mathcal{P}_q \cos\theta_r \cos[2(\theta_r - \pi/3)] \tag{2.17b}$$

$$= -\frac{\mathcal{P}_q + \mathcal{P}_d}{4} + \frac{\mathcal{P}_q - \mathcal{P}_d}{2}\cos[2(\theta_r - \pi/3)] \tag{2.17c}$$

$$= -\frac{1}{2}\mathcal{K}_1 - \mathcal{K}_2 \cos[2(\theta_r - \pi/3)] \qquad [\mathrm{Wb}]\ . \tag{2.17d}$$

The total mutual inductance is thus of the form

$$\ell_{ab} = -\left[\ell_{ab0} + \ell_{aa2}\cos[2(\theta_r - \pi/3)]\right] \qquad [\mathrm{H}]\ . \tag{2.18}$$

The variable part of the mutual inductance is of exactly the same amplitude as that of the variable part of the self-inductance and the constant part has a magnitude close to the half that of the constant part of the self-inductance [31].

Finally, all stator mutual inductances may be written as

$$\ell_{ab} = \ell_{ba} = -\left[\ell_{ab0} + \ell_{aa2}\cos[2(\theta_r - \pi/3)]\right] \qquad [\mathrm{H}]\ , \tag{2.19a}$$

$$\ell_{bc} = \ell_{cb} = -\left[\ell_{ab0} + \ell_{aa2}\cos[2(\theta_r + \pi)]\right] \qquad [\mathrm{H}]\ , \tag{2.19b}$$

$$\ell_{ca} = \ell_{ac} = -\left[\ell_{ab0} + \ell_{aa2}\cos[2(\theta_r + \pi/3)]\right] \qquad [\mathrm{H}]\ . \tag{2.19c}$$

2.4.2.3 Rotor self-inductances

Considering assumption (viii) and neglecting saturation effects, the rotor self-inductances $\ell_{fdfd}, \ell_{kdkd}, \ell_{kqkq}$ are constants.

2.4.2.4 Rotor mutual inductances

All mutual inductances between any two circuits in the direct-axis and between any two circuits both in the quadrature-axis are constant. Because of assumption (i) of rotor symmetry, there is no mutual inductance between any direct- and any quadrature-axis circuit. Thus,

$$\ell_{fdkq} = \ell_{kdkq} = \ell_{kqfd} = \ell_{kqkd} = 0, \text{ etc.} \qquad [\mathrm{H}] \ . \tag{2.20}$$

2.4.2.5 Mutual inductances between stator and rotor circuits

By considering current in each rotor winding in turn and recalling that only the space-fundamental component of the flux produced will link the sinusoidally distributed stator – under assumption (vii) – all stator-rotor mutual inductances vary sinusoidally with angle and are maximum when the two coils under analysis are aligned with one another. Thus:

$$\ell_{afd} = \ell_{fad} = \ell_{afd} \sin \theta_r \qquad [\mathrm{H}] \ , \tag{2.21a}$$
$$\ell_{bfd} = \ell_{fbd} = \ell_{afd} \sin (\theta_r - 2\pi/3) \qquad [\mathrm{H}] \ , \tag{2.21b}$$
$$\ell_{cfd} = \ell_{fcd} = \ell_{afd} \sin (\theta_r - 4\pi/3) \qquad [\mathrm{H}] \ , \tag{2.21c}$$

$$\ell_{akd} = \ell_{kda} = \ell_{akd} \sin \theta_r \qquad [\mathrm{H}] \ , \tag{2.21d}$$
$$\ell_{bkd} = \ell_{kdb} = \ell_{akd} \sin (\theta_r - 2\pi/3) \qquad [\mathrm{H}] \ , \tag{2.21e}$$
$$\ell_{ckd} = \ell_{kdc} = \ell_{akd} \sin (\theta_r - 4\pi/3) \qquad [\mathrm{H}] \ , \tag{2.21f}$$

$$\ell_{akq} = \ell_{kqa} = \ell_{akq} \cos \theta_r \qquad [\mathrm{H}] \ , \tag{2.21g}$$
$$\ell_{bkq} = \ell_{kqb} = \ell_{akq} \cos (\theta_r - 2\pi/3) \qquad [\mathrm{H}] \ , \tag{2.21h}$$
$$\ell_{ckq} = \ell_{kqc} = \ell_{akq} \cos (\theta_r - 4\pi/3) \qquad [\mathrm{H}] \ . \tag{2.21i}$$

Altogether, it is important to observe that all inductance elements can be expressed in terms of a set of six positive inductance parameters $\ell_{aa0}, \ell_{aa2}, \ell_{ab0}, \ell_{akq}, \ell_{afd}, \ell_{akd}$ and the rotor position angle, θ_r. Also, in all above expressions, the angle θ_r must be understood to represent the electrical angle [32]. As shown in Figure 2.5, the electrical and mechanical angles are identical for a one-pole-pair machine. For a generic ρ-pole-pair machine, the electrical angle corresponds to ρ times the mechanical angle.

Following a notation that will be useful when the machine equations are treated in the state space, the following equations present the inductances previously developed in matrix notation. Also, the ℓ's adopted for them will now be replaced by the corresponding symbols:

$$\ell_{aa0} = L_{ls} + L_A, \quad \ell_{aa2} = -L_B, \quad \ell_{ab0} = \frac{1}{2} L_A, \quad \ell_{akq} = L_{akq}, \quad \ell_{afd} = L_{afd},$$

$$\ell_{akd} = L_{akd}, \quad \ell_{kqkq} = L_{lkq} + L_{mkq}, \quad \ell_{fdfd} = L_{lfd} + L_{mfd}, \quad \ell_{kdkd} = L_{lkd} + L_{mkd}.$$

The stator inductance matrix $\mathbf{L}_s \in \mathbb{R}^{3\times 3}$ is

$$\mathbf{L}_s = \begin{bmatrix} L_{ls} + L_A - L_B \cos 2\theta_r & -\frac{1}{2}L_A - L_B \cos[2(\theta_r - \pi/3)] & -\frac{1}{2}L_A - L_B \cos[2(\theta_r + \pi/3)] \\ -\frac{1}{2}L_A - L_B \cos[2(\theta_r - \pi/3)] & L_{ls} + L_A - L_B \cos[2(\theta_r - 2\pi/3)] & -\frac{1}{2}L_A - L_B \cos[2(\theta_r + \pi)] \\ -\frac{1}{2}L_A - L_B \cos[2(\theta_r + \pi/3)] & -\frac{1}{2}L_A - L_B \cos[2(\theta_r + \pi)] & L_{ls} + L_A - L_B \cos[2(\theta_r - 4\pi/3)] \end{bmatrix} \quad [\mathrm{H}]. \tag{2.22}$$

The stator-rotor inductance matrix $\mathbf{L}_{sr} \in \mathbb{R}^{3\times(k+1)}$, where k is the number of damper windings, is

$$\mathbf{L}_{sr} = \begin{bmatrix} L_{akq}\cos\theta_r & L_{afd}\sin\theta_r & L_{akd}\sin\theta_r \\ L_{akq}\cos(\theta_r - 2\pi/3) & L_{afd}\sin(\theta_r - 2\pi/3) & L_{akd}\sin(\theta_r - 2\pi/3) \\ L_{akq}\cos(\theta_r - 4\pi/3) & L_{afd}\sin(\theta_r - 4\pi/3) & L_{akd}\sin(\theta_r - 4\pi/3) \end{bmatrix} \quad [\mathrm{H}]. \tag{2.23}$$

Finally, the rotor inductance matrix is $\mathbf{L}_r \in \mathbb{R}^{(k+1)\times(k+1)}$ is:

$$\mathbf{L}_r = \begin{bmatrix} L_{lkq} + L_{mkq} & 0 & 0 \\ 0 & L_{lfd} + L_{mfd} & L_{fdkd} \\ 0 & L_{fdkd} & L_{lkd} + L_{mkd} \end{bmatrix} \quad [\mathrm{H}]. \tag{2.24}$$

In (2.22), $L_A > L_B$ and $L_B = 0$ for round rotor machine. In (2.22) and (2.24), the subscript l denotes the leakage inductances and, in (2.23), the subscripts akq, afd, and akd denote mutual inductances between stator and rotor windings.

The equivalent circuit that has been obtained is still rather complex, since the mutual inductance matrix, $\mathbf{L}_{sr}$, is non-symmetrical. In order to establish a simpler representation, a change of variables, such that the magnetizing inductances corresponding to each mesh current in a given axis are identical to each other, becomes necessary. Defining the magnetizing inductances as

$$L_{mq} = \frac{3}{2}(L_A - L_B) \quad [\mathrm{H}], \tag{2.25a}$$

$$L_{md} = \frac{3}{2}(L_A + L_B) \quad [\mathrm{H}], \tag{2.25b}$$

it can be shown [19] that

$$L_{akq} = \frac{2}{3}\left(\frac{N_{kq}}{N_s}\right) L_{mq} \quad [\mathrm{H}], \tag{2.26a}$$

$$L_{afd} = \frac{2}{3}\left(\frac{N_{fd}}{N_s}\right) L_{md} \quad [\mathrm{H}], \tag{2.26b}$$

$$L_{akd} = \frac{2}{3}\left(\frac{N_{kd}}{N_s}\right) L_{md} \quad [\mathrm{H}], \tag{2.26c}$$

$$L_{mkq} = \frac{2}{3}\left(\frac{N_{kq}}{N_s}\right)^2 L_{mq} \quad [\mathrm{H}], \tag{2.26d}$$

$$L_{mfd} = \frac{2}{3}\left(\frac{N_{fd}}{N_s}\right)^2 L_{md} \quad [\mathrm{H}], \tag{2.26e}$$

$$L_{mkd} = \frac{2}{3}\left(\frac{N_{kd}}{N_s}\right)^2 L_{md} \qquad [\mathrm{H}] \ , \tag{2.26f}$$

$$L_{fdkd} = \frac{2}{3}\left(\frac{N_{fd}}{N_{kd}}\right) L_{md} \qquad [\mathrm{H}] \ . \tag{2.26g}$$

In order to refer the rotor resistances and inductances to the stator windings, another set of variables transformation must be considered:

$$r'_j = \frac{3}{2}\left(\frac{N_s}{N_j}\right)^2 r_j \, , \tag{2.27}$$

$$L'_{lj} = \frac{3}{2}\left(\frac{N_s}{N_j}\right)^2 L_{lj} \, , \tag{2.28}$$

where j may be kq, fd, or kd.

The flux linkages may now be written in terms of machine variables referred to the stator windings as

$$\begin{bmatrix} \boldsymbol{\psi}_{abcs} \\ \boldsymbol{\psi'}_{qdr} \end{bmatrix} = \begin{bmatrix} \mathbf{L}_s & \mathbf{L}'_{sr} \\ \frac{2}{3}(\mathbf{L'}_{sr})^{\mathrm{T}} & \mathbf{L}'_r \end{bmatrix} \begin{bmatrix} \mathbf{i}_{abcs} \\ \mathbf{i'}_{qdr} \end{bmatrix} \qquad [\mathrm{Wb\text{-}t}] \ , \tag{2.29}$$

where $\mathbf{L}_s$ is defined by (2.22),

$$\mathbf{L'}_{sr} = \begin{bmatrix} L_{mq}\cos\theta_r & L_{md}\sin\theta_r & L_{md}\sin\theta_r \\ L_{mq}\cos(\theta_r - 2\pi/3) & L_{md}\sin(\theta_r - 2\pi/3) & L_{md}\sin(\theta_r - 2\pi/3) \\ L_{mq}\cos(\theta_r - 4\pi/3) & L_{md}\sin(\theta_r - 4\pi/3) & L_{md}\sin(\theta_r - 4\pi/3) \end{bmatrix} \qquad [\mathrm{H}] \ , \tag{2.30}$$

and

$$\mathbf{L'}_r = \begin{bmatrix} L'_{lkq} + L_{mq} & 0 & 0 \\ 0 & L'_{lfd} + L_{md} & L_{md} \\ 0 & L_{md} & L'_{lkd} + L_{md} \end{bmatrix} \qquad [\mathrm{H}] \ . \tag{2.31}$$

As it may be immediately ascertained, all inductance elements, with the single exception of $\mathbf{L}'_r$, depend upon the position of the rotor and are, therefore, functions of the time-varying angle θ_r.

Another important concept related to inductance elements – thus a form of representing them – is the inductive reactance. The quantity ωL, called reactance (from the word reaction) of an inductor, is symbolically represented by x_L and is measured in ohms.

Inductive reactance is the opposition to the flow of current, which results in the continual interchange of energy between the source and the magnetic field of the inductor. In other words, inductive reactance, unlike resistance (which dissipates energy in the form of heat), does not dissipate electrical energy [47].

Therefore, there are corresponding leakage reactances x_l and magnetizing reactances x_m associated with the previously shown inductances. They will be further explored in detail.

Besides inductance and reactance, another electrical characteristic of a coil is its resistance. At this point, it is important to observe that stator resistances, r_a, r_b, and r_c, are defined by project as equal to each other and relatively small. Both $\mathbf{r}_s \in \mathbb{R}^{3\times3}$ and $\mathbf{r}_r \in \mathbb{R}^{(k+1)\times(k+1)}$ are diagonal matrices; in particular

$$\mathbf{r}_s = \operatorname{diag}\begin{bmatrix} r_s & r_s & r_s \end{bmatrix} \qquad [\Omega] \;, \tag{2.32}$$

$$\mathbf{r}_r = \operatorname{diag}\begin{bmatrix} r_{kq} & r_{fd} & r_{kd} \end{bmatrix} \qquad [\Omega] \;. \tag{2.33}$$

■

A final comment on this subsection may be borrowed from IEEE [13]:

> Synchronous machine electrical parameters are used in a variety of power system problems. In the steady-state, a knowledge of the direct-axis synchronous reactance, x_d, and the quadrature-axis synchronous reactance, x_q, is required to determine, after appropriate adjustments for saturation, the maximum value of reactive power output, Q, for certain armature terminal conditions. Such maximum reactive power outputs are basically a function of the field excitation. The reactive-power output capabilities of generators are used in load-flow studies for control of power systems voltages and supply of load reactive powers. As a corollary to this, the above mentioned synchronous reactances are used to determine the approximate values of reactive power, which can be absorbed by a synchronous machine. This is sometimes studied in load-flow studies under system minimum-load conditions [13, p. 91].

2.4.3 Voltage equations in machine variables

To proceed further, let the follow set of three-phase currents:

$$i_a = \sqrt{2}\,|I_a| \sin\left(w_e t - \phi_a\right) \qquad [\mathrm{A}] \;, \tag{2.34a}$$

$$i_b = \sqrt{2}\,|I_b| \sin\left(w_e t - \phi_a - 2\pi/3\right) \qquad [\mathrm{A}] \;, \tag{2.34b}$$

$$i_c = \sqrt{2}\,|I_c| \sin\left(w_e t - \phi_a - 4\pi/3\right) \qquad [\mathrm{A}] \;, \tag{2.34c}$$

where $|I_a|$, $|I_b|$, and $|I_c|$ are the root mean square (rms) value of each phase current, in amperes; w_e is the angular frequency of the induced emf, in electrical radians per second; and ϕ_a is the phase angle of phase-a current, in electrical radians.

Considering a balanced three-phase system, $|I_a| = |I_b| = |I_c|$, it is possible to relate $\sqrt{2}\,|I_a| = \sqrt{2}\,|I_b| = \sqrt{2}\,|I_c| = I_{max}$. Hence, the equations above may be re-written without loss of generality as:

$$i_a = I_{max} \sin\left(w_e t - \phi_a\right) \qquad [\mathrm{A}] \;, \tag{2.35a}$$

$$i_b = I_{max} \sin\left(w_e t - \phi_a - 2\pi/3\right) \qquad [\mathrm{A}] \;, \tag{2.35b}$$

$$i_c = I_{max} \sin\left(w_e t - \phi_a - 4\pi/3\right) \qquad [\mathrm{A}] \;. \tag{2.35c}$$

It is convenient to begin this development with the stator voltage equations.

Using Ohm[16]'s and Faraday's laws, the stator voltage equations are readily expressed

$$v_a = r_s i_a + p\psi_a \qquad [\mathrm{V}] \ , \tag{2.36a}$$

$$v_b = r_s i_b + p\psi_b \qquad [\mathrm{V}] \ , \tag{2.36b}$$

$$v_c = r_s i_c + p\psi_c \qquad [\mathrm{V}] \ , \tag{2.36c}$$

where $p = d/dt$ is the derivative of a function with respect to time.

The voltages equations in machine variables may be expressed in matrix form as

$$\mathbf{v}_{abcs} = \mathbf{r}_s \mathbf{i}_{abcs} + p\boldsymbol{\psi}_{abcs} \qquad [\mathrm{V}] \ , \tag{2.37a}$$

$$\mathbf{v}_{qdr} = \mathbf{r}_r \mathbf{i}_{qdr} + p\boldsymbol{\psi}_{qdr} \qquad [\mathrm{V}] \ , \tag{2.37b}$$

where $\mathbf{v}_{abcs} = \left[\, v_a \;\; v_b \;\; v_c \,\right]^{\mathrm{T}} \in \mathbb{R}^3$; $\mathbf{i}_{abcs} = \left[\, i_a \;\; i_b \;\; i_c \,\right]^{\mathrm{T}} \in \mathbb{R}^3$; $\boldsymbol{\psi}_{abcs} = \left[\, \psi_a \;\; \psi_b \;\; \psi_c \,\right]^{\mathrm{T}} \in \mathbb{R}^3$; $\mathbf{v}_{qdr} = \left[\, v_{kq} \;\; v_{fd} \;\; v_{kd} \,\right]^{\mathrm{T}} \in \mathbb{R}^3$; $\mathbf{i}_{qdr} = \left[\, i_{kq} \;\; i_{fd} \;\; i_{kd} \,\right]^{\mathrm{T}} \in \mathbb{R}^3$; and $\boldsymbol{\psi}_{qdr} = \left[\, \psi_{kq} \;\; \psi_{fd} \;\; \psi_{kd} \,\right]^{\mathrm{T}} \in \mathbb{R}^3$.

Each term in the equation above is obtained by determining the voltage induced in a particular circuit when current flows in one circuit only, in the same way as in ordinary circuit theory. The equation of any circuit is obtained by superimposing all the induced voltages and the resistance drop and equating to the impressed voltage.

The next step is to use the flux linkage equations that relate the stator and field flux linkages to the stator and field currents. As developed in the previous subsection, associating with the flux linkages, shown in (2.29), leads to the following terminal voltage equations in matrix notation

$$\begin{bmatrix} \mathbf{v}_{abcs} \\ \mathbf{v}'_{qdr} \end{bmatrix} = \begin{bmatrix} \mathbf{r}_s + p\mathbf{L}_s & p\mathbf{L}'_{sr} \\ \frac{2}{3}p(\mathbf{L}'_{sr})^{\mathrm{T}} & \mathbf{r}'_r + p\mathbf{L}'_r \end{bmatrix} \begin{bmatrix} \mathbf{i}_{abcs} \\ \mathbf{i}'_{qdr} \end{bmatrix} \qquad [\mathrm{V}] \ , \tag{2.38}$$

where $\mathbf{r}'_j$ is defined in (2.27).

The same results can be achieved by means of a magnetic-field point-of-view development [34, 44].

2.4.4 Power equations in machine variables

Let phase-a voltage, defined as

$$v_a = \sqrt{2}\,|V_a|\sin(w_e t) \qquad [\mathrm{V}] \ , \tag{2.39}$$

and the previously defined phase-a current,

$$i_a = \sqrt{2}\,|I_a|\sin(w_e t - \phi_a) \qquad [\mathrm{A}] \ . \tag{2.40}$$

[16] The law was named after Georg Simon Ohm (1789-1854) – a German physicist and mathematician – who, in a treatise published in 1827, using equipment of his own creation, found that there is a direct proportionality between the potential difference applied across a conductor and the resultant electric current.

The associated instantaneous electrical power is[17]

$$
\begin{aligned}
\mathcal{P}_{1\phi} &= v_a i_a && (2.41a)\\
&= 2\,|V_a|\,|I_a|\sin\left(w_e t\right)\sin\left(w_e t-\phi_a\right) && (2.41b)\\
&= |V_a|\,|I_a|\left[\cos\left(w_e t-w_e t-\phi_a\right)-\cos\left(w_e t+w_e t-\phi_a\right)\right] && (2.41c)\\
&= |V_a|\,|I_a|\left[\cos\left(\phi_a\right)-\cos\left(2w_e t-\phi_a\right)\right] && (2.41d)\\
&= |V_a|\,|I_a|\cos\left(\phi_a\right)+|V_a|\,|I_a|\cos\left(2w_e t-\phi_a\right) \qquad [\mathrm{W}]\ . && (2.41e)
\end{aligned}
$$

As it can be noticed, the instantaneous output power pulsates around an average power, $|V_a|\,|I_a|\cos\left(\phi_a\right)$, at double radian frequency $2\omega_e$. During certain periods, the power is actually negative, indicating that the energy flow during these intervals in the negative direction [32]. Equation (2.41e) can be transformed into

$$
\mathcal{P}_{1\phi} = |V_a|\,|I_a|\cos\phi_a\left[\left(1-\cos 2\omega_e t\right)\right]-|V_a|\,|I_a|\sin\left(\phi_a\right)\sin\left(2\omega_e t\right) \qquad [\mathrm{W}]\ . \tag{2.42}
$$

The first term on the right-hand side of (2.42) pulsates around the same average value as before but never goes negative, and the second one has a zero-average value. Accordingly, by defining two quantities

$$
\begin{aligned}
P &\triangleq |V_a|\,|I_a|\cos\phi_a && [\mathrm{W}] && \text{real, or active, power}\,, && (2.43a)\\
Q &\triangleq |V_a|\,|I_a|\sin\phi_a && [\mathrm{VAR}] && \text{reactive, or nonactive, power}\,, && (2.43b)
\end{aligned}
$$

Equation (2.42) can be more compactly written as

$$
\mathcal{P}_{1\phi} = P(1-\cos 2\omega_e t)-Q\sin 2\omega_e t \qquad [\mathrm{W}]\ . \tag{2.44}
$$

The real power, P, is defined as the average value of $v_a i_a$ and, therefore, physically means the useful power being transmitted. Its magnitude depends very strongly on the power factor, $\cos\phi_a$.

The reactive power[18], Q, is, by definition, equal to the peak value of that power component that travels back and forth, resulting in zero average, and therefore capable of no useful work.

For a three-phase system, the three-phase real power, $P_{3\phi}$, equals the sum of the individual phase powers[19] [32]:

$$
\begin{aligned}
P_{3\phi} &= 3\,|V_a|\,|I_a|\cos\phi_a && (2.45a)\\
&= 3P \qquad [\mathrm{W}]\ , && (2.45b)
\end{aligned}
$$

where the voltages are defined by (2.36); the currents by (2.34); and P, $|V_a|$, and $|I_a|$ represent per-phase values.

---

[17] A list of trigonometric relationships is found in Annex A.

[18] Both P and Q have dimension of watts, but to emphasize the fact that the latter represents a nonactive, or reactive, power, it is measured in volt-ampere reactive (VAR) [32].

[19] It is a direct result from the law of conservation of energy [48].

The three-phase reactive power $Q_{3\phi}$ is

$$Q_{3\phi} = 3Q \qquad [\mathrm{VAR}] \ . \tag{2.46}$$

The total instantaneous power of a three-phase $\mathcal{P}_{abcs}$ system may be expressed in *abc* variables as

$$\mathcal{P}_{abcs} = \langle \mathbf{v}_{abcs}, \mathbf{i}_{abcs} \rangle \tag{2.47a}$$

$$= v_a i_a + v_b i_b + v_c i_c \qquad [\mathrm{VA}] \ , \tag{2.47b}$$

where $\langle \cdot, \cdot \rangle$ is the inner product between two vectors.

2.4.5 Torque equation in machine variables

In addition to the electrical equations given in the previous sections, modeling a synchronous machine requires an expression for the electromechanical torque to be used in the calculation of the machine mechanical dynamics [49]. In general, the electrical torque is produced by the interaction between the three stator circuits, the field current, and other circuits such as the damper windings [22].

The flux linking each circuit in the machine depends upon the exciter output voltage, the loading of the magnetic circuit (saturation), and the current in different windings. Whether the machine is operating at synchronous speed or asynchronously affects all the above factors [22]. If the instantaneous values of these flux linkages and currents are known, the correct instantaneous value of the electrical torque may be determined.

It is important to note that useful torque is obtained if the armature mmf has the same velocity of the field mmf in relation to a common reference frame [28]. Therefore, the condition to produce torque is that both mmfs are stationary between them.

As stated by Anderson and Fouad [22], the electrical torque may be divided into the synchronous torque and a second component that includes all other electrical torques:

1. Synchronous torque: it is the most important component of the electrical torque and is produced by the interaction of the stator windings with the fundamental component of the air gap flux. It is dependent upon the machine terminal voltage, the rotor angle, the machine reactances, and the so-called quadrature-axis emf, which may be thought of as an effective rotor emf that is dependent on the armature and rotor currents and is a function of the exciter response;

2. Other electrical torques: during a transient, other extraneous electrical torques are developed in a synchronous machine. The most important is associated with the damper windings. Although these asynchronous torques are usually small in magnitude, their effect on stability studies may not be negligible.

At this point, only the synchronous torque will be considered.

The energy $\mathcal{W}_f$, stored in the coupling field of a synchronous machine, may be expressed as

$$\mathcal{W}_f = \frac{1}{2}(\mathbf{i}_{abcs})^{\mathrm{T}}\mathbf{L}_s\mathbf{i}_{abcs} + (\mathbf{i}_{abcs})^{\mathrm{T}}\mathbf{L'}_{sr}\mathbf{i}'_{qdr} + \left(\frac{1}{2}\right)\left(\frac{3}{2}\right)(\mathbf{i}'_{qdr})^{\mathrm{T}}\mathbf{L'}_r\mathbf{i}'_{qdr} \qquad [\mathrm{J}] \ . \tag{2.48}$$

Assuming the magnetic system to be linear and using the fact that

$$\theta_r = \rho\theta_m \qquad [\text{electrical rad}] \ , \tag{2.49}$$

the torque may be expressed in terms of electrical rotor position as

$$\mathcal{T}_e = \rho\left\{\frac{1}{2}(\mathbf{i}_{abcs})^{\mathrm{T}}\frac{\partial}{\partial\theta_r}\left[\mathbf{L}_s\right]\mathbf{i}_{abcs} + (\mathbf{i}_{abcs})^{\mathrm{T}}\frac{\partial}{\partial\theta_r}\left[\mathbf{L'}_{sr}\right]\mathbf{i}'_{qdr}\right\} \qquad [\mathrm{N.m}] \ . \tag{2.50}$$

Neglecting the shaft torsional effects, the torque and the rotor speed are related by

$$\mathcal{T}_e = \mathcal{T}_m - J\left(\frac{1}{\rho}\right)p\omega_r \qquad [\mathrm{N.m}] \ , \tag{2.51}$$

where J is the inertia expressed in kilogram meter squared (kg.m^2) or Joule second squared (J.s^2); and $\mathcal{T}_m$ is the net mechanical shaft torque, in Newton meter (N.m).

2.4.6 Motion equations in machine variables

If the rotor speed varies, it interacts with the electromagnetic changes to produce electromechanical dynamic effects. The time scale associated with these dynamics is sufficiently long for them to be influenced by the turbine and the generator control systems [50].

Therefore, the equations of central importance in power system stability analysis are the rotational inertia equations describing the unbalance between the electromagnetic torque and the mechanical torque [33].

When considering free-body rotation, the shaft can be assumed to be rigid when the total inertia of the rotor J is simply the sum of the individual inertias [50]. Any unbalance torque acting on the rotor will result in the acceleration or deceleration of the rotor as a complete unit according to Newton's second law and expressed in (2.51).

Although the turbine torque changes relatively slowly, the electromagnetic torque $\mathcal{T}_e$ may change its value almost instantaneously. The net mechanical shaft torque $\mathcal{T}_m$, which is the turbine torque less the rotational losses at synchronous speed, is the one converted into electromagnetic torque. If, due to some disturbance, $\mathcal{T}_m > \mathcal{T}_e$, the rotor accelerates; if $\mathcal{T}_m < \mathcal{T}_e$, then it decelerates.

At this point, an important definition becomes necessary:

> **rotor displacement angle (rotating machinery)** (load angle): the displacement caused by load between the terminal voltage and the armature voltage generated by that component of flux produced by the field current [37, p. 992].

Another entry in the IEEE [37]:

> **load angle (synchronous machine)**: the angular displacement, at a specified load, of the center line of a field pole from the axis of the armature magnetomotive force (mmf) pattern [37, p. 630].

The rotor velocity can be expressed as

$$\omega_m = \omega_{sm} + \Delta\omega_m \tag{2.52a}$$

$$= \omega_{sm} + p\delta_m \qquad [\text{mechanical rad/s}] \ , \tag{2.52b}$$

where δ_m is the load angle, expressed in mechanical radians; and $\Delta\omega_m = p\delta_m$ is the speed deviation, in mechanical radians per second.

Transforming the mechanical quantities into electrical quantities, (2.52b) becomes

$$\omega_r = \omega_s + p\delta \qquad [\text{electrical rad/s}] \ , \tag{2.53}$$

where ω_s is the synchronous speed, in electrical radians per second; and δ is the load angle, in electrical radians. Recall that $\omega_s = \omega_{sm}/\rho$ and $\delta = \delta_m/\rho$.

Furthermore, from (2.52b) [51]:

$$\delta = \int (\omega_r - \omega_s)\, dt \tag{2.54a}$$

$$= (\omega_r t + \theta_{r0}) - (\omega_s t + \theta_{s0}) \tag{2.54b}$$

$$= \omega_r t - \omega_s t + \delta_0 \qquad [\text{electrical rad}] \ , \tag{2.54c}$$

where $\delta_0 = \theta_{r0} - \theta_{s0}$ is the load angle value at $t = 0$, in electrical radians.

Through Figure 2.6, it is possible to visualize an arbitrary sinusoidal time-varying phase voltage v_s and how the angles $\delta, \theta_r, \theta_s$ and respective velocities are related to each other. The qd axes spin anticlockwise at the angular rotor speed ω_r and the QD, at the angular synchronous speed ω_s. Accordingly, the load angle δ is defined to show the difference angle between the q-axis and the space vector v_s.

Equation (2.54c) provides an important definition for the load angle: if the initial condition is known, it is possible to compute it from frequency measurements, both network's and rotor's. In Chapter 6, a simplified approach will be presented to estimate the quadrature- and direct-axis magnetizing reactances; it considers a specific initial load condition and applies it to the integral computation.

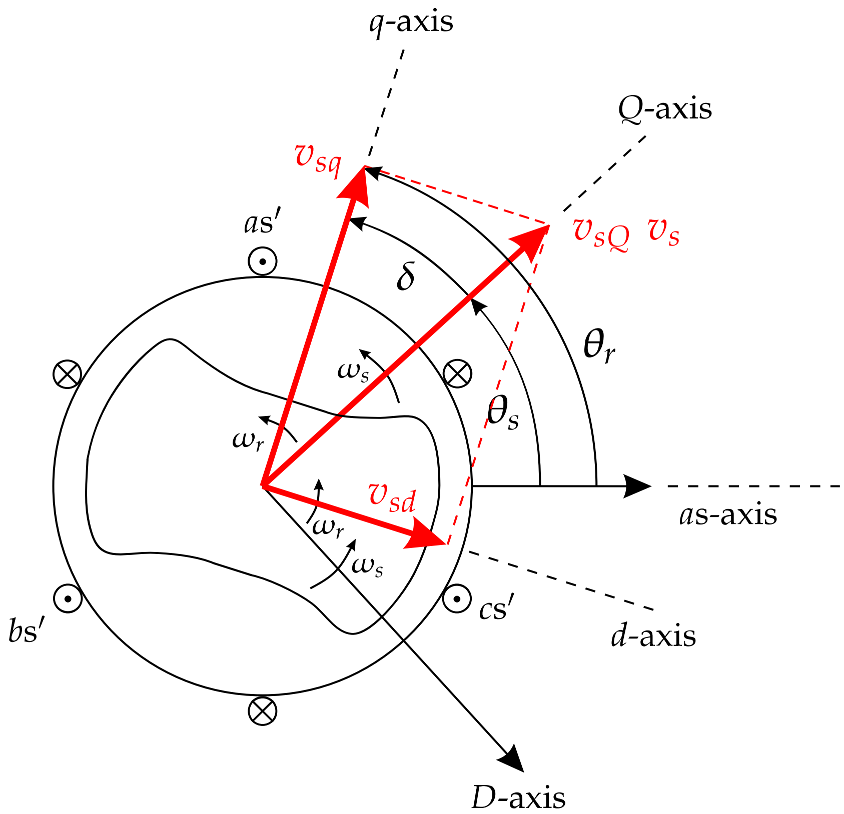

Figure 2.6: A visual description on the angles, speeds, and reference frames in a simplified salient-pole synchronous machine. Adapted from Malekpour et al. [51].

2.5 A change of variables

An introduction to this section may be the following text given by Lipo [36, p. 78], which elucidates the change of variables under an interesting point of view:

> It might be stated that one lives in a world of reference frames. The world as one perceives it, is observed in a reference frame fixed by our senses. As we seat ourselves in the family car, one can say that we change reference frames and attach ourselves to a reference frame fixed in the automobile. Changes of reference frames are clearly an everyday experience. In most cases the reference frame to which we attach ourselves is associated with linear rather than rotational motion. [...] it should come as no surprise that rotational reference frames are a part of life since the Earth itself is a rotating reference frame. Rotating reference frames are of central importance in the analysis of electric machines.

In the last section, the synchronous machine voltage equations for the stator and the rotor in machine variables were presented in (2.37). They may be re-written as:

$$\begin{cases} p\mathbf{i} &= \mathbf{L}^{-1}\left[-\mathbf{ri} - p\theta_r \dfrac{\partial \mathbf{L}}{\partial \theta_r}\mathbf{i} - \mathbf{v}\right], \\ \boldsymbol{\psi} &= \mathbf{Li} \end{cases} \tag{2.55}$$

where

$$\mathbf{L} = \begin{bmatrix} \mathbf{L}_s & \mathbf{L}'_{sr} \\ \frac{2}{3}(\mathbf{L}'_{sr})^{\mathrm{T}} & \mathbf{L}'_r \end{bmatrix}, \quad \mathbf{r} = \mathrm{diag}\begin{bmatrix} \mathbf{r}_s & \mathbf{r}_r \end{bmatrix}, \quad \boldsymbol{\psi} = \begin{bmatrix} \boldsymbol{\psi}^{\mathrm{T}}_{abcs} & \boldsymbol{\psi}'^{\mathrm{T}}_{qdr} \end{bmatrix}^{\mathrm{T}},$$

$$\mathbf{i} = \begin{bmatrix} \mathbf{i}^{\mathrm{T}}_{abcs} & \mathbf{i}^{\mathrm{T}}_{qdr} \end{bmatrix}^{\mathrm{T}}, \quad \mathbf{v} = \begin{bmatrix} \mathbf{v}^{\mathrm{T}}_{abcs} & \mathbf{v}^{\mathrm{T}}_{qdr} \end{bmatrix}^{\mathrm{T}}.$$

Although it is possible to solve (2.55) numerically, it is almost impossible to obtain an analytical solution even when $p\theta_r$ is constant [35]. This is due to the fact that the inductance matrix, $\mathbf{L}$, is time-varying[20] and the computation of its inverse, $\mathbf{L}^{-1}$, is required.

It would be advantageous if the time-varying machine equations could be transformed to a time-invariant set. This would result in the simplification of the calculations both for steady-state and transient conditions.

It was shown that some of the machine inductances are functions of rotor position, whereupon the coefficients of the differential-voltage equations that describe the behavior of these machines are rotor-position dependent. These complexities may be reduced by means of a change of variables that eliminates them [19, 43, 45, 52–55].

The choice of a reference frame must be wisely taken. However, it was found that the varying inductances of a synchronous machine are eliminated if, and only if, the reference frame rotates at the electrical velocity of the rotor [19].

Bim [28] summarizes the most common reference frames used in the analysis of electric machines: the $\alpha\beta0$ frame, a reference frame fixed in the stator-physical structure [56]; the $mn0$ frame, fixed in the rotor-physical structure [53]; and the $qd0$ frame, fixed in the synchronous rotating magnetic field [43].

In a paper published in the late 1920s, R. H. Park[21] [43] formulated a change of variables – known as Park's Transformation (PT) or, also commonly, as direct-quadrature-zero transformation – that in effect replaced the variables (voltages, currents, and flux linkages) associated with the stator windings of a synchronous machine with variables associated with fictitious windings rotating at the electrical velocity of the rotor.

PT revolutionized electric machine analysis and has the unique property of eliminating all rotor position-dependent inductances from the voltage equations of the synchronous machine that occur due to (i) electric circuits in relative motion; and (ii) electric circuits with varying magnetic reluctance [19].

[20] The inductance terms vary with angle θ_r, which, in turn, varies with time.

[21] Robert H. Park (1902–1994) was an American electrical engineer and inventor, best known for the Park's Transformation (PT), used to simplify the analysis of three-phase electric circuits. His related 1929 concept paper ranked second, when looking at the impact of all twentieth century power engineering papers.

2.5.1 Park's Transformation

The idea behind PT is an old one, stemming from the work of Andre Blondel[22] [57] in France; the technique is sometimes referred to as the *Blondel's Two-Reaction Method*. However, much of the development of the method was carried out by R. E. Doherty[23] and C. A. Nickel[24] in [58–61], and R. H. Park in [43, 62].

Since the air gap of a salient-pole synchronous machine is non-uniform because of the presence of a large inter-polar air space, the resultant field wave obtained due to the resultant mmf – between the armature-reaction and field-winding ones – will be unsymmetrical and will contain higher-order harmonics of significant magnitude [34]. In order to quantitatively analyze the effects of armature reaction, the aforementioned mmfs are considered to create independent fluxes, which induce independent emf in the stator windings – therefore the name *two-reaction method*.

From electric machine theory, the stator currents due to the spatial distribution of the stator winding give rise to a mmf that is proportional to such currents. In a reference frame fixed with respect to the stator and having its origin coinciding with the axis of the phase-a winding, the mmf caused by current i_a is therefore directly proportional to itself.

For a reference frame fixed with respect to the rotor and having the origin coinciding with the midpoint of the pole, the same mmf wave has an intensity proportional to $i_a \cos\theta_r$. From this point of view, the current i_d is therefore a measure of the total mmf as measured in the midpole direction. Similarly, i_q gives the mmf in the quadrature pole direction [32].

The technique defines a new set of stator variables such as currents, voltages, or flux linkages in terms of the actual winding variables. The new quantities are obtained by projecting the actual variables onto three axes: one along the direct-axis of the rotor field winding, called the direct-axis (d); a second along the neutral axis of the field winding, called the quadrature-axis (q); and a third on a stationary axis, called the zero axis (0) [22].

These Park – or Blondel, as one may prefer – currents are defined as follows:

$$i_q \triangleq k_q \left\{ i_a \cos\theta_r + i_b \cos\left(\theta_r - 2\pi/3\right) + i_c \cos\left(\theta_r - 4\pi/3\right) \right\} \qquad [\mathrm{A}]\ , \tag{2.56a}$$

$$i_d \triangleq k_d \left\{ i_a \sin\theta_r + i_b \sin\left(\theta_r - 2\pi/3\right) + i_c \sin\left(\theta_r - 4\pi/3\right) \right\} \qquad [\mathrm{A}]\ , \tag{2.56b}$$

$$i_0 \triangleq k_0 \left\{ i_a + i_b + i_c \right\} \qquad [\mathrm{A}]\ , \tag{2.56c}$$

where i_a, i_b, and i_c are defined in (2.34); θ_r is the angular position of the reference frame relative to phase-a axis, as shown in Figure 2.7; and k_q, k_d, and k_0 are arbitrary constants.

---

[22] André-Eugène Blondel (1863–1938) was a French engineer and physicist. He is the inventor of the electromechanical oscillograph and a system of photometric units of measurement.

[23] Robert E. Doherty (1885–1950) was an American electrical engineer. He became dean of the Yale School of Engineering & Applied Science in 1932.

[24] C. A. Nickel was an General Electric engineer and an American Institute of Electrical Engineers associate. No further information was found about him.

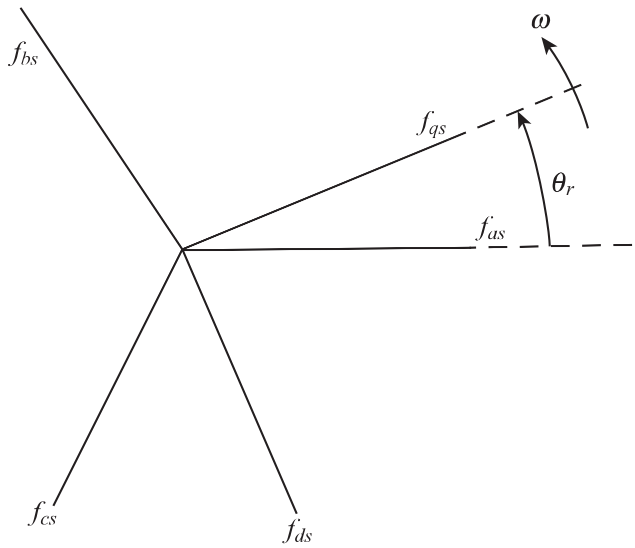

Figure 2.7: Transformation for stationary circuits portrayed by trigonometric relationships. Adapted from Krause et al. [19, p. 90].

The term "zero-sequence current" is adopted from the analogy with the "zero-sequence component" used in Fortescue[25]'s symmetrical components theory [63], however i_0 is an instantaneous value of a stationary current, which may vary with time in any manner. It may be visualized physically as the magnitude of each of a set of equal currents, flowing in all three phases and therefore producing no resultant air gap mmf [22, 30].

Although the transformation to a reference frame is simply a change of variables and needs no physical connotation, it is often convenient to visualize the transformation equations as trigonometric relationships between variables as shown in Figure 2.7. As mentioned earlier in this text, f can represent either voltages, currents, flux linkages, or electric charges. At this point, the subscript s indicates the variables associated with stationary circuits. The angular displacement θ_r must be continuous. The new frame of reference rotates at the electrical velocity of the rotor.

Considering Figure 2.7, a change of variables that formulates a transformation of the three-phase currents from (2.34) of a stationary circuit elements to currents fixed in the *qd0* reference frame from (2.56) may be expressed as

$$\mathbf{i}_{qd0s} = \mathbf{K}\,\mathbf{i}_{abcs} \qquad [\mathrm{A}]\,, \tag{2.57}$$

where $\mathbf{i}_{qd0s} = \left[\, i_q \;\; i_d \;\; i_0 \,\right]^{\mathrm{T}} \in \mathbb{R}^3$, $\mathbf{i}_{abcs} = \left[\, i_a \;\; i_b \;\; i_c \,\right]^{\mathrm{T}} \in \mathbb{R}^3$, and

$$\mathbf{K} = \begin{bmatrix} k_q \cos\theta_r & k_q \cos\left(\theta_r - 2\pi/3\right) & k_q \cos\left(\theta_r - 4\pi/3\right) \\ k_d \sin\theta_r & k_d \sin\left(\theta_r - 2\pi/3\right) & k_d \sin\left(\theta_r - 4\pi/3\right) \\ k_0 & k_0 & k_0 \end{bmatrix}, \tag{2.58}$$

[25]Charles LeGeyt Fortescue (1876-1936) was an Canadian electrical engineer. He was one of the authors of a paper on measurement of high voltage by the breakdown of a gap between two conductive spheres, which is a technique still used in high-voltage laboratories today. Although, he is most famous because of his paper presented in 1918, in which he demonstrated that any set of N unbalanced phasors could be expressed as the sum of N symmetrical sets of balanced phasors known as *symmetrical components*. The paper was judged to be the most important power engineering paper in the twentieth century.

where the constants k_q, k_d, and k_0 may be chosen to simplify numerical coefficients in performance equations[26] [33].

By solving for the stator currents in accordance with

$$\mathbf{i}_{abcs} = \mathbf{K}^{-1}\,\mathbf{i}_{qd0s} \qquad [\text{A}] \;, \tag{2.59}$$

it can be shown that the inverse transformation matrix is

$$\mathbf{K}^{-1} = \begin{bmatrix} k_1 \cos\theta_r & k_2 \sin\theta_r & k_3 \\ k_1 \cos(\theta_r - 2\pi/3) & k_2 \sin(\theta_r - 2\pi/3) & k_3 \\ k_1 \cos(\theta_r - 4\pi/3) & k_2 \sin(\theta_r - 4\pi/3) & k_3 \end{bmatrix} , \tag{2.60}$$

where

$$k_1 = \frac{2}{3k_q} , \quad k_2 = \frac{2}{3k_d} , \quad \text{and } k_3 = \frac{1}{3k_0} .$$

In most of the literature on synchronous machines theory [30, 64–68], k_q and k_d are taken as $2/3$ and k_0 as $1/3$. Therefore, $k_1 = k_2 = k_3 = 1$. Several different alternatives have been proposed. Some analysis, notably Lewis [69], have also suggested the use of $k_q = k_d = \sqrt{2/3}$ instead of $2/3$, and the zero-sequence as $\sqrt{1/2}$ instead of $1/3$.

The latter choice of constants results in an orthogonal matrix, i.e., the inverse of the transformation matrix is equal to its transpose $\mathbf{K}^{-1} = \mathbf{K}^{\mathrm{T}}$. This also means that the transformation is power invariant:

$$\mathcal{P}_{qd0s} = \mathcal{P}_{abcs} \tag{2.61a}$$
$$= v_a i_a + v_b i_b + v_c i_c \tag{2.61b}$$
$$= v_q i_q + v_d i_d + v_0 i_0 \qquad [\text{VA}] \;. \tag{2.61c}$$

In addition, with this transformation, all mutual inductances would be reciprocal. However, Harris, Lawrenson, and Stephenson [70] showed that this transformation has several drawbacks, which appear to override its advantages. The orthogonal transformation does not correspond to any particular physical situation. With $k_q = k_d = \sqrt{2/3}$, the equivalent quadrature- and direct-axis coils would have $\sqrt{3/2}$ times the number of turns as *abc* coils. This removes the unit-to-unit relationship between *abc* and *qd*0 variables that exists when $k_q = k_d = 2/3$.

With $k_q = k_d = 2/3$, for balanced sinusoidal conditions, the peak values of i_q and i_d are equal to the peak value of the stator current. From (2.56),

$$i_q = k_q \left\{ i_a \cos\theta_r + i_b \cos(\theta_r - 2\pi/3) + i_c \cos(\theta_r - 4\pi/3) \right\}$$
$$= k_q \frac{3}{2} I_{max} \sin(\omega_e t - \theta_r) \qquad [\text{A}] \;, \tag{2.62}$$

[26] At this point, it is important to notice that the difference in matrix **K** from the one presented in (2.58) to the ones from other texts is due to measuring angle θ_r relative to the quadrature-axis instead of the direct-axis, as mentioned previously. Also, note the use of *qd*0 instead of *dq*0.

$$i_d = k_d \left\{ i_a \sin\theta_r + i_b \sin\left(\theta_r - 2\pi/3\right) + i_c \sin\left(\theta_r - 4\pi/3\right) \right\}$$

$$= k_d \frac{3}{2} I_{max} \cos\left(\omega_e t - \theta_r\right) \qquad [\mathrm{A}] \ , \tag{2.63}$$

$$i_0 = k_0 \left\{ i_a + i_b + i_c \right\}$$

$$= 0 \qquad [\mathrm{A}] \ . \tag{2.64}$$

Another important point refers to the zero-sequence component. It is required to yield a unique transformation of the three stator-phase quantities; it corresponds to components of armature current which produce no net air gap flux and hence no net flux linking the rotor circuits. Under balanced-three-phase conditions, there are no zero-sequence components [44].

Given all previous considerations, this work considers the following direct $\mathbf{K}$ and inverse $\mathbf{K}^{-1}$ transformation matrices:

$$\mathbf{K} = \frac{2}{3} \begin{bmatrix} \cos\theta_r & \cos\left(\theta_r - 2\pi/3\right) & \cos\left(\theta_r - 4\pi/3\right) \\ \sin\theta_r & \sin\left(\theta_r - 2\pi/3\right) & \sin\left(\theta_r - 4\pi/3\right) \\ 1/2 & 1/2 & 1/2 \end{bmatrix} \tag{2.65}$$

and

$$\mathbf{K}^{-1} = \begin{bmatrix} \cos\theta_r & \sin\theta_r & 1 \\ \cos\left(\theta_r - 2\pi/3\right) & \sin\left(\theta_r - 2\pi/3\right) & 1 \\ \cos\left(\theta_r - 4\pi/3\right) & \sin\left(\theta_r - 4\pi/3\right) & 1 \end{bmatrix} . \tag{2.66}$$

Furthermore, PT is applied to instantaneous rather than rms quantities values.

The transformation of the stator currents is defined in (2.57). Similarly, the transformed stator voltages are

$$\mathbf{v}_{qd0s} = \mathbf{K}\, \mathbf{v}_{abcs} \qquad [\mathrm{V}] \ , \tag{2.67}$$

with the associated inverse transformation

$$\mathbf{v}_{abcs} = \mathbf{K}^{-1}\, \mathbf{v}_{qd0s} \qquad [\mathrm{V}] \ , \tag{2.68}$$

where $\mathbf{v}_{qd0s} = \left[\, v_q \;\; v_d \;\; v_0 \,\right]^{\mathrm{T}} \in \mathbb{R}^3$; and $\mathbf{v}_{abcs} = \left[\, v_a \;\; v_b \;\; v_c \,\right]^{\mathrm{T}} \in \mathbb{R}^3$.

The transformed flux linkages are

$$\boldsymbol{\psi}_{qd0s} = \mathbf{K}\, \boldsymbol{\psi}_{abcs} \qquad [\mathrm{Wb\text{-}t}] \ , \tag{2.69}$$

with the associated inverse transformation

$$\boldsymbol{\psi}_{abcs} = \mathbf{K}^{-1}\, \boldsymbol{\psi}_{qd0s} \qquad [\mathrm{Wb\text{-}t}] \ , \tag{2.70}$$

where $\boldsymbol{\psi}_{qd0s} = \left[\, \psi_q \;\; \psi_d \;\; \psi_0 \,\right]^{\mathrm{T}} \in \mathbb{R}^3$; and $\boldsymbol{\psi}_{abcs} = \left[\, \psi_a \;\; \psi_b \;\; \psi_c \,\right]^{\mathrm{T}} \in \mathbb{R}^3$.

2.5.2 Voltage equations in rotor reference-frame variables

The voltage equation for a salient-pole synchronous machine in the *abc* reference frame is presented in (2.37a). Upon substituting the voltages, currents, and flux linkages for Park ones,

$$\mathbf{v}_{abcs} = \mathbf{r}_s \mathbf{i}_{abcs} + p\boldsymbol{\psi}_{abcs} \tag{2.71a}$$

$$\mathbf{K}^{-1}\mathbf{v}_{qd0s} = \mathbf{r}_s \mathbf{K}^{-1}\mathbf{i}_{qd0s} + p\mathbf{K}^{-1}\boldsymbol{\psi}_{qd0s} \qquad [\mathrm{V}]\ . \tag{2.71b}$$

Upon premultiplying (2.71b) by $\mathbf{K}$, it becomes

$$\mathbf{v}_{qd0s} = \mathbf{K}\mathbf{r}_s \mathbf{K}^{-1}\mathbf{i}_{qd0s} + \mathbf{K}p\mathbf{K}^{-1}\boldsymbol{\psi}_{qd0s} \qquad [\mathrm{V}]\ . \tag{2.72}$$

All stator phase windings of a synchronous machine are designed to have the same resistance. If the nonzero elements of the diagonal matrix $\mathbf{r}_s$ are equal, then

$$\mathbf{K}\mathbf{r}_s\mathbf{K}^{-1} = \mathbf{r}_s \qquad [\Omega]\ . \tag{2.73}$$

Thus, the resistance matrix associated with the *qd0* reference frame variables equals the resistance matrix associated with the actual variables if each phase of the actual circuit has the same resistance [19].

Furthermore, applying the product rule for derivatives,

$$\mathbf{K}p\mathbf{K}^{-1}\boldsymbol{\psi}_{qd0s} = \mathbf{K}p\left(\mathbf{K}^{-1}\right)\boldsymbol{\psi}_{qd0s} + \mathbf{K}\mathbf{K}^{-1}p\boldsymbol{\psi}_{qd0s} \tag{2.74a}$$

$$= \mathbf{K}p\left(\mathbf{K}^{-1}\right)\boldsymbol{\psi}_{qd0s} + p\boldsymbol{\psi}_{qd0s} \qquad [\mathrm{Wb/s}]\ , \tag{2.74b}$$

where

$$p\left(\mathbf{K}^{-1}\right) = \omega_r \begin{bmatrix} -\sin\theta_r & \cos\theta_r & 0 \\ -\sin\left(\theta_r - 2\pi/3\right) & \cos\left(\theta_r - 2\pi/3\right) & 0 \\ -\sin\left(\theta_r - 4\pi/3\right) & \cos\left(\theta_r - 4\pi/3\right) & 0 \end{bmatrix} \qquad [\text{electrical rad/s}]\ , \tag{2.75a}$$

and, therefore[27]

$$\mathbf{K}p\left(\mathbf{K}^{-1}\right) = \omega_r \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad [\text{electrical rad/s}]\ . \tag{2.75b}$$

Considering the results in equations (2.73), (2.74b), and (2.75b), (2.72) becomes

$$\mathbf{v}_{qd0s} = \mathbf{r}_s\mathbf{i}_{qd0s} + p\boldsymbol{\psi}_{qd0s} + \omega_r\begin{bmatrix}\psi_d & 0 & 0\end{bmatrix}^{\mathrm{T}} - \omega_r\begin{bmatrix}0 & \psi_q & 0\end{bmatrix}^{\mathrm{T}} \tag{2.76a}$$

$$= \mathbf{r}_s\mathbf{i}_{qd0s} + p\boldsymbol{\psi}_{qd0s} + \omega_r\boldsymbol{\psi}_{dqs} \qquad [\mathrm{V}]\ , \tag{2.76b}$$

where $\boldsymbol{\psi}_{dqs} = \begin{bmatrix}\psi_d & -\psi_q & 0\end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^3$.

[27] Refer to Annex A for useful trigonometric relationships.

Alternatively, (2.76) can be expressed in an expanded form as

$$v_q = r_s i_q + p\psi_q + \omega_r \psi_d \qquad [\mathrm{V}] \, , \tag{2.77a}$$
$$v_d = r_s i_d - \omega_r \psi_q + p\psi_d \qquad [\mathrm{V}] \, , \tag{2.77b}$$
$$v_0 = r_s i_0 + p\psi_0 \qquad [\mathrm{V}] \, . \tag{2.77c}$$

In (2.77), some terms contain the differential operator p and represent voltages due to changing currents in coils on the axis under analysis. They are called *transformer voltages* and are present even when the machine is stationary. The terms containing the angular frequency of the generated voltage represent voltages induced by rotation in the flux set up by the current in a coil on the other axis. Such voltages are called *rotation voltages*. When the coils carry steady DC currents, there are no transformer voltages, but the rotation voltages are still present [30].

Since (2.37a) is valid in general, it follows that (2.77) is valid regardless if the system is magnetically linear or nonlinear. If the system is magnetically linear, (2.77) is valid regardless of the form of the inductance matrix [19].

The rotor windings of a synchronous machine are asymmetrical; therefore, changing variables offers no advantages in the analysis of rotor circuits. Since the rotor variables are not transformed, the rotor voltage equations are expressed only in the rotor reference frame. Hence, from (2.37b), with appropriate turns ratios included (being indicated by primes $(\cdot)'$), the rotor voltage equations are [19]:

$$\mathbf{v}'_{qdr} = \mathbf{r}'_r \mathbf{i}'_{qdr} + p\boldsymbol{\psi}'_{qdr} \qquad [\mathrm{V}] \, . \tag{2.78}$$

More readily, as for the stator voltage equations, the matrix equation above may be expressed in an expanded form,

$$v'_{kq} = r'_{kq} i'_{kq} + p\psi'_{kq} \qquad [\mathrm{V}] \, , \tag{2.79a}$$
$$v'_{fd} = r'_{fd} i'_{fd} + p\psi'_{fd} \qquad [\mathrm{V}] \, , \tag{2.79b}$$
$$v'_{kd} = r'_{kd} i'_{kd} + p\psi'_{kd} \qquad [\mathrm{V}] \, . \tag{2.79c}$$

In summary, all voltage equations in the *qd0* reference frame and matrix notation are:

$$\mathbf{v}_{qd0s} = \mathbf{r}_s \mathbf{i}_{qd0s} + p\boldsymbol{\psi}_{qd0s} + \omega_r \boldsymbol{\psi}_{dqs} \qquad [\mathrm{V}] \, , \tag{2.80a}$$
$$\mathbf{v}'_{qdr} = \mathbf{r}'_r \mathbf{i}'_{qdr} + p\boldsymbol{\psi}'_{qdr} \qquad [\mathrm{V}] \, . \tag{2.80b}$$

2.5.3 Flux linkage equations in rotor reference-frame variables

For a magnetically linear system, the flux linkage equations may be expressed from (2.29) and transforming the stator variables to the rotor reference frame:

$$\begin{bmatrix} \boldsymbol{\psi}_{qd0s} \\ \boldsymbol{\psi}'_{qdr} \end{bmatrix} = \begin{bmatrix} \mathbf{K}\mathbf{L}_s\mathbf{K}^{-1} & \mathbf{K}\mathbf{L}'_{sr} \\ \frac{2}{3}(\mathbf{L}'_{sr})^{\mathrm{T}}\mathbf{K}^{-1} & \mathbf{L}'_r \end{bmatrix} \begin{bmatrix} \mathbf{i}_{qd0s} \\ \mathbf{i}'^{r}_{qdr} \end{bmatrix} \qquad [\mathrm{Wb\text{-}t}] \, , \tag{2.81}$$

which was obtained by means of trigonometric identities and matrix multiplication, similarly to the procedure used for voltage equations.

Furthermore,

$$\mathbf{K}\mathbf{L}_s\mathbf{K}^{-1} = \begin{bmatrix} L_{ls}+L_{mq} & 0 & 0 \\ 0 & L_{ls}+L_{md} & 0 \\ 0 & 0 & L_{ls} \end{bmatrix} \qquad [\mathrm{H}] \ , \tag{2.82a}$$

$$\mathbf{K}\mathbf{L}'_{sr} = \begin{bmatrix} L_{mq} & 0 & 0 \\ 0 & L_{md} & L_{md} \\ 0 & 0 & 0 \end{bmatrix} \qquad [\mathrm{H}] \ , \tag{2.82b}$$

$$\frac{2}{3}(\mathbf{L}'_{sr})^{\mathrm{T}}\mathbf{K}^{-1} = \begin{bmatrix} L_{mq} & 0 & 0 \\ 0 & L_{md} & 0 \\ 0 & L_{md} & 0 \end{bmatrix} \qquad [\mathrm{H}] \ . \tag{2.82c}$$

Substituting equations (2.31) and (2.82) into (2.81) yields the expressions for the flux linkages. In an expanded form

$$\psi_q = L_{ls}i_q + L_{mq}\left(i_q + i'_{kq}\right) \qquad [\mathrm{Wb\text{-}t}] \ , \tag{2.83a}$$

$$\psi_d = L_{ls}i_d + L_{md}\left(i_d + i'_{fd} + i'_{kd}\right) \qquad [\mathrm{Wb\text{-}t}] \ , \tag{2.83b}$$

$$\psi_0 = L_{ls}i_0 \qquad [\mathrm{Wb\text{-}t}] \ , \tag{2.83c}$$

$$\psi'_{kq} = L'_{kq}i'_{kq} + L_{mq}\left(i_q + i'_{kq}\right) \qquad [\mathrm{Wb\text{-}t}] \ , \tag{2.83d}$$

$$\psi'_{fd} = L'_{fd}i'_{fd} + L_{md}\left(i_d + i'_{fd} + i'_{kd}\right) \qquad [\mathrm{Wb\text{-}t}] \ , \tag{2.83e}$$

$$\psi'_{kd} = L'_{kd}i'_{kd} + L_{md}\left(i_d + i'_{fd} + i'_{kd}\right) \qquad [\mathrm{Wb\text{-}t}] \ . \tag{2.83f}$$

Again, all the inductances are seen to be constant – i.e., they are independent of the rotor positions. It should be noticed, however, that saturation effects are not considered here. The variations in inductances due to saturation are of a different nature and must be treated separately [33].

It is also interesting to notice that i_0 does not appear in the rotor flux linkage equations. This is because zero-sequence components of armature current do not produce net mmf across the air gap.

In order to improve the visualization of the previous equation, an expanded matrix form becomes necessary:

$$\begin{bmatrix} \psi_q \\ \psi_d \\ \psi_0 \\ \hline \psi'_{kq} \\ \psi'_{fd} \\ \psi'_{kd} \end{bmatrix} = \left[\begin{array}{ccc|ccc} L_{ls}+L_{mq} & 0 & 0 & L_{mq} & 0 & 0 \\ 0 & L_{ls}+L_{md} & 0 & 0 & L_{md} & L_{md} \\ 0 & 0 & L_{ls} & 0 & 0 & 0 \\ \hline L_{mq} & 0 & 0 & L'_{lkq}+L_{mq} & 0 & 0 \\ 0 & L_{md} & 0 & 0 & L'_{lfd}+L_{md} & L_{md} \\ 0 & L_{md} & 0 & 0 & L_{md} & L'_{lkd}+L_{md} \end{array}\right] \begin{bmatrix} i_q \\ i_d \\ i_0 \\ \hline i'_{kq} \\ i'_{fd} \\ i'_{kd} \end{bmatrix} \qquad [\mathrm{Wb\text{-}t}] \ . \tag{2.84}$$

Park's Transformation (PT) has reduced the complexity of machine equations in two ways [32]:

1. It has transformed a set of differential equations with time-varying coefficients to a set characterized by constant parameters. This means, for example, that the equations are now made amenable to the extremely powerful Laplace transform analysis technique;
2. The new transformed equations contain comparatively few terms. This means that the parameters matrices contain many zeros, i.e., these matrices are sparse. Whereas the physical stator currents are strongly coupled to each other, Park currents are only weakly coupled.

After presenting the change of variables, an important remark may be borrowed from Padiyar [35]:

> Although the physical interpretation of Park's Transformation is useful in gaining an intuitive understanding of its implications, it must be understood that it is not essential in the mathematical analysis of the synchronous machine. This is true of any mathematical transformation whose main objective is to simplify the analysis. From this point of view, the major benefit of Park's Transformation is to obtain the machine equations in time-invariant form which simplifies the analysis. [35, p. 82]

2.5.4 Voltage and flux-linkage equations in terms of reactances

It is often convenient to express the voltage and flux linkage equations in terms of reactances rather than inductances [19]. Hence, from the definition of flux linkages per second

$$\Psi_q = x_{ls} i_q + x_{mq}\left(i_q + i'_{kq}\right) \qquad [\mathrm{V}]\ , \tag{2.85a}$$

$$\Psi_d = x_{ls} i_q + x_{md}\left(i_d + i'_{fd} + i'_{kd}\right) \qquad [\mathrm{V}]\ , \tag{2.85b}$$

$$\Psi_0 = x_{ls} i_0 \qquad [\mathrm{V}]\ , \tag{2.85c}$$

$$\Psi'_{kq} = x'_{lkq} i'_{kq} + x_{mq}\left(i_q + i'_{kq}\right) \qquad [\mathrm{V}]\ , \tag{2.85d}$$

$$\Psi'_{fd} = x'_{lfd} i'_{fd} + x_{md}\left(i_d + i'_{fd} + i'_{kd}\right) \qquad [\mathrm{V}]\ , \tag{2.85e}$$

$$\Psi'_{kd} = x'_{lkd} i'_{kd} + x_{md}\left(i_d + i'_{fd} + i'_{kd}\right) \qquad [\mathrm{V}]\ . \tag{2.85f}$$

Equation (2.77) may be written as

$$v_q = r_s i_q + \frac{\omega_r}{\omega_b}\Psi_d + \frac{p}{\omega_b}\Psi_q \qquad [\mathrm{V}]\ , \tag{2.86a}$$

$$v_d = r_s i_d - \frac{\omega_r}{\omega_b}\Psi_q + \frac{p}{\omega_b}\Psi_d \qquad [\mathrm{V}]\ , \tag{2.86b}$$

$$v_0 = r_s i_0 + \frac{p}{\omega_b}\Psi_0 \qquad [\mathrm{V}]\ . \tag{2.86c}$$

In the same manner, (2.79) may become

$$v'_{kq} = r'_{kq} i'_{kq} + \frac{p}{\omega_b}\Psi'_{kq} \qquad [\mathrm{V}]\ , \tag{2.87a}$$

$$v'_{fd} = r'_{fd} i'_{fd} + \frac{p}{\omega_b} \Psi'_{fd} \qquad [\text{V}] \ , \tag{2.87b}$$

$$v'_{kd} = r'_{kd} i'_{kd} + \frac{p}{\omega_b} \Psi'_{kd} \qquad [\text{V}] \ , \tag{2.87c}$$

where ω_b is the base electrical angular velocity used to calculate the inductive reactances.

The expanded matrix form is

$$\left[\begin{array}{c} v_q \\ v_d \\ v_0 \\ \hline v'_{kq} \\ v'_{fd} \\ v'_{kd} \end{array}\right] = \left[\begin{array}{ccc|ccc} r_s & 0 & 0 & 0 & 0 & 0 \\ 0 & r_s & 0 & 0 & 0 & 0 \\ 0 & 0 & r_s & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & r'_{lkq} & 0 & 0 \\ 0 & 0 & 0 & 0 & r'_{lfd} & 0 \\ 0 & 0 & 0 & 0 & 0 & r'_{lkd} \end{array}\right] \left[\begin{array}{c} i_q \\ i_d \\ i_0 \\ \hline i'_{kq} \\ i'_{fd} \\ i'_{kd} \end{array}\right] + \frac{\omega_r}{\omega_b} \left[\begin{array}{c} \Psi_d \\ -\Psi_q \\ 0 \\ \hline 0 \\ 0 \\ 0 \end{array}\right] + \frac{1}{\omega_b} \left[\begin{array}{c} p\Psi_q \\ p\Psi_d \\ p\Psi_0 \\ \hline p\Psi'_{kq} \\ p\Psi'_{fd} \\ p\Psi'_{kd} \end{array}\right] \qquad [\text{V}] \ . \tag{2.88}$$

2.5.5 Power equations in rotor reference-frame variables

The total power expressed in the $qd0$ variables, $\mathcal{P}_{qd0s}$, must equal the total power expressed in the abc variables, $\mathcal{P}_{abcs}$:

$$\mathcal{P}_{qd0s} = \mathcal{P}_{abcs} \qquad [\text{VA}] \ . \tag{2.89}$$

Therefore, let the three-phase real power as stated before:

$$\mathcal{P}_{qd0s} = \langle \mathbf{v}_{abcs}, \mathbf{i}_{abcs} \rangle \qquad [\text{VA}] \ . \tag{2.90}$$

Using (2.68) and (2.59) to replace the actual currents and voltages in the equation above, the three-phase real power becomes

$$\mathcal{P}_{qd0s} = \langle \mathbf{K}^{-1} \mathbf{v}_{qd0s}, \mathbf{K}^{-1} \mathbf{i}_{qd0s} \rangle \tag{2.91a}$$

$$= \left[\mathbf{K}^{-1} \mathbf{v}_{qd0s} \right]^{\mathrm{T}} \mathbf{K}^{-1} \mathbf{i}_{qd0s} \tag{2.91b}$$

$$= \left[\mathbf{v}_{qd0s} \right]^{\mathrm{T}} \left[\mathbf{K}^{-1} \right]^{\mathrm{T}} \mathbf{K}^{-1} \mathbf{i}_{qd0s} \tag{2.91c}$$

$$= \left[\mathbf{v}_{qd0s} \right]^{\mathrm{T}} \begin{bmatrix} 3/2 & 0 & 0 \\ 0 & 3/2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \mathbf{i}_{qd0s} \tag{2.91d}$$

$$= \frac{3}{2} \left(v_q i_q + v_d i_d + 2 v_0 i_0 \right) \qquad [\text{VA}] \ . \tag{2.91e}$$

The factor $3/2$ comes about due to the choice of constants k_q, k_d, and k_0 used when the transformation is firstly defined, in (2.58).

The instantaneous power may be also expressed in a more expanded form as:

$$\mathcal{P}_{qd0s} = \left\{ \omega_r \frac{3}{2} \left(\psi_d i_q - \psi_q i_d \right) \right\} + \left\{ \frac{3}{2} \left(p\psi_d i_d - p\psi_q i_q + 2 p\psi_0 i_0 \right) \right\} + \left\{ \frac{3}{2} r_s \left(i_q^2 + i_d^2 + 2 i_0^2 \right) \right\} \quad [\text{VA}] \ , \tag{2.92}$$

which has three principal parts. From the right is dissipation in the armature resistance, then a set of terms that relate to energy stored in magnetic fields, or more precisely, rate of change of energy stored in magnetic fields. The leftmost term, which is proportional to rotational speed, must be energy conversion [71] – i.e., the power transferred across the air gap [33].

2.5.6 Torque equation in rotor reference-frame variables

The expression for electromagnetic torque in terms of rotor reference-frame variables may be obtained by substituting the equation of transformation into (2.50). Hence

$$\mathcal{T}_e = \rho \left[\mathbf{K}^{-1}\mathbf{i}_{qd0s}\right]^{\mathrm{T}} \left\{\frac{1}{2}\frac{\partial}{\partial\theta_r}\left[\mathbf{L}_s\right]\mathbf{K}^{-1}\mathbf{i}_{qd0s} + \frac{\partial}{\partial\theta_r}\left[\mathbf{L}'_{sr}\right]\mathbf{i}'_{qdr}\right\} \qquad [\mathrm{N.m}] \;, \tag{2.93}$$

which, after some considerable work, reduces to

$$\mathcal{T}_e = \frac{3}{2}\rho\left[L_{md}\left(i_d + i_{fd} + i_{kd}\right) i_q - L_{mq}\left(i_q + i_{kq}\right) i_d\right] \qquad [\mathrm{N.m}] \;. \tag{2.94}$$

Furthermore, (2.94) can be equivalently expressed as

$$\mathcal{T}_e = \frac{3}{2}\rho\left(\psi_d i_q - \psi_q i_d\right) \qquad [\mathrm{N.m}] \;, \tag{2.95}$$

or, in terms of flux linkages per second and currents,

$$\mathcal{T}_e = \frac{3}{2}\rho\left(\frac{1}{\omega_b}\right)\left(\Psi_d i_q - \Psi_q i_d\right) \qquad [\mathrm{N.m}] \;. \tag{2.96}$$

If either $\mathbf{v}_{qds}$ or $\mathbf{i}_{qds}$ is an unsymmetrical or unbalanced function of θ_r, then other coefficients of ω_r could arise in addition to (2.95).

2.6 Per-unitized equations

At long last, the equations for a synchronous machine may be written in pu where base voltage is generally selected as the rms value of the rated phase voltage for the *abc* variables and the peak value for the *qd*0 variables. Although, the same base value may be considered when comparing *abc* and *qd*0 variables [19].

The per-unit system is of great benefit in making design calculations for machines, because it makes the comparison between different machines very much easier. Corresponding quantities are of the same order of magnitude even for widely different designs [30].

The presented Park's equations written in terms of flux linkages per second and reactances are readily per unitized by dividing each term by the peak of the base voltage (or the peak value of the base current times base impedance). The form of these equations remains unchanged as a result of per unitizing.

2.6.1 Base quantities

Up to this point, all machine equations have been written in terms of actual units. For purposes of analysis, it is convenient to convert these equations to a normalized or pu form. The choice of reference or base quantities is arbitrary, but it is usually related to the nameplate rating of the machine [36].

The base quantities are defined as follows:

2.6.1.1 Base angular frequency

$$\omega_b = \omega_{base} = 2\pi f_n \qquad [\mathrm{rad/s}] \ , \tag{2.97}$$

where f_n is the rated frequency, in hertz.

2.6.1.2 Base stator current

$$I_b = I_{sbase} = \frac{\mathcal{P}_n\sqrt{2}}{V_n\sqrt{3}} \qquad [\mathrm{A}] \ , \tag{2.98}$$

where $\mathcal{P}_n$ is three-phase rated power, in volt-ampere; and V_n is the rated rms line-to-line voltage, in volts.

2.6.1.3 Base stator voltage

$$V_b = V_{sbase} = \frac{V_n\sqrt{2}}{\sqrt{3}} \qquad [\mathrm{V}] \ , \tag{2.99}$$

where V_n is the rated rms line-to-line voltage, in volts. Note that the stator voltage base value is the peak rated line-to-neutral voltage.

2.6.1.4 Base power

Considering equations (2.98) and (2.99),

$$\mathcal{P}_n = \frac{V_n I_b\sqrt{3}}{\sqrt{2}} \tag{2.100a}$$

$$= \frac{V_b I_b\sqrt{3}\sqrt{3}}{\sqrt{2}\sqrt{2}} \tag{2.100b}$$

$$= \frac{3}{2} V_b I_b \qquad [\mathrm{VA}] \ , \tag{2.100c}$$

where V_n is the rated rms line-to-line voltage, in volts; I_b is the base stator current, in amperes; and V_b is the base stator voltage, in volts.

Thus,

$$\mathcal{P}_b = \mathcal{P}_{base} = \mathcal{P}_n = \frac{3}{2} V_b I_b \qquad [\mathrm{VA}] \ . \tag{2.101}$$

2.6.1.5 Base stator impedance

$$Z_b = Z_{sbase} = \frac{V_{sbase}}{I_{sbase}} = \frac{V_n^2}{\mathcal{P}_n} \qquad [\Omega] \ , \tag{2.102}$$

where V_n is the rated rms line-to-line voltage, in volts; and $\mathcal{P}_n$ is three-phase rated power, in volt-ampere.

2.6.1.6 Base stator inductance

$$L_b = L_{sbase} = \frac{Z_{sbase}}{\omega_{base}} \qquad [\mathrm{H}] \ , \tag{2.103}$$

where Z_{sbase} is the base stator impedance, in ohms; and ω_{base} is the base angular frequency, in radians per second.

2.6.1.7 Base torque

$$\mathcal{T}_b = \mathcal{T}_{base} = \frac{\mathcal{P}_b}{(1/\rho)\omega_{base}} \qquad [\mathrm{N.m}] \ , \tag{2.104}$$

where $\mathcal{P}_b$ is the base power, in volt-ampere; ω_{base} is the base angular frequency, in radians per second; and ρ is the number of pole pairs.

2.6.1.8 Base field current

$$I_{fbase} = I_{fn} \qquad [\mathrm{A}] \ , \tag{2.105}$$

where I_{fn} is the field current that produces rated stator voltage at no load, in amperes.

2.6.1.9 Base field voltage

$$V_{fbase} = \frac{\mathcal{P}_n}{I_{fbase}} \qquad [\mathrm{V}] \ , \tag{2.106}$$

where $\mathcal{P}_n$ is three-phase rated power, in volt-ampere; and I_{fbase} is the base field current, in amperes.

2.6.1.10 Base field impedance

$$Z_{fbase} = \frac{V_{fbase}}{I_{fbase}} \qquad [\Omega] \; , \tag{2.107}$$

where V_{fbase} is the base field voltage, in volts; and I_{fbase} is the base field current, in amperes.

2.6.1.11 Base field inductance

$$L_{fbase} = \frac{Z_{fbase}}{\omega_{base}} \qquad [\mathrm{H}] \; , \tag{2.108}$$

where Z_{fbase} is the base field impedance, in ohms; and ω_{base} is the base angular frequency, in radians per second.

2.6.2 Voltage equations in per-unit

The voltage equations in per-unit form are obtained by dividing all of the voltage differential equations by V_b or $Z_b I_b$ as appropriate. As an example, the q-axis voltage equation can be written as

$$v_q = r_s i_q + \frac{\omega_r}{\omega_b}\Psi_d + \frac{p}{\omega_b}\Psi_q \qquad [\mathrm{V}] \; , \tag{2.109a}$$

$$\frac{v_q}{V_b} = \frac{r_s i_q}{Z_b I_b} + \frac{\omega_r}{\omega_b}\frac{\Psi_d}{V_b} + \frac{p}{\omega_b}\frac{\Psi_q}{V_b} \qquad [\mathrm{pu}] \; , \tag{2.109b}$$

$$\bar{v}_q = \bar{r}_s \bar{i}_q + \overline{\omega}_r \overline{\Psi}_d + \frac{p}{\omega_b}\overline{\Psi}_q \qquad [\mathrm{pu}] \; , \tag{2.109c}$$

where the $\overline{\text{bars}}$ indicate per-unitized quantities.

The other voltage equations can be handled in the same manner.

All voltage equations are summarized as follows:

$$\bar{v}_q = \bar{r}_s \bar{i}_q + \overline{\omega}_r \overline{\Psi}_d + \frac{p}{\omega_b}\overline{\Psi}_q \qquad [\mathrm{pu}] \; , \tag{2.110a}$$

$$\bar{v}_d = \bar{r}_s \bar{i}_d - \overline{\omega}_r \overline{\Psi}_q + \frac{p}{\omega_b}\overline{\Psi}_d \qquad [\mathrm{pu}] \; , \tag{2.110b}$$

$$\bar{v}_0 = \bar{r}_s \bar{i}_0 + \frac{p}{\omega_b}\overline{\Psi}_0 \qquad [\mathrm{pu}] \; , \tag{2.110c}$$

$$\bar{v}'_{kq} = \bar{r}'_{kq} \bar{i}'_{kq} + \frac{p}{\omega_b}\overline{\Psi}'_{kq} \qquad [\mathrm{pu}] \; , \tag{2.110d}$$

$$\bar{v}'_{fd} = \bar{r}'_{fd} \bar{i}'_{fd} + \frac{p}{\omega_b}\overline{\Psi}'_{fd} \qquad [\mathrm{pu}] \; , \tag{2.110e}$$

$$\bar{v}'_{kd} = \bar{r}'_{kd} \bar{i}'_{kd} + \frac{p}{\omega_b}\overline{\Psi}'_{kd} \qquad [\mathrm{pu}] \; . \tag{2.110f}$$

2.6.3 Flux linkage equations in per-unit

Considering the q-axis flux linkage equation, its per-unitized version is obtained by dividing the proper quantities by V_b or $Z_b I_b$:

$$\Psi_q = x_{ls} i_q + x_{mq}\left(i_q + i'_{kq}\right) \qquad \text{[Wb-t]} \tag{2.111a}$$

$$\frac{\Psi_q}{V_b} = \frac{x_{ls} i_q}{Z_b I_b} + \frac{x_{mq}}{Z_b}\left(\frac{i_q}{I_b} + \frac{i'_{kq}}{I_b}\right) \qquad \text{[pu]} \tag{2.111b}$$

$$\overline{\Psi}_q = \overline{x}_{ls} \overline{i}_q + \overline{x}_{mq}\left(\overline{i}_q + \overline{i}'_{kq}\right) \qquad \text{[pu]} \; . \tag{2.111c}$$

The other flux linkage equations can be handled in the same manner.

All flux linkage equations are summarized as follows:

$$\overline{\Psi}_q = \overline{x}_{ls} \overline{i}_q + \overline{x}_{mq}\left(\overline{i}_q + \overline{i}'_{kq}\right) \qquad \text{[pu]} \; , \tag{2.112a}$$

$$\overline{\Psi}_d = \overline{x}_{ls} \overline{i}_d + \overline{x}_{md}\left(\overline{i}_d + \overline{i}'_{fd} + \overline{i}'_{kd}\right) \qquad \text{[pu]} \; , \tag{2.112b}$$

$$\overline{\Psi}_0 = \overline{x}_{ls} \overline{i}_q \qquad \text{[pu]} \; , \tag{2.112c}$$

$$\overline{\Psi}'_{kq} = \overline{x}'_{lkq} \overline{i}'_{kq} + \overline{x}_{mq}\left(\overline{i}_q + \overline{i}'_{kq}\right) \qquad \text{[pu]} \; , \tag{2.112d}$$

$$\overline{\Psi}'_{fd} = \overline{x}'_{lfd} \overline{i}'_{fd} + \overline{x}_{md}\left(\overline{i}_d + \overline{i}'_{fd} + \overline{i}'_{kd}\right) \qquad \text{[pu]} \; , \tag{2.112e}$$

$$\overline{\Psi}'_{kd} = \overline{x}'_{lkd} \overline{i}'_{kd} + \overline{x}_{md}\left(\overline{i}_d + \overline{i}'_{fd} + \overline{i}'_{kd}\right) \qquad \text{[pu]} \; . \tag{2.112f}$$

2.6.4 Power equations in per-unit

Dividing the instantaneous power equation by the power base and converting all other quantities to pu yields to

$$\mathcal{P}_{qd0s} = \frac{3}{2}\left(v_q i_q + v_d i_d + 2 v_0 i_0\right) \qquad \text{[VA]} \tag{2.113a}$$

$$\frac{\mathcal{P}_{qd0s}}{\mathcal{P}_b} = \frac{3/2 V_b I_b \left(\overline{v}_q \overline{i}_q + \overline{v}_d \overline{i}_d + 2\overline{v}_0 \overline{i}_0\right)}{3/2 V_b I_b} \qquad \text{[pu]} \tag{2.113b}$$

$$\overline{\mathcal{P}}_{qd0s} = \overline{v}_q \overline{i}_q + \overline{v}_d \overline{i}_d + 2\overline{v}_0 \overline{i}_0 \qquad \text{[pu]} \; . \tag{2.113c}$$

2.6.5 Torque equation in per-unit

Base torque is the base power divided by the synchronous speed of the rotor. With all quantities expressed in pu, (2.96) becomes

$$\overline{\mathcal{T}}_e = \overline{\Psi}_d \overline{i}_q - \overline{\Psi}_q \overline{i}_d \qquad \text{[pu]} \; . \tag{2.114}$$

Equation (2.51), which relates torque and rotor speed, is expressed in pu as

$$\overline{\mathcal{T}}_e = \overline{\mathcal{T}_m} - 2Hp\overline{\omega}_r \qquad [\text{pu}] \ , \tag{2.115}$$

where ω_b corresponds to rated or base frequency, in rad/s; and the inertia constant

$$H = \frac{1}{2\rho^2} J \frac{\omega_b^2}{\mathcal{P}_b} \qquad [\text{s}] \tag{2.116}$$

is expressed in seconds. In (2.116), $\mathcal{P}_b$ is the base power, in volt-ampere.

2.6.6 Motion equations in per-unit

The motion equations are easily transformed into per-unit. Considering the base definitions, (2.53) becomes

$$\omega_r = \omega_s + p\delta \qquad [\text{electrical rad/s}] \tag{2.117a}$$

$$\omega_b\overline{\omega}_r = \omega_b\overline{\omega}_s + p\delta \qquad [\text{pu}] \ . \tag{2.117b}$$

where $\overline{\omega}_s$ is the synchronous speed, in pu; $\overline{\omega}_r$ is the rotor speed, in pu; and δ is the load angle, in electrical radians.

As before, from (2.117b):

$$\delta = \int \left(\omega_b\overline{\omega}_r - \omega_b\overline{\omega}_s\right) dt \tag{2.118a}$$

$$= \omega_b\left(\overline{\omega}_r t - \overline{\omega}_s t\right) + \delta_0 \qquad [\text{electrical rad}] \ , \tag{2.118b}$$

where δ_0 is the load angle value at $t = 0$, in electrical radians.

According to Kundur [33], it is often desirable to include a component of damping torque, not accounted for in the calculation of $\mathcal{T}_e$, separately. This is accomplished by adding a term proportional to speed deviation in the previous equations as follows:

$$\frac{2H}{\omega_b} p^2\delta = \overline{\mathcal{T}_m} - \overline{\mathcal{T}_e} - \kappa_d \Delta\overline{\omega}_r \qquad [\text{pu}] \ , \tag{2.119}$$

where κ_d is the damping-torque coefficient, in newton-meter-second (N.m.s) and accounts for the mechanical rotational loss due to windage and friction; and $\Delta\omega_r = p\delta$ is the speed deviation, in electrical radians per second.

A couple important remarks:

1. The time derivative of the load angle δ is not the speed itself, but the speed deviation:

$$\Delta\overline{\omega}_r = \frac{\Delta\omega_r}{\omega_b} = \frac{p}{\omega_b}\delta \qquad [\text{electrical rad/s}] \ ; \tag{2.120}$$

2. Equation 2.119 represents the equation of motion of a synchronous machine. It is commonly referred to as the *swing equation*, because it represents swings in rotor angle δ during disturbances [33, 50];

3. It can be shown that when $\overline{\mathcal{T}_m} = 0$ and rated torque $\overline{\mathcal{T}_e} = 1$ is exerted by the machine, the time required to accelerate the rotor shaft from zero to rated speed is identically equal to $2H$ [36].

2.7 Electrical equivalent circuits

The modeling concept used in this work forms the basis for all but the simplest of synchronous machine models. Most of the models available are based upon direct- and quadrature-axis representations of the synchronous machine. These representations may take a number of forms: equivalent circuits, transfer functions, flux-current and voltage relationships, state-space equations, among others. However, all these forms are equivalent and provide the same results [49].

Another important aspect to notice is that the direct- and quadrature-axis models derived here represent these axes as being magnetically uncoupled. This representation is based upon the assumption that currents in one axis do not produce flux in the other axis – i.e., produce no changes in the flux in the other axes. In reality, magnetic nonlinearities, e.g., magnetic saturation, will produce some degree of coupling between the axes. Although models that neglect this coupling have been found to be adequate for many studies [72, 73], work is currently underway to develop techniques for incorporating the effects of magnetic nonlinearities in both steady-state and transient analyses [74–79].

While the equations derived in previous sections can be used directly to determine synchronous machine performance, it is a common practice to use equivalent circuits to provide visual description of the machine model.

The q-axis equivalent circuit is shown in Figure 2.8; the d-axis equivalent circuit, in Figure 2.9; and the zero-sequence equivalent circuit, in Figure 2.10. In these equivalent circuits, voltages, as well as flux linkages in terms of their time derivatives, appear.

2.7.1 Quadrature-axis equivalent circuit

Because there is no rotor winding with terminals on the quadrature-axis, the quadrature-axis equivalent circuit needs to be represented only as a single-port network. Although Figure 2.8 includes two terminal ports, it is important to realize that $v'_{kq} = 0$, as it is a short-circuited winding.

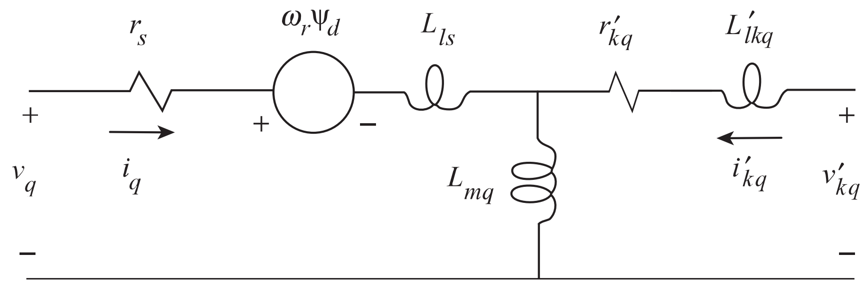

Figure 2.8: Quadrature-axis equivalent circuit of a three-phase synchronous machine with the reference frame fixed in rotor: Park equations. Adapted from Krause et al. [19, p. 153].

Furthermore, there is no field winding and the single damper winding represents the overall effects of the amortisseur-winding and eddy-currents paths. Therefore, it is reasonable to assume that the armature and damper circuits all link a single ideal mutual flux represented by L_{mq} [33].

The quadrature- and direct-axis circuits are not decoupled because of the speed voltage terms – represented by those controlled sources. The quadrature-axis speed voltage depends on the direct-axis currents, and vice-versa. These speed voltages also depend on the shaft speed, ω_r, which is not constant under transient conditions [42]. Hence, the speed voltage terms are nonlinear.

2.7.2 Direct-axis equivalent circuit

The direct-axis of a synchronous machine includes three terminal ports. These ports correspond to the direct-axis equivalent armature winding, the field winding, and the amortisseur winding. As mentioned before, although the amortisseur winding is shown with provisions to apply a voltage, it is, in fact, a short-circuited winding that represents the path for an induced rotor current [19].

Figure 2.9 shows the equivalent-circuit representation for the direct-axis model with a single damper winding. The variables v'_{fd}, i'_{fd}, v'_{kd}, and i'_{kd} correspond to the values of field voltage and current, damper-winding voltage and current, respectively, reflected to the armature winding through the equivalent winding turns ratio.

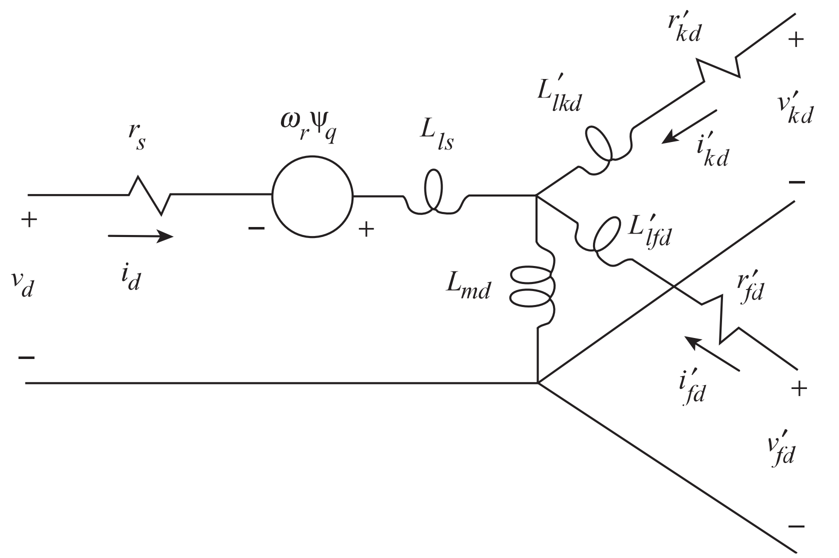

Figure 2.9: Direct-axis equivalent circuit of a three-phase synchronous machine with the reference frame fixed in rotor: Park equations. Adapted from Krause et al. [19, p. 153].

It is also important to notice that the leakage inductance L'_{lfd} accounts for the fact that the mutual inductance between the field winding and the armature winding is not necessarily equal to that between the field winding and the amortisseur winding L'_{lkd}. IEEE-1110 [49] states that for turbo-generators, L'_{lfd} is often found to be positive while for salient-pole machines, L'_{lfd} is usually negative. This reflects the different physical couplings between the field circuit and the equivalent rotor body circuits in turbo-generators as compared to hydro-generators.

2.7.3 Zero-sequence equivalent circuit

The zero-sequence equivalent circuit, show in Figure 2.10, has no mutual coupling with either quadrature- or direct-axis circuits and is therefore in quadrature with the quadrature and direct axes; it must be orthogonal to them. Therefore, it magnetizes an axis that lies along the rotor center-line or rotational axis and is perpendicular to the plane formed by the quadrature and direct axes [42].

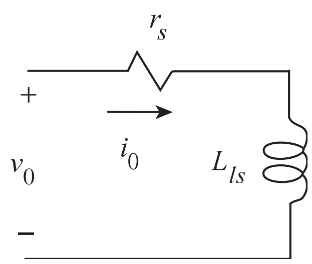

Figure 2.10: Zero-sequence equivalent circuit of a three-phase synchronous machine with the reference frame fixed in rotor: Park equations. Adapted from Krause et al. [19, p. 153].

The zero-sequence equivalent circuit plays a relatively minor role in stability studies – in fact, no role at all in studies which assume balanced operating conditions [49]. Differently from the other circuits, the zero-sequence circuit has no speed voltage and can be neglected when studying balanced conditions [42].

2.7.4 Equivalent circuits coupling

Another interesting way of representing the voltage equations (2.86) is the $qd0$ equivalent circuit shown in Figure 2.11, where the damper winding driving voltages are zero – these voltages were carried symbolically in previous equations for the sake of completeness of notation.

All inductances are constant and the zero-sequence network is completely decoupled from the other ones. The circuit from Figure 2.11 also presents the speed voltage terms previously mentioned, and the coupling between the circuits.

2.8 Steady-state analysis

The performance of synchronous machines under balanced steady-state conditions may be readily analyzed by applying the per unit equations summarized in Section 2.6. For balanced conditions, the zero-sequence quantities are zero. For balanced steady-state conditions, the electrical angular velocity of the rotor is constant and equal to ω_s, whereupon the electrical

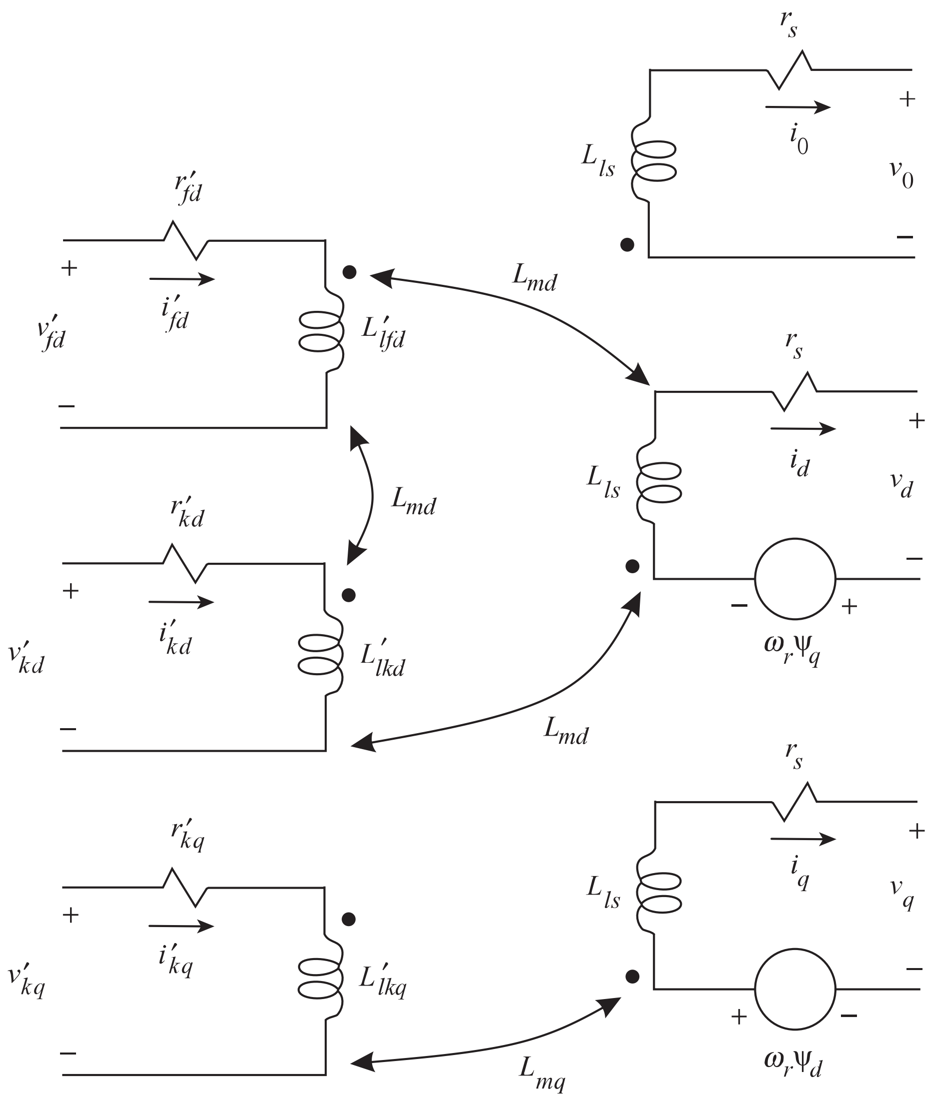

Figure 2.11: Coupling circuit representation of the synchronous machine with the reference frame fixed in the rotor. Adapted from Anderson, Agrawal, and Ness [42, p. 52].

angular velocity of the rotor reference frame becomes the electrical angular velocity of the synchronously rotating reference frame.

In this mode of operation, the rotor windings do not experience a change of flux linkages, hence current is not flowing in the short-circuited damper windings. Thus, with $\overline{\omega}_r$ set equal to $\overline{\omega}_s$ and the time rate of change of all flux linkages set equal to zero, the steady-state versions of (2.110) become:

$$\bar{v}_q = \bar{r}_s\bar{i}_q + \overline{\omega}_s\overline{\Psi}_d \qquad [\text{pu}] \;, \tag{2.121a}$$

$$\bar{v}_d = \bar{r}_s\bar{i}_d - \overline{\omega}_s\overline{\Psi}_q \qquad [\text{pu}] \;, \tag{2.121b}$$

$$\bar{v}'_{fd} = \bar{r}'_{fd}\bar{i}'_{fd} \qquad [\text{pu}] \;. \tag{2.121c}$$

Moreover, considering (2.112),

$$\bar{v}_q = \bar{r}_s\bar{i}_q + \overline{\omega}_s\left[\bar{x}_{ls}\bar{i}_d + \bar{x}_{md}\left(\bar{i}_d + \bar{i}'_{fd}\right)\right] \qquad [\text{pu}] \tag{2.122a}$$

$$= \bar{r}_s \dot{i}_q + \overline{\omega}_s \left(\bar{x}_{ls} + \bar{x}_{md} \right) \dot{i}_d + \overline{\omega}_s \bar{x}_{md} \dot{i}'_{fd} \qquad [\mathrm{pu}] \tag{2.122b}$$

$$= \bar{r}_s \dot{i}_q + \overline{\omega}_s \bar{x}_d \dot{i}_d + \overline{\omega}_s \bar{x}_{md} \dot{i}'_{fd} \qquad [\mathrm{pu}] \ , \tag{2.122c}$$

$$\bar{v}_d = \bar{r}_s \dot{i}_d - \overline{\omega}_s \left[\bar{x}_{ls} \dot{i}_q + \bar{x}_{mq} \dot{i}_q \right] \qquad [\mathrm{pu}] \tag{2.122d}$$

$$= \bar{r}_s \dot{i}_d - \overline{\omega}_s \bar{x}_q \dot{i}_q \qquad [\mathrm{pu}] \ , \tag{2.122e}$$

$$\bar{v}'_{fd} = \bar{r}'_{fd} \dot{i}'_{fd} \qquad [\mathrm{pu}] \ . \tag{2.122f}$$

2.8.1 Phasor diagrams

The phasor diagram is of very importance for analyzing working conditions in a synchronous machine. By using it, it is possible to determine the operating conditions of a machine without actually applying the load, which would become especially difficult for large rating machines.

Furthermore, the phasor diagram allows to determine the load angle, δ, between the emf produced by the excitation field and the voltage across the terminals. The angle, δ, plays a very important role in the analysis of torque and power developed by a machine in both steady-state and transient conditions.

The vector difference between the emf, $\tilde{E}_f$, due to the excitation flux, and the terminal voltage, $\tilde{V}_s$, depends on the effect of the armature reaction and on the voltage drop in the active resistance and leakage inductive reactance of the armature winding.

Since armature reaction depends, to a great extent, on the type of the machine (salient-pole or non-salient-pole), load characteristics (inductive, active, or capacitive), and on the degree of load symmetry (balanced or unbalanced), all these factors must be properly considered when plotting a phasor diagram.

It is necessary to bear in mind that all the emfs and voltages that participate as components in the phasor diagram should correspond to its fundamental frequency; therefore, all the emfs must, preliminarily, be resolved into harmonics and, from each of them, the fundamental wave must be taken separately.

Since the vector summation of fluxes and the corresponding emfs induced by them by the superposition method is legitimate only when the reluctances are constant in all sections of the magnetic circuit of the machine, this method is directly applicable to the unsaturated magnetic circuit of a synchronous machine.

The method that is of greatest interest is the Blondel two-reaction theory [57], according to which all fluxes due to the load current, including the leakage flux, are solved along the quadrature and direct axes.

Moreover, for balanced conditions, assuming that the parameters of all phases are equal, the diagram construction may be restricted for one phase only. It should also be noticed that the vector diagrams constructed for a synchronous machine operating as a generator may be readily extended to its operation as a synchronous motor and a synchronous condenser.

2.8.1.1 Phasor diagram for salient-pole synchronous machines

The following development was performed by Kostenko and Piotrovsky [34], to which one should refer for further detail.

In a salient-pole machine, the fundamental wave of the armature-reaction mmf rotates in step with the rotor and, owing to the non-uniformity of the air gap between the rotor and the stator, produces a non-sinusoidal armature-reaction magnetic flux which induces, in turn, a non-sinusoidal armature-reaction emf. To include the armature-reaction emf in the phasor diagram, the fundamental wave must be separated from it. This is achieved with the aid of the method based on the Blondel two-reaction theory.

According to his method, the fundamental wave of the armature reaction is resolved into two components: the quadrature- and direct-axis reaction components. Separating from the fluxes,

$$\tilde{\mathcal{F}}_{mq} = \tilde{\mathcal{F}}_{as} \cos\gamma\,, \tag{2.123a}$$

$$\tilde{\mathcal{F}}_{md} = \tilde{\mathcal{F}}_{as} \sin\gamma\,, \tag{2.123b}$$

where $\tilde{\mathcal{F}}_{mq}$ is the quadrature-axis component of the armature-reaction mmf $\tilde{\mathcal{F}}_{as}$; $\tilde{\mathcal{F}}_{md}$ is the direct-axis component of the armature-reaction mmf $\tilde{\mathcal{F}}_{as}$; the angle γ indicates the space displacement of the conductors carrying maximum current I_s relative to the conductors which have the maximum emf E_f and are opposite the pole axis. By this same angle γ, current $\tilde{I}_s$ lags behind the emf $\tilde{E}_f$ in time phase.

These armature-reaction components, $\tilde{\mathcal{F}}_{mq}$ and $\tilde{\mathcal{F}}_{md}$, will produce the fundamental-wave fluxes, $\tilde{\Phi}_{mq}$ and $\tilde{\Phi}_{md}$, which induce the armature-reaction emfs, $\tilde{E}_{mq}$ and $\tilde{E}_{md}$.

■

Inductive loading case

Considering a three-phase salient-pole synchronous generator for the case of an inductive load, when $0 < \gamma < \pi/2$, the phasor diagram is obtained by drawing vectors in the following sequence:

1. The emf vector $(\tilde{E}_f)$ produced by the magnetic excitation flux $(\tilde{\Phi}_f)$ on the positive direction of the quadrature-axis;
2. The magnetic excitation flux vector $(\tilde{\Phi}_f)$ on the negative direction of the direct-axis;
3. The current vector $(\tilde{I}_{as})$, as lagging $(\tilde{E}_f)$ by $0 < \gamma < \pi/2$;
4. The $(\tilde{I}_{as})$ components: $(\tilde{I}_q)$, on the quadrature-axis, and $(\tilde{I}_d)$, on the direct-axis;

5. The flux of armature-reaction ($\tilde{\Phi}_{as}$) components: ($\tilde{\Phi}_{mq}$) in phase with ($\tilde{I}_q$) and ($\tilde{\Phi}_d$) in phase with ($\tilde{I}_d$).

In the stator winding, the armature-reaction fluxes ($\tilde{\Phi}_{mq}$) and ($\tilde{\Phi}_{md}$) induce the emfs ($\tilde{E}_{mq}$) and ($\tilde{E}_{md}$), which have fundamental frequency and lag the corresponding fluxes by $\pi/2$. If the magnetic circuit is not saturated, ($\tilde{E}_{mq} = -j\tilde{I}_q x_{mq}$) and ($\tilde{E}_{md} = -j\tilde{I}_d x_{md}$). Therefore,

6. Draw the voltages drops ($-j\tilde{I}_q x_{mq}$) and ($-j\tilde{I}_d x_{md}$).

The armature-reaction emf vector $\left(\tilde{E}_{as} = \sqrt{E_{mq}^2 + E_{md}^2}\right)$ lags the armature current ($\tilde{I}_{as}$) by a time phase angle other than $\pi/2$ – in the case of non-salient-pole machines, ($\tilde{E}_{as}$) lags ($\tilde{I}_{as}$) by exactly $\pi/2$. With the armature-reaction flux ($\tilde{\Phi}_{as}$), there is the stator-winding leakage flux ($\tilde{\Phi}_{ls}$), whose vector is in phase with current ($\tilde{I}_{as}$) and creates, in the stator winding, a leakage emf of fundamental frequency ($\tilde{E}_{ls} = -j\tilde{I}_{as} x_{ls}$). Then,

7. Draw the leakage reactance (x_{ls}) voltage drop ($-j\tilde{I}_{as} x_{ls}$).

At long last, the generator terminal voltage ($\tilde{V}_{as}$) is obtained upon vector adding ($\tilde{E}_f$), ($\tilde{E}_{mq}$), ($\tilde{E}_{md}$), ($\tilde{E}_{ls}$), and ($\tilde{E}_r = -r_s\tilde{I}_{as}$). Finally,

8. Draw the resistance voltage drop ($-r_s\tilde{I}_{as}$); and
9. Obtain ($\tilde{V}_{as}$) by performing the aforementioned vector addition.

The resulting phasor diagram is shown in Figure 2.12.

■

Capacitive loading case

Considering a three-phase salient-pole synchronous generator for the case of a capacitive load, when $-\pi/2 < \gamma < 0$, the phasor diagram is obtained upon changing the following step from the previous development (inductive loading case):

3. The current vector ($\tilde{I}_{as}$), as leading ($\tilde{E}_f$) by $0 < \gamma < \pi/2$;

The resulting phasor diagram is shown in Figure 2.13.

■

Comparing the diagrams in Figure 2.12 and Figure 2.13, the armature reaction produces a demagnetizing effect on the excitation system with the inductive load, whereas with a capacitive load, on the contrary, it produces a magnetizing effect. Therefore, in the first case, the resulting flux – which actually exists in the generator air gap and determines the saturation of its magnetic circuit – ($\tilde{\Phi}_\delta < \tilde{\Phi}_{as}$), and, in the second, ($\tilde{\Phi}_\delta > \tilde{\Phi}_{as}$); accordingly, in the first case, the resulting emf ($\tilde{E}_\delta < \tilde{E}_{as}$), and, in the second, ($\tilde{E}_\delta > \tilde{E}_{as}$).

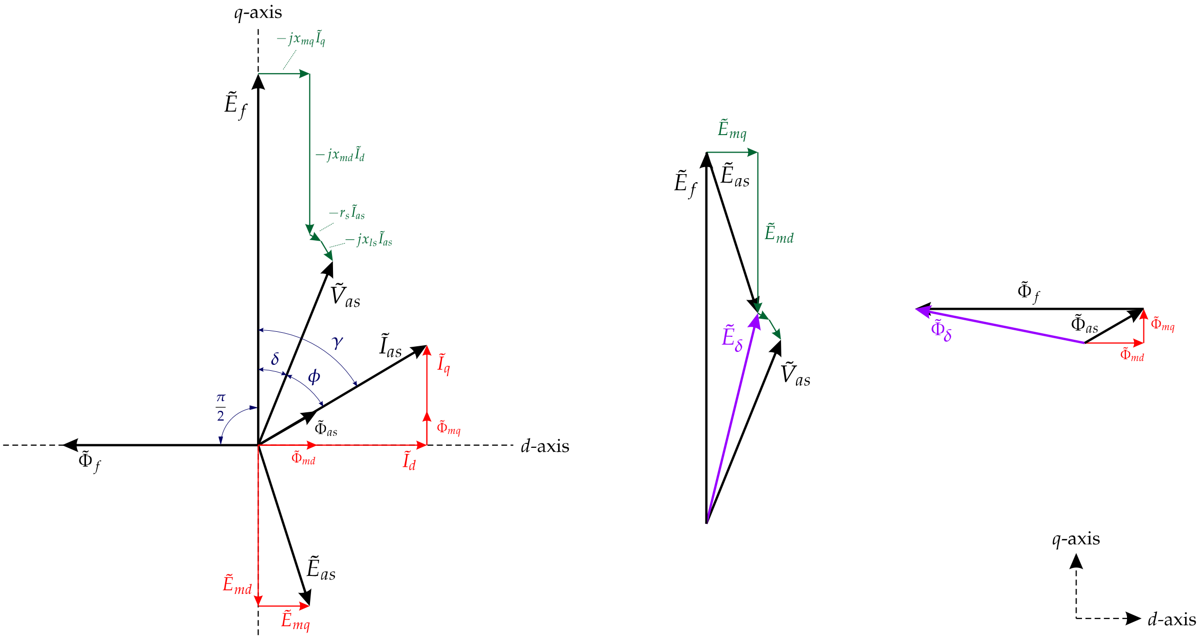

Figure 2.12: Phasor diagram of a three-phase salient-pole synchronous machine for the case of an inductive load.

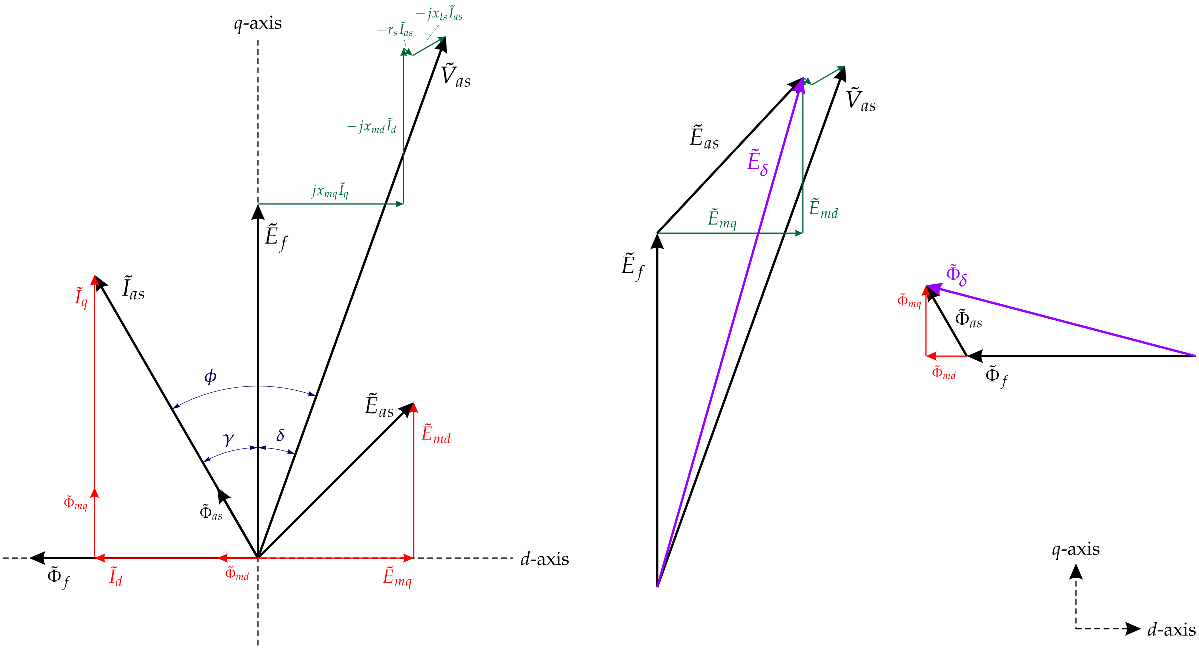

Figure 2.13: Phasor diagram of a three-phase salient-pole synchronous machine for the case of a capacitive load.

2.9 Standard synchronous machine reactances and time constants

It is instructive to set forth the commonly used reactances for the three-winding rotor synchronous machine and to relate these reactances to the operational impedances[28] whenever appropriate [19]. The following equations are derived in many works, such as Adkins [30] and Krause et al. [19].

The quadrature- and direct-axis reactances are

$$\bar{x}_q = \bar{x}_{ls} + \bar{x}_{mq} \qquad [\text{pu}] \tag{2.124}$$

and

$$\bar{x}_d = \bar{x}_{ls} + \bar{x}_{md} \qquad [\text{pu}]\ . \tag{2.125}$$

These reactances were previously defined and characterize the machine during balanced steady-state operation whereupon variables in the rotor reference frame are constants. The zero-frequency value of $\bar{x}_q(p)$ or $\bar{x}_d(p)$ is found by replacing the operator p with zero. Hence, the operational impedances for balanced steady-state operation are

$$\bar{x}_q(0) = \bar{x}_q \qquad [\text{pu}] \tag{2.126}$$

and

$$\bar{x}_d(0) = \bar{x}_d \qquad [\text{pu}]\ . \tag{2.127}$$

Similarly, the steady-state value of the transfer function is

$$G(0) = \frac{\bar{x}_{md}}{\bar{r}'_{fd}} \qquad [1]\ . \tag{2.128}$$

The direct-axis transient reactance[29] is defined as

$$\bar{x}'_d \triangleq \bar{x}_d \frac{\tau'_d}{\tau'_{d0}} = \bar{x}_{ls} + \frac{\bar{x}_{md}\bar{x}'_{lfd}}{\bar{x}_{md} + \bar{x}'_{lfd}} \qquad [\text{pu}]\ . \tag{2.129}$$

The quadrature- and direct-axis sub-transient reactances are

$$\bar{x}''_q \triangleq \bar{x}_q \frac{\tau''_q}{\tau''_{q0}} = \bar{x}_{ls} + \frac{\bar{x}_{mq}\bar{x}'_{lkq}}{\bar{x}_{mq} + \bar{x}'_{lkq}} \qquad [\text{pu}]\ , \tag{2.130}$$

and

$$\bar{x}''_d \triangleq \bar{x}_d \frac{\tau''_d}{\tau''_{d0}} = \bar{x}_{ls} + \frac{\bar{x}_{md}\bar{x}'_{lfd}\bar{x}'_{lkd}}{\bar{x}_{md}\bar{x}_{ls} + \bar{x}_{md}\bar{x}'_{lfd} + \bar{x}_{ls}\bar{x}'_{lfd}} \qquad [\text{pu}]\ . \tag{2.131}$$

[28] A brief description of operational impedances and associate circuits are presented in Appendix I.

[29] Primes are used to denote transient and sub-transient quantities, which can be confused with rotor quantities referred to the stator windings by a turns ratio. It is important to note, therefore, that x'_d is the only single-primed parameter that is not a referred impedance.

These reactances are the high-frequency asymptotes of the operational impedances, which means that the high-frequency response of the machine is characterized by them. That is

$$\lim_{p\to\infty} \bar{x}_q(p) = \bar{x}_q'' \qquad [\text{pu}] \tag{2.132}$$

and

$$\lim_{p\to\infty} \bar{x}_d(p) = \bar{x}_d'' \qquad [\text{pu}]\ . \tag{2.133}$$

It is interesting that $\lim_{p\to\infty} G(p)$ is zero, which indicates that the stator flux linkages are essentially insensitive to high-frequency changes in field voltage [19].

A final and important comment about the operational impedances, the transient and sub-transient reactances is given by Krause et al. [19]:

> Although the steady-state and sub-transient reactances can be related to the operational impedances, this is not the case with the transient reactances. It appears that the d-axis transient reactance evolved from the development [found in] [58] of approximate transient torque-angle characteristic where the effects of d-axis damper windings are neglected. The q-axis transient reactance has come into use when it became desirable to portray more accurately the dynamic characteristics of the solid iron rotor machine in transient stability studies [therefore not considered in this work]. [...] It is perhaps apparent that the sub-transient reactances characterize the equivalent reactances of the machine during a very short period of time following an electrical disturbance. After a period, of perhaps a few milliseconds, the machine equivalent reactances approach the values of the transient reactances, and even though they are not directly related to $\bar{x}_q(p)$ and $\bar{x}_d(p)$, their values lie between the sub-transient and steady-state values. As more time elapses after a disturbance, the transient reactances give way to the steady-state reactances.

Moreover, the use of transient and sub-transient quantities to portray the behavior of the machine over specific time intervals was a direct result of the need to simplify the machine equations so that pre-computer computational techniques could be used.

2.9.1 Summary

Table 2.2 presents a list of the fundamental constants. All are per-unit values; ω_b is the base electrical angular velocity, in electrical radians per second, used to calculate the inductive reactances. Against each quantity is given the name by which it is known in the usually accepted terminology.

The machine time constants are presented in Table 2.3.

All derived reactances are presented in Table 2.4.

Table 2.2: Fundamental salient-pole synchronous machine constants.

Constant	Meaning
$\bar{r}_s$	armature resistance
$\bar{r}'_{kq}$	quadrature-axis damper resistance
$\bar{r}'_{fd}$	field resistance
$\bar{r}'_{kd}$	direct-axis damper resistance
$x_{mq} = \omega_b L_{mq}$	quadrature-axis magnetizing reactance
$x_{md} = \omega_b L_{md}$	direct-axis magnetizing reactance
$x_{ls} = \omega_b L_{ls}$	armature leakage reactance
$x'_{lkq} = \omega_b L'_{lkq}$	quadrature-axis damper leakage reactance
$x'_{lfd} = \omega_b L'_{lfd}$	field leakage reactance
$x'_{lkd} = \omega_b L'_{lkd}$	direct-axis damper leakage reactance

Table 2.3: Salient-pole synchronous machine time constants.

Time constant	Meaning
τ'_{d0}	direct-axis transient open-circuit time constant
τ'_{d}	direct-axis transient short-circuit time constant
τ''_{d0}	direct-axis sub-transient open-circuit time constant
τ''_{d}	direct-axis sub-transient short-circuit time constant
τ''_{q0}	quadrature-axis sub-transient open-circuit time constant
τ''_{q}	quadrature-axis sub-transient short-circuit time constant
τ_{kd}	direct-axis damper leakage time constant

Table 2.4: Salient-pole synchronous machine derived reactances.

Derived reactance	Meaning
$x_d = x_{ls} + x_{md}$	direct-axis synchronous reactance
x'_d	direct-axis transient reactance
x''_d	direct-axis sub-transient reactance
$x_q = x_{ls} + x_{mq}$	quadrature-axis synchronous reactance
x''_q	quadrature-axis sub-transient reactance

2.10 The load rejection test

Tests for transient and sub-transient parameters involve sudden changes to any, or all, of the three-phase circuits at, or electrically near, the machine armature terminals. Sudden changes to the field electrical current are also included. Changes at, or near, the armature terminals could result from single or multiple faults between phases, or faults from one or more phases to the machine neutral, or active and reactive load rejections [13].

As described in the IEEE [13]:

> The characteristic values of transient and sub-transient reactances (and time constants) of synchronous machines have been used for about [89] years and for many purposes. Initially, such reactances and time constants were calculated to give, both machine designers and users of synchronous machines, first-hand knowledge of short-circuit currents magnitudes and their rate of change or decay [13, p. 117].

Original analysis of short-circuit currents by machine designers [10, 80], commencing about 89 years ago, indicated that there are basically two periods during which the rates of current decay may be easily identified. The initial and shortest period is named the sub-transient regime. The subsequent and much longer period is called the transient regime. Such regimes can be associated with a time constant. This characteristic value can be identified as the time taken for exponentially decaying current or voltage to change to $1/e$, or 0.368, of its original value [13].

Unlike the short-circuit test, the load rejection presents favorable characteristics, since does not depend on special equipment and does not produce hazardous forces in the machine [66]. The load rejection test consists in opening the generator main breaker while the generator is initially carrying some reactive and/or active power. It is a particular type of decrement test, while the field voltage is kept constant, for determining generator parameters.

To obtain direct-axis parameters, the appropriate loads are purely reactive and, to obtain quadrature-axis parameters, the appropriate loads are the ones under which there is alignment between the armature current and the quadrature-axis. From the test results, it is possible to estimate synchronous, transient, and sub-transient reactances, as well as the time constants for both axes [18].

To reach the pure quadrature-axis armature current, it is possible to use a trial and error procedure, as proposed by de Mello and Ribeiro [65]. That procedure is time consuming and, for that reason, disadvantageous for power plants operators. Certain alternative procedures have been developed to replace pure quadrature-axis load rejection. To apply those methods, it is necessary to either measure or estimate the load angle, δ, or to use more advanced parameter estimation procedures. Another possibility, developed by Giesbrecht [81] for salient-pole machines and extended by Giesbrecht and Meneses [18], is to use analytical methods to estimate the steady-state condition under which the armature current is on the quadrature-axis. Those methods are based only on estimates of the machine synchronous reactances and measurements of its armature voltage and are robust to errors in initial estimates.

While the intention of this section is to develop graphical and analytical conditions to read the test data, further review on the literature is presented in Chapter 5.

2.10.1 Preliminary considerations

The operational description presented in Appendix I is a common approach to analyze the electrical responses of a synchronous machine to perturbations [13]. The transient regime after a load rejection is studied by evaluating the flux-linkage per second variations in both quadrature- and direct-axis.

Recalling the $qd0$-voltages equations, as defined by Park [43],

$$\bar{v}_q = -\bar{r}_s\bar{i}_q + \overline{\omega}_r\overline{\Psi}_d + \frac{p}{\omega_b}\overline{\Psi}_q \qquad [\text{pu}]\ , \tag{2.134a}$$

$$\bar{v}_d = -\bar{r}_s\bar{i}_d - \overline{\omega}_r\overline{\Psi}_q + \frac{p}{\omega_b}\overline{\Psi}_d \qquad [\text{pu}]\ , \tag{2.134b}$$

$$\bar{v}_0 = -\bar{r}_s\bar{i}_0 + \frac{p}{\omega_b}\overline{\Psi}_0 \qquad [\text{pu}]\ , \tag{2.134c}$$

where

$$\overline{\Psi}_q = -\bar{x}_q(p)\bar{i}_q \qquad [\text{pu}]\ , \tag{2.135a}$$

$$\overline{\Psi}_d = -\bar{x}_d(p)\bar{i}_d + G(p)\bar{v}'_{fd} \qquad [\text{pu}]\ , \tag{2.135b}$$

$$\overline{\Psi}_0 = -\bar{x}_{ls}\bar{i}_0 \qquad [\text{pu}]\ . \tag{2.135c}$$

From Park's inverse transform (2.66):

$$\bar{v}_a(t) = \bar{v}_q\cos\theta_r + \bar{v}_d\sin\theta_r + \bar{v}_0 \qquad [\text{pu}]\ . \tag{2.136}$$

2.10.1.1 Before the load rejection

Prior to the load rejection itself, the salient-pole synchronous machine is operating at steady-state – which was already analyzed in Section 2.8. Being extremely small, the voltage drops across the armature resistance, $\bar{r}_s\bar{i}_q$ and $\bar{r}_s\bar{i}_d$, may be neglected without much affecting desired results.

Furthermore, as, in steady-state, $\overline{\omega}_s = 1$ pu, (2.134) may be re-written simply as

$$\bar{v}_q^0 = \overline{\Psi}_d^0 \qquad [\text{pu}]\ , \tag{2.137a}$$

$$\bar{v}_d^0 = -\overline{\Psi}_q^0 \qquad [\text{pu}]\ ,$$

where $\bar{v}_q^0$ and $\bar{v}_d^0$ are, respectively, the quadrature-axis and direct-axis steady-state armature voltages; and $\overline{\Psi}_q^0$ and $\overline{\Psi}_d^0$ are, respectively, the quadrature-axis and direct-axis steady-state flux linkage per second. These quantities are taken right before the load rejection.

2.10.1.2 Following the load rejection

Let the flux-linkage per second variations be defined as:

$$\Delta\bar{\Psi}_q(p) = -\bar{x}_q(p)\Delta\bar{i}_q(p) \qquad [\text{pu}] \; , \tag{2.138a}$$

$$\Delta\bar{\Psi}_d(p) = -\bar{x}_d(p)\Delta\bar{i}_d(p) + G(p)\Delta\bar{v}'_{fd}(p) \qquad [\text{pu}] \; . \tag{2.138b}$$

Above equations may be initially manipulated by considering that:

1. If the field voltage is held fixed at its pre-rejection value, then the Laplace transform of the change in the field voltage is zero;
2. Following the load rejection, both quadrature- and direct-axis currents become zero. Therefore, they may be modeled as decreasing steps:

$$\Delta\bar{i}_q(p) = \mathcal{L}\left\{-\bar{i}_q^0\theta^H(t)\right\} = -\frac{\bar{i}_q^0}{p} \, , \tag{2.139a}$$

$$\Delta\bar{i}_d(p) = \mathcal{L}\left\{-\bar{i}_d^0\theta^H(t)\right\} = -\frac{\bar{i}_d^0}{p} \, , \tag{2.139b}$$

where $\bar{i}_q^0$ and $\bar{i}_d^0$ are the quadrature-axis and direct-axis steady-state currents before the load rejection, respectively; $\mathcal{L}\{\cdot\}$ is the Laplace transform operator; and $\theta^H(t)$ is the Heaviside step function.

Thus, (2.138) becomes:

$$\Delta\bar{\Psi}_q(p) = \bar{x}_q(p)\frac{\bar{i}_q^0}{p} \qquad [\text{pu}] \; , \tag{2.140a}$$

$$\Delta\bar{\Psi}_d(p) = \bar{x}_d(p)\frac{\bar{i}_d^0}{p} \qquad [\text{pu}] \; . \tag{2.140b}$$

Considering (2.140), (8.13), (8.14), (8.6), and (8.7), the quadrature-axis flux linkage per second transient in the time domain is given by

$$\Delta\bar{\Psi}_q(t) = \mathcal{L}^{-1}\left\{\bar{x}_q(p)\frac{\bar{i}_q^0}{p}\right\} = \mathcal{L}^{-1}\left\{\bar{x}_q\frac{(1+\tau''_q p)}{(1+\tau''_{q0}p)}\frac{\bar{i}_q^0}{p}\right\} \tag{2.141a}$$

$$= \bar{x}_q\bar{i}_q^0 + \bar{x}_q\frac{\tau''_q}{\tau''_{q0}}\bar{i}_q^0\exp\left\{-\frac{t}{\tau''_{q0}}\right\} - \bar{x}_q\bar{i}_q^0\exp\left\{-\frac{t}{\tau''_{q0}}\right\} \tag{2.141b}$$

$$= \bar{x}_q\bar{i}_q^0 + \left(\bar{x}''_q - \bar{x}_q\right)\bar{i}_q^0\exp\left\{-\frac{t}{\tau''_{q0}}\right\} \qquad [\text{pu}] \; , \tag{2.141c}$$

and the direct-axis flux linkage per second transient in the time domain, by

$$\Delta\bar{\Psi}_d(t) = \mathcal{L}^{-1}\left\{\bar{x}_d(p)\frac{\bar{i}_d^0}{p}\right\} = \mathcal{L}^{-1}\left\{\bar{x}_d\frac{(1+\tau'_d p)(1+\tau''_d p)}{(1+\tau'_{d0}p)(1+\tau''_{d0}p)}\frac{\bar{i}_d^0}{p}\right\} \tag{2.142a}$$

$$= \bar{x}_d\bar{i}_d^0 + \left(\bar{x}'_d - \bar{x}_d\right)\bar{i}_d^0\exp\left\{-\frac{t}{\tau'_{d0}}\right\} + \left(\bar{x}''_d - \bar{x}'_d\right)\bar{i}_d^0\exp\left\{-\frac{t}{\tau''_{d0}}\right\} \quad [\text{pu}] \; . \tag{2.142b}$$

2.10.1.3 Complete behavior

When the load is suddenly changed, the transformer effects are represented by $p\overline{\Psi}_q$ and $p\overline{\Psi}_d$ and produce nonperiodic and second-harmonic components in the armature quantities. These phenomena can be neglected without including large errors [55].

The synchronous machine behavior throughout the load rejection process can be obtained by composing the phenomena before and following the load rejection itself. Therefore, the quadrature-axis flux-linkage per second in the time-domain is:

$$\overline{\Psi}_q(t) = \overline{\Psi}_q^0 + \Delta\overline{\Psi}_q(t) \tag{2.143a}$$

$$= -\bar{v}_d^0 + \bar{x}_q \bar{i}_q^0 + \left(\bar{x}_q'' - \bar{x}_q\right) \bar{i}_q^0 \exp\left\{-\frac{t}{\tau_{q0}''}\right\} \qquad [\text{pu}] \ , \tag{2.143b}$$

and the direct-axis flux linkage per second is:

$$\overline{\Psi}_d(t) = \overline{\Psi}_d^0 + \Delta\overline{\Psi}_d(t) \tag{2.144a}$$

$$= \bar{v}_q^0 + \bar{x}_d \bar{i}_d^0 + \left(\bar{x}_d' - \bar{x}_d\right) \bar{i}_d^0 \exp\left\{-\frac{t}{\tau_{d0}'}\right\} + \left(\bar{x}_d'' - \bar{x}_d'\right) \bar{i}_d^0 \exp\left\{-\frac{t}{\tau_{d0}''}\right\} \quad [\text{pu}] \ . \tag{2.144b}$$

2.10.2 Direct-axis load rejection

The direct-axis load rejection test is performed when the generator is at zero active power flow, but drawing or supplying reactive power. In such load condition, flux exists only in the direct-axis (i.e., $\overline{\Psi}_q = 0 \to \bar{i}_q = 0$) [65]. For balanced load conditions, $\overline{\Psi}_0 = 0 \to \bar{i}_0 = 0$. Thus, let the power-factor angle be $\phi = \pi/2$ rad.

Since the armature voltage is:

$$\bar{v} = \sqrt{\bar{v}_d^2 + \bar{v}_q^2} = \bar{v}_q \qquad [\text{pu}] \ , \tag{2.145}$$

there is no angular displacement between $\bar{v}$ and the quadrature-axis, which implies that the load angle $\delta = 0$. From (2.136),

$$\bar{v}_a(t) = \bar{v}_q \cos\theta_r \qquad [\text{pu}] \ . \tag{2.146}$$

Therefore,

$$\bar{v}_q(t) = \overline{\Psi}_d(t) \tag{2.147a}$$

$$= \bar{v}_q^0 + \bar{x}_d \bar{i}_d^0 + \left(\bar{x}_d' - \bar{x}_d\right) \bar{i}_d^0 \exp\left\{-\frac{t}{\tau_{d0}'}\right\} + \left(\bar{x}_d'' - \bar{x}_d'\right) \bar{i}_d^0 \exp\left\{-\frac{t}{\tau_{d0}''}\right\} \quad [\text{pu}] \ . \tag{2.147b}$$

At the initial condition, prior to the load rejection, the synchronous machine phasor diagram is presented in Figure 2.14; it is possible to visualize that all the flux is on the direct-axis.

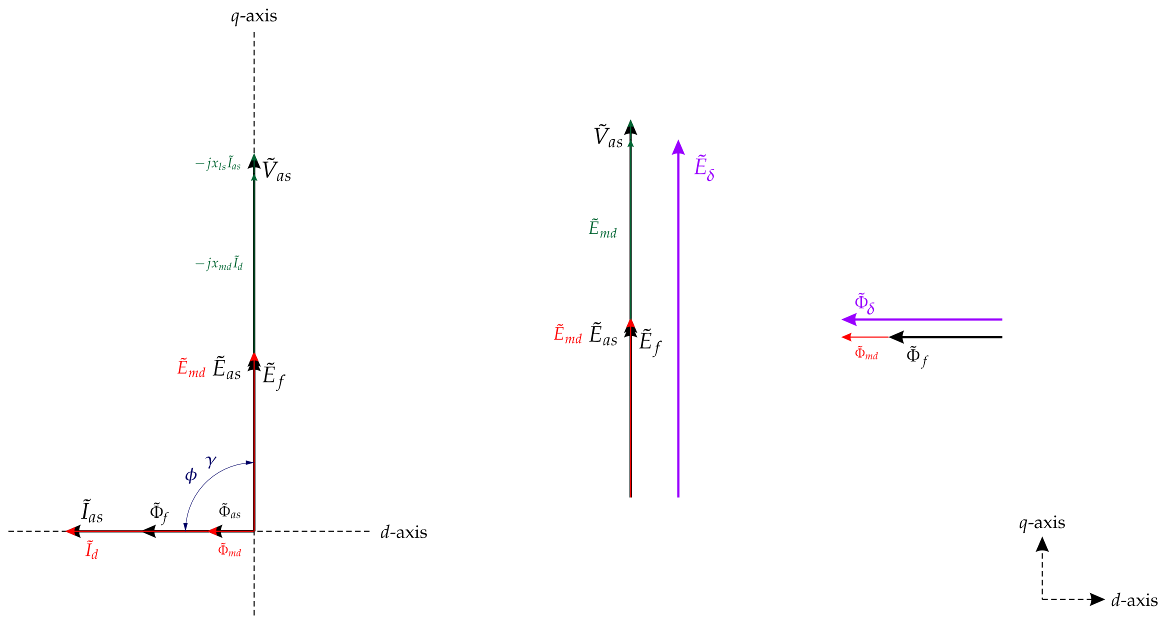

Figure 2.14: Phasor diagram of a three-phase salient-pole synchronous machine when armature magnetic-flux is exclusively on the direct-axis.

2.10.3 Quadrature-axis load rejection

In order to perform the quadrature-axis load rejection test, the generator armature flux must be only on the quadrature-axis. Therefore, the angle $\gamma = 0$, which means that $\phi = -\delta$, or, similarly, $|\phi| = |\delta|$ [81].

As in the direct-axis load rejection, it is desirable to keep the machine underexcited to avoid dangerous voltage levels. Such condition is reached for capacitive loads – i.e., when the armature current phasor leads the armature voltage phasor.

At the initial condition, prior to the load rejection, the synchronous machine phasor diagram is presented in Figure 2.15; it is possible to visualize that all the armature flux is on the quadrature-axis.

From Figure 2.15, the armature voltage has components on both quadrature and direct axes. After the quadrature-axis load rejection, there is no armature-voltage component on the direct-axis, as shown in Figure 2.16. It can also be seen that the armature-voltage quadrature-axis component does not vary from before the load rejection to after it. Therefore, the influence of the quadrature-axis circuit can be analyzed by means of the armature-voltage direct-axis component.

The armature voltage is:

$$\bar{v} = \sqrt{\bar{v}_d^2 + \bar{v}_q^2} \qquad [\mathrm{pu}] \ , \tag{2.148}$$

and, from (2.136),

$$\bar{v}_a(t) = \bar{v}_q \cos\theta_r + \bar{v}_d \sin\theta_r \qquad [\mathrm{pu}] \ . \tag{2.149}$$

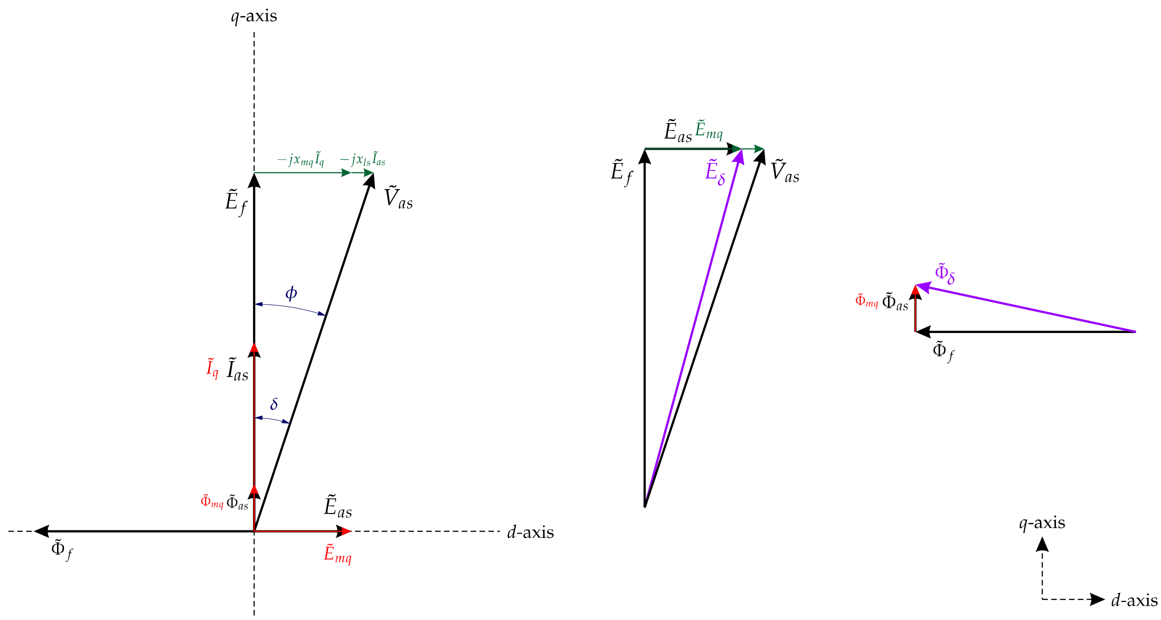

Figure 2.15: Phasor diagram of a three-phase salient-pole synchronous machine when armature magnetic-flux is exclusively on the quadrature-axis.

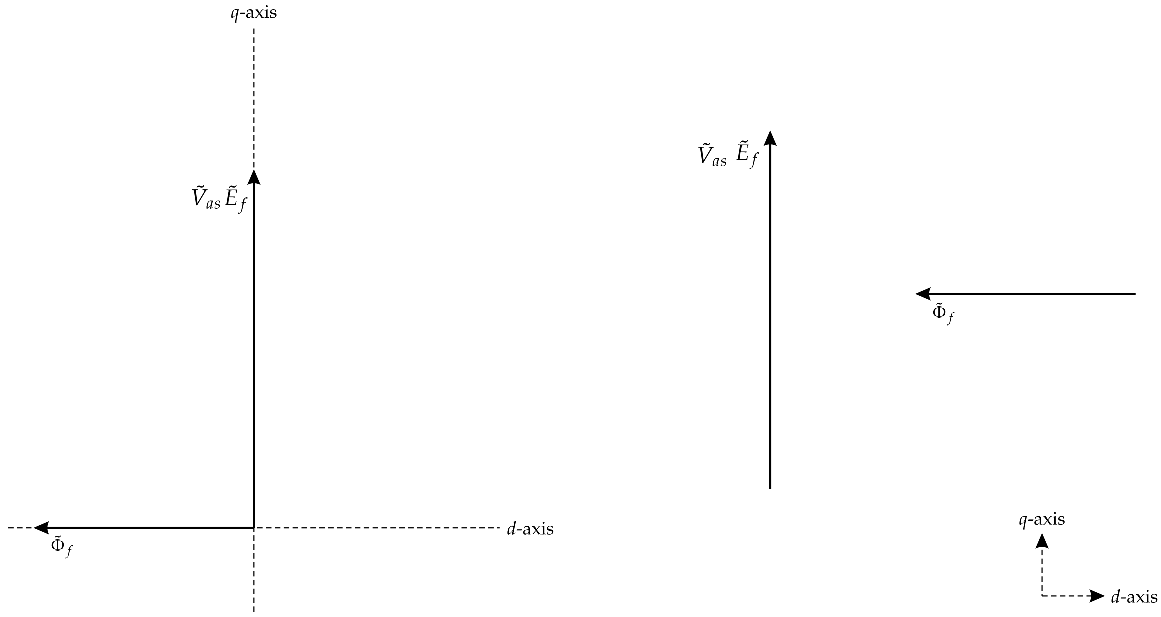

Figure 2.16: Phasor diagram of a three-phase salient-pole synchronous machine after the quadrature-axis load rejection.

The armature voltage direct-axis component is

$$\bar{v}_d(t) = -\overline{\Psi}_q(t) \tag{2.150a}$$

$$= \bar{v}_d^0 - \bar{x}_q \bar{i}_q^0 + \left(\bar{x}_q - \bar{x}_q''\right) \bar{i}_q^0 \exp\left\{-\frac{t}{\tau_{q0}''}\right\} \qquad [\mathrm{pu}]\ . \tag{2.150b}$$

Chapter 3

Concepts on System Identification and System Theory and their Applications to Synchronous Machines

"It is difficult to do justice to a subject as complex as system theory in a compass of a few printed pages. [...] I believe that system theory is here to stay, and that the coming years will witness its evolution into a respectable and active area of scientific endeavor."

— Lofti A. Zadeh[1], *From Circuit Theory to System Theory* [82]

The salient-pole synchronous machine equations, in both machine and rotor reference-frame variables, have already been derived. This chapter aims at adapting the machine equations into the state-space representation, which is a very useful tool for states and parameters estimation.

The majority of concepts involved in Section 3.1 are borrowed from the fundamental work of Sarachik [83]. Furthermore, the doctoral thesis of Barreto [84] is a great and important compiled of concepts and works on System Identification, Data Modeling, and Time Series; one should refer to it for additional details on these subjects.

3.1 Preliminary concepts

The first definition to be presented comprehends the concept of physical systems.

Definition 3.1: Physical system

*A **physical system** is an interconnection of physical components that perform a specific function.*

[1] Lofti Aliasker Zadeh (1921–2017) was a mathematician, computer scientist, electrical engineer, artificial intelligence researcher, and professor emeritus of computer science at the University of California, Berkeley. Zadeh is best known for proposing fuzzy mathematics and for pioneering the development of the Z-Transform method in discrete time signal processing and analysis.

These components may be electrical, mechanical, hydraulic, thermal, and so forth.

A great variety of physical quantities may be associated with every system. Some of the signals can be directly changed with time in order to indirectly accomplish desired changes in some other signals of the systems, which happen to be of particular interest. The former set of signals is called set of *inputs* or *excitations*; the latter is called set of *outputs*, *measurements*, or *responses*.

The set of inputs will be denoted by an m-dimensional vector $\mathbf{u} \in \mathbb{R}^m$; and the set of measurements by an l-dimensional vector $\mathbf{y} \in \mathbb{R}^l$. It is an essential part of the system concept that changes of the input are followed by changes to the output [83]. A mathematical relation between them may be written as

$$\mathbf{y}(t) = \mathcal{S}\left\{\mathbf{u}(t)\right\} , \tag{3.1}$$

where $\mathcal{S}$ is an operator if the mapping of inputs $\mathbf{u}(t)$ into outputs $\mathbf{y}(t)$ is unique [85] .

The term *system*, as applied to general analysis, was originated as a recognition that meaningful investigation of a particular phenomenon can often only be achieved by explicitly accounting for its environment [86]. Accordingly, mathematical models of systems are likely to involve a large number of interrelated variables – and this is emphasized by describing such situations as *multivariable systems*.

Definition 3.2: Physical realizability

A system $\mathcal{S}$ with the input-output relation (3.1) *is called* ***physically realizable*** *if a physical system can be built whose inputs and outputs are related via* (3.1).

Definition 3.3: Causal system

A system $\mathcal{S}$ is called ***causal*** *(or is said to be* ***nonantecipative****) if for any t, $\mathbf{y}(t)$ does not depend on any $\mathbf{u}(t')$ for $t' > t$ (i.e., if $\mathbf{y}(t)$ does not depend on future values of $\mathbf{u}$). Otherwise, it is called* ***noncausal*** *or* ***antecipative****.*

From Definition 3.3, a noncausal system is not realizable.

Definition 3.4: Dynamic system

A system $\mathcal{S}$ is called ***dynamic*** *(or is said to* ***have memory****) if $\mathbf{y}(t_0)$ depends on some values of $\mathbf{u}(t)$ for $t \neq t_0$. A system for which $\mathbf{y}(t_0)$ does not depend on $\mathbf{u}(t)$ for $t \neq t_0$ is called* ***instantaneous*** *(or is said to* ***have zero memory****).*

As stated by Luenberger [86],

> the term *dynamic* refers to phenomena that produce time-changing patterns, the characteristics of the pattern at one time being interrelated with those at other times. The term is nearly synonymous with time-evolution or pattern

of change. It refers to the unfolding of events in a continuing evolutionary process [86, p. 1].

Definition 3.5: State of a system

*The **state** of a nonantecipative dynamic system at time t_0 is the smallest set of quantities $\mathbf{x}(t_0)$, which summarize all information about $\mathbf{u}(-\infty, t_0)$ needed to determine $\mathbf{y}[t_0, t_1]$ when $\mathbf{u}[t_0, t_1]$ is known.*

A little digression on notation: $\mathbf{u}[t_0, t_1] \triangleq \mathbf{u}(t)$ for all $t_0 \leq t \leq t_1$ (i.e., the entire time function defined over $[t_0, t_1]$). The same notation is valid for other quantities.

The input-output relation of a nonantecipative dynamic system of (3.1) can be modified to include the state as follows:

$$\mathbf{y}(t) = \mathcal{C}\left\{\mathbf{x}(t_0); \mathbf{u}[t_0, t]\right\} \quad \text{for } t \geq t_0 \,, \tag{3.2}$$

and is now called the *input-output-state relation* or the *measurement equation* of the system.

Inherent in the concept of state is the requirement that for any $t_1 \in [t_0, t]$, it must be possible to define $\mathbf{x}(t_1)$ such that the state itself at t_1 must be uniquely determined by an earlier state at t_0 and the input $\mathbf{u}[t_0, t_1]$. It implies that for consistency, a condition of the form:

$$\mathbf{x}(t_1) = \mathcal{A}\left\{\mathbf{x}(t_0); \mathbf{u}[t_0, t_1]\right\} \quad \text{for } t \geq t_0 \,, \tag{3.3}$$

must be satisfied. Equation (3.3) is called the *state transition equation* of the system.

In general, the state equations are differential equations of the form:

$$\dot{\mathbf{x}}(t) = \mathbf{f}\left(\mathbf{x}(t), \mathbf{u}(t), t\right) \,, \tag{3.4a}$$

$$\mathbf{y}(t) = \mathbf{h}\left(\mathbf{x}(t), \mathbf{u}(t), t\right) \,, \tag{3.4b}$$

or difference equations of the form:

$$\mathbf{x}(k+1) = \mathbf{f}\left(\mathbf{x}(k), \mathbf{u}(k), k\right) \,, \tag{3.5a}$$

$$\mathbf{y}(k) = \mathbf{h}\left(\mathbf{x}(k), \mathbf{u}(k), k\right) \,. \tag{3.5b}$$

These are called the *state (differential or difference) equations*.

Definition 3.6: Zero state

*The **zero state** θ of a dynamic system is the state for which $\mathbf{y}[t_0, \infty] = \mathbf{0}$ when $\mathbf{x}(t_0)$ and the input $\mathbf{y}[t_0, \infty] = \mathbf{0}$ (i.e., $\mathcal{C}\left\{\theta; 0\right\} = 0$).*

Definition 3.7: Homogeneity and additivity

*A zero-memory system is **homogeneous** if*

$$\mathcal{S}\left\{k\mathbf{u}\right\} = k\mathcal{S}\left\{\mathbf{u}\right\} \quad \textit{for all } k \in \mathbb{R} \textit{ and } \mathbf{u} \in \mathbb{C}^m\,. \tag{3.6}$$

*A zero-memory system is **additive** if*

$$\mathcal{S}\left\{\mathbf{u}_1 + \mathbf{u}_2\right\} = \mathcal{S}\left\{\mathbf{u}_1\right\} + \mathcal{S}\left\{\mathbf{u}_2\right\} \quad \textit{for any } \mathbf{u}_1 \textit{ and } \mathbf{u}_2 \in \mathbb{C}^m\,. \tag{3.7}$$

Remark.

1. When homogeneous or additive is used in referring to dynamic system, it is implied that the term refers to zero-state response;
2. (3.7) implies (3.6) for k rational.

This is the entire definition of linearity for zero-memory systems. However, if a system is dynamic, the concept of linearity is a bit more complex [83].

Definition 3.8: Zero-state linear

*A dynamic system is called **zero-state linear** if* $\mathcal{C}\left\{\theta; k_1\mathbf{u}_1 + k_2\mathbf{u}_2\right\} = k_1\mathcal{C}\left\{\theta; \mathbf{u}_1\right\} + k_2\mathcal{C}\left\{\theta; \mathbf{u}_2\right\}$ *for all* $k_1, k_2 \in \mathbb{R}$ *and any* $\mathbf{u}_1, \mathbf{u}_2 \in \mathbb{C}^m$ *(i.e., if it is zero-state additive and zero-state homogeneous).*

Definition 3.9: Decomposition property

*A dynamic system is said to have the **decomposition property** if* $\mathcal{C}\left\{\mathbf{x}(t_0); \mathbf{u}\right\} = \mathcal{C}\left\{\mathbf{x}(t_0); \mathbf{0}\right\} + \mathcal{C}\left\{\theta; \mathbf{u}\right\}$ *for all* $\mathbf{x}(t_0)$ *and all* $\mathbf{u}$.

Definition 3.10: Zero-input linear

*A dynamic system is called **zero-input linear** if it is zero-input homogeneous and additive.*

Definition 3.11: Linear dynamic system

*A dynamic system is called **linear** if:*

1. *it is zero-state linear;*
2. *it has the decomposition property;*
3. *it is zero-input linear.*

*A system that is not linear is called **nonlinear**.*

The systems from (3.4) and (3.5) are sometimes called lumped systems to distinguish them from distributed systems which are describable by partial differential equations or differential-difference equations and whose states can only be expressed as functions rather than by finite dimensional vectors [83].

Definition 3.12: Order of a system

*The number of components of the state vector $\mathbf{x}(t)$ is called the **order of the system**, and it is designated by letter n.*

In addition to the state vector, the input $\mathbf{u}(t)$ and output $\mathbf{y}(t)$ have m and l components, respectively.

The salient-pole synchronous machine is a physically realizable, nonantecipative, dynamic, nonlinear system. Further details will be provided throughout this chapter.

3.2 Observability

The analysis of the interaction between input and state, on one hand, and between state and output, on the other hand, has proved of fundamental importance in understanding the possibility of solving a large number of relevant control problems, including eigenvalues assignment via feedback, minimization of quadratic cost criteria, disturbance rejection, asymptotic output regulation, etc. Key tools for the analysis of such interactions are the notions of reachability and observability and the corresponding decomposition of the control system into reachable/unreachable and, respectively, observable/unobservable parts [87].

Perhaps the most important definition within systems theory is the concept of observability. It came from answering the following questions: How much information about the state of the system is contained in the data? Can the state be determined from the data?

Intuitively, it would seem that the answers to such questions are related to the system model itself, and indeed this is so. The importance of such questions is obvious. If little is to be gained from filtering, then one should consider remodeling the system. This might involve taking additional or alternate measurements or redesigning the dynamics of the system [88].

Such concept is due to Kalman[2] [89, 90]. Kalman [89] introduced the notion of observability, but as a mere dual of controllability. Contemporaneously, Kalman [90] provided an alternative, more satisfactory definition, where observability is defined in a more intrinsic way in terms of the possibility of deducing the state trajectory from input/output measurements.

[2]Rudolf Emil Kálmán (1930–2016) was an Hungarian-American electrical engineer, mathematician, and inventor. He was most noted for his co-invention and development of the Kalman filter, a mathematical algorithm that is widely used in signal processing, control systems, and guidance, navigation and control. For this work, U.S. President Barack Obama awarded Kálmán the National Medal of Science on October 7, 2009.

Let $\boldsymbol{X}^{\star}$ be the dual vector space of the state space $\boldsymbol{X}$, i.e., the space of all linear functions of $\boldsymbol{x}$. An element $\boldsymbol{z}^{\star}$ of $\boldsymbol{X}^{\star}$ is called a costate.

Definition 3.13: Observability

*A costate $\boldsymbol{z}^{\star}$ of a plant (linear, stationary, discrete-time, free, and single output) is said to be **observable** if its exact value $[\boldsymbol{z}^{\star}, \boldsymbol{x}]$ at any state $\boldsymbol{x}$ at time 0 can be determined from measurements of the output signal over the finite interval $0 \geq t \geq t_2$. The time t_2 will depend, in general, on $\boldsymbol{z}^{\star}$. If every costate is observable, the plant is said to be **completely observable**.*

How well the state is known is measured by the estimation error covariance matrix. But it depends on its initial condition (the initial data) and does not reflect the uncertainty in the estimate by virtue of filtering the data alone.

Although the concepts and properties of observability and controllability are completely defined for linear systems [83, 84, 88–90], how does Definition 3.13 apply to nonlinear systems? In a similar manner, it must express that there is indeed a possibility that the purpose of an observer can be achieved.

For nonlinear systems, local observability conditions are reported in the work of Lee and Markus [91]. Furthermore, necessary conditions and a sufficient condition for observability have been proven by Griffith and Kumar [92].

Krener [93, 94] and Sussmann and Jurdjevic [95] developed the nonlinear analog of linear observability in terms of the Lie algebra[3] of vector fields on the state space generated by vector fields. It was shown that if the dimension of the Lie algebra is constant, or if the system is analytic, then there exist a unique maximal submanifold $\boldsymbol{X}^{\star}$ of $\boldsymbol{X}$ through $\boldsymbol{x}_0$ which carries all the trajectories of the system passing through $\boldsymbol{x}_0$ such that any point of this submanifold can be reached from $\boldsymbol{x}_0$ going forward and backward along the trajectories of the system [96].

Even though it would be a great contribution to dedicate a few pages of this work to prove and to analytically demonstrate the observability conditions to the salient-pole synchronous machine model, due to its complexity and the time-demand required to do so, it will be left for further works.

However, a more superficial approach can still be considered. By means of linearization, it is possible to apply the so-called *Rank Condition Test* to locally analyze observability.

[3] In mathematics, a Lie algebra is a vector space together with a non-associative operation called the Lie bracket, an alternating bilinear map, satisfying the Jacobi identity. Lie algebras were introduced to study the concept of infinitesimal transformations by Marius Sophus Lie in the 1870s, and independently discovered by Wilhelm Killing in the 1880s. The name Lie algebra was given by Hermann Weyl in the 1930s; in older texts, the term infinitesimal group is used.

3.2.1 Rank condition test

Let a general linear invariant-time system be

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{B}\boldsymbol{u}(t)\,, \tag{3.8a}$$

$$\boldsymbol{y}(t) = \boldsymbol{C}\mathbf{x}(t)\,, \tag{3.8b}$$

where $\boldsymbol{x} \in \mathbb{R}^n$ is the state vector, $\boldsymbol{y} \in \mathbb{R}^l$ is the measurement vector, $\boldsymbol{u} \in \mathbb{R}^m$ is the input vector, and $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$ are constant matrices known respectively as dynamics matrix, the input or control matrix, and the output or measurement matrix.

One way of testing whether the system (3.8) is observable is to define the observability matrix:

$$\mathcal{O} \triangleq \begin{bmatrix} \boldsymbol{C} \\ \boldsymbol{C}\boldsymbol{A} \\ \boldsymbol{C}\boldsymbol{A}^2 \\ \vdots \\ \boldsymbol{C}\boldsymbol{A}^{n-1} \end{bmatrix} \in \mathbb{R}^{nl \times n}\,. \tag{3.9}$$

Therefore, the system (3.8) is observable if matrix $\mathcal{O}$ has a full column rank, that is, if $\texttt{rank}\ \mathcal{O} = n$. This is known as Kalman's rank condition for observability and, according to it, a pair $\{\boldsymbol{A}, \boldsymbol{C}\}$ is either observable or not.

The concept was extended to nonlinear systems in the 1970s, e.g. Hermann and Krener [96]. The pair $\{\boldsymbol{f}, \boldsymbol{h}\}$ is said to be observable if $\texttt{rank}\ \mathcal{O}(\boldsymbol{x}) = n$, $\forall \boldsymbol{x} \in \mathbb{R}^n$, which is the counterpart of Kalman's rank condition for linear systems [97].

The following procedure is presented by Nahar, Liu, and Shah [98]. Hereafter, each process equation will be denoted f_i, for $i = 1, \ldots, n$ and each measurement equation, h_j, for $j = 1, \ldots, l$.

Let $\boldsymbol{F}(k)$ be the Jacobian matrix:

$$\boldsymbol{F}(k) \triangleq \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix} \in \mathbb{C}^{n \times n} \tag{3.10}$$

and $\boldsymbol{H}(k)$ be the Jacobian matrix:

$$\boldsymbol{H}(k) \triangleq \begin{bmatrix} \frac{\partial h_1}{\partial x_1} & \frac{\partial h_1}{\partial x_2} & \cdots & \frac{\partial h_1}{\partial x_n} \\ \frac{\partial h_2}{\partial x_1} & \frac{\partial h_2}{\partial x_2} & \cdots & \frac{\partial h_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial h_l}{\partial x_1} & \frac{\partial h_l}{\partial x_2} & \cdots & \frac{\partial h_l}{\partial x_n} \end{bmatrix} \in \mathbb{C}^{l \times n}\,. \tag{3.11}$$

The system is considered to be observable at time k if the observability matrix $\mathcal{O}(k)$ has a full column rank, $\texttt{rank}\ \mathcal{O}(k) = n$, where the observability matrix is given by:

$$\mathcal{O}(k) \triangleq \begin{bmatrix} \boldsymbol{H} \\ \boldsymbol{HF} \\ \boldsymbol{HF}^2 \\ \vdots \\ \boldsymbol{HF}^{n-1} \end{bmatrix} \in \mathbb{C}^{nl \times n} . \tag{3.12}$$

The rank test provides on whether the system is observable or not and does not provide any information on how strongly or weakly observable it is [98].

3.3 State-space representation

One of the major references of this work is the brilliant work of Candy [99]. It provides a clear comment on the importance of state-space representation:

> The interesting property of the state-space representation is to realize that these models represent a complete generic form for almost any physical systems. [...] Systems theory, which is essentially the study of dynamic systems, is based on the study of state-space models and is rich with theoretical results exposing the underlying properties of the dynamic system under investigation. This is one of the major reasons why state-space models are employed [...], especially when the system is multivariable having multiple inputs and multiple outputs [as in the case of salient-pole synchronous machines] [99, p. 99].

Unfortunately, in practice, complete state measurements are rarely realistic. Therefore, at least some states are unknown and, thus, the nonlinear state-space model cannot be directly applied in reality. The states have to be considered as unknown quantities and must be estimated as well. This leads to modeling approaches with internal states. They are subsumed under the class of the so-called *internal dynamics models* [100].

The main objective of this work is to investigate methods for estimating states and parameters of salient-pole synchronous machines. Therefore, it is mandatory to develop their state-space models. A set of first-order differential equations is displayed in (2.88). However, flux linkages and currents are represented as variables in it. Since these two quantities are mutually dependent, both cannot simultaneously be independent or state variables [19].

Although numerous possibilities for state variables are available [19, 22, 33, 35, 49, 50], the two most commonly applied [19] are:

i) a set based on currents as state variables,

$$\mathbf{x} = \begin{bmatrix} i_q & i_d & i_0 & i'_{kq} & i'_{fd} & i'_{kd} \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^6 , \tag{3.13}$$

which has the advantage of offering simple relations between voltages v_d and v_q and state variables (through the power network connected to the machine terminals);

ii) a set based on flux linkages – or flux linkages per second – as state variables, where the particular set to be chosen depends upon how conveniently they can be expressed in terms of the machine currents and stator voltages. For example,

$$\mathbf{x} = \begin{bmatrix} \Psi_q & \Psi_d & \Psi_0 & \Psi'_{kq} & \Psi'_{fd} & \Psi'_{kd} \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^6 . \tag{3.14}$$

Being expressed as time derivatives in (2.88), flux linkages per second could be considered to be the natural set of state variables to be solved by time-step integration. By virtue of some auxiliary equations [36], the flux linkages per second can be eliminated from Park's equations and be replaced by currents as state variables. However, flux linkages per second in a synchronous machine change slowly, limited by the open-circuit time constants, while currents change rapidly, limited by short-circuit time constants[4]. Hence, Park's equations are most efficiently solved with flux linkages per second as the state variables, in which case currents are best eliminated from these equations [36].

3.3.1 Flux-linkage per second state-space model

For salient-pole synchronous machines, the most widely used model is derived from the voltage equations expressed in the rotor reference frame with stator and rotor flux linkages per second as state variables [19]. This model was first developed by C. H. Thomas[5] [101]. At that time, he was interested in developing a block diagram of synchronous machines for analog-computer simulations, which are shown in Figure B.1 and Figure B.2. In particular, these representations apply to salient-pole machines, where saturation occurs principally along the main-pole axis [102].

The development of Thomas' flux-linkage per second state-space model is given in the following way.

Let the stator and rotor voltage equations (2.110) and flux linkages per second equations (2.112) in per unit (pu). Defining the quadrature- and direct-axis magnetizing flux linkages as

$$\overline{\Psi}_{mq} = \bar{x}_{mq} \left(\bar{i}_q + \bar{i}'_{kq} \right) \qquad [\mathrm{pu}] \, , \tag{3.15a}$$

$$\overline{\Psi}_{md} = \bar{x}_{md} \left(\bar{i}_d + \bar{i}'_{fd} + \bar{i}'_{kd} \right) \qquad [\mathrm{pu}] \, , \tag{3.15b}$$

and using them in (2.112), the winding currents can be expressed in terms of winding and magnetizing flux linkages as

$$\bar{i}_q = \frac{1}{\bar{x}_{ls}} \left(\overline{\Psi}_q - \overline{\Psi}_{mq} \right) \qquad [\mathrm{pu}] \, , \tag{3.16a}$$

---

[4] Refer to Appendix I for further details.

[5] Charles H. Thomas was an instructor of electrical engineering at Harvard University, Cambridge/Massachusetts, and former employee of Allis-Chalmers, a machinery manufacturer from Milwaukee/Wisconsin. He worked with P. C. Krause and C. Concordia throughout his life. No further information was found about him.

$$\bar{i}_d = \frac{1}{\bar{x}_{ls}} \left(\overline{\Psi}_d - \overline{\Psi}_{md} \right) \qquad \text{[pu]} \ , \tag{3.16b}$$

$$\bar{i}_0 = \frac{1}{\bar{x}_{ls}} \overline{\Psi}_0 \qquad \text{[pu]} \ , \tag{3.16c}$$

$$\bar{i}'_{kq} = \frac{1}{\bar{x}'_{lkq}} \left(\overline{\Psi}'_{kq} - \overline{\Psi}_{mq} \right) \qquad \text{[pu]} \ , \tag{3.16d}$$

$$\bar{i}'_{fd} = \frac{1}{\bar{x}'_{lfd}} \left(\overline{\Psi}'_{fd} - \overline{\Psi}_{md} \right) \qquad \text{[pu]} \ , \tag{3.16e}$$

$$\bar{i}'_{kd} = \frac{1}{\bar{x}'_{lkd}} \left(\overline{\Psi}'_{kd} - \overline{\Psi}_{md} \right) \qquad \text{[pu]} \ . \tag{3.16f}$$

Substituting (3.16) into (2.110) yields the state equations of stator and rotor windings:

$$\frac{p}{\omega_b} \overline{\Psi}_q = \bar{v}_q - \overline{\omega}_r \overline{\Psi}_d + \frac{\bar{r}_s}{\bar{x}_{ls}} \left(\overline{\Psi}_{mq} - \overline{\Psi}_q \right) \qquad \text{[pu]} \ , \tag{3.17a}$$

$$\frac{p}{\omega_b} \overline{\Psi}_d = \bar{v}_d + \overline{\omega}_r \overline{\Psi}_q + \frac{\bar{r}_s}{\bar{x}_{ls}} \left(\overline{\Psi}_{md} - \overline{\Psi}_d \right) \qquad \text{[pu]} \ , \tag{3.17b}$$

$$\frac{p}{\omega_b} \overline{\Psi}_0 = \bar{v}_0 - \frac{\bar{r}_s}{\bar{x}_{ls}} \overline{\Psi}_0 \qquad \text{[pu]} \ , \tag{3.17c}$$

$$\frac{p}{\omega_b} \overline{\Psi}'_{kq} = \bar{v}'_{kq} + \frac{\bar{r}'_{kq}}{\bar{x}'_{lkq}} \left(\overline{\Psi}_{mq} - \overline{\Psi}'_{kq} \right) \qquad \text{[pu]} \ , \tag{3.17d}$$

$$\frac{p}{\omega_b} \overline{\Psi}'_{fd} = \bar{v}'_{fd} + \frac{\bar{r}'_{fd}}{\bar{x}'_{lfd}} \left(\overline{\Psi}_{md} - \overline{\Psi}'_{fd} \right) \qquad \text{[pu]} \ , \tag{3.17e}$$

$$\frac{p}{\omega_b} \overline{\Psi}'_{kd} = \bar{v}'_{kd} + \frac{\bar{r}'_{kd}}{\bar{x}'_{lkd}} \left(\overline{\Psi}_{md} - \overline{\Psi}'_{kd} \right) \qquad \text{[pu]} \ . \tag{3.17f}$$

Before concluding, it should be mentioned that, in accordance to Lipo [36], the urge to per unitize sometimes extends to normalizing time as well. Therefore,

$$\frac{1}{\omega_b} p = \frac{1}{\omega_b} \frac{d}{dt} \tag{3.18a}$$

$$= \frac{d}{d(\omega_b t)} \tag{3.18b}$$

$$= \frac{d}{dT} \ . \tag{3.18c}$$

That being the case, for a 60 Hz base system, one pu time corresponds to approximately 1/377 second. Including the normalization factor inside the operator p, (3.17) is simply written as:

$$p \overline{\Psi}_q = \bar{v}_q - \overline{\omega}_r \overline{\Psi}_d + \frac{\bar{r}_s}{\bar{x}_{ls}} \left(\overline{\Psi}_{mq} - \overline{\Psi}_q \right) \qquad \text{[pu]} \ , \tag{3.19a}$$

$$p \overline{\Psi}_d = \bar{v}_d + \overline{\omega}_r \overline{\Psi}_q + \frac{\bar{r}_s}{\bar{x}_{ls}} \left(\overline{\Psi}_{md} - \overline{\Psi}_d \right) \qquad \text{[pu]} \ , \tag{3.19b}$$

$$p\overline{\Psi}_0 = \bar{v}_0 - \frac{\bar{r}_s}{\bar{x}_{ls}}\overline{\Psi}_0 \qquad [\text{pu}] \ , \tag{3.19c}$$

$$p\overline{\Psi}'_{kq} = \bar{v}'_{kq} + \frac{\bar{r}'_{kq}}{\bar{x}'_{lkq}}\left(\overline{\Psi}_{mq} - \overline{\Psi}'_{kq}\right) \qquad [\text{pu}] \ , \tag{3.19d}$$

$$p\overline{\Psi}'_{fd} = \bar{v}'_{fd} + \frac{\bar{r}'_{fd}}{\bar{x}'_{lfd}}\left(\overline{\Psi}_{md} - \overline{\Psi}'_{fd}\right) \qquad [\text{pu}] \ , \tag{3.19e}$$

$$p\overline{\Psi}'_{kd} = \bar{v}'_{kd} + \frac{\bar{r}'_{kd}}{\bar{x}'_{lkd}}\left(\overline{\Psi}_{md} - \overline{\Psi}'_{kd}\right) \qquad [\text{pu}] \ . \tag{3.19f}$$

It is also noted that to have a proper state model, the magnetizing flux linkages must be expressed in terms of winding flux linkages per second (which are the states). Manipulating the replacement of (3.16) into (3.15) yields

$$\overline{\Psi}_{mq} = \left(\frac{1}{\bar{x}_{mq}} + \frac{1}{\bar{x}_{ls}} + \frac{1}{\bar{x}'_{lkq}}\right)^{-1}\left(\frac{\overline{\Psi}_q}{\bar{x}_{ls}} + \frac{\overline{\Psi}'_{kq}}{\bar{x}'_{lkq}}\right) \qquad [\text{pu}] \ , \tag{3.20a}$$

$$\overline{\Psi}_{md} = \left(\frac{1}{\bar{x}_{md}} + \frac{1}{\bar{x}_{ls}} + \frac{1}{\bar{x}'_{lfd}} + \frac{1}{\bar{x}'_{lkd}}\right)^{-1}\left(\frac{\overline{\Psi}_d}{\bar{x}_{ls}} + \frac{\overline{\Psi}'_{fd}}{\bar{x}'_{lfd}} + \frac{\overline{\Psi}'_{kd}}{\bar{x}'_{lkd}}\right) \qquad [\text{pu}] \ . \tag{3.20b}$$

■

Another way of writing the flux-linkage per second model is also found in the work of Krause et al. [19]. Its development begins by isolating the flux linkages per second from (2.112):

$$\begin{bmatrix} \overline{\Psi}_q \\ \overline{\Psi}_d \\ \overline{\Psi}_0 \\ \overline{\Psi}'_{kq} \\ \overline{\Psi}'_{fd} \\ \overline{\Psi}'_{kd} \end{bmatrix} = \begin{bmatrix} \bar{x}_q & 0 & 0 & \bar{x}_{mq} & 0 & 0 \\ 0 & \bar{x}_d & 0 & 0 & \bar{x}_{md} & \bar{x}_{mq} \\ 0 & 0 & \bar{x}_{ls} & 0 & 0 & 0 \\ \bar{x}_{mq} & 0 & 0 & \bar{x}'_{kq} & 0 & 0 \\ 0 & \bar{x}_{md} & 0 & 0 & \bar{x}'_{fd} & \bar{x}_{md} \\ 0 & \bar{x}_{md} & 0 & 0 & \bar{x}_{md} & \bar{x}'_{kd} \end{bmatrix} \begin{bmatrix} \bar{i}_q \\ \bar{i}_d \\ \bar{i}_0 \\ \bar{i}'_{kq} \\ \bar{i}'_{fd} \\ \bar{i}'_{kd} \end{bmatrix} \qquad [\text{pu}] \ . \tag{3.21}$$

However, it is more convenient to write (3.21) as

$$\begin{bmatrix} \overline{\Psi}_q \\ \overline{\Psi}'_{kq} \end{bmatrix} = \begin{bmatrix} \bar{x}_q & \bar{x}_{mq} \\ \bar{x}_{mq} & \bar{x}'_{kq} \end{bmatrix} \begin{bmatrix} \bar{i}_q \\ \bar{i}'_{kq} \end{bmatrix} \qquad [\text{pu}] \ , \tag{3.22}$$

$$\begin{bmatrix} \overline{\Psi}_d \\ \overline{\Psi}'_{fd} \\ \overline{\Psi}'_{kd} \end{bmatrix} = \begin{bmatrix} \bar{x}_d & \bar{x}_{md} & \bar{x}_{mq} \\ \bar{x}_{md} & \bar{x}'_{fd} & \bar{x}_{md} \\ \bar{x}_{md} & \bar{x}_{md} & \bar{x}'_{kd} \end{bmatrix} \begin{bmatrix} \bar{i}_d \\ \bar{i}'_{fd} \\ \bar{i}'_{kd} \end{bmatrix} \qquad [\text{pu}] \ , \tag{3.23}$$

and

$$\overline{\Psi}_0 = \bar{x}_{ls}\bar{i}_0 \qquad [\text{pu}] \ . \tag{3.24}$$

Solving (3.22)–(3.24) for currents yields

$$\begin{bmatrix} \bar{i}_q \\ \bar{i}'_{kq} \end{bmatrix} = \frac{1}{\Delta_q}\begin{bmatrix} \bar{x}'_{kq} & -\bar{x}_{mq} \\ -\bar{x}_{mq} & \bar{x}_q \end{bmatrix} \qquad [\text{pu}] \ , \tag{3.25}$$

$$\begin{bmatrix} \bar{i}_d \\ \bar{i}'_{fd} \\ \bar{i}'_{kd} \end{bmatrix} = \frac{1}{\Delta_d} \begin{bmatrix} \bar{x}'_{fd}\bar{x}'_{kd} - \bar{x}^2_{md} & -\bar{x}_{md}\bar{x}'_{kd} + \bar{x}^2_{md} & -\bar{x}_{md}\bar{x}'_{fd} + \bar{x}^2_{md} \\ -\bar{x}_{md}\bar{x}'_{kd} + \bar{x}^2_{md} & \bar{x}_d\bar{x}'_{kd} - \bar{x}^2_{md} & -\bar{x}_d\bar{x}_{md} + \bar{x}^2_{md} \\ -\bar{x}_{md}\bar{x}'_{fd} + \bar{x}^2_{md} & -\bar{x}_d\bar{x}_{md} + \bar{x}^2_{md} & \bar{x}_d\bar{x}'_{fd} - \bar{x}^2_{md} \end{bmatrix} \quad [\mathrm{pu}] \,, \tag{3.26}$$

and

$$\bar{i}_0 = \frac{1}{\bar{x}_{ls}}\overline{\Psi}_0 \qquad [\mathrm{pu}] \,, \tag{3.27}$$

where

$$\Delta_q = \bar{x}_q\bar{x}'_{kq} - \bar{x}^2_{mq} \qquad [\mathrm{pu}] \,, \tag{3.28a}$$

$$\Delta_d = -\bar{x}^2_{md}\left(\bar{x}_d - 2\bar{x}_{md} + \bar{x}'_{fd} + \bar{x}'_{kd}\right) + \bar{x}_d\bar{x}'_{fd}\bar{x}'_{kd} \qquad [\mathrm{pu}] \,. \tag{3.28b}$$

The flux-linkage per second equations are obtained from substituting (3.25)–(3.27) into voltage equations (2.110). Therefore,

$$p\overline{\Psi}_q = \bar{v}_q - \overline{\omega}_r\overline{\Psi}_d - \bar{r}_s\left(\bar{x}'_{kq}\overline{\Psi}_q + \bar{x}_{mq}\overline{\Psi}'_{kq}\right) \qquad [\mathrm{pu}] \,, \tag{3.29a}$$

$$p\overline{\Psi}_d = \bar{v}_d + \overline{\omega}_r\overline{\Psi}_q - \bar{r}_s\left(\Xi_{11}\overline{\Psi}_d - \Xi_{12}\overline{\Psi}'_{fd} - \Xi_{13}\overline{\Psi}'_{kd}\right) \qquad [\mathrm{pu}] \,, \tag{3.29b}$$

$$p\overline{\Psi}_0 = \bar{v}_0 - \frac{\bar{r}_s}{\bar{x}_{ls}}\overline{\Psi}_0 \qquad [\mathrm{pu}] \,, \tag{3.29c}$$

$$p\overline{\Psi}'_{kq} = \bar{v}'_{kq} + \bar{r}'_{kq}\left(\bar{x}_{mq}\overline{\Psi}_q - \bar{x}_q\overline{\Psi}_{kq}\right) \qquad [\mathrm{pu}] \,, \tag{3.29d}$$

$$p\overline{\Psi}'_{fd} = \bar{v}'_{fd} - \bar{r}'_{fd}\left(\Xi_{21}\overline{\Psi}_d + \Xi_{22}\overline{\Psi}_{fd} + \Xi_{23}\overline{\Psi}'_{kd}\right) \qquad [\mathrm{pu}] \,, \tag{3.29e}$$

$$p\overline{\Psi}'_{kd} = \bar{v}'_{kd} - \bar{r}'_{kd}\left(\Xi_{31}\overline{\Psi}_d + \Xi_{32}\overline{\Psi}'_{fd} + \Xi_{33}\overline{\Psi}'_{kd}\right) \qquad [\mathrm{pu}] \,, \tag{3.29f}$$

where Ξ_{rc} is the element in the rth row and cth column of the 3×3 matrix in (3.26):

$$\boldsymbol{\Xi} = \begin{bmatrix} \bar{x}'_{fd}\bar{x}'_{kd} - \bar{x}^2_{md} & -\bar{x}_{md}\bar{x}'_{kd} + \bar{x}^2_{md} & -\bar{x}_{md}\bar{x}'_{fd} + \bar{x}^2_{md} \\ -\bar{x}_{md}\bar{x}'_{kd} + \bar{x}^2_{md} & \bar{x}_d\bar{x}'_{kd} - \bar{x}^2_{md} & -\bar{x}_d\bar{x}_{md} + \bar{x}^2_{md} \\ -\bar{x}_{md}\bar{x}'_{fd} + \bar{x}^2_{md} & -\bar{x}_d\bar{x}_{md} + \bar{x}^2_{md} & \bar{x}_d\bar{x}'_{fd} - \bar{x}^2_{md} \end{bmatrix} \quad [\mathrm{pu}] \,. \tag{3.30}$$

In summary, the afore-developed model considers the following state-, input-, and measurement-vector:

$$\boldsymbol{x} = \begin{bmatrix} \overline{\Psi}_q & \overline{\Psi}_d & \overline{\Psi}_0 & \overline{\Psi}'_{kq} & \overline{\Psi}'_{fd} & \overline{\Psi}'_{kd} \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^6 \,, \tag{3.31a}$$

$$\boldsymbol{u} = \begin{bmatrix} \bar{v}_q & \bar{v}_d & \bar{v}_0 & \bar{v}'_{kq} & v'_{fd} & \bar{v}'_{kd} \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^6 \,, \tag{3.31b}$$

$$\boldsymbol{y} = \begin{bmatrix} \bar{i}_q & \bar{i}_d & \bar{i}_0 & \bar{i}'_{kq} & \bar{i}'_{fd} & \bar{i}'_{kd} \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^6 \,. \tag{3.31c}$$

Chapter 4

Bayesian State-Space Processors

"He sat, continuing to look down the nave, when suddenly the solution to the problem just seemed to present itself. It was so simple, so obvious, he just started to laugh."
— Paul Doherty[1], *Satan in St Mary's (1986)*

An enlightening introduction to this chapter may be borrowed from the work of Robert and Casella [103], in which the use of Bayesian processors is justified:

> Until the advent of powerful and accessible computing methods, the experimenter was often confronted with a difficult choice. Either describe an accurate model of a phenomenon, which would usually preclude the computation of explicit answers, or choose a standard model which would allow this computation but may not be a close representation of a realistic model. This dilemma is present in many branches of statistical applications, for example, in electrical engineering, aeronautics, biology, networks, and astronomy. To use realistic models, the researchers in these disciplines have often developed original approaches for model fitting that are customized for their own problems. This is particularly true for physicists, the originators of Markov chain Monte Carlo methods. Traditional methods of analysis, such as the usual numerical analysis techniques, are not well adapted for such settings.

If the data are modeled by a linear Gaussian state-space model, it is possible to derive an exact analytical expression to compute the evolving sequence of posterior distributions. This recursion is the well-known and widespread Kalman Filter (KF). If the data are modeled as a partially observed, finite state-space Markov chain, it is also possible to obtain an analytical solution, which is known as the hidden Markov model filter [99].

The aforementioned filters rely on various assumptions to ensure mathematical tractability. However, real data can be very complex, typically involving elements of non-Gaussianity, high dimensionality and nonlinearity, which conditions usually preclude analytic solution. This is a problem of fundamental importance that permeates most disciplines of science [103]. The problem appears under many different names, including Bayesian filtering [99], optimal nonlinear filtering [104], stochastic filtering [105], and on-line inference and learning [106].

[1]Paul Charles Dominic Doherty (1946–Present) is an award-winning English author, educator, lecturer, and historian. He is also the Headmaster of Trinity Catholic High School in London, England. Doherty is a prolific writer, has produced dozens of historical novels and a number of nonfiction history books.

For over thirty years, many approximations schemes, such as the Extended Kalman Filter (EKF), Gaussian sum approximations, and grid-based filters have been proposed to surmount this problem. The first two methods fail to take into account all the salient statistical features of the processes under consideration, leading quite often to poor results. Grid-based filters, based on deterministic numerical integration methods, can lead to accurate results, but are difficult to implement and too computationally expensive to be of any practical use in high dimensions [107].

Sequential Monte Carlo (SMC) methods are a set of simulation-based methods that provide a convenient and attractive approach to computing the posterior distributions. Unlike grid-based methods, SMC methods are very flexible, easy to implement, parallelizable, and applicable in very general settings [107].

The majority of concepts involved in this chapter are based on the works of Candy [99], Robert and Casella [103], Simon [105], Doucet, Freitas, and Gordon [107], Hogg, McKean, and Craig [108], and Maybeck [109] and the doctoral thesis of Giesbrecht [110].

The approach developed in this work to estimating states and parameters of salient-pole synchronous machines is based on the PF, which is a probability-based, SMC processor. Therefore, it is mandatory to begin this chapter with the foundation for the derivation of the PF – the Bayesian approach to state estimation.

4.1 Preliminary concepts

Modern probability theory is rigorously based on an axiomatic definition of probability. The axiomatic definition must still be a valid mathematical model of empirically observed frequencies of occurrence, but it is meant to extract the essence of the ideas involved and to deal with them in a precise, rather than heuristic, manner [109].

The first definition to be presented refers to sample space.

Definition 4.1: Sample space

*To describe an experiment in precise terms, let Ω be the fundamental **sample space** containing all possible outcomes of a conducted experiment. Each single elementary outcome of this experiment is denoted as an ω; these ω's are then the elements of Ω: $\omega \in \Omega$.*

In other words, the sample space is just the collection of possible outcomes of an experiment, each of which begin thought of as a point in Ω. Within every sample space, there are, at least, two subsets: the empty set, denoted by $\oslash$; and the sample space itself Ω.

The sample space Ω can be discrete, with a finite or countably infinite number of elements. It could also be continuous, with an uncountable number of elements.

Definition 4.2: Event

*Let F be defined as a specific **event** of interest, a specific set of outcomes of the experiment. Thus, each such event F is a subset of Ω: $F \subset \Omega$. An event F is said to occur if the observed outcome ω is an element of F, if $\omega \in F$.*

The structure of a sample space Ω, composed of elements ω_i, whose subsets are denoted as F_i, is shown in Figure 4.1.

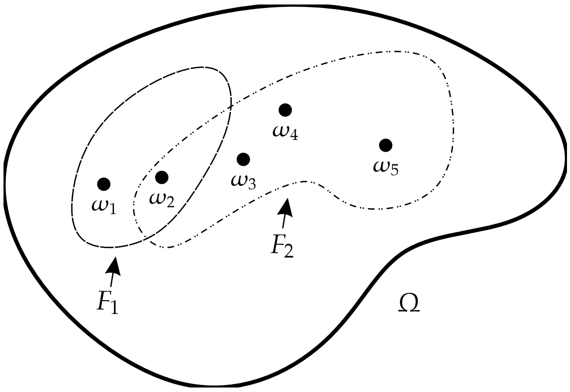

Figure 4.1: Representation of an arbitrary sample space Ω subsets F_i and elements ω_i. Adapted from Maybeck [109, p. 62].

Definition 4.3: Probability function (or probability measure)

Let $\Pr(\cdot)$ *be a real scalar-valued function defined on $\mathcal{F}$ that assigns a value,* $\Pr(F)$, *to each F which is a member of $\mathcal{F}(F \in \mathcal{F})$ such that:*

1. $\Pr(F) \geq 0$ *for all $F \in \mathcal{F}$;*
2. $\Pr(\Omega) = 1$;
3. *If $F_1, F_2, \cdots$ are elements of $\mathcal{F}$ and are disjoint, or mutually exclusive: i.e., if*

$$F_i \cap F_j = \oslash \qquad \text{for all } i \neq j$$

then

$$\Pr\left(\bigcup_{i=1}^{N} F_i\right) = \sum_{i=1}^{N} \Pr(F_i)$$

for all finite and countably infinite N.

A probability value between 0 and 1 (Pr is a mapping from $\mathcal{F}$ into $[0,1]$) is assigned to each set of interest (i.e., each $F \in \mathcal{F}$). Moreover, if F_1 is a subset of F_2, then the probability of set F_2 is at least as great as the probability of F_1, as expected. The set F_2 can be decomposed into two disjoint sets: F_1 and $F_1^c \cap F_2$. Then, according to part (3) of the definition of probability function,

$\Pr(F_2) = \Pr(F_1) + \Pr(F_1^c \cap F_2)$. From part (1), $\Pr(F_1^c \cap F_2) \geq 0$, and so $\Pr(F_2) \geq \Pr(F_1)$ as desired.

A certain class of sets F_i is of main interest: the broad class called a $\sigma-$algebra, denoted as $\mathcal{F}$. In other words, the sets $F_1, F_2, \cdots,$ – admissible for consideration – will be elements of the class $\mathcal{F} : F_i \in \mathcal{F}$.

Definition 4.4: $\sigma - algebra$

A $\boldsymbol{\sigma}-$algebra is a class of sets F_i, each of which is a subset of Ω ($F_i \subset \Omega$), such that if F_i is an element of $\mathcal{F}$ (i.e, if $F_i \in \mathcal{F}$), then:

1. *$F_i^c \in \mathcal{F}$, where F_i^c is the complement of F_i, $F_i^c = \Omega - F_i$;*
2. *$\Omega \in \mathcal{F}$ – and then the empty set $\oslash \in \mathcal{F}$ also, due to the preceding (1);*
3. *if $F_1, F_2, \cdots, \in \mathcal{F}$, then their union and intersection are also in $\mathcal{F}$:*

$$\bigcup_{i=1}^{\infty} F_i \in \mathcal{F} \qquad \text{and} \qquad \bigcap_{i=1}^{\infty} F_i \in \mathcal{F}\,.$$

Let the sample space Ω be the set of points in n-dimensional Euclidean space $\mathbb{R}^n$ and let $\mathcal{F}$ be the class of sets generated by the sets of the form (each of which is a subset of Ω):

$$F = \{\boldsymbol{\omega} : \boldsymbol{\omega} \leq \boldsymbol{\xi}, \quad \boldsymbol{\omega} \in \Omega\}\ , \tag{4.1}$$

and their complements, unions, and intersections.

A little digression on notation: F is the set of $\boldsymbol{\omega}$'s that are elements of Ω – vectors in the n-dimensional Euclidean space, and thus, the boldfacing of ω to denote vector quantity – such that $\boldsymbol{\omega} \leq \boldsymbol{\xi}$, where $\boldsymbol{\omega}$ and $\boldsymbol{\xi}$ are n-dimensional vectors and $\boldsymbol{\xi}$ is specified. Furthermore, $\boldsymbol{\omega} \leq \boldsymbol{\xi}$ is to be interpreted as componentwise: $\boldsymbol{\omega} \leq \boldsymbol{\xi}$ means $\omega_1 \leq \xi_1, \omega_2 \leq \xi_2, \cdots, \omega_n \leq \xi_n$ for the n components ω_i and ξ_i of $\boldsymbol{\omega}$ and $\boldsymbol{\xi}$, respectively.

This particular $\sigma-$algebra is of sufficient interest to have acquired its own name, and it is called a *Borel field*, denoted as $\mathcal{B}$. Taking complements, unions, and intersections of sets described by (4.1) leads to finite intervals (open, closed, or half open) and point values along each of the n dimensions. Thus, a Borel field is virtually composed of all subsets of Euclidean n-space ($\mathbb{R}^n$) that might be of interest in describing a probability problem associated with $\Omega = \mathbb{R}^n$.

At this point, it is possible do define the *probability space.*

Definition 4.5: Probability space

*The **probability space** is defined by the triplet $\{\Omega, \mathcal{F}, \Pr\}$ of the sample space, the underlying $\sigma-$algebra, and the probability function, all axiomatically defined as in the preceding definitions.*

The sample space Ω defines the possible outcomes of the experiment, $\mathcal{F}$ is the collection of events (sets) of interests, and Pr assigns a probability to every one of these events. However, for quantitative analysis, there is the need of a mapping from the sample space, Ω, to the real numbers, $\mathbb{R}$. This is achieved by the random variables.

Definition 4.6: Scalar random variable (or, simply, random variable)

*A **scalar random variable** $\mathcal{X}(\cdot)$ is a real-valued point* function *which assigns a real scalar value to each point $\omega \in \Omega$, denoted as $\mathcal{X}(\omega) = x$, such that every set $F \subset \Omega$ of the form*

$$F = \{\omega \in \Omega : \mathcal{X}(\omega) \leq \xi\}$$

for any ξ value on the real line ($\xi \in \mathbb{R}$), $\mathbb{R} = \{x : -\infty < x < \infty\}$, is an element of the $\sigma-$algebra $\mathcal{F}$ (i.e., $F \in \mathcal{F}$). For a discrete random variable, the subset is a finite or countably infinite set of points.

As highlighted by Maybeck [109], and agreed by the author:

> the name random variable is unfortunate in that it does not seem to imply the fact that we are talking about a function, as opposed to values the functions can assume. In fact, $\mathcal{X}(\cdot)$ is a function, or mapping, from Ω to $\mathbb{R}$.

The notation $\{\omega \in \Omega : \mathcal{X}(\omega) \leq \xi\}$ – or, similarly, $\{\omega : \mathcal{X}(\omega) \leq \xi\}$, is meant to be read *"the set of ω in Ω that the values assumed by the random variable function $\mathcal{X}(\cdot)$, for those ω as its argument, $\mathcal{X}(\omega) = x$, are less than or equal to the given number ξ on the real line"*.

A capital calligraphic letter $\mathcal{X}$ will denote the random variable, and a lowercase letter x, its value.

Definition 4.7: Vector random variable (or, simply, random vector)

*A **vector random variable** $\boldsymbol{\mathcal{X}}(\cdot)$ is a generalization of the random variable concept to the vector case: a real-valued point function which assigns a real vector to each point ω in Ω, denoted as $\boldsymbol{\mathcal{X}}(\omega)$, such that every set F of the form*

$$F = \{\omega : \boldsymbol{\mathcal{X}}(\omega) \leq \boldsymbol{\xi}\}$$

for any $\boldsymbol{\xi} \in \mathbb{R}^n$, is an element of the $\sigma-$algebra $\mathcal{F}$.

Scalar random variables are specifically mappings from Ω to $\mathbb{R}$, such that inverse images of half-open intervals of the form $(-\infty, \xi]$ in $\mathbb{R}$ are events in Ω that belong to $\mathcal{F}$. Vector random variables are simply extensions of the same idea: mappings from Ω into $\mathbb{R}^n$ such that the inverse images of sets of the form $\{\boldsymbol{\mathcal{X}}(\omega) \in \mathbb{R}^n : -\infty < \mathcal{X}_i(\omega) \leq \xi_i\ ;\ i = 1, 2, \cdots, n\}$ are the events in Ω to which probabilities have been ascribed[2].

[2]From a measure theoretic point of view, this just says that random variables are measurable functions [109, 110].

From now on, let the sample space Ω be $\mathbb{R}^n$ itself and the underlying $\sigma-$algebra be the Borel field $\mathcal{B}$ generated by sets of the form $F_i = \{\boldsymbol{\omega} : \boldsymbol{\omega} \leq \boldsymbol{f},\ \boldsymbol{\omega} \in \Omega\}$. An appropriate random variable definition for this case is simply the identity mapping:

$$\boldsymbol{\mathcal{X}}(\omega) = \boldsymbol{\omega}\,. \tag{4.2}$$

An element in the sample space, $\Omega = \mathbb{R}^n$, is just a single point in the space (a single vector), and the random variable just mentioned does map each such element into a single vector in $\mathbb{R}^n$. Thus, each realization $\boldsymbol{\mathcal{X}}(\omega)$ is an n-dimensional vector, whose components can take on any value within $(-\infty, \infty)$.

The elements of the set Ω that are contained in the event $\{\boldsymbol{\mathcal{X}} \leq \boldsymbol{\xi}\}$ change as ξ takes various values. The probability $\Pr(\{\omega : \boldsymbol{\mathcal{X}}(\omega) \leq \boldsymbol{\xi}\})$ of the event $\{\boldsymbol{\mathcal{X}} \leq \boldsymbol{\xi}\}$ is, therefore, a number that depends on ξ. This number is denoted by $F_{\boldsymbol{\mathcal{X}}}(\boldsymbol{\xi})$ and is called the cumulative density function (cdf) of the random variable $\boldsymbol{\mathcal{X}}$.

Definition 4.8: (Cumulative) distribution function

*The **(cumulative) distribution function** $F_{\boldsymbol{\mathcal{X}}}(\boldsymbol{\xi})$ is a real scalar-valued function defined by:*

$$\begin{aligned} F_{\boldsymbol{\mathcal{X}}}(\boldsymbol{\xi}) &= \Pr(\{\omega : \boldsymbol{\mathcal{X}}(\omega) \leq \boldsymbol{\xi}\}) && \text{(4.3a)} \\ &= \text{``}\Pr(\boldsymbol{\mathcal{X}} \leq \boldsymbol{\xi})\text{''} && \text{(4.3b)} \\ &= \text{``}\Pr(x_1 \leq \xi_1, x_2 \leq \xi_2, \cdots, x_n \leq \xi_n)\text{''} && \text{(4.3c)} \end{aligned}$$

that always exists.

The quotation marks in (4.3) are meant to emphasize that such notation, very typical in probability theory literature, should be interpreted in terms of the probability of a set of ω's in the original sample space Ω. Moreover, since

$$F_{\boldsymbol{\mathcal{X}}}(\boldsymbol{\xi}) = F_{x_1, x_2, \cdots, x_n}(\xi_1, \xi_2, \cdots, \xi_n)\,, \tag{4.4}$$

this is sometimes called the *joint distribution function* of $x_1, x_2, \cdots$, and x_n.

The cdf is a basic entity associated with any random variable that allows the generation of probabilities of sets of interest. Its existence is assured [109]. On the other hand, the existence of its derivative – the so-called probability density function (pdf) – everywhere is not assured. Some properties can be obtained from its definition [105]:

$$F_{\boldsymbol{\mathcal{X}}}(\boldsymbol{\xi}) \in [0, 1]\,, \tag{4.5}$$

$$F_{\boldsymbol{\mathcal{X}}}(-\infty) = 0\,, \tag{4.6}$$

$$F_{\boldsymbol{\mathcal{X}}}(\infty) = 1\,, \tag{4.7}$$

$$F_{\boldsymbol{\mathcal{X}}}(\xi_1) \leq F_{\boldsymbol{\mathcal{X}}}(\xi_2), \quad \text{if } \xi_1 \leq \xi_2\,, \tag{4.8}$$

$$\Pr(\xi_1 < \mathcal{X} \leq \xi_2) = F_{\boldsymbol{\mathcal{X}}}(\xi_2) - F_{\boldsymbol{\mathcal{X}}}(\xi_1)\,. \tag{4.9}$$

Figure 4.2 summarizes the majority of concepts discussed up to this point.

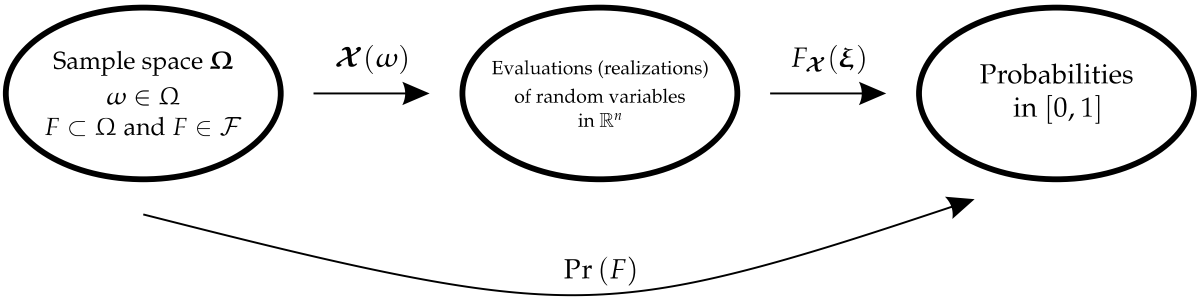

Figure 4.2: Probabilities and random variables.
Adapted from Maybeck [109, p. 70].

Definition 4.9: Probability density function

*The **probability density function** $f_{\boldsymbol{\mathcal{X}}}(\boldsymbol{\xi})$ is defined as the derivative of the cdf:*

$$f_{\boldsymbol{\mathcal{X}}}(\boldsymbol{\xi}) = \frac{F_{\boldsymbol{\mathcal{X}}}(\boldsymbol{\xi})}{d\xi}. \tag{4.10}$$

If $F_{\boldsymbol{\mathcal{X}}}(\boldsymbol{\xi})$ is absolutely continuous[3], then the pdf does exist. If such pdf exists, then $\boldsymbol{\mathcal{X}}$ is termed *continuous random variable*.

Some properties of the pdf that can be obtained from its definition are [105]:

$$F_{\boldsymbol{\mathcal{X}}}(\xi) = \int_{-\infty}^{\xi} f_{\boldsymbol{\mathcal{X}}}(z)dz\,, \tag{4.11}$$

$$f_{\boldsymbol{\mathcal{X}}}(\boldsymbol{\xi}) \geq 0\,. \tag{4.12}$$

Furthermore, since $F_{\boldsymbol{\mathcal{X}}}(\infty) = 1$, (4.11) yields

$$\int_{-\infty}^{\infty} f_{\boldsymbol{\mathcal{X}}}(z)dz = 1\,, \tag{4.13}$$

which justifies its name as the density function. Also, from (4.11),

$$\Pr\left[\xi_1 < \boldsymbol{\mathcal{X}}(\xi) \leq \xi_2\right] = F_{\boldsymbol{\mathcal{X}}}(\xi_2) - F_{\boldsymbol{\mathcal{X}}}(\xi_1) = \int_{\xi_1}^{\xi_2} f_{\boldsymbol{\mathcal{X}}}(z)dz\,. \tag{4.14}$$

Thus, the area under $f_{\boldsymbol{\mathcal{X}}}(\xi)$ in the interval (xi_1, xi_2) represents the probability that the random variable $\boldsymbol{\mathcal{X}}$ in such interval.

■

The probability of an event F_1, assuming the occurrence of F_2, is given by

$$\Pr(F_1 \mid F_2) = \frac{\Pr(F_1, F_2)}{\Pr(F_2)}, \quad \Pr(F_2) \neq 0\,, \tag{4.15}$$

[3] Absolute continuity can be define rigorously by means of measure theory. Basically, a function is absolutely continuous if the number of points where it is not differentiable is countable [109].

where $\Pr(F_1 \mid F_2)$ is the conditional probability of F_1 given F_2, that is, the probability that F_1 occurs given the fact that F_2 occurred. $\Pr(F_1, F_2)$ is the joint probability of F_1 and F_2, that is, the probability that both events F_1 and F_2 occur. The probability of a single event is called an *a priori* probability because it applies to the probability of an event apart from any previously known information. A conditional probability is called a *a posteriori* probability because it applies to a probability given the fact that some information about a possibly related event is already known.

Definition 4.10: Conditional distribution

The ***conditional distribution*** $F_{\boldsymbol{\mathcal{X}}}(x|F)$ *of a random variable* $\boldsymbol{\mathcal{X}}$*, assuming F is defined as the conditional probability of the event* $\{\boldsymbol{\mathcal{X}} \leq x\}$*, is:*

$$F_{\boldsymbol{\mathcal{X}}}(x|F) = \Pr(\boldsymbol{\mathcal{X}} \leq x \mid F) \tag{4.16a}$$

$$= \frac{\Pr(\boldsymbol{\mathcal{X}} \leq x, F)}{\Pr(F)} \tag{4.16b}$$

The set $\{\boldsymbol{\mathcal{X}} \leq x \mid F\}$ is the intersection of the events $\{(\boldsymbol{\mathcal{X}} \leq x)\}$ and F, that is, the event consisting of all outcomes ξ such that $\boldsymbol{\mathcal{X}}(\xi) \leq x$ and $\xi \in F$.

Definition 4.10 is the same as Definition 4.8, provided that all probabilities are replaced by conditional probabilities. From that, it follows that $F_{\boldsymbol{\mathcal{X}}}(x|F)$ has the same properties as $F_{\boldsymbol{\mathcal{X}}}(x)$.

Definition 4.11: Conditional density

The ***conditional density*** $f_{\boldsymbol{\mathcal{X}}}(x|F)$ *is the derivative of* $F_{\boldsymbol{\mathcal{X}}}(x|F)$*:*

$$f_{\boldsymbol{\mathcal{X}}}(x|F) = \frac{F_{\boldsymbol{\mathcal{X}}}(x|F)}{dx} = \lim_{\Delta x \to 0} \frac{F_{\boldsymbol{\mathcal{X}}}(x \leq x \leq x + \Delta x \mid F)}{\Delta x}. \tag{4.17}$$

This function is nonnegative and its area equals 1.

To summarize, Papoulis and Pillai [111] add:

> if the pdf of a random variable $\boldsymbol{\mathcal{X}}$ is unknown, one should make noncommittal judgment about its *a priori* pdf, $f_{\boldsymbol{\mathcal{X}}}$. Usually, the uniform distribution is a reasonable assumption in the absence of any other information. Then experiments results, F, are obtained, and the knowledge about $\boldsymbol{\mathcal{X}}$ is updated reflecting this new information. Bayes' rule helps to obtain the *a posteriori* pdf of $\boldsymbol{\mathcal{X}}$ given F. From that point on, this *a posteriori* pdf, $f_{\boldsymbol{\mathcal{X}}}(x|F)$, should be used to make further predictions and calculations [111, p. 105].

The distribution or density function of a random variable is the entity of fundamental interest in Bayesian estimation, embodying all information known about such variable. Once it is generated, an optimal estimate can be defined using some chosen criteria. Similarly, it can be used to compute the expected value of some function, where this expected value is just the average value one would obtain over the ensemble of outcomes of an experiment. The expected value of particular functions will generate moments of a random variable, which are parameters (statistics) that characterize the distribution or density function [109]. Although one would like to portray these functions completely through estimation, it is generally more feasible to evaluate expressions for a finite number of moments instead, thereby generating a partial description of the functions. In the case of Gaussian random variables, it will turn out that the specification of only the first two moments will completely describe the distribution (cdf) or the density function (pdf).

Definition 4.12: Expected value

Let $\boldsymbol{\mathcal{X}}$ be an n-dimensional vector random variable described by means of a density function $f_{\boldsymbol{\mathcal{X}}}(\xi)$. Also, let $\boldsymbol{y}$ be an m-dimensional vector function of $\boldsymbol{\mathcal{X}}$:

$$\boldsymbol{y}(\cdot) = \boldsymbol{g}\left\{\boldsymbol{\mathcal{X}}(\cdot)\right\}, \tag{4.18}$$

*where $\boldsymbol{g}(\cdot)$ is continuous. Then, the **expected value** of $\boldsymbol{y}$ is:*

$$\mathrm{E}\left[\boldsymbol{y}\right] = \int_{-\infty}^{\infty} \boldsymbol{g}(\xi) f_{\boldsymbol{\mathcal{X}}}(\xi) d\xi\,. \tag{4.19}$$

Since the expected value is, by definition, an integration, it is a linear operation. The aforementioned moments are defined in the following.

Definition 4.13: First moment of a random variable

*The **first moment** of $\boldsymbol{\mathcal{X}}$, or the mean of $\boldsymbol{\mathcal{X}}$, is generated by considering $\boldsymbol{g}(\boldsymbol{\mathcal{X}}) = \boldsymbol{\mathcal{X}}$:*

$$\boldsymbol{m} \triangleq \mathrm{E}\left[\boldsymbol{\mathcal{X}}\right] = \int_{-\infty}^{\infty} \boldsymbol{\xi} f_{\boldsymbol{\mathcal{X}}}(\boldsymbol{\xi}) d\boldsymbol{\xi}\,. \tag{4.20}$$

Definition 4.14: Second noncentral moment of a random variable

*The **second noncentral moment** of $\boldsymbol{\mathcal{X}}$, or the autocorrelation matrix of $\boldsymbol{\mathcal{X}}$, is generated by considering $\boldsymbol{g}(\boldsymbol{\mathcal{X}}) = \boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{X}}^T$:*

$$\boldsymbol{\Psi} \triangleq \mathrm{E}\left[\boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{X}}^T\right] = \int_{-\infty}^{\infty} \boldsymbol{\xi}\boldsymbol{\xi}^T f_{\boldsymbol{\mathcal{X}}}(\boldsymbol{\xi}) d\boldsymbol{\xi}\,. \tag{4.21}$$

Definition 4.15: Second central moment of a random variable

*The **second central moment** of $\boldsymbol{\mathcal{X}}$, or the covariance matrix of $\boldsymbol{\mathcal{X}}$, is generated by considering $\boldsymbol{g}(\boldsymbol{\mathcal{X}}) = \left[(\boldsymbol{\mathcal{X}} - \boldsymbol{m})(\boldsymbol{\mathcal{X}} - \boldsymbol{m})^T\right]$. It defines the $n \times n$ matrix $\boldsymbol{P}$, whose ij component is the covariance of x_i and x_j:*

$$\boldsymbol{P} \triangleq \mathrm{E}\left[(\boldsymbol{\mathcal{X}} - \boldsymbol{m})(\boldsymbol{\mathcal{X}} - \boldsymbol{m})^T\right] = \int_{-\infty}^{\infty} (\boldsymbol{\xi} - \boldsymbol{m})(\boldsymbol{\xi} - \boldsymbol{m})^T f_{\boldsymbol{\mathcal{X}}}(\xi) d\xi\,. \tag{4.22}$$

The matrix $\boldsymbol{P}$ is a symmetric, positive semidefinite matrix (its eigenvalues are nonnegative). The *variances* of the separate components of $\boldsymbol{\mathcal{X}}$ are along the diagonal:

$$P_{ii} \triangleq \mathrm{E}\left[(x_i - m_i)^2\right], \tag{4.23}$$

whose square roots are termed the *standard deviation* of x_i, denoted as σ_i.

It will be useful to generalize the concept of the second moment of a single random variable $\boldsymbol{\mathcal{X}}$ to the second moment relationship between two random variables: $\boldsymbol{\mathcal{X}}$ and $\boldsymbol{\mathcal{Z}}$.

Definition 4.16: Cross-correlation matrix

Let $\boldsymbol{\mathcal{X}}$ be an n-dimensional random vector and $\boldsymbol{\mathcal{Z}}$ be an m-dimensional random vector. The cross-correlation matrix of $\boldsymbol{\mathcal{X}}$ and $\boldsymbol{\mathcal{Z}}$ is the $n \times m$ matrix $\boldsymbol{\Psi}_{xz}$:

$$\boldsymbol{\Psi}_{xz} \triangleq \mathrm{E}\left[\boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{Z}}^T\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \boldsymbol{\xi}\boldsymbol{\rho}^T f_{\boldsymbol{\mathcal{X}},\boldsymbol{\mathcal{Z}}}(\boldsymbol{\xi},\boldsymbol{\rho})d\boldsymbol{\xi}d\boldsymbol{\rho}\,. \tag{4.24}$$

Similarly, the second central moment generalizes to the *cross-covariance matrix* of $\boldsymbol{\mathcal{X}}$ and $\boldsymbol{\mathcal{Z}}$:

$$\boldsymbol{P}_{xz} \triangleq \mathrm{E}\left[(\boldsymbol{\mathcal{X}} - \boldsymbol{m}_x)(\boldsymbol{\mathcal{Z}} - \boldsymbol{m}_z)^{\mathrm{T}}\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (\boldsymbol{\xi} - \boldsymbol{m}_x)(\boldsymbol{\rho} - \boldsymbol{m}_z)^{\mathrm{T}} f_{\boldsymbol{\mathcal{X}},\boldsymbol{\mathcal{Z}}}(\boldsymbol{\xi},\boldsymbol{\rho})d\boldsymbol{\xi}d\boldsymbol{\rho}\,. \tag{4.25}$$

Two random vectors $\boldsymbol{\mathcal{X}}$ and $\boldsymbol{\mathcal{Z}}$ are termed *uncorrelated* if their correlation matrix is equal to the outer product of their first order moments, i.e., if

$$\mathrm{E}\left[\boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{Z}}^{\mathrm{T}}\right] = \mathrm{E}\left[\boldsymbol{\mathcal{X}}\right]\mathrm{E}\left[\boldsymbol{\mathcal{Z}}^{\mathrm{T}}\right] = \boldsymbol{m}_x\boldsymbol{m}_z^{\mathrm{T}}, \tag{4.26}$$

or

$$\mathrm{E}\left[x_i z_j\right] = \mathrm{E}\left[x_i\right]\mathrm{E}\left[z_j\right] \quad \text{for all } i \text{ and } j\,, \tag{4.27}$$

which is equivalent to the condition that $\mathrm{E}\left[(x_i - m_{x_i})\left(z_j - m_{z_j}\right)\right] = 0$ for all i and j.

Whereas uncorrelatedness is a condition under which generalized second moments can be expressed as products of first order moments, independence is a condition under which the entire joint distribution or density function can be expressed as product or marginal functions. As might be expected then, if $\boldsymbol{\mathcal{X}}$ and $\boldsymbol{\mathcal{Z}}$ are independent, then they are uncorrelated, but not necessarily vice versa. This implication can be expressed simply as:

$$\boldsymbol{\mathcal{X}} \text{ and } \boldsymbol{\mathcal{Z}} \text{ independent} \rightarrow \boldsymbol{\mathcal{X}} \text{ and } \boldsymbol{\mathcal{Z}} \text{ uncorrelated.}$$

Definition 4.17: Orthogonality

*Two random vectors $\boldsymbol{\mathcal{X}}$ and $\boldsymbol{\mathcal{Z}}$ are termed **orthogonal** if their correlation matrix is the zero matrix. This concept is interrelated with $\boldsymbol{\mathcal{X}}$ and $\boldsymbol{\mathcal{Z}}$ being uncorrelated.*

Some important remarks:

1. *If either $\boldsymbol{\mathcal{X}}$ or $\boldsymbol{\mathcal{Z}}$ (or both) is zero-mean, then orthogonality and uncorrelatedness of $\boldsymbol{\mathcal{X}}$ and $\boldsymbol{\mathcal{Z}}$ imply each other;*

2. *If neither $\mathcal{X}$ nor $\mathcal{Z}$ is zero-mean, then $\mathcal{X}$ and $\mathcal{Z}$ may be uncorrelated or orthogonal or neither, but they cannot be both orthogonal and uncorrelated.*

Orthogonality provides a means to define an optimal estimate: if an estimate $\hat{x}$ of $\mathcal{X}$ is generated based on measurement data $\mathcal{Y}$, then that estimate can be termed optimal if the error $(\hat{x} - \mathcal{X})$ is orthogonal to the data. This geometrical concept is instrumental in deriving optimal estimators by means of *orthogonal projections*, the original means of the KF derivation [109].

Let $\mathcal{X}$ and $\mathcal{W}$ be random variables mapping Ω into $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively, and let $\mathcal{Z}$ be a continuous function of $\mathcal{X}$

$$\mathcal{Z} = g\left[\mathcal{X}(\cdot)\right] , \tag{4.28}$$

so that $\mathcal{Z}$ is itself a random variable mapping Ω into $\mathbb{R}^r$.

Definition 4.18: Conditional expected value

*The **conditional expected value**, or **conditional mean**, of $\mathcal{W}$, conditioned on the fact that $\mathcal{W}$ has assumed the realization $w \in \mathbb{R}^m$, i.e., $\mathcal{W}(\omega) = w$, is*

$$\mathrm{E}_{\mathcal{X}}\left[\mathcal{Z} \mid \mathcal{W} = w\right] = \int_{-\infty}^{\infty} g(\xi) f_{\mathcal{X}|\mathcal{W}}(\xi|w) d\xi . \tag{4.29}$$

The subscript $\mathcal{X}$ on $\mathrm{E}_{\mathcal{X}}\left[\mathcal{Z} \mid \mathcal{W} = w\right]$ denotes that the expectation operation (integration) is performed over the possible values of $\mathcal{X}$. For a given value $w \in \mathbb{R}^m$, $\mathrm{E}_{\mathcal{X}}\left[\mathcal{Z} \mid \mathcal{W} = w\right]$ is a vector in $\mathbb{R}^r$. Thus, $\mathrm{E}_{\mathcal{X}}\left[\mathcal{Z} \mid \mathcal{W} = \cdot\,\right]$ is a mapping from $\mathbb{R}^m$ into $\mathbb{R}^r$, a function of the values $w \in \mathbb{R}^m$. If these w values are realizations of the random variable $\mathcal{W}$, then the conditional expectation can be viewed as a random variable. These interrelationships are depicted in Figure 4.3.

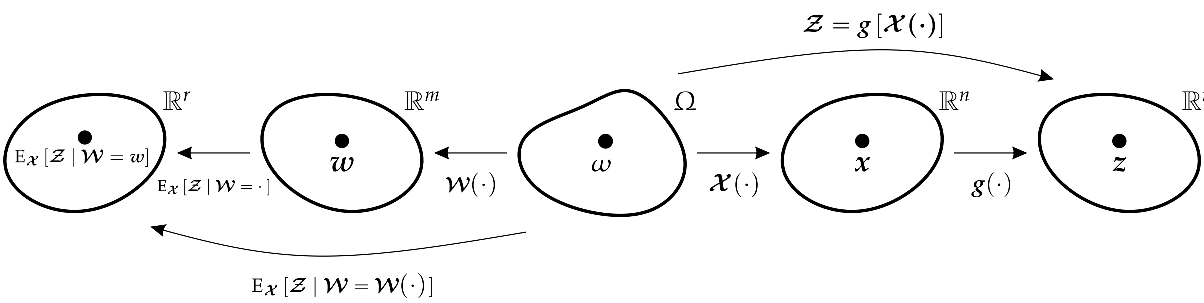

Figure 4.3: Conditional expectation functional relationships.
Adapted from Maybeck [109, p. 96].

Moreover, the random variable $\mathrm{E}_{\mathcal{X}}\left[\mathcal{Z} \mid \mathcal{W} = \mathcal{W}(\cdot)\right]$ is unique and has the property that

$$\mathrm{E}_{\mathcal{W}}\left\{\mathrm{E}_{\mathcal{X}}\left[\mathcal{Z} \mid \mathcal{W} = \mathcal{W}(\cdot)\right]\right\} = \mathrm{E}_{\mathcal{X}}\left[\mathcal{Z}\right] . \tag{4.30}$$

Conceptually, this is reasonable. Consider the conditional expectation of $\mathcal{Z}$, conditioned on a realized value of $\mathcal{W}$, and the expected value over all possible realizations of $\mathcal{W}$. Then, the result is the unconditional expectation of $\mathcal{W}$. A proof of this statement is found in Maybeck [109].

Definition 4.19: Conditional covariance

*The **conditional covariance** of $\boldsymbol{\mathcal{X}}$, given that $\boldsymbol{\mathcal{W}}(\omega) = \boldsymbol{w}$, is defined as*

$$\boldsymbol{P}_{x|w} = \mathrm{E}_{\boldsymbol{\mathcal{X}}}\left\{ \left(\boldsymbol{\mathcal{X}} - \mathrm{E}_{\boldsymbol{\mathcal{X}}}\left[\boldsymbol{\mathcal{X}} \mid \boldsymbol{\mathcal{W}} = \boldsymbol{w}\right]\right)\left(\boldsymbol{\mathcal{X}} - \mathrm{E}_{\boldsymbol{\mathcal{X}}}\left[\boldsymbol{\mathcal{X}} \mid \boldsymbol{\mathcal{W}} = \boldsymbol{w}\right]\right)^T \mid \boldsymbol{\mathcal{W}}(\omega) = \boldsymbol{w} \right\} \quad (4.31a)$$

$$= \int_{-\infty}^{\infty} \left(\boldsymbol{\mathcal{X}} - \mathrm{E}_{\boldsymbol{\mathcal{X}}}\left[\boldsymbol{\mathcal{X}} \mid \boldsymbol{\mathcal{W}} = \boldsymbol{w}\right]\right)\left(\boldsymbol{\mathcal{X}} - \mathrm{E}_{\boldsymbol{\mathcal{X}}}\left[\boldsymbol{\mathcal{X}} \mid \boldsymbol{\mathcal{W}} = \boldsymbol{w}\right]\right)^T f_{\boldsymbol{\mathcal{X}}|\boldsymbol{\mathcal{W}}}(\boldsymbol{\xi}|\boldsymbol{w}) d\boldsymbol{\xi}\,. \quad (4.31b)$$

If an estimate of $\boldsymbol{\mathcal{X}}$ using measurement data $\boldsymbol{\mathcal{W}}(\omega) = \boldsymbol{w}$ is to be generated, one possible estimator that is optimal with respect to many criteria is the random variable $\mathrm{E}_{\boldsymbol{\mathcal{X}}}\left[\boldsymbol{\mathcal{X}} \mid \boldsymbol{\mathcal{W}} = \boldsymbol{\mathcal{W}}(\cdot)\right]$. Then, $\left(\boldsymbol{\mathcal{X}} - \mathrm{E}_{\boldsymbol{\mathcal{X}}}\left[\boldsymbol{\mathcal{X}} \mid \boldsymbol{\mathcal{W}} = \boldsymbol{\mathcal{W}}(\cdot)\right]\right)$ can be interpreted as the random variable to model the error in the estimate: the difference between $\boldsymbol{\mathcal{X}}$ and its estimate. The conditional mean of this error vector would be zero. Consequently, $\boldsymbol{P}_{x|w}$ would be not only the conditional covariance of $\boldsymbol{\mathcal{X}}$, but also the conditional covariance of the error in the estimate of the value of $\boldsymbol{\mathcal{X}}$.

A particular random variable of significance is the Gaussian, or normal, vector-valued random variable. Firstly, it provides an adequate model of the random behavior exhibited in many phenomena observed in nature. Secondly, Gaussian random variables yield tractable mathematical models upon which to base estimators and controllers [109].

Definition 4.20: Gaussian random vector

*The random n-dimensional vector $\boldsymbol{\mathcal{X}}$ is said to be a **Gaussian (normal) random vector**, or a normally distributed vector-valued random variable, if it can be described by means of a pdf of the form*

$$f_{\boldsymbol{\mathcal{X}}} = \frac{1}{(2\pi)^{n/2} \left|\boldsymbol{P}\right|^{1/2}} \exp\left\{ -\frac{1}{2} \left(\boldsymbol{\xi} - \boldsymbol{m}\right)^T \boldsymbol{P}^{-1} \left(\boldsymbol{\xi} - \boldsymbol{m}\right) \right\}, \quad (4.32)$$

where $\boldsymbol{P}$ is a positive definite $n \times n$ matrix, $|\cdot|$ denotes the determinant of a matrix, and $\exp\{\cdot\}$ *denotes exponential.*

In Definition 4.20, the matrix $\boldsymbol{P}$ must be assumed positive definite to be assured the existence of its inverse. Note that the density function in (4.32) is completely defined by the two parameters $\boldsymbol{m}$ and $\boldsymbol{P}$. These parameters are, in fact, the mean vector and covariance matrix, respectively. Therefore, unlike most other density functions, higher order moments are not required to generate a complete description of the density function.

It was mentioned previously that Gaussian random variables are of engineering importance because they provide adequate models of many random phenomena observed empirically. The basic justification for this statement is embodied in the *central limit theorem*.

Theorem 4.1: Central limit

Let $\boldsymbol{\mathcal{X}}_i$, $i = 1, 2, \cdots, N$ be a set of independent random n-vectors which are identically distributed *with means and covariance matrices $\boldsymbol{m}_i$ and $\boldsymbol{P}_i$, respectively. Define the random vector $\boldsymbol{\mathcal{Z}}$ as their sum:*

$$\boldsymbol{\mathcal{Z}} = \sum_{i=1}^{N} \boldsymbol{\mathcal{X}}_i \tag{4.33}$$

and also define $\boldsymbol{\mathcal{Z}}$ as the (zero-mean) normalized sum random variable:

$$\boldsymbol{\mathcal{Z}} = \left[\boldsymbol{P}_{\mathcal{Z},\mathcal{Z}}\right]^{-1/2} \left\{\boldsymbol{\mathcal{Z}} - \mathrm{E}\left[\boldsymbol{\mathcal{Z}}\right]\right\} , \tag{4.34}$$

where

$$\mathrm{E}\left[\boldsymbol{\mathcal{Z}}\right] = \sum_{i=1}^{N} \boldsymbol{m}_i \qquad \text{and } \boldsymbol{P}_{\mathcal{Z},\mathcal{Z}} = \sum_{i=1}^{N} \boldsymbol{P}_i \, .$$

Then, in the limit as $N \to \infty$, $\boldsymbol{\mathcal{Z}}$ becomes a zero-mean Gaussian random n-vector with a covariance matrix equal to the identity matrix:

$$\lim_{N\to\infty} f_{\mathcal{Z}}(\boldsymbol{\xi}) = \frac{1}{(2\pi)^{n/2}} \exp\left\{\frac{1}{2}\boldsymbol{\xi}^T\boldsymbol{\xi}\right\} . \tag{4.35}$$

Essentially, Theorem 4.1 states that if the observed random phenomenon is generated as the sum of effects of many independent random phenomena, then the distribution of the observed phenomenon approaches a Gaussian distribution as more random effects are assumed, regardless of the distribution of each individual phenomenon [111]. In practice, however, the assumptions in the theorem are seldom verifiable. Rather, if there are a large number of additive contributing effects to a random phenomenon, then one suspects that a Gaussian distribution is a reasonable approximation to the actual distribution [109].

Another important concept is the *conditional Gaussian density*.

Let $\boldsymbol{\mathcal{X}}$ and $\boldsymbol{\mathcal{Z}}$ be jointly Gaussian vectors mapping Ω into $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively, so that $f_{\mathcal{X},\mathcal{Z}}(\boldsymbol{\xi},\boldsymbol{\rho})$ can be written as

$$\begin{aligned} f_{\mathcal{X},\mathcal{Z}}(\boldsymbol{\xi},\boldsymbol{\rho}) = &\left[(2\pi)^{(n+m)/2}\left|\begin{bmatrix}\boldsymbol{P}_{xx} & \boldsymbol{P}_{xz}\\ \boldsymbol{P}_{zx} & \boldsymbol{P}_{zz}\end{bmatrix}\right|\right]^{-1} \\ &\times \exp\left\{\frac{1}{2}\begin{bmatrix}\boldsymbol{\xi}-\boldsymbol{m}_x\\ \boldsymbol{\rho}-\boldsymbol{m}_z\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}\boldsymbol{P}_{xx} & \boldsymbol{P}_{xz}\\ \boldsymbol{P}_{zx} & \boldsymbol{P}_{zz}\end{bmatrix}^{-1}\begin{bmatrix}\boldsymbol{\xi}-\boldsymbol{m}_x\\ \boldsymbol{\rho}-\boldsymbol{m}_z\end{bmatrix}\right\} , \end{aligned} \tag{4.36}$$

where the covariance matrix is assumed to be positive definite. $\boldsymbol{\mathcal{X}}$ is a Gaussian n-vector of mean $\boldsymbol{m}_x$ and covariance $\boldsymbol{P}_{xx}$, and $\boldsymbol{\mathcal{Z}}$ is a Gaussian m-vector of mean $\boldsymbol{m}_z$ and covariance $\boldsymbol{P}_{zz}$.

Definition 4.21: Conditional Gaussian density

Considering the jointly Gaussian density $f_{\mathcal{X},\mathcal{Z}}(\boldsymbol{\xi},\boldsymbol{\rho})$ in (4.36), the conditional density $f_{\mathcal{X}|\mathcal{Z}}(\boldsymbol{\xi}|\boldsymbol{\rho})$ is obtained by means of Bayes' rule:

$$f_{\mathcal{X}|\mathcal{Z}}(\boldsymbol{\xi},\boldsymbol{\rho}) = \frac{f_{\mathcal{X},\mathcal{Z}}(\boldsymbol{\xi},\boldsymbol{\rho})}{f_{\mathcal{Z}}(\boldsymbol{\rho})} , \tag{4.37}$$

where $f_{\mathcal{Z}}(\boldsymbol{\rho})$ is Gaussian with moments $\boldsymbol{m}_z$ and covariance $\boldsymbol{P}_{zz}$. Performing algebraic reduction *yields the result as:*

$$f_{\mathcal{X}|\mathcal{Z}}(\boldsymbol{\xi}|\boldsymbol{\rho}) = \frac{1}{(2\pi)^{n/2}\left|\boldsymbol{P}_{x|z}\right|^{1/2}} \exp\left\{-\frac{1}{2}\left(\boldsymbol{\xi} - \boldsymbol{m}_{x|z}\right)^T \boldsymbol{P}_{x|z}^{-1}\left(\boldsymbol{\xi} - \boldsymbol{m}_{x|z}\right)\right\}, \tag{4.38}$$

where

$$\boldsymbol{m}_{x|z} = \boldsymbol{m}_x + \boldsymbol{P}_{xz}\boldsymbol{P}_{zz}^{-1}(\boldsymbol{\rho} - \boldsymbol{m}_z), \tag{4.39}$$

and

$$\boldsymbol{P}_{x|z} = \boldsymbol{P}_{xx} - \boldsymbol{P}_{xz}\boldsymbol{P}_{zz}^{-1}\boldsymbol{P}_{zx}. \tag{4.40}$$

From (4.39), $\mathrm{E}_{\mathcal{X}}\left[\mathcal{X}|\mathcal{Z} = \cdot\,\right] = \boldsymbol{m}_{x|z}$ can be seen to be an explicit function of the realizations $\boldsymbol{z}$ of $\mathcal{Z}$.

4.1.1 Stochastic processes

At this point, dynamics will be added to the system model developed previously, thereby allowing consideration of a much large class of problems of interest.

Definition 4.22: Stochastic process

*Let $\boldsymbol{\Omega}$ be a fundamental sample space and $\mathcal{T}$ be a subset of the real line denoting a time set of interest. Then, a **stochastic process** can be defined as a real-valued function $\mathcal{X}(\cdot,\cdot)$ defined on the product space $\mathcal{T} \times \boldsymbol{\Omega}$ (i.e., a function of two arguments, the first of which is an element of $\mathcal{T}$ and the second, an element of $\boldsymbol{\Omega}$), such that for any fixed $t \in \mathcal{T}$, $\mathcal{X}(t,\cdot)$ is a random variable. A scalar random process assumes values $x(t,\omega) \in \mathbb{R}$, whereas a vector random process assumes values $\boldsymbol{x}(t,\omega) \in \mathbb{R}^n$.*

In other words, $\mathcal{X}(\cdot,\cdot)$ is a stochastic process if all sets of the form

$$\mathcal{A} = \{\omega \in \boldsymbol{\Omega} : \boldsymbol{x}(t,\omega) \leq \boldsymbol{\xi}\}, \tag{4.41}$$

for any $t \in \mathcal{T}$ and $\boldsymbol{\xi} \in \mathbb{R}^n$ are in the underlying $\sigma-$algebra $\mathcal{F}$.

If the second argument of $\mathcal{X}(\cdot,\cdot)$ is fixed instead of the first, it is said that to each point $\omega_i \in \boldsymbol{\Omega}$ there can be associated a time function $\boldsymbol{x}(\cdot,\omega_i) = \boldsymbol{x}(\cdot)$, each of which is a sample from the stochastic process.

Two particular forms of $\mathcal{T}$ will be important. If $\mathcal{T}$ is a sequence $\{t_1, t_2, t_3, \cdots\}$, not necessarily equally spaced, then $\{\mathcal{X}_1(t_1,\cdot), \mathcal{X}_2(t_2,\cdot), \mathcal{X}_3(t_3,\cdot), \cdots\}$ becomes a sequence of random variables. The stochastic process $\mathcal{X}(t,\cdot)$ is then called a discrete-parameter stochastic process, or a discrete-time stochastic process. Instead, if $\mathcal{T}$ is an interval of $\mathbb{R}$, then $\mathcal{X}(\cdot,\cdot)$ becomes a continuous-parameter family of random variables, or a continuous-time stochastic process. For each ω, the sample is a function defined on the interval $\mathcal{T}$.

In Figure 4.4, a portray of samples from a discrete-time stochastic process and from a continuous-time stochastic process is presented.

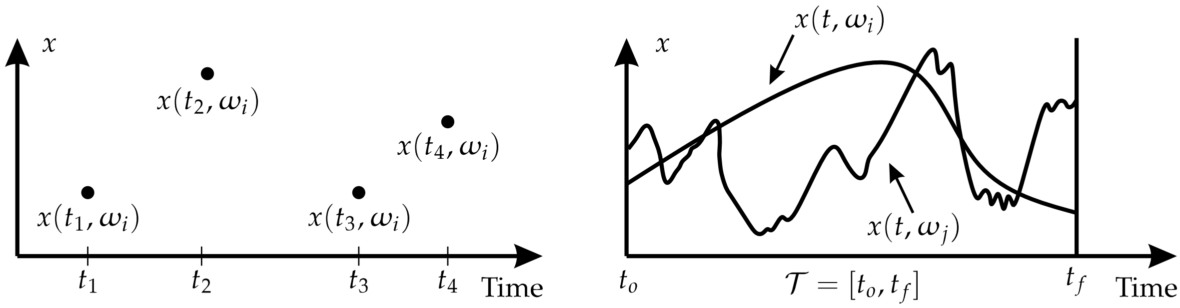

Figure 4.4: A portray of samples from a discrete-time stochastic process, on the left side; and from a continuous-time stochastic process, on the right side.
Adapted from Maybeck [109, p. 134].

If $\mathcal{T}$ is of the discrete form of a finite sequence of N points along the real line, the set of random variables $\{\boldsymbol{\mathcal{X}}_1(t_1,\cdot), \boldsymbol{\mathcal{X}}_2(t_2,\cdot), \cdots, \boldsymbol{\mathcal{X}}_N(t_N,\cdot)\}$ can be characterized by the joint cdf

$$F_{\boldsymbol{\mathcal{X}}(t_1),\cdots,\boldsymbol{\mathcal{X}}(t_N)}(\boldsymbol{\xi}_1, \cdots, \boldsymbol{\xi}_N) = \Pr\left(\omega : \boldsymbol{\mathcal{X}}(t_1,\omega) \leq \boldsymbol{\xi}_1, \cdots, \boldsymbol{\mathcal{X}}(t_N,\omega) \leq \boldsymbol{\xi}_N\right) \tag{4.42}$$

or the joint density function (if it exists):

$$f_{\boldsymbol{\mathcal{X}}(t_1),\cdots,\boldsymbol{\mathcal{X}}(t_N)}(\boldsymbol{\xi}_1, \cdots, \boldsymbol{\xi}_N) = \frac{\partial^{Nn} F_{\boldsymbol{\mathcal{X}}(t_1),\cdots,\boldsymbol{\mathcal{X}}(t_N)}(\boldsymbol{\xi}_1, \cdots, \boldsymbol{\xi}_N)}{\partial \xi_{11} \cdots \partial \xi_{1n} \cdots \partial \xi_{N1} \cdots \partial \xi_{Nn}} . \tag{4.43}$$

Other concepts also readily translate from probability theory, but care must be taken to avoid such ambiguities as the meaning of *independent processes* and *uncorrelated processes*.

Definition 4.23: Process independent in time

A process $\boldsymbol{\mathcal{X}}(\cdot,\cdot)$ is ***independent in time or white*** *if, for any choice of $t_1, \cdots, t_N \in \mathcal{T}$, $\boldsymbol{\mathcal{X}}(t_1), \cdots, \boldsymbol{\mathcal{X}}(t_N)$ are a set of independent random vectors; i.e.,*

$$\Pr\left(\omega : \boldsymbol{\mathcal{X}}(t_1,\omega) \leq \boldsymbol{\xi}_1, \cdots, \boldsymbol{\mathcal{X}}(t_N,\omega) \leq \boldsymbol{\xi}_N\right) = \prod_{i=1}^{N} \Pr\left(\omega : \boldsymbol{\mathcal{X}}(t_i,\omega) \leq \boldsymbol{\xi}_i\right), \tag{4.44}$$

or equivalently,

$$F_{\boldsymbol{\mathcal{X}}(t_1),\cdots,\boldsymbol{\mathcal{X}}(t_N)}(\boldsymbol{\xi}_1, \cdots, \boldsymbol{\xi}_N) = \prod_{i=1}^{N} F_{\boldsymbol{\mathcal{X}}(t_i)}(\boldsymbol{\xi}_i), \tag{4.45}$$

or, if the densities exist,

$$f_{\boldsymbol{\mathcal{X}}(t_1),\cdots,\boldsymbol{\mathcal{X}}(t_N)}(\boldsymbol{\xi}_1, \cdots, \boldsymbol{\xi}_N) = \prod_{i=1}^{N} f_{\boldsymbol{\mathcal{X}}(t_i)}(\boldsymbol{\xi}_i). \tag{4.46}$$

Definition 4.24: Processes independent from each other

Two processes $\boldsymbol{\mathcal{X}}(\cdot,\cdot)$ and $\boldsymbol{\mathcal{Y}}(\cdot,\cdot)$ are said to be ***independent from each other*** *if, for any choice of*

$t_1, \cdots, t_N \in \mathcal{T}$,

$$\begin{aligned}
&\Pr(\omega : \boldsymbol{\mathcal{X}}(t_1, \omega) \le \boldsymbol{\xi}_1, \cdots, \boldsymbol{\mathcal{X}}(t_N, \omega) \le \boldsymbol{\xi}_N, \boldsymbol{\mathcal{Y}}(t_1, \omega) \le \boldsymbol{\rho}_1, \cdots, \boldsymbol{\mathcal{Y}}(t_N, \omega) \le \boldsymbol{\rho}_N) \\
&\quad = \Pr(\omega : \boldsymbol{\mathcal{X}}(t_1, \omega) \le \boldsymbol{\xi}_1, \cdots, \boldsymbol{\mathcal{X}}(t_N, \omega) \le \boldsymbol{\xi}_N) \\
&\quad \times \Pr(\omega : \boldsymbol{\mathcal{Y}}(t_1, \omega) \le \boldsymbol{\rho}_1, \cdots, \boldsymbol{\mathcal{Y}}(t_N, \omega) \le \boldsymbol{\rho}_N) . \qquad (4.47a)
\end{aligned}$$

Therefore, *two independent processes* could mean two processes, each of which were independent in time, or two processes independent of each other, or some combination of these. The term *white* will be used to clarify this issue.

Definition 4.25: Process uncorrelated in time

A process $\boldsymbol{\mathcal{X}}(\cdot, \cdot)$ *is* ***uncorrelated in time*** *if, for all* $t_1, t_2 \in \mathcal{T}$, *except for* $t_1 = t_2$,

$$\boldsymbol{\Psi}_{xx}(t_1, t_2) = \mathrm{E}\left[\boldsymbol{\mathcal{X}}(t_1)\boldsymbol{\mathcal{X}}^T(t_2)\right] = \mathrm{E}\left[\boldsymbol{\mathcal{X}}(t_1)\right]\mathrm{E}\left[\boldsymbol{\mathcal{X}}^T(t_2)\right] \qquad (4.48)$$

or,

$$\boldsymbol{P}_{xx}(t_1, t_2) = \boldsymbol{0} . \qquad (4.49)$$

Definition 4.26: Processes uncorrelated from each other

Two processes $\boldsymbol{\mathcal{X}}(\cdot, \cdot)$ *and* $\boldsymbol{\mathcal{Y}}(\cdot, \cdot)$ *are said to be* ***uncorrelated from each other*** *if, for any choice of* $t_1, \cdots, t_N \in \mathcal{T}$, *including for* $t_1 = t_2$,

$$\boldsymbol{\Psi}_{xy}(t_1, t_2) = \mathrm{E}\left[\boldsymbol{\mathcal{X}}(t_1)\boldsymbol{\mathcal{Y}}^T(t_2)\right] = \mathrm{E}\left[\boldsymbol{\mathcal{X}}(t_1)\right]\mathrm{E}\left[\boldsymbol{\mathcal{Y}}^T(t_2)\right] \qquad (4.50)$$

or,

$$\boldsymbol{P}_{xy}(t_1, t_2) = \boldsymbol{0} . \qquad (4.51)$$

As shown previously, independence implies uncorrelatedness (which restricts attention to only the second moments), but the opposite implication is not true, except in such special cases as Gaussian processes. The term *white* if often accepted to mean uncorrelated in time rather than independent in time; the distinction between these definitions disappears for the important case of white Gaussian processes [109].

Definition 4.27: Gaussian process

A process $\boldsymbol{\mathcal{X}}(\cdot, \cdot)$ *is a* ***Gaussian process*** *if all finite joint distribution functions for* $\boldsymbol{\mathcal{X}}_1(t_1, \cdot)$, $\boldsymbol{\mathcal{X}}_2(t_2, \cdot), \cdots, \boldsymbol{\mathcal{X}}_N(t_N, \cdot)$ *are Gaussian for any choice of* $t_1, t_2, \cdots, t_N$. *For instance, if* $\boldsymbol{\mathcal{X}}(\cdot, \cdot)$ *is Gaussian and the appropriate densities exist, then any choice of* $t_1, t_2 \in \mathcal{T}$,

$$F_{\boldsymbol{\mathcal{X}}(t_1), \boldsymbol{\mathcal{X}}(t_2)}(\boldsymbol{\xi}) = \frac{1}{(2\pi)^{n/2} |\boldsymbol{P}|^{1/2}} \exp\left\{-\frac{1}{2}(\boldsymbol{\xi} - \boldsymbol{m})^T \boldsymbol{P}^{-1} (\boldsymbol{\xi} - \boldsymbol{m})\right\} , \qquad (4.52)$$

where

$$\boldsymbol{m} = \begin{bmatrix} \boldsymbol{m}_x(t_1) \\ \boldsymbol{m}_x(t_2) \end{bmatrix} = \begin{bmatrix} \mathrm{E}\left[\boldsymbol{\mathcal{X}}_x(t_1)\right] \\ \mathrm{E}\left[\boldsymbol{\mathcal{X}}_x(t_2)\right] \end{bmatrix}, \tag{4.53}$$

$$\boldsymbol{P} = \begin{bmatrix} \mathrm{E}\left[\boldsymbol{\mathcal{X}}_x(t_1)\boldsymbol{\mathcal{X}}_x^T(t_1)\right] - \boldsymbol{m}_x(t_1)\boldsymbol{m}_x^T(t_1) & \mathrm{E}\left[\boldsymbol{\mathcal{X}}_x(t_1)\boldsymbol{\mathcal{X}}_x^T(t_2)\right] - \boldsymbol{m}_x(t_1)\boldsymbol{m}_x^T(t_2) \\ \mathrm{E}\left[\boldsymbol{\mathcal{X}}_x(t_2)\boldsymbol{\mathcal{X}}_x^T(t_1)\right] - \boldsymbol{m}_x(t_2)\boldsymbol{m}_x^T(t_1) & \mathrm{E}\left[\boldsymbol{\mathcal{X}}_x(t_2)\boldsymbol{\mathcal{X}}_x^T(t_2)\right] - \boldsymbol{m}_x(t_2)\boldsymbol{m}_x^T(t_2) \end{bmatrix}. \tag{4.54}$$

4.1.1.1 Summary

Up to this point, a series of important definitions and results have been presented. Before continuing to develop other concepts, it is important to make some remarks concerning the nomenclature. Therefore, let $\boldsymbol{\mathcal{X}}$ be an n-dimensional random vector and $\boldsymbol{\mathcal{Y}}$ be an m-dimensional random vector.

The objective is to estimate the random parameter $\boldsymbol{\mathcal{X}}$ from noisy data $\boldsymbol{\mathcal{Y}} = \boldsymbol{y}$. The associated conditional distribution, $\Pr(\boldsymbol{\mathcal{X}}|\boldsymbol{\mathcal{Y}} = \boldsymbol{y})$, is called posterior distribution because the estimate is conditioned after the measurements have been acquired. Estimators based on this *a posteriori* distribution are usually called Bayesian because they are constructed from Bayes' rule, since it is difficult to obtain $\Pr(\boldsymbol{\mathcal{X}}|\boldsymbol{\mathcal{Y}})$ directly.

On the other hand, $\Pr(\boldsymbol{\mathcal{X}})$ is called prior distribution (before measurement); $\Pr(\boldsymbol{\mathcal{Y}}|\boldsymbol{\mathcal{X}})$ is called the likelihood (more likely to be true); and $\Pr(\boldsymbol{\mathcal{Y}})$ is called evidence, or normalizing factor – it scales the posterior to assure its integral is unity.

Bayesian methods view the sought-after parameter as random possessing a known *a priori* distribution. As measurements are made, the prior is transformed to the posterior distribution function adjusting the parameter estimates. If the *a priori* distribution is unknown, it is possible to adopt a generic enough distribution function to describe the process. In fact, that is the principle used to diffusely initialize stochastic processes in many SMC methods.

4.1.2 Problem statement

Although SMC methods can be applied to a more general setting, the following statement, provided by Doucet, Freitas, and Gordon [107], is restricted to signals modeled as Markovian, nonlinear, non-Gaussian state-space models. The unobserved signals (hidden, or internal [87], states) $\{\boldsymbol{x}(k), k \in \mathbb{N}\}$, $\boldsymbol{x}(k) \in \mathbb{X}$, are modeled as a Markov process of initial distribution $\Pr[\boldsymbol{x}(0)]$ and transition equation $\Pr[\boldsymbol{x}(k) \mid \boldsymbol{x}(k-1)]$. The observations $\{\boldsymbol{y}(k), k \in \mathbb{N}^{\star}\}$, $\boldsymbol{y}(k) \in \mathcal{Y}$, are assumed to be conditionally independent given the process $\{\boldsymbol{x}(k), k \in \mathbb{N}\}$ and of marginal distribution $\Pr[\boldsymbol{y}(k) \mid \boldsymbol{x}(k)]$.

To sum up, the model is described by:

$$\Pr[\boldsymbol{x}(k) \mid \boldsymbol{x}(k-1)], \tag{4.55}$$

$$\Pr[\boldsymbol{y}(k) \mid \boldsymbol{x}(k)], \tag{4.56}$$

$$\Pr\left[\boldsymbol{x}(0)\right] . \tag{4.57}$$

Similarly, it can be written as:

$$\mathbf{x}(k+1) = \mathbf{f}\left(\mathbf{x}(k), \mathbf{u}(k), k\right) \ , \tag{4.58}$$

$$\mathbf{y}(k) = \mathbf{h}\left(\mathbf{x}(k), \mathbf{u}(k), k\right) \ . \tag{4.59}$$

The objective is to recursively estimate in time the posterior distribution $\Pr\left\{\boldsymbol{x}[0,k] \mid \boldsymbol{y}[1,k]\right\}$, its associated features (including the marginal distribution $\Pr\left\{\boldsymbol{x}(k) \mid \boldsymbol{y}[1,k]\right\}$, known as the filtering distribution), and the expectations

$$I\left[g(k)\right] = \mathrm{E}\left[g(\boldsymbol{x}[0,k])\right] \triangleq \int g(\boldsymbol{x}[0,k]) \Pr(\boldsymbol{x}[0,k] \mid \boldsymbol{y}[1,k]) d\boldsymbol{x}[0,k] \tag{4.60}$$

for some function of interest $g(k)$ integrable with respect to $\Pr\left\{\boldsymbol{x}(k) \mid \boldsymbol{y}[1,k]\right\}$. Examples of appropriate functions, as shown in the previous section, include the conditional mean and the conditional covariance of $\boldsymbol{x}(k)$.

As before, $\mathbf{u}[t_0, t_1] \triangleq \mathbf{u}(t)$ for all $t_0 \leq t \leq t_1$ – i.e., the entire time function defined over $[t_0, t_1]$. The same notation is valid for other quantities.

At any time k, the posterior distribution is given by Bayes's theorem:

$$\Pr(\boldsymbol{x}[0,k] \mid \boldsymbol{y}[1,k]) = \frac{\Pr(\boldsymbol{y}[1,k] \mid \boldsymbol{x}[0,k]) \Pr(\boldsymbol{x}[0,k])}{\int \Pr(\boldsymbol{y}[1,k] \mid \boldsymbol{x}[0,k]) \Pr(\boldsymbol{x}[0,k]) d\boldsymbol{x}[0,k]} . \tag{4.61}$$

It is possible to obtain straightforwardly a recursive formula for this joint distribution [107]:

$$\Pr(\boldsymbol{x}[0,k+1] \mid \boldsymbol{y}[1,k+1]) = \Pr\left[\boldsymbol{x}[0,k] \mid \boldsymbol{y}[1,k]\right] \frac{\Pr\left[\boldsymbol{y}(k+1) \mid \boldsymbol{x}(k+1)\right] \Pr\left[\boldsymbol{x}(k+1) \mid \boldsymbol{x}(k)\right]}{\Pr\left[\boldsymbol{y}(k+1) \mid \boldsymbol{y}[1,k]\right]} . \tag{4.62}$$

The marginal distribution, $\Pr\left[\boldsymbol{x}(k) \mid \boldsymbol{y}[1,k]\right]$ also satisfies the following recursion. For prediction:

$$\Pr\left[\boldsymbol{x}(k) \mid \boldsymbol{y}[1,k]\right] = \int \Pr\left[\boldsymbol{x}(k) \mid \boldsymbol{x}(k-1)\right] \Pr\left[\boldsymbol{x}(k-1) \mid \boldsymbol{y}[1,k]\right] d\boldsymbol{x}(k-1) , \tag{4.63}$$

and for updating:

$$\Pr\left[\boldsymbol{x}(k) \mid \boldsymbol{y}[1,k]\right] = \frac{\Pr\left[\boldsymbol{y}(k) \mid \boldsymbol{x}(k)\right] \Pr\left[\boldsymbol{x}(k) \mid \boldsymbol{y}[1,k-1]\right]}{\int \Pr\left[\boldsymbol{y}(k) \mid \boldsymbol{x}(k)\right] \Pr\left[\boldsymbol{x}(k) \mid \boldsymbol{y}[1,k-1]\right] d\boldsymbol{x}(k)} . \tag{4.64}$$

These expressions and recursions are deceptively simple because one cannot typically compute the normalizing constant $\Pr\left[\boldsymbol{y}[1,k]\right]$, the marginals of the posterior $\Pr\left[\boldsymbol{x}[0,k] \mid \boldsymbol{y}[1,k]\right]$ – in particular, $\Pr\left[\boldsymbol{x}(k) \mid \boldsymbol{y}(k)\right]$ –, and $I\left[g(k)\right]$ since they require the evaluation of complex high-dimensional integrals.

To address those problems, many scientific and engineering disciplines have recently devoted a considerable effort towards the study and development of Monte Carlo (MC) integration methods. These methods have the great advantage of not being subject to any linearity or Gaussianity constraints on the model, and they also have appealing convergence properties.

4.1.3 The method of Monte Carlo

When a large number of samples are drawn from the required posterior distributions, it is not difficult to approximate the intractable integrals appearing in equations (4.62)–(4.64). However, it is seldom possible to directly obtain samples from these distributions. Therefore, one has to resort to alternative Monte Carlo (MC) methods, such as importance sampling – which will be presented in Section 4.5.

The act of generating observations from a specified distribution or sample is called MC generation [112]. This technique has been used for simulating complicated processes and investigating finite sample properties of statistical methodology for some time. However, in the last 30 years, it has become a very important concept in modern statistics in the realm of inference based on the bootstrap (resampling) and modern Bayesian methods.

As stated by Candy [99], the MC approach to solving Bayesian estimation problems is to replace complex analytic or unknown probability distributions with sample-based representations to solve a variety of unsolvable problems in inference, optimization, statistical mechanics, and nuclear physics [113–115].

4.1.3.1 Example

As a numerical example, let $f(x) = 4\sqrt{1-x^2}$, for $0 < x < 1$. It is desired to use the method of MC integration to estimate π. Then[4],

$$\pi = \int_0^1 f(x)dx = \mathrm{E}\left[f(x)\right] , \tag{4.65}$$

where $\mathcal{X}$ has the uniform $(0, 1)$ distribution. First, N random samples $x_1, \cdots, x_N$ are generated from the uniform $(0, 1)$ distribution and form $y_i = 4\sqrt{1-x_i^2}$.

An unbiased estimator of π is $\bar{y}$. With 95% confidence[5], the estimate is given by

$$\bar{y} - 1.96\frac{\sigma}{\sqrt{N}} , \bar{y} + 1.96\frac{\sigma}{\sqrt{N}} , \tag{4.66}$$

where σ is the value of the sample standard deviation. The algorithm was coded in MATLAB® and the results are summarized in Table 4.1 for different sample sizes. It is notable that for each experiment, the confidence interval trapped the true value of π.

The selection of a confidence level for an interval determines the probability that the confidence interval produced will contain the true parameter value. Common choices for the confidence level are 0.90, 0.95, and 0.99 [112]. These levels correspond to percentages of the area of the normal-density curve.

[4] The area of an unit circle is π, if the objective is to estimate π by integrating one quarter of a circle, the integration of $4\sqrt{1-x^2}$, for $0 < x < 1$ leads to π.

[5] A confidence interval gives an estimated range of values which is likely to include an unknown population parameter, the estimated range being calculated from a given set of sample data.

Table 4.1: Results for estimates of π for various runs of different sample sizes along with the confidence intervals.

Sample size N	100	1,000	10,000	100,000	100,000,000
$\bar{y}$	3.0497	3.1893	3.1464	3.1416	3.1417
$\bar{y} - 1.96\frac{\sigma}{\sqrt{N}}$	2.8690	3.1366	3.1291	3.1361	3.1415
$\bar{y} + 1.96\frac{\sigma}{\sqrt{N}}$	3.2304	3.2420	3.1637	3.1472	3.1419

For example, a 95% confidence interval covers 95% of the normal curve – the probability of observing a value outside of this area is less than 0.05. As the normal curve is symmetric, half of the area is in the left tail of the curve, and the other half of the area is in the right tail of the curve. For a confidence interval with an arbitrary level c, the area in each tail of the curve is equal to $(1 - c)/2$. For a 95% confidence interval, the area in each tail is equal to $0.05/2 = 0.025$.

Let $\mathcal{Z}$ be a normal-distributed random variable (zero mean and unit variance). The integrals

$$\int_{-\infty}^{-z} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}z^2\right\} = 0.025 \quad \text{and} \quad \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}z^2\right\} = 0.975 \tag{4.67}$$

imply that the value for z, such as

$$\Pr(\mathcal{Z} > z) = 0.025 \quad \text{and} \quad \Pr(\mathcal{Z} \leq z) = 0.975\,, \tag{4.68}$$

is $z = 1.96$. Therefore, a 95% confidence interval for the normal distribution is the interval $(-1.96, 1.96)$, since 95% of the area under the curve falls within this interval.

In general, a confidence interval for the population mean, based on a simple random sample of size N, is

$$\left(\bar{x} - z\frac{\sigma}{\sqrt{N}}, \quad \bar{x} + z\frac{\sigma}{\sqrt{N}}\right). \tag{4.69}$$

■

Numerical integration techniques have evolved over the last 30 years. However, the simplicity of MC integration still makes it a powerful technique [112].

4.1.4 Bootstrap procedure

In the last few years, Monte Carlo (MC) procedures have become increasingly used in statistical inference. In this subsection, a general method called the bootstrap procedure, which is a resampling procedure, is presented [112]. It was proposed by Efron [116]. Informative discussions of such procedures can be found in Efron and Tibshirani [117] and Davison and Hinkley [118].

In practice, the cdf of an estimate is not known. Therefore, the previous confidence interval cannot be obtained. On the other hand, suppose it is possible to take an infinite number of samples, obtain the estimates for each sample, and then form a histogram of these estimates. Although it seems impossible, since only one sample is available, it is the idea behind bootstrap procedures.

Although these methods do not call for a simulation-based implementation, in many cases where their use is particularly important, intensive simulation is required. The basic idea of the bootstrap[6] is to evaluate the properties of an arbitrary estimator, through the empirical cdf of the sample, instead of the theoretical one.

Let $\mathcal{X}$ be a random variable, $\boldsymbol{\mathcal{X}} = \{\mathcal{X}_1, \ldots, \mathcal{X}_n\}$ be a random sample on $\mathcal{X}$, and $\hat{\theta} = \hat{\theta}(\boldsymbol{\mathcal{X}})$ an arbitrary estimator. Furthermore, let

$$F_{\boldsymbol{\mathcal{X}}}(\boldsymbol{\xi}) = \frac{1}{n} \sum_{i=1}^{n} f_{\boldsymbol{\mathcal{X}}}(\boldsymbol{\xi}) \,, \tag{4.70}$$

be the empirical cdf of samples $\boldsymbol{\mathcal{X}}$. More precisely, if an estimate of $\theta(F) = \int h(\xi) dF(\xi)$, where F is the theoretical cdf, is desired, an obvious candidate is $\theta(F_{\boldsymbol{\mathcal{X}}}) = \int h(\xi) dF_{\boldsymbol{\mathcal{X}}}(\xi)$. When all $\mathcal{X}_i$ are independent and identically distributed, the *Glivenko-Cantelli Theorem*[7] guarantees the superior norm convergence of $F_{\boldsymbol{\mathcal{X}}}$ to F, and hence guarantees that $\theta(F_{\boldsymbol{\mathcal{X}}})$ is a consistent estimator of $\theta(F)$. The bootstrap provides an *automatic* method of computing $\theta(F_{\boldsymbol{\mathcal{X}}})$ by resampling the data.

Based on drawing $\boldsymbol{\mathcal{X}}^{\star,1}, \boldsymbol{\mathcal{X}}^{\star,2}, \ldots, \boldsymbol{\mathcal{X}}^{\star,B}$, where B is the number of bootstrap replications (i.e., the number of resamples) and

$$\boldsymbol{\mathcal{X}}^{\star,1} = \left\{ \mathcal{X}_1^{\star,1}, \ldots, \mathcal{X}_n^{\star,1} \right\} \sim F_{\boldsymbol{\mathcal{X}}} \,, \tag{4.71}$$

$\theta(F_{\boldsymbol{\mathcal{X}}})$ can be approximated by the bootstrap estimator

$$\hat{\theta}(F_{\boldsymbol{\mathcal{X}}}) \approx \frac{1}{B} \sum_{i=1}^{B} h(\boldsymbol{\mathcal{X}}^{\star,i}) \,, \tag{4.72}$$

with the approximation becoming more accurate as B increases.

If $\hat{\theta}$ is an arbitrary estimator of $\theta(F)$, the bias, the variance, or even the error distribution, of $\hat{\theta}$ can then be approximated by replacing F with $F_{\boldsymbol{\mathcal{X}}}$. Although the direct computation of $\hat{\theta}$ is possible in some particular cases, most setups require simulation to approximate the distribution of $\hat{\theta} - \theta(F_{\boldsymbol{\mathcal{X}}})$. In practice, $B \geq 3000$ [112].

4.2 Bayesian estimation

Bayes' rule provides the foundation of all Bayesian estimation techniques. When it comes to *Bayesian signal processing*, one is concerned with the estimation of the underlying probability distribution of a random signal in order to perform statistical inferences [119]. These inferences enable the extraction of the signal from noisy uncertain measurement data.

[6] This name comes from the German novel *Adventures of Baron Munchausen* by Rudolph Raspe, where the hero saves himself from drowning by pulling on his own bootstraps [103].

[7] Refer to Papoulis and Pillai [111].

4.2.1 Different estimators

To solve the estimation problem, the first step requires the determination of the *a posteriori* distribution. A logical solution to this problem leads to finding the *most probable* value of $\Pr(\mathcal{X}|\mathcal{Y})$ – its maximum. The Maximum *a posteriori* (MAP) estimate is the value of x that maximizes the posterior density, that is,

$$\hat{\mathcal{X}}_{MAP} = \max_{\mathcal{X}} \Pr(\mathcal{X}|\mathcal{Y}) \,. \tag{4.73}$$

The optimization is carried out in the usual manner by differentiating, setting the result to zero, and solving the resulting equation. Since many problems are based on the exponential class of densities, the $\ln \Pr(\mathcal{X}|\mathcal{Y})$ is considered instead. Since the logarithm is a monotonic function, the maximum of $\Pr(\mathcal{X}|\mathcal{Y})$ and $\ln \Pr(\mathcal{X}|\mathcal{Y})$ occur at the same value of $\mathcal{X}$.

Another important estimate is the Maximum likelihood (ML) estimate – it can be considered heuristically as that value of the parameter that best explains the measured data giving the most likely estimation. Searching over all $\mathcal{X}$ and selecting that value of $\mathcal{X}$ that is maximum leads to the ML estimate given by

$$\hat{\mathcal{X}}_{ML} = \arg\max_{\mathcal{X}} \Pr(\mathcal{X}|\mathcal{Y}) = \arg\max_{\mathcal{X}} \ln \Pr(\mathcal{X}|\mathcal{Y}) \,, \tag{4.74}$$

where $\ln \Pr(\mathcal{X}|\mathcal{Y})$ represents the *log-likelihood function*.

What makes the ML estimator popular is the fact that it enjoys some very desirable properties (proofs are found in Van Trees [120]):

1. ML estimates are consistent;
2. ML estimates are asymptotically Gaussian;
3. ML estimates of the sufficient statistic are equivalent to the ML estimates over the original data.

These properties are asymptotic and therefore imply that a large amount of data must be available for processing [99].

The main point is to note that the MAP estimate provides a mechanism to incorporate the *a priori* information, while the ML does not. Therefore, for some problems, MAP is the efficient estimator.

The most natural criterion to consider when constructing an estimate is one that minimizes the error between the true parameter and its estimate based on the measured data. The error-variance criterion is defined by

$$J(\mathcal{X}) = \mathrm{E}_{\mathcal{X}} \left\{ [\mathcal{X} - \hat{\mathcal{X}}(\mathcal{Y})]^{\mathrm{T}} [\mathcal{X} - \hat{\mathcal{X}}(\mathcal{Y})] \;|\; \mathcal{Y} \right\} , \tag{4.75}$$

where $\mathcal{X}$ is the true random n-dimensional vector; $\mathcal{Y}$ is the measured random m-dimensional vector (data); and $\hat{\mathcal{X}}$ is the estimate of $\mathcal{X}$ given $\mathcal{Y}$.

Minimizing $J(\mathcal{X})$ leads to the Minimum variance (MV) estimator. Thus,

$$\nabla_{\mathcal{X}} J(\mathcal{X}) = \mathrm{E}_{\mathcal{X}} \left\{ \nabla_{\mathcal{X}} \left[\mathcal{X} - \hat{\mathcal{X}}(\mathcal{Y})\right]^{\mathrm{T}} \left[\mathcal{X} - \hat{\mathcal{X}}(\mathcal{Y})\right] \mid \mathcal{Y} \right\} \tag{4.76a}$$

$$= \mathrm{E}_{\mathcal{X}} \left\{ \left[-\mathcal{X} - \hat{\mathcal{X}}(\mathcal{Y})\right] - \left[\mathcal{X} - \hat{\mathcal{X}}(\mathcal{Y})\right] \mid \mathcal{Y} \right\} \tag{4.76b}$$

$$= -2\, \mathrm{E}_{\mathcal{X}} \left\{ \mathcal{X} - \hat{\mathcal{X}}(\mathcal{Y}) \mid \mathcal{Y} \right\} \tag{4.76c}$$

$$= -2 \left\{ \mathrm{E}_{\mathcal{X}} \left[\mathcal{X} \mid \mathcal{Y}\right] - \hat{\mathcal{X}}(\mathcal{Y}) \right\} . \tag{4.76d}$$

Setting (4.76d) to zero and solving it yields the MV estimate as

$$\hat{\mathcal{X}}_{MV} = \hat{\mathcal{X}}(\mathcal{Y}) = \mathrm{E}_{\mathcal{X}} \left[\mathcal{X} \mid \mathcal{Y}\right] . \tag{4.77}$$

The MV estimator is linear, unconditionally and conditionally unbiased and possesses general orthogonality properties [99].

4.3 Classical Bayesian state-space processors

Bayesian estimation relative to the state-space models is based on extracting the unobserved or hidden dynamic variables from noisy measurement data. The state vector with initial distribution, $\Pr\left[\boldsymbol{x}(0)\right]$, propagates temporally throughout the state-space according to the probabilistic transition distribution, $\Pr\left[\boldsymbol{x}(k) \mid \boldsymbol{x}(k-1)\right]$, while the conditionally independent measurements evolve from the likelihood distribution $\Pr\left[\boldsymbol{y}(k) \mid \boldsymbol{x}(k)\right]$. The dynamic state variable at time k is obtained through the transition probability based on the previous state and the knowledge of the underlying conditional probability. Once propagated to time k, the dynamic state variable is used to update or correct based on the likelihood probability and the new measurement.

From the previous chapter and adding process and measurement noises, the functional discrete state representation is given by

$$\mathbf{x}(k) = \mathbf{f}\left(\mathbf{x}(k-1), \mathbf{u}(k-1), \mathbf{w}(k-1), k-1\right) , \tag{4.78}$$

$$\mathbf{y}(k) = \mathbf{h}\left(\mathbf{x}(k), \mathbf{u}(k), \mathbf{v}(k), k\right) , \tag{4.79}$$

where $\mathbf{w} \in \mathbb{R}^n$ and $\mathbf{v} \in \mathbb{R}^l$ are the respective process and measurement noise sources, with $\mathbf{u} \in \mathbb{R}^m$ a known input. The state vector is $\mathbf{x}(k) \in \mathbb{R}^n$ and the measurement vector is $\mathbf{y}(k) \in \mathbb{R}^l$. Here, $\mathbf{f}(\cdot)$ is a nonlinear (or linear) dynamic state transition function and $\mathbf{h}(\cdot)$, the corresponding measurement function. Both conditional probabilistic distributions embedded within the Bayesian framework are completely specified by these functions and the underlying noise distributions: $\Pr\left[\boldsymbol{w}(k-1)\right]$ and $\Pr\left[\boldsymbol{v}(k)\right]$. That is, the equivalence

$$\mathbf{f}\left(\mathbf{x}(k-1), \mathbf{u}(k-1), \mathbf{w}(k-1), k-1\right) \Rightarrow \Pr\left[\boldsymbol{x}(k) \mid \boldsymbol{x}(k-1)\right] \Leftrightarrow \boldsymbol{\mathcal{A}}\left[\boldsymbol{x}(k) \mid \boldsymbol{x}(k-1)\right] , \tag{4.80}$$

$$\mathbf{h}\left(\mathbf{x}(k), \mathbf{u}(k), \mathbf{v}(k), k\right) \Rightarrow \Pr\left[\boldsymbol{y}(k) \mid \boldsymbol{x}(k)\right] \Leftrightarrow \boldsymbol{\mathcal{C}}\left[\boldsymbol{y}(k) \mid \boldsymbol{x}(k)\right] , \tag{4.81}$$

is implied. This notation is used to emphasize the influence of both process ($\boldsymbol{\mathcal{A}}$) and measurement ($\boldsymbol{\mathcal{C}}$) representations on the conditional distributions.

Thus, the state-space model along with the noise statistics and prior distributions define the required Bayesian representation or probabilistic propagation model which defined the evolution of the states and measurements through the transition probabilities [99].

As represented by Candy [99], the basic dynamic state estimation problem can be stated in the Bayesian framework as:

GIVEN a set of noisy uncertain measurements $\boldsymbol{\mathcal{Y}}(k) = \{\boldsymbol{y}(k)\}$ and known inputs $\{\boldsymbol{u}(k)\}$, $k = 0, \cdots, N$, along with the corresponding prior distributions for the initial state and process and measurement noise sources: $\Pr[\boldsymbol{x}(0)]$, $\Pr[\boldsymbol{w}(k-1)]$, and $\Pr[\boldsymbol{v}(k)]$, as well as the conditional transition and likelihood probability distributions: $\Pr[\boldsymbol{x}(k) \mid \boldsymbol{x}(k-1)]$, $\Pr[\boldsymbol{y}(k) \mid \boldsymbol{x}(k)]$, characterized by the state and measurement models: $\boldsymbol{\mathcal{A}}[\boldsymbol{x}(k) \mid \boldsymbol{x}(k-1)]$, $\boldsymbol{\mathcal{C}}[\boldsymbol{y}(k) \mid \boldsymbol{x}(k)]$,

FIND the *best* estimate of the filtering posterior $\hat{\Pr}[\boldsymbol{x}(k) \mid \boldsymbol{\mathcal{Y}}(k)]$ and its associated statistics.

4.3.1 Linear Bayesian processor (Linear Kalman Filter)

Kalman filtering is the workhorse of state estimation [105].

In the transition to the 1960's, Kalman [121] presented his general theory of control systems and the generalization of Wiener filtering which became Kalman filtering [122]. He introduced the state-space representation and laid the foundations for state-space-based optimal filtering and optimal control theory, with linear-quadratic optimal control and the cornerstone of model-based control design [123].

Various books and papers that deal with Kalman filters present the filter equations in ways that appear very different from one another. It is not always obvious, but these different formulations are actually mathematically equivalent. One remarkable aspect of the Kalman filter is that it is optimal in several different senses.

At this point, the state-space model is constrained to be linear (time-varying). The Bayesian approach is applied to obtain the optimal processor assuming additive Gaussian noise. For the sake of simplicity, inputs are ignored.

Let the prediction equation be

$$\Pr[\boldsymbol{x}(k) \mid \boldsymbol{\mathcal{Y}}(k-1)] = \int \boldsymbol{\mathcal{A}}[\boldsymbol{x}(k) \mid \boldsymbol{x}(k-1)] \times \Pr[\boldsymbol{x}(k-1) \mid \boldsymbol{\mathcal{Y}}(k-1)]\, d\boldsymbol{x}(k-1) \tag{4.82}$$

where the filtered conditional[8] is

$$\Pr[\boldsymbol{x}(k) \mid \boldsymbol{\mathcal{Y}}(k-1)] \sim \mathcal{N}\left\{\boldsymbol{x}(k) \;:\; \hat{\boldsymbol{x}}(k-1|k-1), \tilde{\boldsymbol{P}}(k-1|k-1)\right\} \tag{4.83}$$

and $\tilde{\boldsymbol{P}}$ is the covariance for the state estimation error $\tilde{\boldsymbol{x}}$.

Using the process model,

$$\boldsymbol{\mathcal{A}}[\boldsymbol{x}(k-1) \mid \boldsymbol{x}(k)] \sim \mathcal{N}[\boldsymbol{x}(k-1) \;:\; A(k-1)\hat{\boldsymbol{x}}(k-1|k-1),$$

[8]The notation is defined in terms of conditional means by $\hat{x}(k|k) \triangleq \mathrm{E}\{x(k)|\mathcal{Y}(k)\}$.

$$A(k-1)\tilde{P}(k-1|k-1)A^{\mathrm{T}}(k-1) + P_{ww}(k-1)] \tag{4.84}$$

which follows directly from the linearity of the conditional expectation operator, that is

$$\hat{x}(k|k-1) = \mathrm{E}\left[x(k) \mid \mathcal{Y}(k-1)\right] \tag{4.85a}$$
$$= \mathrm{E}\left[A(k-1)x(k-1) + w(k-1) \mid \mathcal{Y}(k-1)\right] \tag{4.85b}$$
$$= A(k-1)\tilde{x}(k-1|k-1)\,. \tag{4.85c}$$

Using this result, the predicted state estimation error can be obtained as

$$\tilde{x}(k|k-1) = x - \hat{x}(k|k-1) \tag{4.86a}$$
$$= \{A(k-1)x(k-1) + w(k-1)\} - \{A(k-1)\hat{x}(k-1|k-1)\} \tag{4.86b}$$
$$= A(k-1)\hat{x}(k-1|k-1)\,. \tag{4.86c}$$

and the corresponding state error covariance $\tilde{P}(k|k-1) = \mathrm{E}\left[\tilde{x}(k|k-1)\tilde{x}^{\mathrm{T}}(k|k-1)\right]$ is easily derived.

Summarizing, the conditional means and covariances that completely characterize the current Gaussian state evolve according to the following equations:

$$\hat{x}(k|k-1) = A(k-1)\hat{x}(k-1|k-1) \tag{4.87}$$

for prediction, and

$$\tilde{P}(k|k-1) = A(k-1)\tilde{P}(k-1|k-1)A^{\mathrm{T}}(k-1) + P_{ww}(k-1) \tag{4.88}$$

for prediction covariance.

Therefore, (4.82) can be rewritten as:

$$\Pr\left[x(k) \mid \mathcal{Y}(k-1)\right] \sim \mathcal{N}\left\{x(k) \ : \ \hat{x}(k|k-1), \tilde{P}(k|k-1)\right\}\,. \tag{4.89}$$

With the prediction distribution available, the correction distribution obtained from the likelihood and the measurement model is:

$$\Pr\left[x(k) \mid \mathcal{Y}(k)\right] = \frac{\mathcal{C}\left[y(k) \mid x(k)\right] \times \Pr\left[x(k) \mid \mathcal{Y}(k-1)\right]}{\Pr\left[y(k) \mid \mathcal{Y}(k-1)\right]}\,. \tag{4.90}$$

Under the model assumptions, each of the conditional distributions can be expressed in terms of the Gaussian distributions as:

$$\mathcal{C}\left[y(k) \mid x(k)\right] \sim \mathcal{N}\left\{y(k) \ : \ C(k)x(k), P_{vv}(k)\right\}\,, \tag{4.91}$$
$$\Pr\left[x(k) \mid \mathcal{Y}(k-1)\right] \sim \mathcal{N}\left\{x(k) \ : \ \hat{x}(k|k-1), \tilde{P}(k|k-1)\right\}\,, \tag{4.92}$$
$$\Pr\left[y(k) \mid \mathcal{Y}(k-1)\right] \sim \mathcal{N}\left\{y(k) \ : \ \hat{y}(k|k-1), P_{ee}(k)\right\}\,, \tag{4.93}$$

for $P_{ee}(k)$ the innovations covariance with innovations defined by $e(k) \triangleq y(k) - \hat{y}(k|k-1)$ and predicted or filtered measurements given by $\hat{y}(k|k-1) = C(k)\hat{x}(k|k-1)$.

Considering these probabilities and combining all constants into a single constant κ,

$$\begin{aligned}\Pr\left[\boldsymbol{x}(k) \mid \boldsymbol{\mathcal{Y}}(k)\right] = \kappa \times \exp\left\{-\frac{1}{2}\left[\boldsymbol{y}(k) - \boldsymbol{C}(k)\boldsymbol{x}(k)\right]^{\mathrm{T}} \boldsymbol{P}_{vv}^{-1}(k)\left[\boldsymbol{y}(k) - \boldsymbol{C}(k)\boldsymbol{x}(k)\right]\right\} \times \\ \times \exp\left\{-\frac{1}{2}\left[\boldsymbol{x}(k) - \hat{\boldsymbol{x}}(k|k-1)\right]^{\mathrm{T}} \tilde{\boldsymbol{P}}^{-1}(k|k-1)\left[\boldsymbol{x}(k) - \hat{\boldsymbol{x}}(k|k-1)\right]\right\} \times \\ \times \exp\left\{+\frac{1}{2}\left[\boldsymbol{y}(k) - \hat{\boldsymbol{y}}(k|k-1)\right]^{\mathrm{T}} \boldsymbol{P}_{ee}^{-1}(k)\left[\boldsymbol{y}(k) - \hat{\boldsymbol{y}}(k|k-1)\right]\right\} .\end{aligned} \tag{4.94}$$

Recognizing the measurement noise, state estimation error, and innovation in above terms, the posterior probability given in terms of the model is:

$$\begin{aligned}\Pr\left[\boldsymbol{x}(k) \mid \boldsymbol{\mathcal{Y}}(k)\right] = \kappa \times \exp\left\{-\frac{1}{2}\boldsymbol{v}^{\mathrm{T}}(k)\boldsymbol{P}_{vv}^{-1}(k)\boldsymbol{v}(k)\right\} \times \\ \times \exp\left\{-\frac{1}{2}\left[\tilde{\boldsymbol{x}}(k|k-1)\right]^{\mathrm{T}} \tilde{\boldsymbol{P}}^{-1}(k|k-1)\left[\tilde{\boldsymbol{x}}(k|k-1)\right]\right\} \times \\ \times \exp\left\{+\frac{1}{2}\boldsymbol{e}^{\mathrm{T}}(k)\boldsymbol{P}_{ee}^{-1}(k)\boldsymbol{e}(k)\right\} .\end{aligned} \tag{4.95}$$

Therefore, the posterior distribution can be estimated under the multivariate Gaussian assumptions and the corresponding linear (time-varying) model. This is the optimal Bayesian Processor (BP) under these assumptions. In most cases, it is not possible to characterize the distributions in closed form and one must resort to numerical (simulation-based) solutions [99, 105].

Once the posterior is obtained, it is possible to estimate a variety of statistics using it as the basis. In this case, the optimal BP will be the one that maximizes the posterior.

Starting with the MAP equation and setting it to zero,

$$\nabla_x \ln \Pr\left[\boldsymbol{x}(k) \mid \boldsymbol{\mathcal{Y}}(k)\right]\Big|_{\boldsymbol{x}=\hat{\mathcal{X}}_{MAP}} = 0\,, \tag{4.96}$$

leads to

$$\nabla_x \ln \Pr\left[\boldsymbol{x}(k) \mid \boldsymbol{\mathcal{Y}}(k)\right] = \boldsymbol{C}^{\mathrm{T}}(k)\boldsymbol{P}_{vv}^{-1}(k)\left[\boldsymbol{y}(k) - \boldsymbol{C}(k)\boldsymbol{x}(k)\right] - \tilde{\boldsymbol{P}}^{-1}(k|k-1)\tilde{\boldsymbol{x}}(k|k-1)\,, \tag{4.97}$$

that, being solved for $\boldsymbol{x}(k)$ gives the Bayesian MAP estimate

$$\begin{aligned}\hat{\mathcal{X}}_{MAP} = \left[\boldsymbol{C}^{\mathrm{T}}(k)\boldsymbol{P}_{vv}^{-1}(k)\boldsymbol{C}(k) + \tilde{\boldsymbol{P}}^{-1}(k|k-1)\right]^{-1} \times \\ \times \left[\tilde{\boldsymbol{P}}^{-1}(k|k-1)\hat{\boldsymbol{x}}(k|k-1) + \boldsymbol{C}^{\mathrm{T}}(k)\boldsymbol{P}_{vv}^{-1}(k)\boldsymbol{y}(k)\right] .\end{aligned} \tag{4.98}$$

Upon applying matrix inversion properties, the first term becomes

$$\begin{aligned}\left[\boldsymbol{C}^{\mathrm{T}}(k)\boldsymbol{P}_{vv}^{-1}(k)\boldsymbol{C}(k) + \tilde{\boldsymbol{P}}^{-1}(k|k-1)\right]^{-1} &= \tilde{\boldsymbol{P}}^{-1}(k|k-1) - \tilde{\boldsymbol{P}}^{-1}(k|k-1)\boldsymbol{C}^{\mathrm{T}}(k)\times \\ &\quad \times \boldsymbol{P}_{ee}^{-1}(k)\boldsymbol{C}(k)\tilde{\boldsymbol{P}}^{-1}(k|k-1) && (4.99a) \\ &= \left[\boldsymbol{I} - \boldsymbol{K}(k)\boldsymbol{C}(k)\right]\tilde{\boldsymbol{P}}^{-1}(k|k-1)\,, && (4.99b)\end{aligned}$$

where $\boldsymbol{K}(k) = \tilde{\boldsymbol{P}}^{-1}(k|k-1)\boldsymbol{C}^{\mathrm{T}}(k)\boldsymbol{P}_{ee}^{-1}(k)$ is the gain; which is simply the updated error covariance $\tilde{\boldsymbol{P}}(k|k)$ equivalent to

$$\tilde{\boldsymbol{P}}(k|k) \equiv \left[\boldsymbol{C}^{\mathrm{T}}(k)\boldsymbol{P}_{vv}^{-1}(k)\boldsymbol{C}(k) + \tilde{\boldsymbol{P}}^{-1}(k|k-1)\right]^{-1} . \tag{4.100}$$

Thus, $\hat{\mathcal{X}}_{MAP}$ becomes

$$\hat{\mathcal{X}}_{MAP} = \tilde{\boldsymbol{P}}(k|k) \times \left[\tilde{\boldsymbol{P}}^{-1}(k|k-1)\hat{\boldsymbol{x}}(k|k-1) + \boldsymbol{C}^{\mathrm{T}}(k)\boldsymbol{P}_{vv}^{-1}(k)\boldsymbol{y}(k)\right] . \tag{4.101}$$

Multiplying out, regrouping terms, and factoring, the most popular form of the MAP estimate is achieved:

$$\hat{\mathcal{X}}_{MAP} = \hat{\boldsymbol{x}}(k|k) = \hat{\boldsymbol{x}}(k|k-1) + \boldsymbol{K}(k)\boldsymbol{e}(k) . \tag{4.102}$$

In terms of the updated instead of the predicted error covariance, the gain is expressed as:

$$\boldsymbol{K}(k) = \tilde{\boldsymbol{P}}(k|k)\boldsymbol{C}^{\mathrm{T}}(k)\boldsymbol{P}_{vv}^{-1}(k) \equiv \tilde{\boldsymbol{P}}(k|k-1)\boldsymbol{C}^{\mathrm{T}}(k)\boldsymbol{P}_{ee}^{-1}(k) . \tag{4.103}$$

It is important to note that a necessary and sufficient condition that the linear BP is optimal is that the innovation sequence is zero-mean and white or uncorrelated. If this condition does not hold, then the underlying model and assumptions are invalid [99].

4.3.2 Extended Bayesian processor (Extended Kalman Filter)

The Extended Kalman filter was originally proposed by Stanley Schmidt[9] [124] so that the Kalman filter could be applied to nonlinear spacecraft navigation problems. Its idea is based on the linearization of the nonlinear process about the Kalman filter estimate, which is based on a linearized system.

Let the following process model

$$\boldsymbol{x}(k) = \boldsymbol{a}\left[\boldsymbol{x}(k-1)\right] + \boldsymbol{b}\left[\boldsymbol{u}(k-1)\right] + \boldsymbol{w}(k-1) , \tag{4.104}$$

with the corresponding measurement model

$$\boldsymbol{y}(k) = \boldsymbol{c}\left[\boldsymbol{x}(k)\right] + \boldsymbol{v}(k) , \tag{4.105}$$

where $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ are nonlinear vectors functions of $\boldsymbol{x}$ and $\boldsymbol{u}$, with $\boldsymbol{x}, \boldsymbol{a}, \boldsymbol{b}, \boldsymbol{w} \in \mathbb{R}^n$ and $\boldsymbol{y}, \boldsymbol{c}, \boldsymbol{v} \in \mathbb{R}^m$, $\boldsymbol{w} \sim \mathcal{N}\left[\boldsymbol{0}, \boldsymbol{P}_{ww}(k-1)\right]$ and $\boldsymbol{v} \sim \mathcal{N}\left[\boldsymbol{0}, \boldsymbol{P}_{vv}(k)\right]$.

Ignoring the additive noise sources, the process and measurement models may be linearized about a known deterministic reference trajectory defined by $[\boldsymbol{x}^\star(k), \boldsymbol{u}^\star(k)]$, that is

$$\boldsymbol{x}^\star(k) = \boldsymbol{a}\left[\boldsymbol{x}^\star(k-1)\right] + \boldsymbol{b}\left[\boldsymbol{u}^\star(k-1)\right] + \boldsymbol{w}(k-1) , \tag{4.106}$$

$$\boldsymbol{y}(k) = \boldsymbol{c}\left[\boldsymbol{x}^\star(k)\right] . \tag{4.107}$$

Deviations or perturbations from this trajectory are define by

$$\delta\boldsymbol{x}(k) \triangleq \boldsymbol{x}(k) - \boldsymbol{x}^\star(k) , \tag{4.108}$$

[9]Stanley F. Schmidt (1926—2015) was an aerospace engineer who pioneered the Schmidt-Kalman filter used in air and space navigation, most notably Apollo spacecraft.

$$\delta\boldsymbol{u}(k) \triangleq \boldsymbol{u}(k) - \boldsymbol{u}^{\star}(k)\,, \tag{4.109}$$

$$\delta\boldsymbol{y}(k) \triangleq \boldsymbol{y}(k) - \boldsymbol{y}^{\star}(k)\,. \tag{4.110}$$

Substituting (4.108)–(4.110) into (4.106)–(4.107), one can obtain a linearized perturbation model valid for small deviations given by:

$$\delta\boldsymbol{x}(k) = \boldsymbol{A}\left[\boldsymbol{x}^{\star}(k-1)\right]\delta\boldsymbol{x}(k-1) + \boldsymbol{B}\left[\boldsymbol{u}^{\star}(k-1)\right]\delta\boldsymbol{u}(k-1) + \boldsymbol{w}(k-1)\,, \tag{4.111}$$

$$\delta\boldsymbol{y}(k) = \boldsymbol{C}\left[\boldsymbol{x}^{\star}(k)\right]\delta\boldsymbol{x}(k) + \boldsymbol{v}(k)\,, \tag{4.112}$$

with the corresponding

$$\boldsymbol{A}\left[\boldsymbol{x}^{\star}(k-1)\right] \triangleq \frac{\partial\boldsymbol{a}\left[\boldsymbol{x}^{\star}(k-1)\right]}{\partial\boldsymbol{x}^{\star}(k-1)}\,, \tag{4.113}$$

$$\boldsymbol{B}\left[\boldsymbol{u}^{\star}(k-1)\right] \triangleq \frac{\partial\boldsymbol{b}\left[\boldsymbol{u}^{\star}(k-1)\right]}{\partial\boldsymbol{u}^{\star}(k-1)}\,, \tag{4.114}$$

$$\boldsymbol{C}\left[\boldsymbol{x}^{\star}(k)\right] \triangleq \frac{\partial\boldsymbol{c}\left[\boldsymbol{x}^{\star}(k)\right]}{\partial\boldsymbol{x}^{\star}(k)}\,, \tag{4.115}$$

Jacobian matrices and $\boldsymbol{w}$, $\boldsymbol{v}$ zero-mean Gaussian.

Therefore, the state perturbation predicted estimate is simply

$$\delta\hat{\boldsymbol{x}}(k|k-1) = \boldsymbol{A}\left[\boldsymbol{x}^{\star}(k-1)\right]\delta\hat{\boldsymbol{x}}(k-1|k-1) + \boldsymbol{B}\left[\boldsymbol{u}^{\star}(k-1)\right]\delta\boldsymbol{u}(k-1)\,. \tag{4.116}$$

However, the interest is in the state estimate, not its deviation. From the definition of perturbation, it can be shown [109] that

$$\hat{\boldsymbol{x}}(k|k-1) = \delta\boldsymbol{x}(k|k-1) + \boldsymbol{x}(k)\,. \tag{4.117}$$

Considering the process model and (4.116),

$$\begin{aligned}\hat{\boldsymbol{x}}(k|k-1) = {} & \boldsymbol{A}\left[\boldsymbol{x}^{\star}(k-1)\right]\delta\hat{\boldsymbol{x}}(k-1|k-1) + \boldsymbol{B}\left[\boldsymbol{u}^{\star}(k-1)\right]\delta\boldsymbol{u}(k-1) + \cdots \\ & + \boldsymbol{a}\left[\boldsymbol{x}^{\star}(k-1)\right] + \boldsymbol{b}\left[\boldsymbol{u}^{\star}(k-1)\right]\end{aligned} \tag{4.118}$$

Using the linear BP with deterministic Jacobian matrices results in

$$\delta\hat{\boldsymbol{y}}(k|k-1) = \boldsymbol{C}\left[\boldsymbol{x}^{\star}(k)\right]\delta\hat{\boldsymbol{x}}^{\star}(k|k-1) \tag{4.119}$$

and, therefore

$$\hat{\boldsymbol{y}}(k|k-1) = \boldsymbol{y}^{\star}(k) + \boldsymbol{C}\left[\boldsymbol{x}^{\star}(k)\right]\delta\hat{\boldsymbol{x}}^{\star}(k|k-1) \tag{4.120a}$$

$$= \boldsymbol{c}\left[\boldsymbol{y}^{\star}(k)\right] + \boldsymbol{C}\left[\boldsymbol{x}^{\star}(k)\right]\delta\hat{\boldsymbol{x}}^{\star}(k|k-1)\,. \tag{4.120b}$$

In the Extended Kalman filter framework, the reference state $\boldsymbol{x}^{\star}(k)$ is replaced with the most recently available state estimate $\hat{\boldsymbol{x}}(k|k)$. The Jacobians used in the linearization process are deterministic (but time-varying), when a reference or perturbation trajectory is used. However, using the current state estimate is an approximation to the conditional mean, which is random, making these associated Jacobians and subsequent relations random. The reason for choosing to linearize about this estimate is that it represents the best information available about the state and therefore most likely results in a better reference trajectory (state estimate). As a consequence, large initial estimation errors do not propagate; therefore, linearity assumptions are less likely to be violated [99].

The updated state estimate is easily obtained by substituting the predicted estimate for the reference $\hat{\boldsymbol{x}}(k|k-1) \to \boldsymbol{x}^{\star}(k)$:

$$\delta\hat{\boldsymbol{x}}(k|k) = \delta\hat{\boldsymbol{x}}(k|k-1) + \boldsymbol{K}(k)\boldsymbol{e}(k)\,, \tag{4.121}$$

$$[\hat{\boldsymbol{x}}(k|k) - \hat{\boldsymbol{x}}(k|k-1)] = [\hat{\boldsymbol{x}}(k|k-1) - \hat{\boldsymbol{x}}(k|k-1)] + \boldsymbol{K}(k)\boldsymbol{e}(k)\,, \tag{4.122}$$

$$\hat{\boldsymbol{x}}(k|k) = \hat{\boldsymbol{x}}(k|k-1) + \boldsymbol{K}(k)\boldsymbol{e}(k)\,, \tag{4.123}$$

where $\boldsymbol{K}(k) = \tilde{\boldsymbol{P}}(k|k-1)\boldsymbol{C}^{\mathrm{T}}(k)\,[\hat{\boldsymbol{x}}(k|k-1)]\,\boldsymbol{P}_{ee}^{-1}(k)$.

Under the model assumptions, each of the conditional distributions can be expressed in terms of the Gaussian distributions as:

$$\Pr\left[\boldsymbol{y}(k) \mid \boldsymbol{x}(k-1)\right] \sim \mathcal{N}\left\{\boldsymbol{c}(k)\,[\boldsymbol{x}(k)]\,, \boldsymbol{P}_{vv}(k)\right\}\,, \tag{4.124}$$

$$\Pr\left[\boldsymbol{x}(k) \mid \boldsymbol{\mathcal{Y}}(k-1)\right] \sim \mathcal{N}\left\{\hat{\boldsymbol{x}}(k|k-1), \tilde{\boldsymbol{P}}(k|k-1)\right\}\,, \tag{4.125}$$

$$\Pr\left[\boldsymbol{y}(k) \mid \boldsymbol{\mathcal{Y}}(k-1)\right] \sim \mathcal{N}\left\{\hat{\boldsymbol{y}}(k|k-1), \boldsymbol{P}_{ee}(k)\right\}\,. \tag{4.126}$$

Similarly to the procedure applied from (4.96), the MAP estimate is given as

$$\hat{\mathcal{X}}_{MAP} = \hat{\boldsymbol{x}}(k|k) = \hat{\boldsymbol{x}}(k|k-1) + \boldsymbol{K}(k)\boldsymbol{e}(k)\,. \tag{4.127}$$

In summary, the Extended Kalman Filter (EKF) is performed accordingly to the following steps [99, 105, 125]:

1. The nonlinear system is given by:

$$\boldsymbol{x}(k) = \boldsymbol{a}\,[\boldsymbol{x}(k-1)] + \boldsymbol{b}\,[\boldsymbol{u}(k-1)] + \boldsymbol{w}(k-1)\,, \tag{4.128}$$

$$\boldsymbol{y}(k) = \boldsymbol{c}\,[\boldsymbol{x}(k)] + \boldsymbol{v}(k)\,, \tag{4.129}$$

$$\boldsymbol{w} \sim \mathcal{N}\left[\boldsymbol{0}, \boldsymbol{P}_{ww}(k-1)\right]\,, \tag{4.130}$$

$$\boldsymbol{v} \sim \mathcal{N}\left[\boldsymbol{0}, \boldsymbol{P}_{vv}(k)\right]\,. \tag{4.131}$$

2. Initialize the filter as follows:

$$\hat{\boldsymbol{x}}(0|0) = \mathrm{E}\left[\boldsymbol{x}(0)\right]\,, \tag{4.132}$$

$$\tilde{\boldsymbol{P}}(0|0) = \mathrm{E}\left[\left(\boldsymbol{x}_0 - \hat{\boldsymbol{x}}(0|0)\right)\left(\boldsymbol{x}_0 - \hat{\boldsymbol{x}}(0|0)\right)^{\mathrm{T}}\right]\,. \tag{4.133}$$

3. For $k = 1, 2, \cdots,$:

   a) Compute the following Jacobian matrix:

$$\boldsymbol{A}\,[\hat{\boldsymbol{x}}(k|k-1)] = \frac{\partial \boldsymbol{a}\,[\hat{\boldsymbol{x}}(k|k-1)]}{\partial \hat{\boldsymbol{x}}(k|k-1)}\,. \tag{4.134}$$

b) Perform the time update of the state estimate and estimation-error covariance as:

$$\hat{\boldsymbol{x}}(k|k-1) = \boldsymbol{a}\left[\hat{\boldsymbol{x}}(k-1|k-1)\right] + \boldsymbol{b}\left[\boldsymbol{u}(k-1)\right] , \tag{4.135}$$

$$\tilde{\boldsymbol{P}}(k|k-1) = \boldsymbol{A}\left[\hat{\boldsymbol{x}}(k|k-1)\right]\tilde{\boldsymbol{P}}(k-1|k-1)\boldsymbol{A}\left[\hat{\boldsymbol{x}}(k|k-1)\right]^{\mathrm{T}} - \boldsymbol{P}_{ww}(k-1) . \tag{4.136}$$

c) Compute the following Jacobian matrix:

$$\boldsymbol{C}\left[\hat{\boldsymbol{x}}(k|k-1)\right] = \frac{\partial \boldsymbol{c}\left[\hat{\boldsymbol{x}}(k|k-1)\right]}{\partial \hat{\boldsymbol{x}}(k|k-1)} . \tag{4.137}$$

d) Perform the measurement update of the state estimate and estimation-error covariance:

$$\boldsymbol{e}(k) = \boldsymbol{y}(k) - \hat{\boldsymbol{y}}(k|k-1) , \tag{4.138}$$

$$\boldsymbol{P}_{ee}(k) = \boldsymbol{C}\left[\hat{\boldsymbol{x}}(k|k-1)\right]\tilde{\boldsymbol{P}}(k|k-1)\boldsymbol{C}\left[\hat{\boldsymbol{x}}(k|k-1)\right]^{\mathrm{T}} + \boldsymbol{P}_{vv}(k) , \tag{4.139}$$

$$\boldsymbol{K}(k) = \tilde{\boldsymbol{P}}(k|k-1)\boldsymbol{C}\left[\hat{\boldsymbol{x}}(k|k-1)\right]^{\mathrm{T}}\boldsymbol{P}_{ee}{}^{-1}(k) , \tag{4.140}$$

$$\hat{\boldsymbol{x}}(k|k) = \hat{\boldsymbol{x}}(k|k-1) + \boldsymbol{K}(k)\boldsymbol{e}(k) , \tag{4.141}$$

$$\tilde{\boldsymbol{P}}(k|k) = \left[\boldsymbol{I} - \boldsymbol{K}(k)\boldsymbol{C}\left[\hat{\boldsymbol{x}}(k|k-1)\right]\right]\tilde{\boldsymbol{P}}(k|k-1) . \tag{4.142}$$

After all, the KF algorithm attempts to propagate the mean and the covariance of a system using a time-update and a measurement update. If the system is linear, then the mean and covariance can be exactly updated with the KF. If the system is nonlinear, then the mean and the covariance can be approximately updated with the Extended Kalman Filter (EKF) [105].

4.4 Modern Bayesian state-space processors

The Extended Kalman Filter (EKF) is the most widely applied state estimation algorithm for nonlinear systems. However, the EKF can be difficult to tune and often gives unreliable estimates if the system nonlinearities are severe [105]. This is because the EKF relies on linearization to propagate the mean and the covariance of the state.

The problem with nonlinear systems is that it is difficult to transform a pdf through a general nonlinear function. The EKF works on the principle that a linearized transformation of means and covariances is approximately equal to the true nonlinear transformation, but the approximation could be unsatisfactory.

The use of Unscented Kalman Filter (UKF) can provide significant improvement over the EKF. An unscented transformation is based on two fundamental principles [105]. First, it is easy to perform a nonlinear transformation on a single point – rather than an entire pdf. Second, it is not too hard to find a set of individuals points in state space whose sample pdf approximates the true pdf of a state vector.

4.4.1 Sigma-point (unscented) transformation

A completely different approach to nonlinear estimation evolves from the concept of statistical linearization [125–127]. Instead of approximating the nonlinear process and measurement dynamics of the underlying system using Taylor series representation – that leads to the classical forms of estimation, including EKF –, the statistical linearization approximation or equivalently statistical linearization method provides an alternative that takes into account the uncertainty or probabilistic spread of the prior random vector. The basic idea is to approximate (linearize) a nonlinear function of a random vector while preserving its first and second moments [127]; therefore, this approach requires *a priori* knowledge of its distribution resulting in a more statistically accurate transformation.

The Sigma-Point Transformation (SPT), or, equivalently, unscented transformation, is a technique for calculating the statistics of a random vector that has been nonlinearly transformed [99]. The set of samples (the so-called sigma points) are chosen so that they capture the specific properties of the underlying distribution. A Sigma-Point Transformation (SPT) is portrayed in Figure 4.5.

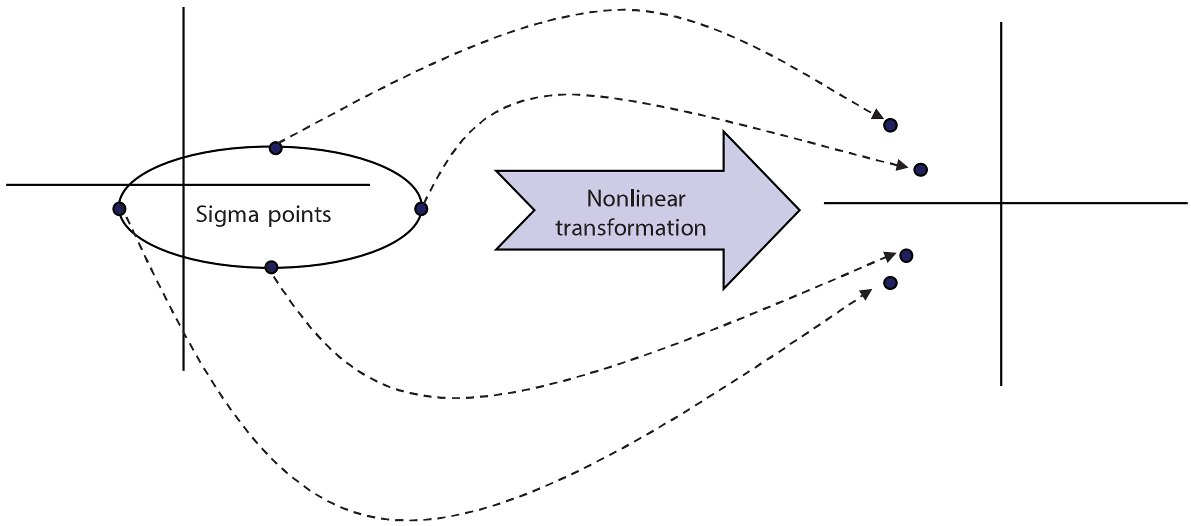

Figure 4.5: Unscented transformation: a set of distribution points shown on an error ellipsoid are selected and transformed into a new space where their underlying statistics are estimated. Adapted from Candy [99, p. 205].

In Figure 4.5, the $f(\boldsymbol{\mathcal{X}})$ is considered to be a two-dimensional Gaussian, so that the σ-points are located along the major and minor axes of the covariance ellipse capturing the essence of this distribution. In general, the goal is to construct a set of σ-points possessing the same statistics as the original distribution such that when the nonlinearity is transformed to the new space, the new set of points sufficiently capture the posterior statistics.

The transformation occurs on a point-by-point basis, since it is simpler to match statistics of individual points rather than the entire pdf. The statistics of the transformed points are then calculated to provide the desired estimates of the transformed distribution.

The following development is based on the work of Julier and Uhlmann [128], which should be consulted for further details.

As before, let the n-dimensional random vector $\boldsymbol{\mathcal{X}}$ be propagated through an arbitrary nonlinear transformation $\boldsymbol{a}[\cdot]$ to generate a new random vector,

$$\boldsymbol{\mathcal{Y}} = \boldsymbol{a}[\boldsymbol{\mathcal{X}}] \,. \tag{4.143}$$

Let a set of σ-points $\{\boldsymbol{\mathfrak{X}}_i\}$ consists of $n_\sigma + 1$ vectors with appropriate weights $\{w_i\}$ given by

$$\boldsymbol{\Sigma} = \{\boldsymbol{\mathfrak{X}}_i, w_i\} \,, \quad i = 0, \cdots, n_\sigma \,. \tag{4.144}$$

The weights can be positive or negative, but must sum to unity, so that the estimate of the statistics remains unbiased. Then, the problem becomes:

GIVEN the σ-points $\boldsymbol{\Sigma}$ and the nonlinear transformation $\boldsymbol{a}[\cdot]$,

FIND the statistics of the transformed samples:

$$\boldsymbol{m}_y = \mathrm{E}\left[\boldsymbol{\mathcal{Y}}\right] \quad \text{and} \quad \boldsymbol{P}_{yy} = \mathrm{cov}\left[\boldsymbol{\mathcal{Y}}\right] \,. \tag{4.145}$$

As an example [105], a set of σ-points that satisfies all conditions consist of a symmetric set of $n_\sigma + 1$ points that lie on the $\sqrt{n}$-th covariance contour:

$$\begin{aligned}
\boldsymbol{\mathfrak{X}}_0 &= \boldsymbol{m}_x \,, & w_0 &= n^{-1} \,, & (4.146)\\
\boldsymbol{\mathfrak{X}}_i &= \boldsymbol{m}_x + \left(\sqrt{n\boldsymbol{P}_{xx}}\right)_i^{\mathrm{T}} \,, & w_i &= (2n)^{-1} \,, & (4.147)\\
\boldsymbol{\mathfrak{X}}_{i+n} &= \boldsymbol{m}_x - \left(\sqrt{n\boldsymbol{P}_{xx}}\right)_i^{\mathrm{T}} \,, & w_{i+n} &= (2n)^{-1} \,, & (4.148)
\end{aligned}$$

where $\sqrt{n\boldsymbol{P}_{xx}}$ is the matrix square root of $n\boldsymbol{P}_{xx}$ such that $\left(\sqrt{n\boldsymbol{P}_{xx}}\right)^{\mathrm{T}}\left(\sqrt{n\boldsymbol{P}_{xx}}\right) = n\boldsymbol{P}_{xx}$; and $\left(\sqrt{n\boldsymbol{P}_{xx}}\right)_i$ is the ith row of $\sqrt{n\boldsymbol{P}_{xx}}$.[10]

Therefore, the SPT can be considered a statistical linearization method that provides an optimal linear approximation to a general nonlinear transformation considering the prior second-order statistics of the underlying random variable – its mean and covariance [129].

To be more precise and parallel, it is preferable to approximate the underlying Gaussian distribution rather than approximate its resultant nonlinear transformation, in contrast to the EKF.

It is important to recognize that the SPT has specific properties when the underlying distribution is Gaussian:

1. Since the distribution is symmetric, the σ-points can be selected with this symmetry;

2. The problem of approximating $\boldsymbol{\mathcal{X}}$ with an arbitrary mean and covariance can be reduced to that of a standard zero-mean, unit-variance Gaussian, since

$$\boldsymbol{\mathcal{X}} = \boldsymbol{m}_x + \boldsymbol{U}\boldsymbol{\mathcal{S}} \,, \tag{4.149}$$

[10] MATLAB®'s Cholesky factorization routine CHOL can be used to find a matrix square root.

where $\boldsymbol{U}$ is the matrix square root of $\boldsymbol{P}_{xx}$; and $\boldsymbol{\mathcal{S}} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$.

Therefore, in the Gaussian case, the second-order SPT uses a set of σ-points which correctly captures the first two moments of $\boldsymbol{\mathcal{S}}$; that is, they must capture the mean, covariance, and symmetry. Also, from the symmetry properties of the distribution, all odd-ordered moments are zero [99].

In summary[11], the sigma-point processor under a multivariate Gaussian assumption relies on:

1. Determine the set of $2n+1$ σ-points from the rows or columns of $\pm\sqrt{(n+\kappa)\boldsymbol{P}_{xx}}$, where κ is a scaling factor. For the nonzero-mean case, compute $\boldsymbol{\mathfrak{X}}_i = \sigma + \boldsymbol{m}_x$;

$$\boldsymbol{\mathfrak{X}}_0 = \boldsymbol{m}_x\,, \qquad \boldsymbol{w}_0 = \kappa(n+\kappa)^{-1}\,, \tag{4.150}$$

$$\boldsymbol{\mathfrak{X}}_i = \boldsymbol{m}_x + \left(\sqrt{(n+\kappa)\boldsymbol{P}_{xx}}\right)_i^{\mathrm{T}}\,, \qquad \boldsymbol{w}_i = (2n+2\kappa)^{-1}\,, \tag{4.151}$$

$$\boldsymbol{\mathfrak{X}}_{i+n} = \boldsymbol{m}_x - \left(\sqrt{(n+\kappa)\boldsymbol{P}_{xx}}\right)_i^{\mathrm{T}}\,, \qquad \boldsymbol{w}_{i+n} = (2n+2\kappa)^{-1}\,, \tag{4.152}$$

where κ is a scalar; $\left(\sqrt{(n+\kappa)\boldsymbol{P}_{xx}}\right)_i$ is the ith row or column of the matrix square root of $(n+\kappa)\boldsymbol{P}_{xx}$; and $\boldsymbol{w}_i$ is the weight associated with the ith σ-point;

2. Nonlinearly transform each point to obtain the set of the new σ-points: $\boldsymbol{\mathcal{Y}} = \boldsymbol{a}[\boldsymbol{\mathcal{X}}]$;

3. Estimate the posterior mean of the new samples by its weighted average (regression):

$$\boldsymbol{m}_x = \sum_{i=0}^{2n} w_i \boldsymbol{\mathfrak{X}}_i\,; \tag{4.153}$$

4. Estimate the posterior covariance of the new samples by its weighted outer product (regression):

$$\boldsymbol{P}_{yy} = \sum_{i=0}^{2n} w_i \left(\boldsymbol{\mathfrak{Y}}_i - \boldsymbol{m}_y\right)\left(\boldsymbol{\mathfrak{Y}}_i - \boldsymbol{m}_y\right)^{\mathrm{T}}\,. \tag{4.154}$$

4.4.2 Sigma-point Bayesian processor (Unscented Kalman Filter)

The Unscented Kalman Filter (UKF) is a recursive processor developed to eliminate some of the deficiencies created by the failure of the first-order (Taylor series) linearization process in solving the state estimation problem.

Uhlmann[12] started to develop the UKF in the 1990s [125] by analyzing the performance of different sigma-point sampling and weighting strategies on a nonlinear estimation problem

[11] For further details, please refer to Candy [99].

[12] Jeffrey Uhlmann is an American research scientist who is probably best known for his mathematical generalizations of the Kalman filter. Most of his publications and patents have been in the field of data fusion. He is also known for being a cult filmmaker and former recording artist.

in robotics. The results of this and collaborative studies [128, 129] was a suit of nonlinear Kalman Filter (KF) extensions, all based on a core methodology for approximating nonlinear transformations of the mean and the covariance.

An important, although intriguing, parenthesis *must* be made on a fun-fact about the name *Unscented*. On an interview [130], Jeffrey Uhlmann said:

> Initially I only referred to it as the "new filter". Needing a more specific name, people in my lab began referring to it as the "Uhlmann filter", which obviously isn't a name that I could use, so I had to come up with an official term. One evening everyone else in the lab was at the Royal Opera House, and as I was working I noticed someone's deodorant on a desk. The word "unscented" caught my eye as the perfect technical term. At first people in the lab thought it was absurd – which is okay because absurdity is my guiding principle – and that it wouldn't catch on. My claim was that people simply accept technical terms as technical terms: for example, does anyone think about why a tree is called a tree? Within a few months we had a speaker visit from another university who talked about his work with the "unscented filter". Clearly, he had never thought about the origin of the term. The cover of the issue of the March 2004 Proceedings we're discussing right now has "Unscented" in large letters on the cover, which shows that it has been accepted as the technical term for that approach.

Differently from the EKF, the sigma-point processor does not approximate the nonlinear process and measurement models; it employs the true nonlinear models and approximates the underlying Gaussian distribution function of the state variable using a statistical linearization approach leading to a set of regression equations for the states and measurements.

Therefore, the EKF equations are simply replaced with the SPT to obtain the UKF algorithm [99, 105, 125]:

1. The nonlinear system is given by:

$$\boldsymbol{x}(k) = \boldsymbol{a}\left[\boldsymbol{x}(k-1)\right] + \boldsymbol{b}\left[\boldsymbol{u}(k-1)\right] + \boldsymbol{w}(k-1)\,, \tag{4.155}$$

$$\boldsymbol{y}(k) = \boldsymbol{c}\left[\boldsymbol{x}(k)\right] + \boldsymbol{v}(k)\,, \tag{4.156}$$

$$\boldsymbol{w} \sim \mathcal{N}\left[\boldsymbol{0}, \boldsymbol{P}_{ww}(k-1)\right]\,, \tag{4.157}$$

$$\boldsymbol{v} \sim \mathcal{N}\left[\boldsymbol{0}, \boldsymbol{P}_{vv}(k)\right]\,. \tag{4.158}$$

2. The UKF is initialized as follows:

$$\hat{\boldsymbol{x}}(0|0) = \mathrm{E}\left[\boldsymbol{x}(0)\right]\,, \tag{4.159}$$

$$\tilde{\boldsymbol{P}}(0|0) = \mathrm{E}\left[\left(\boldsymbol{x}_0 - \hat{\boldsymbol{x}}(0|0)\right)\left(\boldsymbol{x}_0 - \hat{\boldsymbol{x}}(0|0)\right)^{\mathrm{T}}\right]\,. \tag{4.160}$$

3. The following time update equations are used to propagate the state estimate and covariance from one measurement time to the next:

   a) To propagate from time step $(k-1)$ to k, first choose sigma-points $\{\mathfrak{X}_i\}$ as before, with appropriate changes since the current best guess for the mean and the covariance of $\boldsymbol{x}$ are $\hat{\boldsymbol{x}}(k-1|k-1)$ and $\tilde{\boldsymbol{P}}(k-1|k-1)$:

$$\mathfrak{X}_0 = \hat{\boldsymbol{x}}(k-1|k-1)\,, \quad w_0 = \kappa(n+\kappa)^{-1}\,, \tag{4.161}$$

$$\mathfrak{X}_i = \hat{x}(k-1|k-1) + \left(\sqrt{(n+\kappa)\tilde{P}(k-1|k-1)}\right)_i^{\mathrm{T}}, \quad w_i = (2n+2\kappa)^{-1}, \tag{4.162}$$

$$\mathfrak{X}_{i+n} = \hat{x}(k-1|k-1) - \left(\sqrt{(n+\kappa)\tilde{P}(k-1|k-1)}\right)_i^{\mathrm{T}}, \quad w_{i+n} = (2n+2\kappa)^{-1}. \tag{4.163}$$

b) Use the known nonlinear process equation to transform the sigma-points into:

$$\mathfrak{X}_i(k|k-1) = \boldsymbol{a}\left[\mathfrak{X}_i(k-1|k-1)\right] + \boldsymbol{b}\left[\boldsymbol{u}(k-1)\right]. \tag{4.164}$$

c) Combine the $\mathfrak{X}_i(k|k-1)$ vectors to obtain the *a priori* at time k:

$$\hat{x}(k|k-1) = \sum_{i=0}^{2n} w_i \mathfrak{X}_i(k|k-1). \tag{4.165}$$

d) Estimate the *a priori* error covariance:

$$\tilde{P}(k|k-1) = \sum_{i=0}^{2n} w_i \tilde{\mathfrak{X}}_i(k|k-1)\tilde{\mathfrak{X}}_i^{\mathrm{T}}(k|k-1) + \boldsymbol{P}_{ww}(k-1), \tag{4.166}$$

where $\tilde{\mathfrak{X}}_i(k|k-1) = \mathfrak{X}_i(k|k-1) - \hat{x}(k|k-1)$.

4. Now that the time update equations are done, implement the measurement-update equations:

a) Choose sigma-points $\{\mathfrak{X}_i\}$, with appropriate changes since the current best guess for the mean and the covariance of $\boldsymbol{x}$ are $\hat{x}(k|k-1)$ and $\tilde{P}(k|k-1)$:

$$\hat{\mathfrak{X}}_0 = \hat{x}(k|k-1), \quad w_0 = \kappa(n+\kappa)^{-1}, \tag{4.167}$$

$$\hat{\mathfrak{X}}_i = \hat{x}(k|k-1) + \left(\sqrt{(n+\kappa)\tilde{P}(k|k-1)}\right)_i^{\mathrm{T}}, \quad w_i = (2n+2\kappa)^{-1}, \tag{4.168}$$

$$\hat{\mathfrak{X}}_{i+n} = \hat{x}(k|k-1) - \left(\sqrt{(n+\kappa)\tilde{P}(k|k-1)}\right)_i^{\mathrm{T}}, \quad w_{i+n} = (2n+2\kappa)^{-1}. \tag{4.169}$$

This step can be omitted, if desired. Instead of generating new sigma-points, it is possible to reuse the sigma-points that were obtained from the time update.

b) Use the known nonlinear measurement equation to transform the sigma-points into:

$$\mathfrak{X}(k|k-1) = \boldsymbol{c}\left[\hat{\mathfrak{X}}(k|k-1)\right]. \tag{4.170}$$

c) Combine the $\mathfrak{Y}_i(k|k-1)$ vectors to obtain the predicted measurement at time k:

$$\hat{y}(k|k-1) = \sum_{i=0}^{2n} w_i \mathfrak{Y}_i(k|k-1). \tag{4.171}$$

d) Estimate the covariance of the predicted measurement:

$$\boldsymbol{P}_{\xi\xi}(k|k-1) = \sum_{i=0}^{2n} w_i \boldsymbol{\xi}_i(k|k-1)\boldsymbol{\xi}_i^{\mathrm{T}}(k|k-1) + \boldsymbol{P}_{vv}(k), \tag{4.172}$$

where $\boldsymbol{\xi}_i(k|k-1) = \mathfrak{Y}_i(k|k-1) - \hat{y}(k|k-1)$.

e) Estimate the cross-covariance between $\hat{\boldsymbol{x}}(k|k-1)$ and $\boldsymbol{\xi}(k|k-1)$:

$$\boldsymbol{P}_{\tilde{x}\xi}(k|k-1) = \sum_{i=0}^{2n} w_i \tilde{\boldsymbol{\mathfrak{X}}}_i(k|k-1)\boldsymbol{\xi}_i^{\mathrm{T}}(k|k-1) + \boldsymbol{P}_{vv}(k)\,. \tag{4.173}$$

f) The measurement update of the state estimate can be performed using the normal KF equations:

$$\boldsymbol{K}(k) = \boldsymbol{P}_{\tilde{x}\xi}(k|k-1) - \boldsymbol{P}_{\xi\xi}(k|k-1)\,, \tag{4.174}$$

$$\boldsymbol{e}(k) = \boldsymbol{y}(k) - \hat{\boldsymbol{y}}(k|k-1)\,, \tag{4.175}$$

$$\hat{\boldsymbol{x}}(k|k) = \hat{\boldsymbol{x}}(k|k-1) + \boldsymbol{K}(k)\boldsymbol{e}(k)\,, \tag{4.176}$$

$$\tilde{\boldsymbol{P}}(k|k) = \tilde{\boldsymbol{P}}(k|k-1) - \boldsymbol{K}(k)\boldsymbol{P}_{\xi\xi}(k|k-1)\boldsymbol{K}^{\mathrm{T}}(k)\,. \tag{4.177}$$

The algorithm above assumes that the process and measurement equations are linear with respect to the noise [105].

Other modern Bayesian state-space processors include the Quadrature Kalman Filter [131–133], the Gaussian-sum Processor [134, 135], and the Ensemble Kalman Filter [136, 137], which is a hybrid between the UKF and the Particle Filter (PF).

4.5 Particle-based Bayesian state-space processors

As the UKF, the PF transforms a set of points via known nonlinear equations and combine the results to estimate the mean and the covariance of the state. However, in the PF, the points are randomly chosen, whereas in the UKF the points are chosen on the basis of a specific algorithm. Because of this, the number of points in a PF generally needs to be much greater than the number of points in a UKF. Another difference between these two filters is that the estimation error in a UKF does not converge to zero in any sense [105], by the estimation error in a PF does converge to zero as the number of particles approaches infinity.

Particle filters had their beginning in the 1940s with the work of Metropolis and Ulam [138]. Wiener [139] also suggested something much like particle filtering as early as 1940, but only since the 1980s, there has been adequate computational power for their implementation [105]. Although the early implementation occurred in particle physics [140], the term *particle* only dates back to Kitagawa [141], while Carpenter, Clifford, and Fearnhead [142] coined the term Particle Filter (PF). In signal processing, early occurrences of a PF can be traced back to Handschin and Mayne [143].

The particle filter is, according to Simon [105]:

> a statistical, brute-force approach to estimation that often works well for problems that are difficult for the conventional Kalman filter (i.e., systems that are highly nonlinear) [105, p. 461].

It is a sequential Monte Carlo (MC) methodology where the basic idea is the recursive computation of relevant probability distributions using the concepts of importance sampling and approximation of probability distributions with discrete random measures [99, 144, 145]. Sequential MC methods found limited use in the past, except for the last four decades [144], primarily due to their very high computational complexity and the lack of adequate computing resources of the time. The rapid advances of computers in the last several years and the outstanding potential of PF have made them a very active area of research.

Importance sampling plays a crucial role in state-space particle algorithm development. The PF does not involve linearization around current estimates, but rather approximations of the desired distributions by these discrete random measures, in contrast to the KF, which sequentially estimates the conditional mean and covariance used to characterize the filtering posterior.

The key idea is to represent the posterior distribution by a set of N random samples, the particles, with associated weights, $\{\boldsymbol{x}_i(k), \mathcal{W}_i(k); i = 1, \cdots, N\}$, and compute the required MC estimates. Of course, as the number of particles becomes very large, the MC representation becomes an equivalent characterization of the analytical description of the posterior distribution [99].

Particle filtering goes by many other names, including sequential importance sampling [107], bootstrap filtering [146], the condensation algorithm [147, 148], interacting particle approximations [149], Monte Carlo filtering [141], sequential Monte Carlo [150, 151], and survival of the fittest [152]. Furthermore, there is an obvious analogy between population MC and Genetic Algorithm (GA); however, there is an essential difference in the goals of these algorithms [153].

4.5.1 Importance sampling

Monte Carlo (MC) methods involve techniques to estimate the posterior distribution of interest using numerical integration-based methods or sample-based simulation methods which attempt to produce independent identically distributed samples from a targeted posterior distribution and use them to make statistical inferences.

The generation of random samples from a known distribution is essential for simulation. If the distribution is standard and has a closed analytic form (e.g., Gaussian), then it is usually possible to perform this simulation easily. This method is called the direct method because it evolves directly from the analytic form.

One way to mitigate difficulties with the inability to directly sample from a posterior distribution is based on the concept of importance sampling [26, 154, 155], which is a method to compute expectations with respect to one distribution using random samples drawn from another. That is, it is a method for simulating samples from a proposal distribution to be used to approximate a targeted posterior distribution by appropriate weighting [105].

The method is called importance sampling because it is based on so-called importance functions – although it would be more accurate to call it weighted sampling [103].

Importance sampling is a generalization of the MC approach.

Let $F(x)$ be a cdf, then its integral,

$$I = \int_{\mathcal{X}} F(x)dx\,, \tag{4.178}$$

can be rewritten as

$$I = \int_{\mathcal{X}} F(x)dx = \int_{\mathcal{X}} \frac{F(x)}{Q(x)} Q(x)dx\,, \tag{4.179}$$

for $\int Q(x)dx = 1$. The function $Q(x)$ is referred to as the sampling distribution or, more appropriately, the importance sampling distribution, since it samples the target distribution $F(x)$ nonuniformly giving more importance to some values of $F(x)$ than others.

Candy [99] states that the support of $Q(x)$ covers that of $F(x)$, or the samples drawn from $Q(\cdot)$ overlap the same region (or more) corresponding to the samples of $F(\cdot)$. Both functions $F(x)$ and $Q(x)$ are said to have the same support if

$$F(x) > 0 \Rightarrow Q(x) > 0, \quad \forall x \in \mathbb{R}^n\,, \tag{4.180}$$

which is a necessary condition for importance sampling to hold [99].

The integral in (4.179) can be estimated through the following procedure:

1. Drawing N samples from

   $$Q(x) : \mathcal{X}(i) \sim Q(x) \tag{4.181}$$

   and

   $$\hat{Q}(x) \approx \frac{1}{N} \sum_{i=1}^{N} \delta(x - \mathcal{X}(i))\,, \tag{4.182}$$

   where $\delta(\cdot)$ is the Dirac[13] delta function[14].

2. Computing the sample mean

   $$I = \mathrm{E}_Q\left[\frac{F(x)}{Q(x)}\right] \approx \int \left(\frac{F(x)}{Q(x)}\right) \times \frac{1}{N} \sum_{i=1}^{N} \delta(x - \mathcal{X}(i)) = \frac{1}{N} \sum_{i=1}^{N} \frac{F(\mathcal{X}(i))}{Q(\mathcal{X}(i))} \tag{4.183}$$

   with correspond error variance

   $$\mathrm{Var}\left\{\mathrm{E}_Q\left[\frac{F(x)}{Q(x)}\right]\right\} = \int \left(\frac{F(x)}{Q(x)} - I\right)^2 \times Q(x)dx\,. \tag{4.184}$$

[13] Paul Adrien Maurice Dirac (1902–1984) was an English theoretical physicist, who is regarded as one of the most significant physicists of the 20^{th} century. Dirac made fundamental contributions to the early development of both quantum mechanics and quantum electrodynamics. Among other discoveries, he formulated the Dirac equation which describes the behavior of fermions and predicted the existence of antimatter. He also made significant contributions to the reconciliation of general relativity with quantum mechanics.

[14] The Dirac delta function is used to model the density of an idealized point mass or point charge as a function equal to zero everywhere except for zero and whose integral over the entire real line is equal to one.

It is interesting to note that the MC approach provides an unbiased estimator with the corresponding error variance calculated from the above relation [99].

At long last, Robert and Casella [103] points out:

> Importance sampling methods can bring considerable improvement over naive Monte Carlo estimates when implemented with care. However, they can encounter disastrous performances and produce extremely poor estimates when the variance conditions are not met.

4.5.2 Importance sampling distributions

Selection of the importance distribution is a critical part of the design phase in particle filtering. Besides assuring that the distribution covers the posterior, there are a number of properties that can also be satisfied to achieve a robust design.

4.5.2.1 Minimum-variance importance distribution

The generic algorithm presented in Subsection 4.5.4 has a serious flaw: the variance of the importance weights increases over time [144, 156]. Therefore, the algorithm degenerates to a single non-zero weight after a few iterations. One way to limit degeneracy is to choose an importance distribution that minimizes the weight variance based on the available information.

It has been shown by Cappé, Godsill, and Moulines [156] that the minimum-variance importance distribution that minimizes the variance of the set of weights is given by

$$Q_{MV} \to \Pr\left[\boldsymbol{x}(k)|\boldsymbol{x}(k-1), \boldsymbol{y}(k)\right] . \tag{4.185}$$

Furthermore, Candy [99] proved that it can also be expressed as

$$Q_{MV} = \frac{\Pr\left[\boldsymbol{y}(k)|\boldsymbol{x}(k)\right] \times \Pr\left[\boldsymbol{x}(k)|\boldsymbol{x}(k-1)\right]}{\Pr\left[\boldsymbol{y}(k)|\boldsymbol{x}(k-1)\right]} . \tag{4.186}$$

Accordingly, the expression for the weight recursion considering the minimum-variance importance distribution is:

$$q(k) = q(k-1) \times \frac{\Pr\left[\boldsymbol{y}(k)|\boldsymbol{x}(k)\right] \times \Pr\left[\boldsymbol{x}(k)|\boldsymbol{x}(k-1)\right]}{Q_{MV}} , \tag{4.187}$$

which indicates that the importance weights can be calculated before the particles are propagated to time k. From that expression, it is possible to see the problem with the minimum-variance importance function approach: (1) it requires to sample from $\Pr\left[\boldsymbol{x}(k)|\boldsymbol{x}(k-1), \boldsymbol{y}(k)\right]$; and (2) it is necessary to evaluate the integral, which generally has not analytic form.

4.5.2.2 Transition-prior importance distribution

Another choice for an importance distribution if the transition prior. This prior is defined in terms of the state-space representation by

$$\mathcal{A}\left[\boldsymbol{x}(k)|\boldsymbol{x}(k-1)\right] \Leftarrow \boldsymbol{a}\left[\boldsymbol{x}(k-1)\right] + \boldsymbol{b}\left[\boldsymbol{u}(k-1)\right] + \boldsymbol{w}(k-1)\,, \tag{4.188}$$

which is dependent on the known excitation and process noise statistics and is given by

$$Q_{prior} \rightarrow \Pr\left[\boldsymbol{x}(k)|\boldsymbol{x}(k-1)\right]\,. \tag{4.189}$$

This choice leads to

$$\begin{aligned} q(k) &= q(k-1) \times \frac{\Pr\left[\boldsymbol{y}(k)|\boldsymbol{x}^p(k)\right] \times \Pr\left[\boldsymbol{x}(k)|\boldsymbol{x}^p(k-1)\right]}{Q_{prior}}\,, && \text{(4.190a)} \\ &= q(k-1) \times \mathcal{C}\left[\boldsymbol{y}(k)|\boldsymbol{x}^p(k)\right]\,, && \text{(4.190b)} \end{aligned}$$

since the priors cancel.

This choice of importance distribution has two properties: first, the weight does not use the most recent observation $y(k)$; and second, it does not use the past particle $\boldsymbol{x}^p(k-1)$, but only the likelihood. This choice is easily implemented and updated by simply evaluating the measurement likelihood, $\mathcal{C}\left[\boldsymbol{y}(k)|\boldsymbol{x}^p(k)\right], i = 1, \cdots, N$. In contrast to the minimum-variance choice, these weights require the particles to be propagated to time instant k before the weights can be calculated.

However, since the transition prior is not conditioned on the measurement data – especially the most recent – it fails to incorporate the latest available information from the most recent measurement to propose new values for the states and, therefore, leading to only a few particles having significant weights when their likelihood is calculated. The transition prior is a much broader distribution than the likelihood, indicating that only a few particles will be assigned a large weight [99]. Thus, the algorithm will degenerate rapidly and lead to poor performance especially when data outliers occur or measurement noise is small.

The aforementioned conditions lead to a mismatch between the prior prediction and posterior distribution. Techniques such as the Auxiliary Particle Filter [107, 145], as well as Local-Linearized Particle Filters [156, 157], have been developed to drive the particles to regions of high likelihood by incorporating the current measurement.

4.5.3 Resampling

The main objective in simulation-based sampling techniques is to generate independent and identically distributed samples from the targeted posterior distribution in order to perform statistical inferences extracting the desired information. Thus, importance weights are quite critical since they contain probabilistic information about each specific particle. In fact, they provide information about how probable a sample drawn from the target posterior has been.

Therefore, the weights can be considered acceptance probabilities enabling the generation of approximately independent samples from the posterior $\Pr[\boldsymbol{x}(k)|\mathcal{Y}(k)]$.

One of the major problems with importance sampling algorithms is particles depletion [107]. The degeneracy of the particle weights creates a problem that must be resolved before these algorithms can be of any pragmatic use. It occurs because the variance of the importance weights increases in time [99] thereby making it impossible to avoid this weight degradation. Degeneracy implies that a large computational effort is devoted to updating particles whose contribution to the posterior is negligible.

There is a need to, somehow, resolve the degeneracy problem to make the sequential simulation-based techniques viable. The solution is to resample the particles.

Resampling can be thought of as a realization of enhanced particles $\hat{\boldsymbol{x}}^i(k)$ extracted from the original samples $\boldsymbol{x}^p(k)$ based on their acceptance probability $q(k)$ at time k. Statistically,

$$\Pr\left[\hat{\boldsymbol{x}}^i(k) = \boldsymbol{x}^p(k)\right] = q(k)\,, \tag{4.191}$$

for $p = 1, \cdots, N$. The new set of particles $\left\{\hat{\boldsymbol{x}}^i(k)\right\}$ replace the old set $\{\boldsymbol{x}^p(k)\}$.

In summary, the fundamental concept in resampling theory is to preserve particles with large weights while discarding those with small weights. Two steps must occur to resample effectively [99]: (i) a decision, on a weight-by-weight basis, must be made to select the appropriate weights and reject the inappropriate; and (ii) resampling must be performed to minimize degeneracy. This overall strategy coupled with importance sampling is termed sequential sampling-importance-resampling [158].

A reasonable measure of degeneracy is the effective particle sample size based on the coefficient of variation [113] defined by

$$N_{eff}(k) \triangleq \frac{N}{\mathrm{E}\left[q^2(k)\right]} \leq N\,, \tag{4.192}$$

which can be estimated by

$$\hat{N}_{eff}(k) = \frac{1}{\sum_{i=1}^{N} q^2(k)}\,. \tag{4.193}$$

A decision based on the rejection method [111] is made by comparing it to a threshold N_{th}. That is, resampling only is performed when $\hat{N}_{eff}(k)$ is less than N_{th}.

There are a variety of techniques available to implement the basic resampling method [99]. The usual approach is to resample with replacement – the multinomial resampling method. A second more efficient way of generating independent and identical samples from the empirical posterior distribution is the systematic resampling method.

4.5.3.1 The multinomial resampling method

The multinomial resampling method resamples with replacement, since the probability of each particle $\boldsymbol{x}^p(k)$ is given by the normalized weight $q^p(k)$. Therefore, the number of times N^i

each particle in the original set, $\{\boldsymbol{x}^p(k)\}$, is selected follows a binomial distribution, $\mathcal{B}(N, q^p(k))$. The corresponding vector $\left[N^1, \cdots, N^N\right]^{\mathrm{T}}$ is distributed according to a multinomial distribution with parameter N and probability of success $\left[q^1(k), \cdots, q^N(k)\right]^{\mathrm{T}}$.

Within this resampling scheme, particles in the original set with small variance weights are most likely discarded, while those of high weights are replicated in proportion to these weights. The multinomial resampling method is given by:

GIVEN a set of particles and weights at time k, $\{\boldsymbol{x}^p(k), q^p(k)\}\,; p = 1, \cdots, N$;

SAMPLE uniformly $u^m = \mathcal{U}(0,1); m = 1, \cdots, N$;

DETERMINE the index p^m : $p^m = m$ for $\Pr\left[\boldsymbol{x}^{p^m}(k) = \boldsymbol{x}^m(k)\right] = u^m$;

SELECT a new sample $\hat{\boldsymbol{x}}^{p^m} \Rightarrow \boldsymbol{x}^p(k)$ and weight $\hat{q}^{p^m}(k) = 1/N$ based on the new sample index p^m; and

GENERATE a new random (resampled) measure: $\left\{\hat{\boldsymbol{x}}^{p^m}, \hat{q}^{p^m}(k)\right\}; m = 1, \cdots, N$.

The index notation p^m designates the original pth particle or parent and the new mth particle. This sampling scheme is equivalent to drawing $p^m, m = 1, \cdots, N$ samples from a multinomial distribution with parameters $\mathcal{M}(N^{pm}, q^{pm}(k))$ and corresponding statistics.

4.5.3.2 The systematic resampling method

The systematic resampling method is based on an ordered technique in which a set of N-ordered uniform variates are generated [88, 159]. It minimizes the error variance between the original selected sample and its mean. The systematic resampling method is given by:

GIVEN a set of particles and weights at time k, $\{\boldsymbol{x}^p(k), q^p(k)\}\,; p = 1, \cdots, N$;

SAMPLE uniform N-ordered variates: $\hat{u}^m = u^m + \dfrac{m-1}{N}, m = 1, \cdots, N$ and $u^m \sim \mathcal{U}(0,1)$;

DETERMINE the index p^m : $p^m = m$ for $\Pr\left[\boldsymbol{x}^{m-1}(k)\right] < \hat{u}^m < \Pr\left[\boldsymbol{x}^m(k)\right]$;

SELECT

p^m; and

GENERATE a new random (resampled) measure: $\left\{\hat{\boldsymbol{x}}^{p^m}, \hat{q}^{p^m}(k)\right\}; m = 1, \cdots, N$.

Recall that the cdf is given by: $\Pr\left[\boldsymbol{x}^m(k)\right] = \sum_{m=1}^{N} q^m(k)\mu\left[\boldsymbol{x}(k) - \boldsymbol{x}_m(k)\right]$, with $\mu(\cdot)$ is a unit-step function.

■

According to Simon [105], although resampling is a very important technique to decrease the degeneracy problem, it introduces its own problems. Sample impoverishment is one of them. It occurs when the region of state space in which the pdf $f(\boldsymbol{y}(k)|\boldsymbol{x}(k))$ has significant values does not overlap the pdf $f(\boldsymbol{x}(k)|\boldsymbol{y}(k-1))$. This means that if all the *a priori* particles are distributed according to $f(\boldsymbol{x}(k)|\boldsymbol{y}(k-1))$, and then the computed pdf, $f(\boldsymbol{y}(k)|\boldsymbol{x}(k))$, is used to resample the particles, only a few particles will be resampled to become *a posteriori* particles.

Eventually, all particles will collapse to the same value[15].

Sample impoverishment can be overcome by a brute-force method of simply increasing the number of particles, but this can quickly lead to unreasonable computational demands, and often simply delays the inevitable sample impoverishment.

Furthermore, after one resampling step, the simulated trajectories are no longer statistically independent [99]. Therefore, the simple theoretical convergence results under these assumptions lose their validity. Pragmatically, resampling can limit algorithm parallelization because combining particles causes an increase in computational complexity.

Several remedies have been proposed in the literature, including roughening [107], prior editing [146], regularized particle filtering [160, 161], Markov chain Monte Carlo (MC) resampling [162], and auxiliary particle filtering [163].

4.5.3.3 Roughening

Roughening can be used to prevent sample impoverishment. In this method, random noise is added to each particle after the resampling process. This is similar to adding artificial process noise to the KF. In the roughening approach, the *a posteriori* particles are modified as follows:

$$\hat{\boldsymbol{x}}(k|k) = \hat{\boldsymbol{x}}(k|k) + \Xi(k)\,, \tag{4.194a}$$

$$\Xi(k) \sim \mathcal{N}(0, \alpha \boldsymbol{b} N^{-1/n})\,, \tag{4.194b}$$

where Ξ is a zero-mean Gaussian random variable; α is a scalar tuning parameter; N is the number of particles; n is the dimension of the state space; and $\boldsymbol{b}$ is a vector containing the maximum difference between the particle elements before roughening. The vector $\boldsymbol{b}$ is given as

$$\boldsymbol{b} = \max_{p,m} |\hat{\boldsymbol{x}}^p(k|k) - \hat{\boldsymbol{x}}^m(k|k)|\ , \tag{4.195}$$

where p and m are particle numbers. Further, α is a tuning parameter that specifies the amount of jitter that is added to each particle. Gordon, Salmond, and Smith [146] recommends the use of $\alpha = 0.2$.

4.5.3.4 Prior editing

If roughening does not prevent sample impoverishment, the prior editing can be tried [105]. This involves rejection of an *a priori* sample if it is in a region of state-space with small $q^p(k)$. If an *a priori* sample is in a region of small probability, then it can be roughened as many times as necessary, using a procedure like (4.194a)–(4.194b), until it is in a region of significant probability.

[15] This is called the black hole of particle filtering [105].

In Gordon, Salmond, and Smith [146], prior editing is implemented as follows: if $|\boldsymbol{y}(k) - \mathcal{C}\left[\boldsymbol{y}(k)|\boldsymbol{x}^p(k|k-1)\right]|$ is more than six standard deviations of the measurement noise, then it is highly unlikely to be selected as an *a posteriori* particle. In this case, $\boldsymbol{x}^p(k-1|k-1)$ is roughened and then passed through the system equation again to obtain a new $\boldsymbol{x}^p(k|k-1)$. This is repeated as many times as necessary until $\boldsymbol{x}^p(k|k-1)$ is in a region of nonnegligible probability.

4.5.3.5 Regularized particle filtering

Another solution to the diversity problem is to develop a continuous rather than discrete approximation to the empirical posterior distribution using a kernel density estimator [161] and then perform resampling directly from it.

The key idea of the regularized particle filter is the transformation of the discrete empirical posterior distribution in order to resample from an absolutely continuous distribution producing a new set of N-particles with different locations.

4.5.3.6 Markov chain Monte Carlo resampling

Another approach for preventing sample impoverishment is the Markov chain MC move step. This approach moves the *a priori* particle to a new randomly generated state if a uniformly distributed random number is less than an acceptance probability. The acceptance probability is computed as the probability that the *a priori* sample is consistent with the measurement, relative to the probability that the resampled state is consistent with the measurement.

Many different Markov chain MC techniques – Metropolis-Hastings [164, 165], Random Walk Metropolis-Hastings [113], Gibbs [166], Slice [103], and so on – can be used to perform the move step.

4.5.3.7 Auxiliary particle filtering

Another approach to evening out the probability of *a priori* particles (and thus increasing diversity in *a posteriori* particles) is called the auxiliary particle filter. It was first proposed by Pitt and Shephard [163].

This approach was developed by augmenting each *a priori* particle by one element (an auxiliary variable). This increases the dimension of the problem and thus adds a degree of freedom to the choice of the resampling weights, which allows them to be more evenly distributed.

The idea is to perform resampling at time $(k-1)$ using the available measurement at time k before the particles are propagated to time k through the transition and likelihood distributions. The key step is to favor particles at time $(k-1)$ that are likely to survive (largest weights) at the next time step k.

The auxiliary particle filter also attempts to mitigate poor outlier performance and poor posterior tail performance. These problems evolve from the empirical approximation of the filtering posterior which can be considered a mixture distribution.

■

Further details on the aforementioned remedies can be found in the works of Candy [99], Doucet, Freitas, and Gordon [107], and Simon [105].

In this work, the Bootstrap Particle Filter was implemented to deal with the salient-pole synchronous machines states and parameters estimation problem. Along with it, the roughening remedy and the random-walk move were also implemented. Due to schedule issues, the performance of other Markov chain MC techniques and the auxiliary particle filter will be evaluated in future works.

4.5.4 The Bootstrap Particle Filter

The basic bootstrap algorithm developed by Gordon, Salmond, and Smith [146] is one of the first practical implementations of the processor to the tracking problem. It is the most heavily applied of all Particle Filter (PF) techniques due to its simplicity [99]. It is based on sequential sampling-importance-resampling ideas and uses the transition prior of (4.189) as its underlying proposal distribution.

The corresponding weight becomes quite simple and only depends on the likelihood; therefore, it is not even necessary to perform a sequential updating because

$$q(k) = q(k-1) \times \mathcal{C}\left[\boldsymbol{y}(k)|\boldsymbol{x}^p(k)\right] , \tag{4.196}$$

since the filter requires resampling to mitigate variance (weight) increases at each time-step [146]. After resampling, the new weights become

$$q(k) = \frac{1}{N} . \tag{4.197}$$

Furthermore, it is important to mention that, in order to achieve convergence, it is necessary to resample at every time-step.

The Bootstrap Particle Filter can be summarized [99, 105] in the following steps.

1. The system is given by:

$$\boldsymbol{x}(k) = \boldsymbol{a}\left[\boldsymbol{x}(k-1)\right] + \boldsymbol{b}\left[\boldsymbol{u}(k-1)\right] + \boldsymbol{w}(k-1) , \tag{4.198}$$

$$\boldsymbol{y}(k) = \boldsymbol{c}\left[\boldsymbol{x}(k)\right] + \boldsymbol{v}(k) , \tag{4.199}$$

$$\boldsymbol{w} \sim \mathcal{N}\left[\boldsymbol{0}, \boldsymbol{P}_{ww}(k-1)\right] , \tag{4.200}$$

$$\boldsymbol{v} \sim \mathcal{N}\left[\boldsymbol{0}, \boldsymbol{P}_{vv}(k)\right] . \tag{4.201}$$

2. Assuming that the pdf of the initial state is known, $f(\boldsymbol{x}(0))$, randomly generate N initial particles on the basis of $f(\boldsymbol{x}(0))$. These particles are denoted $\boldsymbol{x}(0|0)$. The parameter N is chosen by the user as a trade-off between computational effort and estimation accuracy.

3. For $k = 1, 2, \cdots$, do the following:

   a) Perform the time propagation step to obtain the *a priori* particles $\hat{\boldsymbol{x}}^p(k|k-1)$, using the known process equation and the known pdf of the process noise:

$$\hat{\boldsymbol{x}}^p(k|k-1) = \boldsymbol{a}\left[\hat{\boldsymbol{x}}^p(k-1|k-1)\right] + \boldsymbol{b}\left[\boldsymbol{u}(k-1)\right] + \boldsymbol{w}^p(k-1), \quad p = 1, \ldots, N, \tag{4.202}$$

   where each $\boldsymbol{w}^p(k-1)$ is randomly generated on the basis of the known pdf of $\boldsymbol{w}(k-1)$.

   b) Compute the relative likelihood $q^p(k)$ of each particle $\hat{\boldsymbol{x}}^p(k|k-1)$ conditioned on the measurement $\boldsymbol{y}(k)$. This is done by evaluating the cdf $\Pr\left[y(k)|\hat{\boldsymbol{x}}^p(k|k-1)\right]$ on the basis of the nonlinear measurement equation and the cdf of the measurement noise;

   c) Scale the relative likelihoods obtained in the previous step as follows:

$$q^p(k) = \frac{q^p(k)}{\sum_{j=1}^{N} q^j(k)}. \tag{4.203}$$

   Now, all likelihoods sum to one.

   d) Generate a set of *a posteriori* particles $\hat{\boldsymbol{x}}^p(k|k)$ on the basis of the relative likelihoods $q^p(k)$ by means of the multinomial resampling method or the systematic resampling method described in previous sections.

   e) Now that the set of particles $\hat{\boldsymbol{x}}^p(k|k)$ is distributed according to the cdf $\Pr\left[\hat{\boldsymbol{x}}(k|k)|\boldsymbol{y}(k)\right]$, it is possible to compute any desired statistical measure of this cdf.

4.5.4.1 Example

Let a scalar system be given by the following equations:

$$x(k) = \frac{1}{2}x(k-1) + \frac{25x(k-1)}{1+x^2(k-1)} + 8\cos\left[1.2(k-1)\right] + w(k-1), \tag{4.204}$$

$$y(k) = \frac{1}{20}x^2(k) + v(k), \tag{4.205}$$

where $w(k)$ and $v(k)$ are zero-mean Gaussian white noise sequences, both with variance equal to one. This system has become a benchmark in the nonlinear estimation literature [159, 167]. The high degree of nonlinearity in both process and measurement equations makes this a difficult state estimation problem for a KF.

Let the initial state be $x(0|0) = 0.1000$, the initial state estimate be $\hat{x}(0) = x(0|0)$, and the initial estimation covariance be $P(0|0) = 2$. The EKF and the PF will be used to estimate the state x.

For the sake of simplicity, only the first iteration of each algorithm will be presented. As for the PF, only two particles will be presented.

Extended Kalman Filter

1. Perform the time update of the state estimate:

$$\hat{x}(1|0) = \boldsymbol{a}\left[\hat{\boldsymbol{x}}(0|0)\right] , \tag{4.206a}$$
$$= \frac{1}{2}0.1 + \frac{25 \times 0.1}{1+0.1^2} + 8\cos(0) + w(0) \tag{4.206b}$$
$$= 10.5252 \;+\; 0.5377 = 11.0629\,. \tag{4.206c}$$

2. Compute the following Jacobian matrix:

$$\boldsymbol{A}\left[\hat{x}(1|0)\right] = \frac{da\left[\hat{x}(1|0)\right]}{d\hat{x}(1|0)} \tag{4.207a}$$
$$= \frac{1}{2} - 25\frac{\hat{x}(1|0)^2 - 1}{\left(1+\hat{x}(1|0)^2\right)^2} \tag{4.207b}$$
$$= 0.3007\,. \tag{4.207c}$$

3. Perform the time update of the estimation-error covariance:

$$\tilde{\boldsymbol{P}}(1|0) = \boldsymbol{A}\left[\hat{x}(1|0)\right]\tilde{\boldsymbol{P}}(0|0)\boldsymbol{A}\left[\hat{x}(1|0)\right]^{\mathrm{T}} - \boldsymbol{P}_{ww}(0) \tag{4.208a}$$
$$= 0.3007 \times 2 \times 0.3007 \tag{4.208b}$$
$$= 0.1808\,. \tag{4.208c}$$

4. Compute the following Jacobian matrix:

$$\boldsymbol{C}\left[\hat{x}(1|0)\right] = \frac{dc\left[\hat{x}(1|0)\right]}{d\hat{x}(1|0)} \tag{4.209a}$$
$$= \frac{1}{10}\hat{x}(1|0) + v(1) \tag{4.209b}$$
$$= \frac{1}{10}11.0629 - 2.2588 = -1.1525\,. \tag{4.209c}$$

5. Perform the measurement update of the state estimate and estimation-error covariance:

$$e(1) = y(1) - \hat{y}(1|0) \tag{4.210a}$$
$$= 10 + 1.1525 = 11.1525\,. \tag{4.210b}$$
$$\boldsymbol{P}_{ee}(1) = \boldsymbol{C}\left[\hat{x}(1|0)\right]\tilde{\boldsymbol{P}}(1|0)\boldsymbol{C}\left[\hat{x}(1|0)\right]^{\mathrm{T}} + \boldsymbol{P}_{vv}(1) \tag{4.211a}$$
$$= -1.1525 \times 0.1808 \times (-1.1525) + 0.8622 = 1.1023\,. \tag{4.211b}$$
$$\boldsymbol{K}(1) = \tilde{\boldsymbol{P}}(1|0)\boldsymbol{C}\left[\hat{x}(1|0)\right]^{\mathrm{T}}\boldsymbol{P}_{ee}{}^{-1}(1) \tag{4.212a}$$
$$= 0.1808 \times (-1.1525) \times (1.1023)^{-1} = -0.1890\,. \tag{4.212b}$$
$$\hat{x}(1|1) = \hat{x}(1|0) + \boldsymbol{K}(1)e(1) \tag{4.213a}$$
$$= 11.0629 + (-0.1890) \times (11.1525) = 8.9551\,. \tag{4.213b}$$
$$\tilde{\boldsymbol{P}}(1|1) = \left[\boldsymbol{I} - \boldsymbol{K}(1)\boldsymbol{C}\left[\hat{x}(1|0)\right]\right]\tilde{\boldsymbol{P}}(1|0) \tag{4.214a}$$
$$= \left[1 - (-0.1890) \times (-1.1525)\right]0.1808 = 0.1414\,. \tag{4.214b}$$

Particle Filter

1. Perform the time propagation step to obtain the *a priori* particles:

$$\hat{x}^p(1|0) = \boldsymbol{a}\left[\hat{x}^p(0|0)\right] + \boldsymbol{w}^p(0) \tag{4.215a}$$

$$\hat{x}^1(1|0) = \frac{1}{2}0.1 + \frac{25 \times 0.1}{1 + 0.1^2} + 8\cos(0) + w(0) \tag{4.215b}$$

$$= 10.5252 + 0.3188 = 10.8440 \tag{4.215c}$$

$$\hat{x}^2(1|0) = 10.5252 - 0.4336 = 10.0916 \tag{4.215d}$$

2. Compute the relative likelihood $q^p(1)$ of each particle conditioned on the current measurement $\boldsymbol{y}(1)$:

$$q^1(1) = \exp\left\{-\frac{1}{2}\left(y(1) - \mathcal{C}\left[y(1)|x^1(1)\right]\right) \times 1 \times \left(y(1) - \mathcal{C}\left[y(1)|x^1(1)\right]\right)^{\mathrm{T}}\right\} \tag{4.216a}$$

$$= \exp\left\{-\frac{1}{2}(10 - 15.7080) \times 1 \times (10 - 15.7080)^{\mathrm{T}}\right\} \tag{4.216b}$$

$$= 8.4153 \times 10^{-8} \tag{4.216c}$$

$$q^2(1) = \exp\left\{-\frac{1}{2}(10 - 15.4990) \times 1 \times (10 - 15.4990)^{\mathrm{T}}\right\} \tag{4.216d}$$

$$= 2.7145 \times 10^{-7} \tag{4.216e}$$

3. Scale the relative likelihoods obtained in the previous step as follows:

$$q^p(1) = \frac{q^p(1)}{\sum_{j=1}^{N} q^j(1)} \tag{4.217a}$$

$$q^1(1) = \frac{8.4153 \times 10^{-8}}{3.5560 \times 10^{-7}} = 0.2367 \tag{4.217b}$$

$$q^2(1) = \frac{2.7145 \times 10^{-7}}{3.5560 \times 10^{-7}} = 0.7633 \tag{4.217c}$$

   Now, all likelihoods sum to one.

4. Generate a set of *a posteriori* particles $\hat{\boldsymbol{x}}^p(1|1)$ on the basis of the relative likelihoods $q^p(1)$:

$$\hat{x}^1(1|0) = 10.8440 \tag{4.218a}$$

$$\hat{x}^2(1|0) = 10.0916 \tag{4.218b}$$

5. Compute any desired statistical measure:

$$\hat{x}(1|0) = 0.2367 \times 10.8440 + 0.7633 \times 10.0916 = 10.2697\,. \tag{4.219}$$

Using a simulation length of 50 time-steps and 100 particles in the PF, not only the EKF estimate is poor, but the EKF thinks (on the basis of the computed covariance) that the estimate is much better than it really is. The true state is usually farther away from the estimated state than the 95% confidence measure of the EKF.

On the other hand, the PF does a nice job of estimating the state for this example. The rms estimation errors for the Kalman and the particle filters were 16.3 and 2.6, respectively.

4.5.5 Parameter estimation

Methods based on numerical integration [159] are limited only to models with relatively low state dimension. With the development of algorithms such as the Bootstrap Particle Filter, it became possible to use high-dimensional nonlinear non-Gaussian state-space models for the analysis of complex systems [167]. Nevertheless, a very important question remained [168]: how to operate it without the knowledge of system parameters? Before the development of self-organizing models [167], precise maximum likelihood parameter estimates could only be obtained by using a very large number of particles or by parallel application of many MC filters.

In Kitagawa's [167] proposal, the unknown parameters of the model are appended to the state vector, and both the state and the parameters are estimated simultaneously by the recursive filter. Earlier attempts failed. Anderson and Moore [169] even stated:

> Although [the] Extended Kalman Filter (EKF) approach [to parameter estimation] appears perfectly straightforward, experience has shown that with the usual state-space model, it does not work well in practice [169, p. 284].

Let a non-Gaussian nonlinear state-space model be:

$$\boldsymbol{x}(k) = \boldsymbol{a}\left[\boldsymbol{x}(k-1)\right] + \boldsymbol{b}\left[\boldsymbol{u}(k-1)\right] + \boldsymbol{w}(k-1)\,, \tag{4.220}$$

$$\boldsymbol{y}(k) = \boldsymbol{c}\left[\boldsymbol{x}(k)\right] + \boldsymbol{v}(k)\,, \tag{4.221}$$

$$\boldsymbol{w} \sim \mathcal{N}\left[\boldsymbol{0}, \boldsymbol{P}_{ww}(k-1)\right]\,, \tag{4.222}$$

$$\boldsymbol{v} \sim \mathcal{N}\left[\boldsymbol{0}, \boldsymbol{P}_{vv}(k)\right]\,, \tag{4.223}$$

$$\boldsymbol{x}(0) \sim f(\boldsymbol{x}(0))\,. \tag{4.224}$$

The possibly nonlinear functions $\boldsymbol{a}, \boldsymbol{b}$, and $\boldsymbol{c}$ may contain some parameters. The vector consisting of these unknown parameters is hereafter denoted by $\boldsymbol{\theta}$. An augmented state vector is now considered:

$$\boldsymbol{z}(k) = \begin{bmatrix} \boldsymbol{x}(k) \\ \boldsymbol{\theta}(k) \end{bmatrix}\,. \tag{4.225}$$

Therefore, the state-space model[16] for the augmented state vector $\boldsymbol{z}(k)$ is given by:

$$\boldsymbol{z}(k) = \boldsymbol{A}\left[\boldsymbol{z}(k-1)\right] + \boldsymbol{b}\left[\boldsymbol{u}(k-1)\right] + \boldsymbol{w}(k-1)\,, \tag{4.226}$$

$$\boldsymbol{y}(k) = \boldsymbol{C}\left[\boldsymbol{z}(k)\right] + \boldsymbol{v}(k)\,, \tag{4.227}$$

$$\boldsymbol{w} \sim \mathcal{N}\left[\boldsymbol{0}, \boldsymbol{P}_{ww}(k-1)\right]\,, \tag{4.228}$$

$$\boldsymbol{v} \sim \mathcal{N}\left[\boldsymbol{0}, \boldsymbol{P}_{vv}(k)\right]\,, \tag{4.229}$$

$$\boldsymbol{z}(0) \sim f(\boldsymbol{x}(0), \boldsymbol{\theta}(0))\,. \tag{4.230}$$

Since the original state vector and the parameter vector are included in the augmented state vector, it immediately yields the marginal posterior densities of both the parameter and of the original state [107].

[16] Kitagawa [167] calls this model "self-organizing state-space model".

This method of simultaneous estimation of parameters and states can be easily extended to time-varying parameter situation where the parameter evolves with time. Actually, the original formulation of the self-organizing state-space does not work well when MC filters and smoothers are used. This is because, since the parameters do not have their own system noises, the distribution gradually collapses as time proceeds [107]. In that case, by allowing the parameter to change gradually, namely by assuming the random-walk model

$$\boldsymbol{\theta}(k) = \boldsymbol{\theta}(k-1) + \boldsymbol{\xi}(k-1)\,, \tag{4.231}$$

where $\boldsymbol{\xi}$ is a zero-mean Gaussian random vector, a reasonable estimate of the parameter is achieved.

The key motivating idea is that the artificial evolution provides the mechanism for generating new parameter values at each time step in the simulation analysis, so helping to address sample attrition in reweighting methods that stay with the same sets of parameter points between time steps. However, this method has its drawbacks. If one adopts a model in which all parameters are subject to independent random shocks at each time point, the precision of resulting inferences is inevitably limited – the resulting posteriors are, eventually, far too diffuse relative to the theoretical posteriors for the actual fixed parameters [107].

Despite the aforementioned disadvantages, in this work, the parameters are modeled by means of random-walk models due to their simplicity.

Chapter 5

State of the Art on Synchronous Machine Parameters Estimation

"We are like dwarfs on the shoulders of giants, by whose grace we see farther than they. Our study of the works of the ancients enables us to give fresh life to their finer ideas and rescue them from time's oblivion and man's neglect."
— Peter of Blois[1], *writing in the late twelfth century*

The problem of building a mathematical model of a given system may be, basically, approached by two different ways: modeling and system identification [170]. Modeling is also called white-box approach and depends on prior knowledge of the system and the physical principles that describe it. The models resulted from this approach correspond to a direct-mathematical representation between the inputs and outputs of the system. Due to the complex nature of some problems, unknown equations, and the required time to model them, the white-box modeling is not always feasible [171].

On the other hand, system identification consists in determining a dynamic model that describes the input–output data measured from some process, as well as some parametrization and experimental conditions under which an estimated model would converge to a best approximation of the actual system [172]. The concept of best approximation is relative and depends upon which characteristics the estimated model must represent. The products of an identification problem are a model and a set of parameters, which are just a vehicle for describing the model [123] and may or may not represent the physical parameters of an actual system [173]. The traditional approaches to system identification problems may be grouped under two different categories: black- and gray-box methods [171].

One of the main characteristics of black-box identification (or, empirical modeling) is that it requires little or no prior knowledge of the system. The determined function works only as a mathematical structure capable of describing a cause-and-effect relation with parameters that do not represent the physical parameters of the system [171]. The black-box approach is based on experimental data and results in a description of the data used in the identification.

[1]Peter of Blois (1130–1211) was a French cleric, theologian, poet, and diplomat. He is particularly noted for his corpus of Latin letters.

Therefore, replications of the experiments may result in quite different models [174].

The gray-box method comprehends the determination of parameters of known-structures equations from experimental data and consists in a hybrid method between the white- and black-box approaches [174]. This identification method is advantageous since physical knowledge reduces the model-space that must be searched, which, in turn, preserves the validity of statistical methods and helps to prevent overfitting [175]. The gray-box techniques cover many different methods, from elaborate experiments specifically and carefully designed to yield certain information, to simple transient-response measurement [170].

This work aims at estimating the physical parameters of salient-pole synchronous machines by means of operational and/or experimental data. Therefore, methods under the gray-box identification category are to be considered.

5.1 Important challenges in modeling synchronous machine

The modeling of synchronous machines presents structured and unstructured nonlinearities [176]. Structured nonlinearities correspond to those modeled in the structure of the synchronous machine model, such as the sine and cosine functions of the rotor angle. On the other hand, the unstructured nonlinearities refer to the nonlinearities that are not modeled, such as the magnetic saturation of the iron parts of the rotor and stator [177]. Although some attempts have been made to define some model structures for magnetic saturation [74, 75], no unique nonlinear structure seems to be available to define the system behavior over the full operation range when dealing with a practical synchronous generator with dramatic changes in the operating conditions [176].

Another important challenge refers to determining the load angle. Acquiring accurate measurements of the required load angle is not an easy task. As stated by Giesbrecht and Meneses [18]:

> It is necessary to introduce a system [178] to synchronize armature voltage measurements to the measurements of a shaft positioning system. This device may not be available in some power plants and its installation may require a machine outage, which is certainly not desirable. Simpler techniques that are usually applied to machines with a small number of poles, such as attaching a black and white striped paper around the shaft using an optical sensor to detect the angle, are not accurate for low head hydrogenerators with a large number of poles, where the electrical angle is just a small portion of the mechanical angle [18, p. 5051].

Some works have dealt with this issue by using Phasor Measurement Unit (PMU) measurements [16, 179–184], which dates back to the work of Phadke, Thorp, and Adamiak [185]. According to Ma, Makarov, and Dong [186]:

> The PMU is a digital equipment that records the magnitude and phase angles of currents and voltages in a power system at a very high speed (usually 30 measurements per second). They can be used to provide real-time power system information [...], such information is particularly useful when the system

is in a stressed operating state or subject to potential system instability [186, p. 34].

The synchronous machine load angle may be calculated using voltages and currents measurements obtained from PMUs placed at the terminals of the generator buses [187]. Synchrophasors, which are obtained from PMU measurements, are the state-of-the-art in evaluating power system dynamic performance.

5.2 Synchronous machine identification methods

The determination of synchronous machine parameters is directly associated with its state of operation. In some methods, it is required the machine to be taken out of operation (or, to be taken off-line) so that tests and other procedures can be applied. On the other hand, the most recent approaches aim to avoid the disturbance created by halting the power generation and, therefore, seek to estimate the parameters by means of on-line measurements (that is, with the machine in operation). Both paradigms are discussed in the following.

5.2.1 Off-line procedures

The traditional methods to determine the performance characteristics and parameters of synchronous machines are fully described in IEEE [13]. These methods are off-line approaches to the identification problem. Among them, the most commonly performed are the sudden short-circuit [10, 188], load rejection [18, 66, 189], standstill frequency response [190, 191], and low-slip [192]. These tests, despite of being able to determine the physical parameters of synchronous generators, require a high implementation time, present complex executions, and require the machine to be left out of operation [176].

5.2.1.1 The sudden short-circuit test

The sudden short-circuit test was developed by Wright [10], in 1931, and received enhancements upon the work of Shackshaft [80], in 1974. Shackshaft and Poray [64] and Shackshaft [80] provided important advances regarding new definitions for machine parameters, such as transient reactance and transient time constants, which enabled a more accurate calculation of reactances and resistances.

The sudden short-circuit test is one of the oldest and most familiar methods to obtain information on the transient performance of synchronous machines [193] and its mechanical integrity [13]. Although the sudden short-circuit test is harmful to the machine, it is commonly applied due to its high recommendation in the main standard [13]. Critics of the sudden short-circuit test also underline the difficulty in recovering the quadrature-axis parameters and the complexity of the numerical nonlinear problem arising in the estimation of the time constants [194].

In a recent study [195], the sudden short-circuit test was applied to a wounded-field salient-pole nine-phase synchronous machine and showed good results for all the stator self- and mutual leakage inductances.

5.2.1.2 The load-rejection test

The load-rejection test was developed by de Mello and Ribeiro [65], in 1977, and validated by de Mello and Hannett [66], in 1981. It is a particular type of decrement test for determining generator parameters while the field voltage is kept constant. The convenient time to perform the load-rejection test is when the machine is being taken out of service or for an outage. Although the IEEE standard [13] indicates that the condition for the quadrature-axis load rejection test is difficult to obtain if a power angle indicator is unavailable, recent studies [18, 81, 196] have shown a method to determine the load conditions to ensure that the armature currents are completely on the quadrature-axis.

The theoretical foundation for performing the load rejection test was developed in Section 2.10. Once all measurements are obtained, the traditional methodology [13, 65, 66] consists of a graphical approach for estimating reactances and time constants. A great variety of works [18, 65, 66, 81, 188, 189, 196–205] have dealt with load-rejection procedures.

Wamkeue, Jolette, and Kairous [205] apply the Asymptotic Weighted Least-Squares Estimator to adjust the time-variant responses to the actual data, along with a Newton-type optimization algorithm. Silva, Bortoni, and Rocha [189] use the well-known Genetic Algorithm (GA) to approximate random variables to the machine constants, but it is not able to perform a pure quadrature-load rejection. The most recent work available in the literature [199] attempted to compare short-circuit and load-rejection tests results but failed to obtain an unsaturated operating condition and found inconsistent results for the parameters.

5.2.1.3 The standstill frequency response test

The standstill frequency response test was developed by Coultes and Watson [191], in 1981. The test involves exciting the stator or the field of the machine when the machine is off-line and at standstill. The operational parameters of the machine, which are required to derive the complete model, can be obtained from it.

Among the advantages claimed for the frequency response approach at standstill are that the test is safe, provides information about both quadrature and direct axes, and is inexpensive to perform [206]. Critics of this method have pointed out that the effects of saturation, the centrifugal force on damper windings, and the machine end-winding magnetic are not characterized because the machine is kept stationary and the signal levels are well-below the machine rating [206]. These effects have to be determined under loaded conditions [207]. Furthermore, significant errors result from the data-reduction process of fitting an equivalent circuit to the frequency response data [193].

Furthermore, the standstill frequency response test provides extremely-low currents, does not consider the rotational effects – which implies in coupling between direct- and quadrature-axis –, nor the centrifugal force in the amortisseur windings; therefore, it does not allow observing saturation effects [206].

Until recent years, the majority of available standstill frequency response tests in the literature focused on turbo-generators (i.e., round rotors synchronous machines) [68, 191, 208–211]. Only in 2019, Belqorchi et al. [212] successfully applied the standstill frequency response test to a large salient-pole synchronous machine.

5.2.1.4 The low-slip test

The literature review shows that the first mention to the low-slip test is due to the work of Park and Robertson [213], in 1928. In this test, the machine to be tested is left unexcited and a three-phase voltage is applied to its armature. The rotor is coupled to a driving motor, of sufficient capacity to overcome the reluctance torque at the reduced voltage and is run at a low value of slip. In this manner, the poles are slipped past the mmf wave. The magnetizing current is then a function of the rotor position with respect to the mmf.

However, it is difficult to maintain constant speed when the slip is sufficiently low for an accurate determination of the quadrature-axis synchronous reactance because the effects of salient poles and the currents induced in the amortisseur winding produce a pulsating torque [13]. Some adjustments have been proposed to this test [192, 214] and present accurate information under certain assumptions. This test is also unfeasible for large synchronous generators coupled to hydraulic or steam turbines.

■

Despite the aforementioned disadvantages, a series of studies performs off-line estimation of synchronous generators parameters [77, 188, 189, 201, 203, 215]. Zaker, Gharehpetian, and Karrari [77] use a seventh order model and estimate the parameters by means of two load rejection tests (one for each axis). Wamkeue, Kamwa, and Dai-Do [215] use a generic model to develop a parameter estimation method by combining the Maximum Likelihood estimator and the Kalman Filter predictor; in this method, sudden short-circuit test data are considered. Bortoni and Jardini [188], Silva, Bortoni, and Rocha [189], Wamkeue, Baetscher, and Kamwa [203], and Wamkeue, Christian, and Kamwa [201] perform the parameter estimation by means of load rejection test data. In addition, the former uses the Levenberg-Marquardt method; the second, a Genetic Algorithm; the third, the weighted least squares estimator and Newton's finite differences; and the latter, the exact solution of the state equations.

5.2.2 On-line procedures

The on-line estimation of synchronous generators parameters usually consists in the application of small disturbances followed by output measurements [77]. Using this approach, the works of Huang et al. [27], Shamsollahi and Malik [216], Wamkeue et al. [217], Vermeulen, Strauss, and Shikoana [218], Karrari and Malik [176], Karrari and Malik [177], Valverde et al. [72], Ghahremani and Kamwa [180], Zhou, Meng, and Lu [20], Hosseini, Abdollahi, and Karrari [73], and Monteiro, Vianna, and Giesbrecht [1, 2] stand out.

It is important to observe that Wamkeue et al. [217] estimate the parameters by means of data obtained from a large perturbation test, which consists in abruptly varying the reference voltage. Further, it used a maximum likelihood estimator derived from the generalized least-squares estimator. Vermeulen, Strauss, and Shikoana [218], on the other hand, proposed an estimation technique based on bipolar Pseudorandom Binary Sequence (PRBS) perturbations applied to the voltage regulator reference voltage. In that work, first-order models and Park equations [43] are considered for both quadrature and direct axes. Due to the low amplitude of the PRBS signals applied, the machine is capable of maintaining normal operating conditions at the same time that excitation dynamics are induced in both axes.

A methodology for estimating the physical parameters of a third-order nonlinear model is proposed by Karrari and Malik [176], which is based on the fact that a linear structure – the Heffron-Philips model [219] – and a nonlinear structure are well-defined for the system. While the Recursive Extended Least-Squares, Recursive Instrumental Variable, and Error Prediction methods are used to obtain the system transfer function [176], the Numerical Algorithms for Subspace State Space System Identification (N4SID) [220] is used by Karrari and Malik [177]. With respect to the excitation of the system for identification purposes, the former adds a PRBS signal to the field voltage, whereas the latter adds a random voltage signal.

Valverde et al. [72] estimate the quadrature- and direct-axis reactances, as well as the field winding resistance, of synchronous machines from a highly nonlinear model using the Unscented Kalman Filter (UKF) [129], which effectively filters the noisy measurements. The set of state-space equations used by them only considers the terminal and field voltages as inputs. An Extended Kalman Filter (EKF) with unknown inputs is applied on a fourth-order model by Ghahremani and Kamwa [180], subject to a step perturbation on the field voltage. The extended and unscented versions of the Kalman Filter (KF) assume joint Gaussian distribution of both measurement and states and use the Bayesian approach to derive the Kalman gain.

In contrast, the Particle Filter (PF) is a more general Bayesian approach, which does not rely on Gaussian noise assumptions [99]. The work developed by Zhou, Meng, and Lu [20] proposes an Extended PF to estimate the dynamic states of a fourth-order synchronous machine model after a three-phase fault using PMU data. Recent applications to Smart Grids have also been proposed [73]. In that context, Monteiro, Vianna, and Giesbrecht [1, 2] applied the PF, in its simplest version, to estimate the quadrature- and direct-axis magnetizing reactances [1] and flux linkages per second [2].

In addition to the aforementioned methods, many others have been used to estimate the physical parameters of synchronous generators, such as Artificial Neural Networks [216, 221], Conjugate Gradient Method [27], Hartley Series [222], Maximum Likelihood [215], Orthogonal Series Expansion [223], Piecewise Linear Static Maps [224], and Recursive Least-

Square Method [225].

5.3 This work contributions

Once some of the latest studies in the literature were presented, as well as the advances and areas in which there is more research, it is necessary to point out the contributions of this work:

1. An automated method to analyze the data from load rejection tests;
2. A methodology to estimate the load angle from on-line measurements and from known initial loading conditions;
3. Quadrature- and direct-axis magnetizing reactances estimation from on-line measurements and upon the load angle estimation mentioned in the previous topic;
4. States and parameters estimation by means of the PF.

These contributions will be dealt with appropriate details in Chapter 6 and the related results will be presented in Chapter 7.

Chapter 6

Experiments and Methodology

"Theory attracts practice as the magnet attracts iron."
— Carl Friedrich Gauss[1]

Several concepts on Salient-Pole Synchronous Machines, Systems Theory, Systems Identification, and Bayesian Processors have been developed. At this point, it is mandatory to connect them all and show how this work dealt with its main objective – the estimation of states and physical parameters of salient-pole synchronous machines.

A preliminary analysis of synchronous machines is performed by applying the load-rejection test in a simulated scenario. This work contributes with an automated methodology to separate the sum of exponentials that results upon the armature-current decrement.

Regarding on-line procedures, two different approaches have been presented aiming at estimating the quadrature- and direct-axis magnetizing reactances from on-line measurements: in the first one, a novel approach based on algebraic computations and load angle estimation is used. In the second one, the Bootstrap Particle Filter from Chapter 4 is applied to estimating states, as well.

6.1 Generating data

The subsequent sections present details of the frameworks used to obtain the simulation and real data for the salient-pole synchronous machines.

6.1.1 On-line simulation data

A model of an actual synchronous generator of 126 MVA rated power, 13.8 kV rated terminal voltage, and 8-pole pairs was used to generate data for validating the performance of

[1]Johann Carl Friedrich Gauss (1777–1855) was a German mathematician and physicist who made significant contributions to many fields in mathematics and science. Sometimes referred to as the Princeps mathematicorum (Latin for "the foremost of mathematicians") and "the greatest mathematician since antiquity", Gauss had an exceptional influence in many fields of mathematics and science, and is ranked among history's most influential mathematicians.

the proposed methods as a state and parameter estimator.

The Simulink® simulation framework of Figure 6.1 is compound of two major blocks:

1. The pu fundamental salient-pole synchronous machine block;

2. The infinite-bus block.

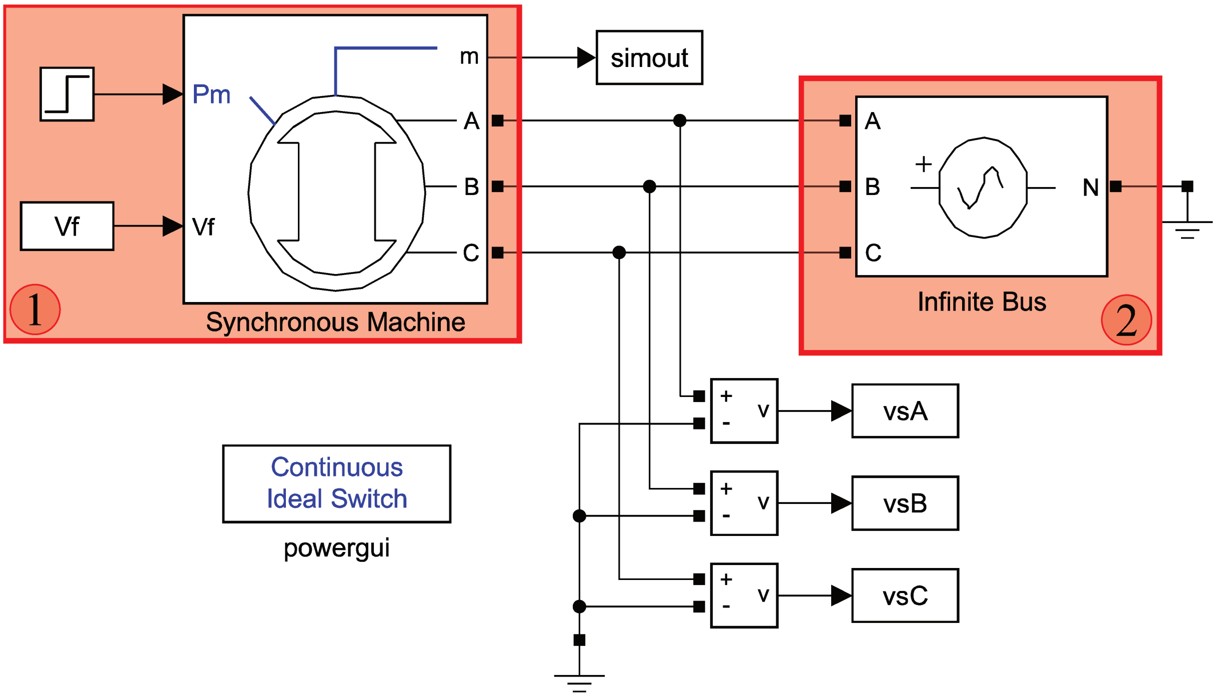

Figure 6.1: Simulink® simulation framework.

6.1.1.1 Salient-pole synchronous machine block

The salient-pole synchronous machine block models the dynamics of a three-phase salient-pole synchronous machine in the $qd0$ reference frame. It considers that the stator windings are connected in wye to an internal neutral point.

In order to perform the simulation of an actual machine, parameters provided by a manufacturer are used as the model's parameters. All adopted parameters are summarized in Table B.1. Saturation effects are not considered.

6.1.1.2 Infinite-bus block

The infinite-bus block is implemented as a three-phase zero-impedance voltage source. The common node (neutral) of the three source is accessible and grounded. It is set as a swing bus with 13.8 kV line-to-line voltage, 60 Hz frequency, and $\pi/2$ rad phase angle, so it matches the machine's phase.

6.1.2 Actual machine data

All information regarding the actual machine used is based on the work of Giesbrecht and Meneses [18].

The proposed methods were tested in Unit 03 of the hydro-power plant Salvajina, located in the city of Suárez, Cauca state, Colombia. This power plant was built on the Cauca river in the beginning of the 1980s and is equipped with three salient-pole synchronous generators manufactured by Toshiba®.

The unit is equipped with a static voltage regulation system, comprised of a Toshiba® analog voltage regulator, which is responsible for controlling the voltage when the unit is synchronized with the power system. The controller presents three limiting functions: over-excitation limiter (OEL), under-excitation limiter (UEL), and power system stabilizer (PSS). The speed regulator system is from Mitsubishi, the analog model EA-5. It offers the function of primary frequency regulation. Additionally, it has a proportional-integral-derivative (PID) regulation characteristic, which performs the primary frequency regulation and implements the statism by refeeding the actuator position. This loop provides an actuator position setpoint that is sent to the position control loop. The position control of the actuator is carried out through a P loop. This regulation system allows the control of P and Q and is able to operate in a constant field voltage.

Due to the great amount of noise, it is mandatory to treat the data before using them in the proposed so-called simplified approach. The first step in data processing was to disregard the period corresponding to the machine synchronization to the network. Thus, only the data from 400 s were considered.

All measured quantities are presented in Appendix II, from Figure II.14 to Figure II.18. In Figure II.19, network's and rotor's angular speed are presented. In order for the load angle to be correctly estimated, the network and machine angular speeds were smoothed – to reduce the amount of noise – and then approached each other in steady-state – so that the load angle estimation reached a constant value.

6.2 Experiments

In this work, the synchronous machine was operated under transient and steady-state conditions, with permanent connection to the network or with load interruption, resulting in three different experiments:

1. The load rejection test: after synchronizing and achieving the desired load condition, the load is rejected;
2. On-line steady-state operation: after synchronizing, the machine is left to operate in steady-state (only the steady-state part is of interest);

3. On-line transient and steady-state operations: after synchronizing, the machine is also left to operate in steady-state (opposing to the previous case, the transient and the steady-state parts are of interest).

In the following subsections, further details will be presented concerning the performed experiments. In Subsection 6.2.1, a new approach based on variable projection is used to analyze the data and obtain the parameters from the well-known load rejection test; in Subsection 6.2.2, a method developed in this work for determining the magnetizing reactances from certain load conditions is presented; and in Subsection 6.2.3, after observability analyses, the PF is used to estimate states and parameters from on-line transient and steady-state load conditions.

6.2.1 Load rejection test

Instead of using the traditional graphical approach, as presented in Chapter 5, the present work applies the variable projection algorithm – initially developed by Golub and Pereya [226], in 1973, and improved by O'Leary and Rust [227][2], in 2012. The variable projection algorithm is used to fit a model to measured data, which is often quite numerically challenging. In fitting exponential models, for example, small changes in the data can make large changes in the estimated parameters. Equally serious is the fact that data fitting problems are most often nonconvex, so a set of parameters can be optimal among nearby sets of parameters without being globally optimal, and software can be fooled into accepting a sub-optimal solution.

Most nonlinear models have some parameters – perhaps quite a few that appear linearly. For example, in fitting a sum of two exponentials, the model for the data observations $\{y(t_1), \cdots, y(t_m)\}$ might be

$$y(t) \approx c_1 \exp\{\alpha_1 t\} + c_2 \exp\{\alpha_2 t\} = \boldsymbol{\eta}(\boldsymbol{\alpha}, \mathbf{c}, t) \,. \tag{6.1}$$

The parameters $\mathbf{c} = [c_1\; c_2]^{\mathrm{T}}$ appear linearly; so, for every choice of nonlinear parameters $\boldsymbol{\alpha} = [\alpha_1\; \alpha_2]^{\mathrm{T}}$, optimal values for $\mathbf{c}$ can be found by solving a linear least-squares problem. Let a nonlinear least-squares problem be:

$$\text{(P1)} \;\Big|\; \textbf{minimize}_{\boldsymbol{\alpha},\mathbf{c}} \quad \|\mathbf{y} - \boldsymbol{\eta}(\boldsymbol{\alpha}, \mathbf{c})\|_2^2 \,. \tag{6.2}$$

Then, the solution to (P1) is the same as the solution to [227]:

$$\text{(P2)} \;\Big|\; \textbf{minimize}_{\boldsymbol{\alpha}} \quad \|\mathbf{y} - \boldsymbol{\eta}(\boldsymbol{\alpha}, \mathbf{c}(\boldsymbol{\alpha}))\|_2^2 \,. \tag{6.3}$$

The beauty of variable projection is that it reduces the number of parameters in the minimization problem, thus improving efficiency and possibly reducing the number of local minimizers. Therefore, convergence to the globally optimal solution is more likely [227].

---

[2] This reference should be consulted for further detail on implementation and theoretical foundations.

6.2.1.1 Procedure for analyzing the direct-axis load rejection data

The procedure for performing the direct-axis load rejection and analyzing the resulting data may be summarized in the following steps:

1. Synchronize the synchronous machine to the network and ensure that there is no active power flow from one to the other and that the machine is under-excited;
2. Measure the steady-state armature voltage[3], $\bar{v}_q^0$, and the steady-state armature current, $\bar{i}_d^0$;
3. Open the switches that connect the machine to the network and record the armature voltage behavior;
4. Once it is complete, supply the armature-voltage data to the variable projection algorithm, which provides the following parameters:
$$y(t) = c_0 + c_1 \exp\left\{-\frac{t}{\tau'_{d0}}\right\} + c_2 \exp\left\{-\frac{t}{\tau''_{d0}}\right\} ; \tag{6.4}$$
5. By comparing (6.4) to (2.147b), the following equations are obtained:
$$c_0 = \bar{v}_q^0 + \bar{x}_d \bar{i}_d^0 \quad \Rightarrow \quad \bar{x}_d = \frac{c_0 - \bar{v}_q^0}{\bar{i}_d^0} , \tag{6.5a}$$
$$c_1 = \left(\bar{x}'_d - \bar{x}_d\right) \bar{i}_d^0 \quad \Rightarrow \quad \bar{x}'_d = \bar{x}_d + \frac{c_1}{\bar{i}_d^0} , \tag{6.5b}$$
$$c_2 = \left(\bar{x}''_d - \bar{x}'_d\right) \bar{i}_d^0 \quad \Rightarrow \quad \bar{x}''_d = \bar{x}'_d + \frac{c_2}{\bar{i}_d^0} . \tag{6.5c}$$

6.2.1.2 Procedure for analyzing the quadrature-axis load rejection data

The procedure for performing the quadrature-axis load rejection and analyzing the resulting data may be summarized in the following steps:

1. Synchronize the synchronous machine to the network and ensure that the armature-current is on the quadrature-axis and that the machine is under-excited;
2. Measure the steady-state direct-axis armature voltage, $\bar{v}_d^0$, and the steady-state armature current, $\bar{i}_q^0$;
3. Open the switches that connect the machine to the network and record the armature voltage behavior;
4. Once it is complete, supply the direct-axis armature-voltage data to the variable projection algorithm, which provides the following parameters:
$$y(t) = c_1 \exp\left\{-\frac{t}{\tau''_{q0}}\right\} ; \tag{6.6}$$

[3] In the test condition, the armature voltage is on the quadrature-axis and the armature current is on the direct-axis; see Figure 2.14.

5. By comparing (6.6) to (2.150b), the following equations are obtained:

$$\bar{x}_q = \frac{\bar{v}_d^0}{\bar{i}_q^0} = \frac{\sqrt{(\bar{v}^0)^2 - (\bar{v}_q^0)^2}}{\bar{i}_q^0}, \tag{6.7a}$$

$$c_1 = \left(\bar{x}_q - \bar{x}_q''\right) \bar{i}_q^0 \quad \Rightarrow \quad \bar{x}_q'' = \bar{x}_q - \frac{c_1}{\bar{i}_q^0}. \tag{6.7b}$$

6.2.2 Simplified approach

The proposed method aims at calculating the quadrature- and direct-axis magnetizing reactances from armature voltages and currents and field current measurements. The first point to be highlighted is that the model considered in this work (refer to Subsection 3.3.1) is written in terms of *qd*0 quantities, which, in practical configurations, are not possible to be directly measured at the machine terminals.

6.2.2.1 The simulation data

While simulating the synchronous machine behavior, the mechanical power at the machine's shaft, `Pm`, is taken as a step input from 0 to 0.2 pu. The field voltage, `Vf`, for instance, is kept constant and equal to 1 pu.

The simulation was performed during 25 s and contemplated transient- and steady-state conditions. The 5 kHz sampling rate resulted in 125,002 samples. The set of measurements includes the Alternating Current (AC) voltages and currents at the terminal of the machine armature; and the Direct Current (DC) voltage and current injected to the field winding.

A complete visual description of all measured quantities[4] is presented in Appendix II, from Figure II.3 to Figure II.8.

■

In Chapter 2, Park's Transformation (PT) was used to transform the synchronous machine equations from the *abc* to the *qd*0 reference frame. As stated before, PT has the unique property of eliminating all rotor position-dependent inductances from the voltage equations of the synchronous machine and, therefore, simplify its analysis – all quantities are dealt as if they were DC quantities.

To transform from one reference frame to the other, PT requires load angle measurements. Recall:

$$\mathbf{f}_{qd0s} = \mathbf{K}\,\mathbf{f}_{abcs}, \tag{6.8}$$

[4] Voltages and currents in the *abc* reference-frame, although measured, are not presented due to the enormous amount of data and consequent impossibility of properly visualizing them.

where f represents either voltages, currents, flux linkages, or electric charges,

$$\mathbf{K} = \frac{2}{3}\begin{bmatrix} \cos\theta_r & \cos(\theta_r - 2\pi/3) & \cos(\theta_r - 4\pi/3) \\ \sin\theta_r & \sin(\theta_r - 2\pi/3) & \sin(\theta_r - 4\pi/3) \\ 1/2 & 1/2 & 1/2 \end{bmatrix}, \tag{6.9}$$

$$\theta_r = \omega_s t + \delta - \pi/2 \qquad [\text{electrical rad}]\ , \tag{6.10}$$

the synchronous speed ω_s, in electrical radians per second; and the load angle δ, in electrical radians. Equation (6.10) differs from (2.13) in the speed used to compute it.

6.2.2.2 Load angle computation

From literature, it is shown that measuring the load angle with desired precision is not a trivial task. When the machine has a large number of pole pairs, as in the case of salient-pole synchronous machines used for hydraulic power generation, the measurement of the load angle becomes especially challenging. Even in the simplest cases, (i.e., for smaller machines), it is necessary to install equipment capable of measuring the aforementioned angle. Therefore, the proposed method attempts to estimate the load angle from frequencies measurements (network's and rotor's).

In Subsection 2.4.6, the motion equations were presented. At the end, the load angle was written in terms of synchronous and rotor speeds as:

$$\delta = \int (\omega_r - \omega_s)\, dt \tag{6.11a}$$

$$= \omega_r t - \omega_s t + \delta_0 \qquad [\text{electrical rad}]\ , \tag{6.11b}$$

where δ_0 is the load angle value at $t = 0$, in electrical radians. Therefore, if the initial condition δ_0 is known, it is possible to compute the load angle.

Once ω_r measurements are available, Euler[5] and Runge[6]–Kutta[7] methods may be applied to numerically compute the integration. Further details on these methods are found in Annex D

Preliminary considerations

The following points are assumed *a priori*:

---

[5]Leonhard Euler (1707–1783) was a Swiss mathematician, physicist, astronomer, geographer, logician and engineer who made important and influential discoveries in many branches of mathematics, such as infinitesimal calculus and graph theory, while also making pioneering contributions to several branches such as topology and analytic number theory. He also introduced much of the modern mathematical terminology and notation, particularly for mathematical analysis, such as the notion of a mathematical function. He is also known for his work in mechanics, fluid dynamics, optics, astronomy, and music theory.

[6]Carl David Tolmé Runge (1856–1927) was a German mathematician, physicist, and spectroscopist.

[7]Martin Wilhelm Kutta (1867–1944) was a German mathematician.

1. The quadrature- and direct-axis magnetizing reactances are unknown;

2. There are no active, nor reactive, power flow in the initial condition.

When the salient-pole synchronous machine operates in steady-state condition, the flux linkages per second do not vary with time – i.e., all flux linkage per second derivatives with respect to time are equal to zero. Therefore, the voltage equations (2.110) become:

$$\bar{v}_q = \bar{r}_s \bar{i}_q + \bar{\omega}_r \bar{\Psi}_d \qquad [\text{pu}] \; , \tag{6.12a}$$

$$\bar{v}_d = \bar{r}_s \bar{i}_d - \bar{\omega}_r \bar{\Psi}_q \qquad [\text{pu}] \; , \tag{6.12b}$$

$$\bar{v}'_{kq} = \bar{r}'_{kq} \bar{i}'_{kq} \qquad [\text{pu}] \; , \tag{6.12c}$$

$$\bar{v}'_{fd} = \bar{r}'_{fd} \bar{i}'_{fd} \qquad [\text{pu}] \; , \tag{6.12d}$$

$$\bar{v}'_{kd} = \bar{r}'_{kd} \bar{i}'_{kd} \qquad [\text{pu}] \; . \tag{6.12e}$$

Equations (6.12c) and (6.12c) refer to the quadrature- and direct-axis amortisseur windings, respectively. Since they are short-circuited windings, $\bar{v}'_{kq} = 0$ and $\bar{v}'_{kd} = 0$. Therefore, as expected, there is no current flow in those windings. The armature flux linkages per second are:

$$\bar{\Psi}_q = -\frac{\bar{v}_d - \bar{r}_s \bar{i}_d}{\bar{\omega}_r} \qquad [\text{pu}] \; , \tag{6.13a}$$

$$\bar{\Psi}_d = \frac{\bar{v}_q - \bar{r}_s \bar{i}_q}{\bar{\omega}_r} \qquad [\text{pu}] \; , \tag{6.13b}$$

which are the only two flux linkages per second required for the present method.

As before, it is convenient to write the magnetizing flux linkages per second:

$$\bar{\Psi}_{mq} = \bar{x}_{mq} \left(\bar{i}_q + \bar{i}'_{kq} \right) = \bar{x}_{mq} \bar{i}_q \qquad [\text{pu}] \; , \tag{6.14a}$$

$$\bar{\Psi}_{md} = \bar{x}_{md} \left(\bar{i}_d + \bar{i}'_{fd} + \bar{i}'_{kd} \right) = \bar{x}_{md} \left(\bar{i}_d + \bar{i}'_{fd} \right) \qquad [\text{pu}] \; . \tag{6.14b}$$

From (2.112) and the previous definitions, the armature currents may be written as:

$$\bar{i}_q = \frac{1}{\bar{x}_{ls}} \left(\bar{\Psi}_q - \bar{\Psi}_{mq} \right) \qquad [\text{pu}] \; , \tag{6.15a}$$

$$\bar{i}_d = \frac{1}{\bar{x}_{ls}} \left(\bar{\Psi}_d - \bar{\Psi}_{md} \right) \qquad [\text{pu}] \; . \tag{6.15b}$$

The method itself

The proposed method consists in the following steps:

1. Measure the following quantities: armature voltages, armature currents, field current, rotor speed (or, similarly, generated voltage frequency), and network frequency;

2. Estimate the load angle from rotor speed and network frequency from (6.11a);

3. Transform the quantities from *abc* to *qd*0 by means of PT;

4. Compute the quadrature- and direct-axis magnetizing reactances:

$$\bar{x}_{mq} = \frac{\overline{\Psi}_{mq}}{\bar{i}_q} \quad [\mathrm{pu}] \ , \tag{6.16a}$$

$$\bar{x}_{md} = \frac{\overline{\Psi}_{md}}{\bar{i}_d + \bar{i}'_{fd}} \quad [\mathrm{pu}] \ . \tag{6.16b}$$

6.2.3 Bayesian approach

When applied in very noisy environments, the simplified approach is not able to provide very satisfactory results. In real scenarios, the amount of noise becomes an important issue. Thus, it is necessary to apply more robust methods, capable of simultaneously filtering highly noisy observations and estimating states with considerable accuracy. The Particle Filter (PF) developed in Section 4.5 becomes very suitable.

6.2.3.1 The simulation data

For the Bayesian approach, the simulation data is quite similar to the one used for the simplified approach.

Although transient- and steady-state conditions were simulated, the proposed method considers only the steady-state condition. Therefore, the whole set of samples was shrunken to 5,000 samples in, approximately, 1 s from 24 s to 25 s, when the machine is operating in steady-state.

Since the collected data are simulated results, they are not contaminated with measurement noise and, therefore, do not represent a real system. To work around this issue, measurement noise $\boldsymbol{v} \sim \mathcal{N}(\mathbf{0}, \mathbf{P}_{vv})$ is added to them. The selection of the covariance matrix $\mathbf{P}_{vv}$ depends on the level of uncertainty of the measurements. In order to use an independent and identically distributed measurement noise, the diagonal elements of $\mathbf{P}_{vv}$ are arbitrarily kept constant at 0.001.

Along with the visual description of the noiseless measured quantities from Figure II.3 to Figure II.8, the quantities with noise added are presented from Figure II.9 to Figure II.13.

■

The proposed method requires the measurement of armature voltages, armature currents, field voltage, field current, rotor speed, and load angle. For each time step, it consists in the following steps:

1. Transform the armature voltages from the *abc* to the *qd*0 reference frame;

2. Use the transformed voltages and the field voltage to compute the *a priori* states – flux linkages per second, rotor speed, and load angle – and torque;

3. Estimate the armature and field currents in the $qd0$ reference frame;

4. Transform these currents to the abc and assign a weight to each particle, a relative likelihood measure;

5. Use normalized weights to resample the particles and obtain the *a posteriori* states estimate.

In Figure 6.2, a simplified schematic diagram on the Bayesian approach for states and parameters estimation of salient-pole synchronous machines is presented.

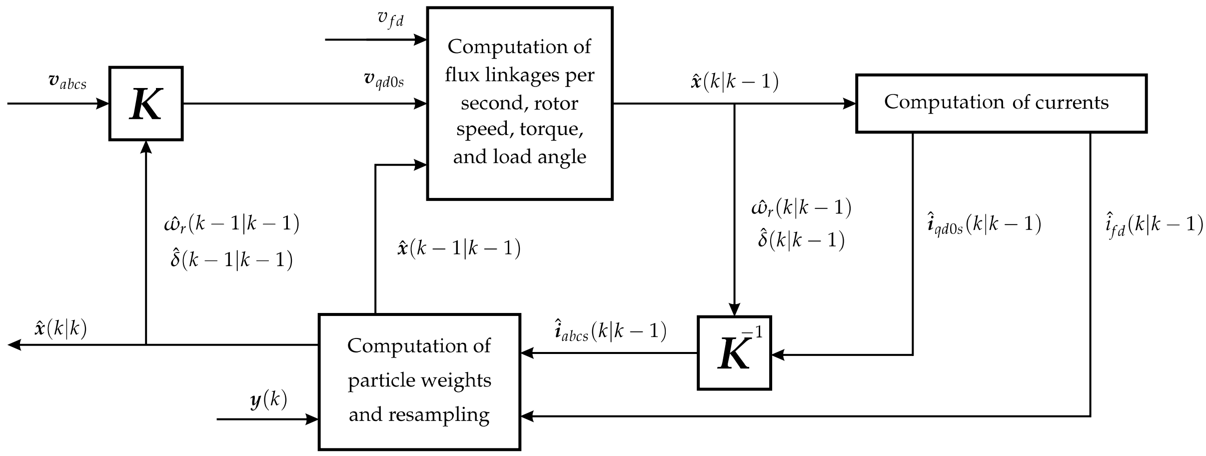

Figure 6.2: A simplified schematic diagram on the Bayesian approach for states and parameters estimation of salient-pole synchronous machines.

Due to observability issues, models of different orders are also considered. The differences among them are presented in the following subsections.

6.2.3.2 Sixth-order model

The sixth-order model, the same as that of C. H. Thomas [101] developed in Subsection 3.3.1, considers the flux linkages per second of armature and field circuits as state variables.

Let the state, input, and measurement vectors respectively be:

$$\boldsymbol{x} = \begin{bmatrix} \overline{\Psi}_q & \overline{\Psi}_d & \overline{\Psi}_0 & \overline{\Psi}'_{kq} & \overline{\Psi}'_{fd} & \overline{\Psi}'_{kd} \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^6 \,, \tag{6.17a}$$

$$\boldsymbol{u} = \begin{bmatrix} \bar{v}_q & \bar{v}_d & \bar{v}_0 & \bar{v}'_{kq} & v'_{fd} & \bar{v}'_{kd} \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^6 \,, \tag{6.17b}$$

$$\boldsymbol{y} = \begin{bmatrix} \bar{i}_q & \bar{i}_d & \bar{i}_0 & \bar{i}'_{kq} & \bar{i}'_{fd} & \bar{i}'_{kd} \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^6 \,. \tag{6.17c}$$

Process equations are simply the voltage equations (3.29) solved for the derivative of the flux linkage per second with respect to time. Hereafter, each process equation will be denoted f_i, for

$i = 1, \ldots, 6$:

$$f_1(\boldsymbol{x}, \boldsymbol{u}, t): \quad p\overline{\Psi}_q = \bar{v}_q - \overline{\omega}_r \overline{\Psi}_d - \bar{r}_s \left(\bar{x}'_{kq} \overline{\Psi}_q + \bar{x}_{mq} \overline{\Psi}'_{kq} \right) \quad [\text{pu}] \;, \tag{6.18a}$$

$$f_2(\boldsymbol{x}, \boldsymbol{u}, t): \quad p\overline{\Psi}_d = \bar{v}_d + \overline{\omega}_r \overline{\Psi}_q - \bar{r}_s \left(\Xi_{11} \overline{\Psi}_d - \Xi_{12} \overline{\Psi}'_{fd} - \Xi_{13} \overline{\Psi}'_{kd} \right) \quad [\text{pu}] \;, \tag{6.18b}$$

$$f_3(\boldsymbol{x}, \boldsymbol{u}, t): \quad p\overline{\Psi}_0 = \bar{v}_0 - \frac{\bar{r}_s}{\bar{x}_{ls}} \overline{\Psi}_0 \quad [\text{pu}] \;, \tag{6.18c}$$

$$f_4(\boldsymbol{x}, \boldsymbol{u}, t): \quad p\overline{\Psi}'_{kq} = \bar{v}'_{kq} + \bar{r}'_{kq} \left(\bar{x}_{mq} \overline{\Psi}_q - \bar{x}_q \overline{\Psi}_{kq} \right) \quad [\text{pu}] \;, \tag{6.18d}$$

$$f_5(\boldsymbol{x}, \boldsymbol{u}, t): \quad p\overline{\Psi}'_{fd} = \bar{v}'_{fd} - \bar{r}'_{fd} \left(\Xi_{21} \overline{\Psi}_d + \Xi_{22} \overline{\Psi}_{fd} + \Xi_{23} \overline{\Psi}'_{kd} \right) \quad [\text{pu}] \;, \tag{6.18e}$$

$$f_6(\boldsymbol{x}, \boldsymbol{u}, t): \quad p\overline{\Psi}'_{kd} = \bar{v}'_{kd} - \bar{r}'_{kd} \left(\Xi_{31} \overline{\Psi}_d + \Xi_{32} \overline{\Psi}'_{fd} + \Xi_{33} \overline{\Psi}'_{kd} \right) \quad [\text{pu}] \;, \tag{6.18f}$$

where Ξ_{rc} is the element in the rth row and cth column of the 3×3 matrix:

$$\Xi = \begin{bmatrix} \bar{x}'_{fd}\bar{x}'_{kd} - \bar{x}^2_{md} & -\bar{x}_{md}\bar{x}'_{kd} + \bar{x}^2_{md} & -\bar{x}_{md}\bar{x}'_{fd} + \bar{x}^2_{md} \\ -\bar{x}_{md}\bar{x}'_{kd} + \bar{x}^2_{md} & \bar{x}_d\bar{x}'_{kd} - \bar{x}^2_{md} & -\bar{x}_d\bar{x}_{md} + \bar{x}^2_{md} \\ -\bar{x}_{md}\bar{x}'_{fd} + \bar{x}^2_{md} & -\bar{x}_d\bar{x}_{md} + \bar{x}^2_{md} & \bar{x}_d\bar{x}'_{fd} - \bar{x}^2_{md} \end{bmatrix} \quad [\text{pu}] \;. \tag{6.19}$$

Measurement equations are the winding currents (3.16). Hereafter, each measurement equation will be denoted h_i, for $i = 1, \ldots, 6$:

$$h_1(\boldsymbol{x}, t): \quad \bar{i}_q = \frac{1}{\bar{x}_{ls}} \left(\overline{\Psi}_q - \overline{\Psi}_{mq} \right) \quad [\text{pu}] \;, \tag{6.20a}$$

$$h_2(\boldsymbol{x}, t): \quad \bar{i}_d = \frac{1}{\bar{x}_{ls}} \left(\overline{\Psi}_d - \overline{\Psi}_{md} \right) \quad [\text{pu}] \;, \tag{6.20b}$$

$$f_3(\boldsymbol{x}, t): \quad \bar{i}_0 = \frac{1}{\bar{x}_{ls}} \overline{\Psi}_0 \quad [\text{pu}] \;, \tag{6.20c}$$

$$h_4(\boldsymbol{x}, t): \quad \bar{i}'_{kq} = \frac{1}{\bar{x}_{lkq}} \left(\overline{\Psi}'_{kq} - \overline{\Psi}_{mq} \right) \quad [\text{pu}] \;, \tag{6.20d}$$

$$h_5(\boldsymbol{x}, t): \quad \bar{i}'_{fd} = \frac{1}{\bar{x}_{lfd}} \left(\overline{\Psi}'_{fd} - \overline{\Psi}_{md} \right) \quad [\text{pu}] \;, \tag{6.20e}$$

$$h_6(\boldsymbol{x}, t): \quad \bar{i}'_{kd} = \frac{1}{\bar{x}_{lkd}} \left(\overline{\Psi}'_{kd} - \overline{\Psi}_{md} \right) \quad [\text{pu}] \;, \tag{6.20f}$$

where

$$\overline{\Psi}_{mq} = \left(\frac{1}{\bar{x}_{mq}} + \frac{1}{\bar{x}_{ls}} + \frac{1}{\bar{x}'_{lkq}} \right)^{-1} \left(\frac{\overline{\Psi}_q}{\bar{x}_{ls}} + \frac{\overline{\Psi}'_{kq}}{\bar{x}'_{lkq}} \right) \quad [\text{pu}] \;, \tag{6.21a}$$

$$\overline{\Psi}_{md} = \left(\frac{1}{\bar{x}_{md}} + \frac{1}{\bar{x}_{ls}} + \frac{1}{\bar{x}'_{lfd}} + \frac{1}{\bar{x}'_{lkd}} \right)^{-1} \left(\frac{\overline{\Psi}_d}{\bar{x}_{ls}} + \frac{\overline{\Psi}'_{fd}}{\bar{x}'_{lfd}} + \frac{\overline{\Psi}'_{kd}}{\bar{x}'_{lkd}} \right) \quad [\text{pu}] \;. \tag{6.21b}$$

6.2.3.3 Seventh-order model

In some cases, rotor speed $\overline{\omega}_r$ is added to the set of states due to the difficulty of accurately measuring it. As an advantage, the rotor speed is estimated with states and it is possible to obtain better results for it.

Let the state, input, and measurement vectors respectively be:

$$\boldsymbol{x} = \begin{bmatrix} \overline{\Psi}_q & \overline{\Psi}_d & \overline{\Psi}_0 & \overline{\Psi}'_{kq} & \overline{\Psi}'_{fd} & \overline{\Psi}'_{kd} & \overline{\omega}_r \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^7 \,, \tag{6.22a}$$

$$\boldsymbol{u} = \begin{bmatrix} \bar{v}_q & \bar{v}_d & \bar{v}_0 & \bar{v}'_{kq} & v'_{fd} & \bar{v}'_{kd} \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^6 \,, \tag{6.22b}$$

$$\boldsymbol{y} = \begin{bmatrix} \bar{i}_q & \bar{i}_d & \bar{i}_0 & \bar{i}'_{kq} & \bar{i}'_{fd} & \bar{i}'_{kd} \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^6 \,. \tag{6.22c}$$

Other than all process and measurement equations – (6.18) and (6.20), respectively – from the sixth-order model, the following is added:

$$f_7(\boldsymbol{x}, t): \quad p\overline{\omega}_r = \frac{\overline{\mathcal{T}}_m - \overline{\mathcal{T}}_e}{2H} \qquad [\text{pu}] \,, \tag{6.23}$$

where ω_b corresponds to rated or base frequency, in rad/s; $\overline{\mathcal{T}}_m$ is the net mechanical shaft torque, in pu; $\overline{\mathcal{T}}_e$ is the electromagnetic torque, in pu; and the inertia constant

$$H = \frac{1}{2\rho^2} J \frac{\omega_b^2}{\mathcal{P}_b} \qquad [\text{s}] \tag{6.24}$$

is expressed in seconds. In (2.116), $\mathcal{P}_b$ is the base power, in volt-ampere.

6.2.3.4 Eighth-order model

In the eighth-order model, the load angle is included in the set of states. Therefore, let the state, input, and measurement vectors respectively be:

$$\boldsymbol{x} = \begin{bmatrix} \overline{\Psi}_q & \overline{\Psi}_d & \overline{\Psi}_0 & \overline{\Psi}'_{kq} & \overline{\Psi}'_{fd} & \overline{\Psi}'_{kd} & \overline{\omega}_r & \delta \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^8 \,, \tag{6.25a}$$

$$\boldsymbol{u} = \begin{bmatrix} \bar{v}_q & \bar{v}_d & \bar{v}_0 & \bar{v}'_{kq} & v'_{fd} & \bar{v}'_{kd} \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^6 \,, \tag{6.25b}$$

$$\boldsymbol{y} = \begin{bmatrix} \bar{i}_q & \bar{i}_d & \bar{i}_0 & \bar{i}'_{kq} & \bar{i}'_{fd} & \bar{i}'_{kd} \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^6 \,. \tag{6.25c}$$

Other than all process and measurement equations from the seventh-order model, the following state equation is included:

$$f_8(\boldsymbol{x}, t): \quad p\delta = \omega_b \left(\overline{\omega}_r - \overline{\omega}_s \right) \qquad [\text{electrical rad}] \,, \tag{6.26}$$

where $\overline{\omega}_s$ is the synchronous speed, in pu; $\overline{\omega}_r$ is the rotor speed, in pu; and δ is the load angle, in electrical radians.

6.2.3.5 Parameter estimation

A common approach to estimate unknown parameters within state-space models consists in adding the desired set of parameters to the set of states, as shown in Subsection 4.5.5.

Therefore, the state-vector is extended with the parameters $\boldsymbol{\theta}$, resulting on the following extended state vector for a salient-pole synchronous machine:

$$\mathbf{z} = \begin{bmatrix} \overline{\Psi}_q & \overline{\Psi}_d & \overline{\Psi}_0 & \overline{\Psi}'_{kq} & \overline{\Psi}'_{fd} & \overline{\Psi}'_{kd} & \boldsymbol{\theta}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}} . \tag{6.27}$$

The parameters transition from k to $k+1$ is represented as

$$\boldsymbol{\theta}(k) = \boldsymbol{\theta}(k-1) + \boldsymbol{\xi}(k-1) , \tag{6.28}$$

which means that the parameters are free to vary during the estimation process until their convergence into the actual values. The resulting state vector gives the best estimate of the parameters that maximize the likelihood for each time step.

According to Valverde et al. [72], when using steady-state conditions, the only parameters that can be truly extracted are $\bar{x}_{mq}$, $\bar{x}_{md}$, and $\bar{r}'_{fd}$. In this work, only the reactances are of interest. Therefore, the aforementioned vectors become

$$\boldsymbol{x} = \begin{bmatrix} \overline{\Psi}_q & \overline{\Psi}_d & \overline{\Psi}_0 & \overline{\Psi}'_{kq} & \overline{\Psi}'_{fd} & \overline{\Psi}'_{kd} & \bar{x}_{mq} & \bar{x}_{md} \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^8 , \tag{6.29a}$$

$$\boldsymbol{u} = \begin{bmatrix} \bar{v}_q & \bar{v}_d & \bar{v}_0 & \bar{v}'_{kq} & v'_{fd} & \bar{v}'_{kd} \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^6 , \tag{6.29b}$$

$$\boldsymbol{y} = \begin{bmatrix} \bar{i}_q & \bar{i}_d & \bar{i}_0 & \bar{i}'_{kq} & \bar{i}'_{fd} & \bar{i}'_{kd} \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^6 . \tag{6.29c}$$

At long last, the following process equations are added to the sixth-order model equations:

$$f_7(\boldsymbol{z}, \boldsymbol{u}, t) : \quad p\bar{x}_{mq} = \bar{x}_{mq} + \xi_1 \qquad [\mathrm{pu}] , \tag{6.30a}$$

$$f_8(\boldsymbol{z}, \boldsymbol{u}, t) : \quad p\bar{x}_{md} = \bar{x}_{md} + \xi_2 \qquad [\mathrm{pu}] . \tag{6.30b}$$

The estimation of parameters of salient-pole synchronous machines is very sensitive to the characteristics of the measurement and the state vectors. In some situations, it is necessary to add the currents of the amortisseur windings to the measurement vector in order to make the system observable. Thus, this work contemplates situations in which these currents appear and not appear in that vector; in Chapter 7, observability and estimation analyzes for these cases will be presented.

Chapter 7

Results and Discussion

"We want now to point out that modern computing machines are extremely well suited to perform the procedures described."
— Nicholas Metropolis and Stanislaw Ulam[1]

All computational experiments were performed within the MATLAB[2]® R2018a environment, in a personal computer with Intel® Xeon™ E3-1270 v6 CPU @3.80 GHz, 62.00 GB RAM, running the Ubuntu 18.04.1 LTS operational system.

In this chapter, the methodologies developed and discussed in Chapter 6 are used to estimate states and/or parameters of salient-pole synchronous machines.

7.1 Publications

Throughout the graduation years, four papers [1–4] have been produced and were sent to the 14^{o} *Simpósio Brasileiro de Automação Inteligente* (14th Brazilian Symposium on Intelligent Automation), to the IEEE Power and Energy Society Innovative Smart Grid Technologies Latin America 2019, and to the IEEE 29th International Symposium on Industrial Electronics – which is still awaiting for technical advice from the respective evaluation committee:

[1] MONTEIRO, I. A.; VIANNA, L. M. S.; GIESBRECHT, M. Nonlinear estimation of salient-pole synchronous machines parameters via Particle Filter. In: 2019 IEEE PES Innovative Smart Grid Technologies Conference – Latin America (ISGT Latin America). Gramado, RS, BR: IEEE, Sept. 2019. P. 1–6. DOI: `10.1109/ISGT-LA.2019.8895417`

[2] MONTEIRO, I. A.; VIANNA, L. M. S.; GIESBRECHT, M. Observador de fluxos, correntes e ângulo de carga de máquinas síncronas por meio da filtragem de partículas. In: ANAIS do XIV Simpósio Brasileiro de Automação Inteligente. Ouro Preto, MG, BR: Galoá, Oct. 2019. v. 1. DOI: `10.17648/sbai-2019-111220`

[1] At Los Alamos, in the 1950s, a group of researchers led by Metropolis, including John von Neumann and Stanislaw Ulam, developed the Monte Carlo method. The citation was extracted from Metropolis and Ulam [138].

[2] MATLAB® is the short form of "MATrix LABoratory". It is a MathWorks software and was first released in 1984.

[3] MONTEIRO, I. A.; MENESES, L.; GIESBRECHT, M. A novel approach on the determination of salient-pole synchronous machine magnetizing reactances from on-line measurements. In: 2020 IEEE 29th International Symposium on Industrial Electronics (ISIE). In Press: [s.n.], 2020

[4] VIANNA, L. et al. Detecção de falhas de alimentação de um motor CC sem escovas via Filtro de Partículas. In: ANAIS do XIV Simpósio Brasileiro de Automação Inteligente. Ouro Preto, MG, Brazil: Galoá, Oct. 2019. v. 1. DOI: 10.17648/sbai-2019-111202

Monteiro, Vianna, and Giesbrecht [1] proposed an estimation method based on particle filtering, where the salient-pole synchronous machine parameters can be simultaneously updated by using on-line measurements of voltages and currents at the machine armature and field terminals. The method can be used in real-time applications because it only depends on a small amount of steady-state data, does not require the usually slow computation of Jacobian matrices, and presents rapid convergence into the reference parameters values. Further, the Particle Filter (PF) allows the modeling of process and observation noises under any kind of probability functions, other than Gaussian.

Monteiro, Vianna, and Giesbrecht [2], although similar to the work of Monteiro, Vianna, and Giesbrecht [1], addressed the problem of the estimation of states of synchronous machines and, therefore, is focused on estimating flux linkages per second, damping windings currents, and load angle by means of particle filtering. The simulation of a synchronous machine operation, which provides the damper windings currents and the load angle, is used to validate the effectiveness of the proposed method.

Monteiro, Meneses, and Giesbrecht [3] considered the new approach developed in Subsection 6.2.2 to the estimation of salient-pole synchronous machines magnetizing reactances. Their results will be presented in the following sections.

Vianna et al. [4] proposed a model-based methodology of fault detection of a Brushless Direct Current (BLDC) motor when one of its phases is lost using the particle filter in a parameter estimation approach similar to the one adopted in the works Monteiro, Vianna, and Giesbrecht [1, 2] for synchronous machines. Although the BLDC motor model is much simpler than the salient-pole synchronous machine's, it was important to study the applicability of the PF on parameter estimation.

The results presented in the following sections are related to ones obtained by Monteiro, Vianna, and Giesbrecht [1, 2] and Monteiro, Meneses, and Giesbrecht [3]. The results from Vianna et al. [4] are further explored in other dissertations produced by the author's research group.

7.2 Parameters estimation by the load rejection tests and the variable projection algorithm

The load-rejection test is divided into purely direct-axis load rejection and purely quadrature-axis load rejection. Therefore, the results will be separately analyzed. In contrast to the simplified approach of Section 7.3 and the Bayesian approach of Section 7.4, in this work, the load rejection test is applied only to the simulated data. The most preeminent difference between the results presented in this section and the other commonly found in the literature is the use of the Variable Projection Algorithm, as detailed in Subsection 6.2.1.

7.2.1 Direct-axis load rejection

In a simulated scenario, quadrature-axis voltages are available and may be used to perform the following analyses. In actual cases, one needs to obtain the envelope from the armature *abc* voltages. The armature-voltage quadrature-axis component is presented in Figure II.1.

Right before the load rejection, the armature voltage and the armature current are measured:

$$\bar{v}_q^0 = 1 \quad [\mathrm{pu}] \,, \tag{7.1a}$$
$$\bar{i}_d^0 = -0.1868 \quad [\mathrm{pu}] \,. \tag{7.1b}$$

The graphical and analytical descriptions for $\bar{v}_q$ may be seen in Figure 7.1 and in (7.2), respectively, which presented a total Mean-squared error (MSE) of 2.8866×10^{-7} [pu].

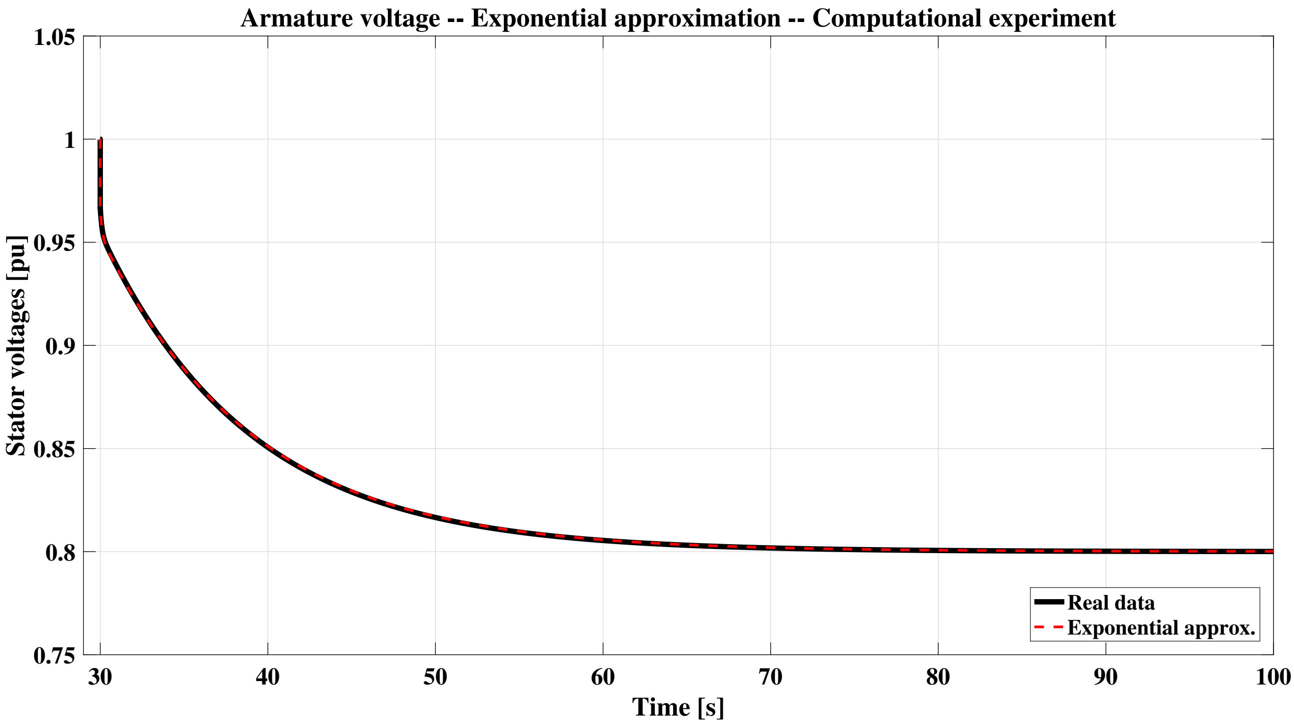

Figure 7.1: Exponential approximation for the armature voltage after the direct-axis load rejection.

$$\bar{v}_q(t) = \bar{v}_q^0 + \bar{x}_d \bar{i}_d^0 + \left(\bar{x}_d' - \bar{x}_d\right) \bar{i}_d^0 \exp\left\{-\frac{t}{\tau_{d0}'}\right\} + \left(\bar{x}_d'' - \bar{x}_d'\right) \bar{i}_d^0 \exp\left\{-\frac{t}{\tau_{d0}''}\right\} \quad [\mathrm{pu}] \,. \tag{7.2a}$$

$$= 0.8 + 0.1532 \exp\left\{-\frac{t}{9.0968}\right\} + 0.0135 \exp\left\{-\frac{t}{0.1139}\right\} \;\; [\text{pu}] \tag{7.2b}$$

By comparing (7.2a) to (7.2b),

$$\bar{v}_q^0 + \bar{x}_d i_d^0 = 0.8\,, \tag{7.3a}$$

$$\left(\bar{x}_d' - \bar{x}_d\right) i_d^0 = 0.1532\,, \tag{7.3b}$$

$$\left(\bar{x}_d'' - \bar{x}_d'\right) i_d^0 = 0.0135\,. \tag{7.3c}$$

Since the armature voltage and current before the rejection are

$$\bar{v}_q^0 = 1 \qquad [\text{pu}]\ , \tag{7.4a}$$

$$i_d^0 = -0.1868 \qquad [\text{pu}]\ , \tag{7.4b}$$

the direct-axis steady-state reactance is, from (7.3a),

$$\bar{x}_d = \frac{0.8 - \bar{v}_q^0}{i_d^0} = \frac{0.8 - 1}{-0.1868} = 1.0705 \qquad [\text{pu}]\ . \tag{7.5}$$

From (7.3b), the direct-axis transient reactance may be obtained:

$$\bar{x}_d' = \bar{x}_d + \frac{0.1532}{i_d^0} = 1.0705 + \frac{0.1532}{-0.1868} = 0.2501 \qquad [\text{pu}]\ . \tag{7.6}$$

At long last, the direct-axis subtransient reactance is calculated from (7.3c)

$$\bar{x}_d'' = \bar{x}_d' + \frac{0.0135}{i_d^0} = 0.2501 + \frac{0.0135}{-0.1868} = 0.1780 \qquad [\text{pu}]\ . \tag{7.7}$$

The transient and subtransient time constants are the moduli of the inverse of the exponential arguments of (7.2b):

$$\tau_{d0}' = 9.0968 \qquad [\text{s}]\ , \tag{7.8a}$$

$$\tau_{d0}'' = 0.1139 \qquad [\text{s}]\ , \tag{7.8b}$$

respectively.

The errors in the estimates are given in Table 7.1. The separation between the transient- and the subtransient states is quite subtle – even small deviations may lead to very different time constants. Therefore, the proposed methodology was not able to efficiently determine the subtransient time constant, τ_{d0}'', resulting in 16.5513% estimation error.

Table 7.1: Comparison between actual and estimated values for the direct-axis load rejection test.

Parameter	Actual value	Estimated value	Estimation error [%]
$\bar{x}_d$	1.0710 [pu]	1.0705 [pu]	0.0472
$\bar{x}'_d$	0.2481 [pu]	0.2501 [pu]	0.8082
$\bar{x}''_d$	0.1775 [s]	0.1780 [pu]	0.2883
τ'_{d0}	8.9530 [s]	9.0968 [s]	1.6059
τ''_{d0}	0.0977 [s]	0.1139 [s]	16.5513

7.2.2 Quadrature-axis load rejection

In a simulated scenario, the direct-axis voltage is available and may be used to perform the following analyses. The direct-axis component of the armature voltage is presented in Figure II.2.

The procedure presented in the work of O'Leary and Rust [227] was used to obtain an analytical expression for $\bar{v}_d$. The result may be seen in Figure 7.2 and in (7.9), which presented a total MSE of 7.9175×10^{-10} [pu].

$$\bar{v}_d(t) = \bar{v}_d^0 - \bar{x}_q \bar{i}_q^0 + \left(\bar{x}_q - \bar{x}''_q\right) \bar{i}_q^0 \exp\left\{-\frac{t}{\tau''_{q0}}\right\} \qquad [\text{pu}] \tag{7.9a}$$

$$= 0.2682 \exp\left\{-\frac{t}{0.1843}\right\} \qquad [\text{pu}] \tag{7.9b}$$

By comparing (7.9a) to (7.9b),

$$\left(\bar{x}_q - \bar{x}''_q\right) \bar{i}_q^0 = 0.2682\,. \tag{7.10}$$

Right before the load rejection, the armature voltages and the armature current are measured:

$$\bar{v}^0 = 1.004 \qquad [\text{pu}]\ , \tag{7.11a}$$

$$\bar{v}_d^0 = 0.4351 \qquad [\text{pu}]\ , \tag{7.11b}$$

$$\bar{v}_q^0 = 0.9008 \qquad [\text{pu}]\ , \tag{7.11c}$$

$$\bar{i}_q^0 = 0.6878 \qquad [\text{pu}]\ . \tag{7.11d}$$

Therefore, the quadrature-axis steady-state reactance is

$$\bar{x}_q = \frac{\bar{v}_d^0}{\bar{i}_q^0} = \frac{\sqrt{\left(\bar{v}^0\right)^2 - \left(\bar{v}_q^0\right)^2}}{\bar{i}_q^0} = \frac{0.4351}{0.6878} = 0.6326 \qquad [\text{pu}]\ . \tag{7.12}$$

From (7.10), the quadrature-axis subtransient reactance may be obtained:

$$\bar{x}''_q = \bar{x}_q - \frac{0.2682}{\bar{i}_q^0} = 0.6326 - \frac{0.2682}{0.6878} = 0.2427 \qquad [\text{pu}]\ . \tag{7.13}$$

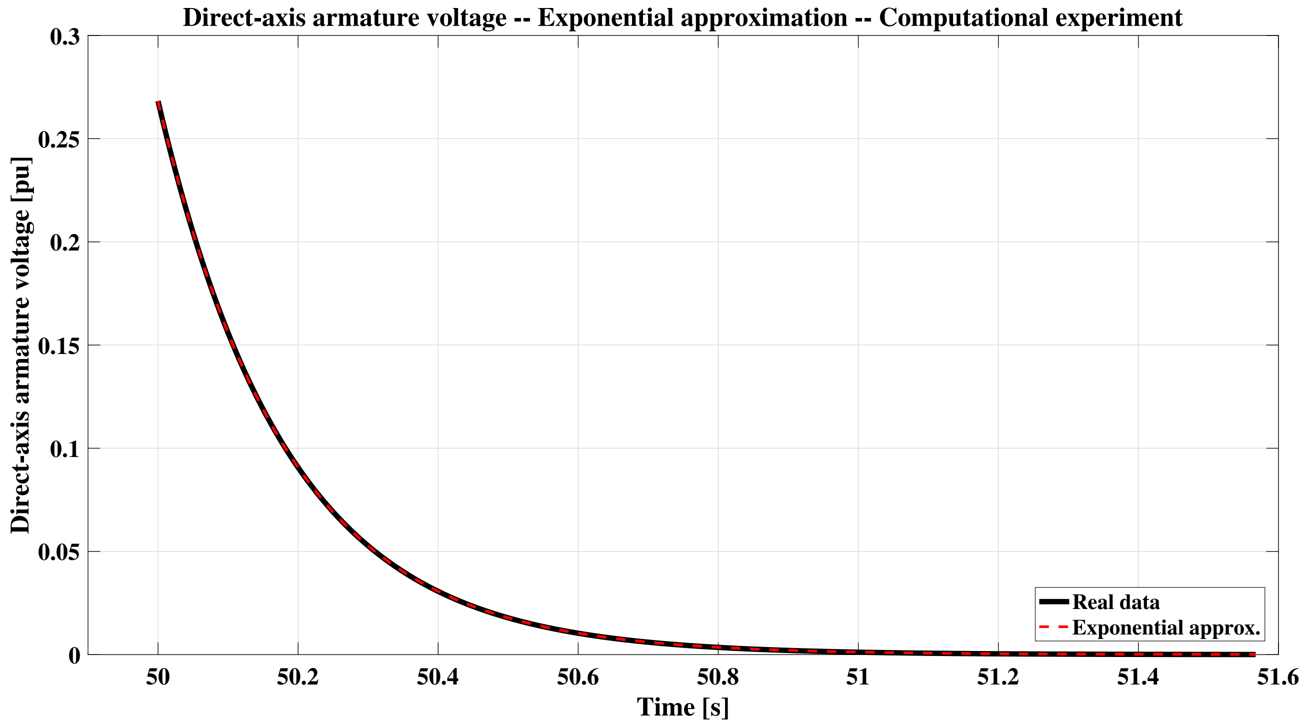

Figure 7.2: Exponential approximation for the direct-axis armature voltage after the quadrature-axis load rejection.

The subtransient time constant is the modulus of the inverse of the exponential arguments of (7.9b):

$$\tau''_{q0} = 0.1843 \qquad [\mathrm{s}]\ . \tag{7.14a}$$

The estimate errors for the quadrature-axis load rejection test are given in Table 7.2. The procedure provided very accurate results for the steady-state and subtransient reactances, as well as for the time constant. It is important to highlight that, in the literature, the results for the quadrature-axis are quite complicated to be achieved.

Table 7.2: Comparison between actual and estimated values for the quadrature-axis load rejection test.

Parameter	Actual value	Estimated value	Estimation error [%]
$\bar{x}_q$	0.6326 [pu]	0.6326 [pu]	0.0004
$\bar{x}''_q$	0.2426 [s]	0.2427 [pu]	0.0278
τ''_{q0}	0.1844 [s]	0.1843 [s]	0.0451

7.3 Simplified approach

The simplified approach described in Subsection 6.2.2 was applied to all two data sets – the computational and the real ones.

7.3.1 Simulation data

Upon the measurement of rotor speed and network frequency, the load angle was estimated by means of both Euler and Runge–Kutta methods. The very satisfactory results are presented in Figure 7.3.

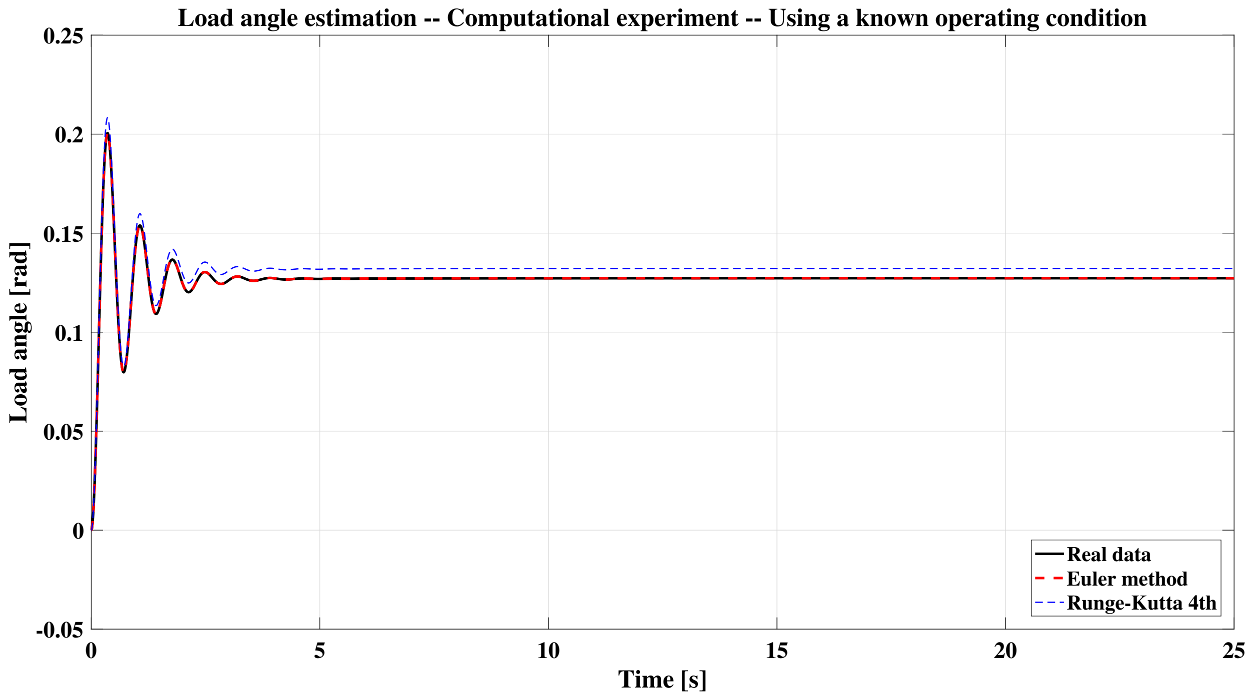

Figure 7.3: Load angle estimation for the computational machine data via Euler's method and $4th$-order Runge–Kutta.

Once the load angle is estimated, all abc quantities are transformed to the $qd0$ reference frame. Further, an arbitrary steady-state point k_1 is selected for the following computations:

1. Quadrature- and direct-axis flux linkages per second:

$$\overline{\Psi}_q(k_1) = -\frac{\bar{v}_d(k_1) - \bar{r}_s \bar{i}_d(k_1)}{\overline{\omega}_r(k_1)} \tag{7.15a}$$

$$= -\frac{0.1269 + 0.0012 \times (-0.0072)}{1} \tag{7.15b}$$

$$= -0.1269 \qquad [\text{pu}] \,, \tag{7.15c}$$

$$\overline{\Psi}_d(k_1) = \frac{\bar{v}_q(k_1) - \bar{r}_s \bar{i}_q(k_1)}{\overline{\omega}_r(k_1)} \tag{7.16a}$$

$$= \frac{0.9920 - 0.0012 \times (-0.2006)}{1} \tag{7.16b}$$

$$= 0.9923 \qquad [\text{pu}] \,. \tag{7.16c}$$

2. Quadrature- and direct-axis magnetizing flux linkages per second:

$$\overline{\Psi}_{mq}(k_1) = \overline{\Psi}_q(k_1) - \bar{i}_q(k_1)\bar{x}_{ls} \tag{7.17a}$$

$$= -0.1269 - (-0.2006) \times 0.1180 \tag{7.17b}$$

$$= -0.1032 \quad [\text{pu}] \ , \tag{7.17c}$$

$$\overline{\Psi}_{md}(k_1) = \overline{\Psi}_d(k_1) - \bar{i}_d(k_1)\bar{x}_{ls} \tag{7.18a}$$

$$= 0.9923 - (-0.0072) \times 0.1180 \tag{7.18b}$$

$$= 0.9931 \quad [\text{pu}] \ . \tag{7.18c}$$

3. Quadrature- and direct-axis magnetizing reactances:

$$\bar{x}_{mq}(k_1) = \frac{\overline{\Psi}_{mq}(k_1)}{\bar{i}_q(k_1)} = \frac{-0.1032}{-0.2006} = 0.5146 \quad [\text{pu}] \ , \tag{7.19}$$

$$\bar{x}_{md}(k_1) = \frac{\overline{\Psi}_{md}(k_1)}{\bar{i}_d(k_1) + \bar{i}'_{fd}(k_1)} = \frac{0.9931}{-0.0072 + 1} = 1.0003 \quad [\text{pu}] \ . \tag{7.20}$$

The simplified approach is a simple methodology for calculating magnetizing reactances. As long as some conditions are met for an accurate load angle estimation, the magnetizing reactances are obtained with great accuracy. As for the computation error, the method encountered 4.5835×10^{-5}% for the quadrature-axis, and 4.9676% for the direct-axis. It is left for future works the investigation of such difference in the direct-axis quantities.

In Table 7.3, a summarizing comparison between the parameters used for generating the data and the estimated values is presented.

Table 7.3: Comparison between actual data and estimated values for the proposed simplified approach.

Parameter	**Manufacturer value** [pu]	**Estimated value** [pu]	**Estimation error** [%]
$\bar{x}_{mq}$	0.5146	0.5146	4.5835×10^{-5}
$\bar{x}_{md}$	0.9530	1.0003	4.9676

7.3.2 Real data

The load angle estimation is presented in Figure 7.4. Euler and Runge–Kutta methods provided the same results.

Differently from the simulated scenario, it is not possible to verify the load angle estimation accuracy. However, it is important to notice that the graph in Figure 7.4 has a different behavior than Figure 7.3.

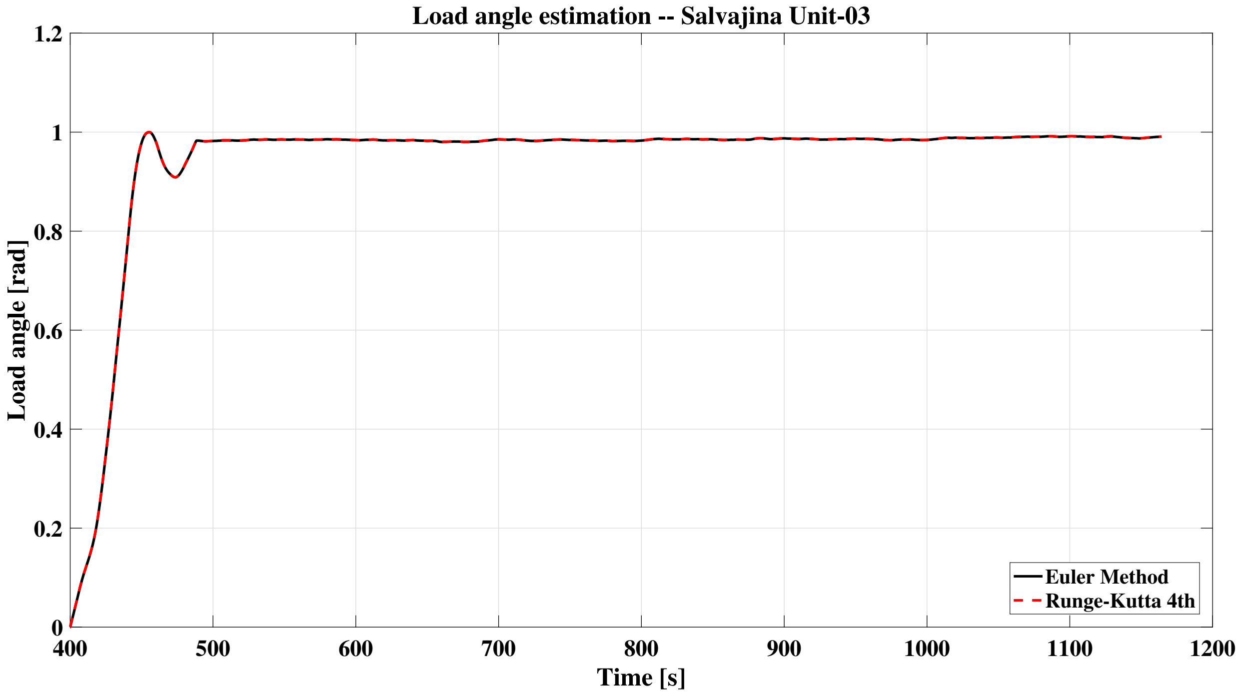

Figure 7.4: Load angle estimation for the actual machine data via Euler's method and 4*th*-order Runge–Kutta.

Once the load angle is estimated, all *abc* quantities are transformed to the $qd0$ reference frame. As before, an arbitrary steady-state point k_1 is selected for the following computations:

1. Quadrature- and direct-axis flux linkages per second:

$$\overline{\Psi}_q(k_1) = -\frac{\bar{v}_d(k_1) - \bar{r}_s\bar{i}_d(k_1)}{\overline{\omega}_r(k_1)} \tag{7.21a}$$

$$= -\frac{-0.1036 + 0.0012 \times (-0.0321)}{1.0002} \tag{7.21b}$$

$$= 0.1035 \qquad [\text{pu}] \ , \tag{7.21c}$$

$$\overline{\Psi}_d(k_1) = \frac{\bar{v}_q(k_1) - \bar{r}_s\bar{i}_q(k_1)}{\overline{\omega}_r(k_1)} \tag{7.22a}$$

$$= \frac{0.9849 - 0.0012 \times (0.1442)}{1.0002} \tag{7.22b}$$

$$= 0.9845 \qquad [\text{pu}] \ . \tag{7.22c}$$

2. Quadrature- and direct-axis magnetizing flux linkages per second:

$$\overline{\Psi}_{mq}(k_1) = \overline{\Psi}_q(k_1) - \bar{i}_q(k_1)\bar{x}_{ls} \tag{7.23a}$$

$$= 0.1035 - 0.1442 \times 0.19 \tag{7.23b}$$

$$= 0.0761 \qquad [\text{pu}] \ , \tag{7.23c}$$

$$\overline{\Psi}_{md}(k_1) = \overline{\Psi}_d(k_1) - \bar{i}_d(k_1)\bar{x}_{ls} \tag{7.24a}$$

$$= 0.9845 - (-0.0321) \times 0.19 \tag{7.24b}$$

$$= 0.9906 \qquad [\text{pu}]\ . \tag{7.24c}$$

3. Quadrature- and direct-axis magnetizing reactances:

$$\bar{x}_{mq}(k_1) = \frac{\overline{\Psi}_{mq}(k_1)}{\bar{i}_q(k_1)} = \frac{0.0761}{0.1442} = 0.5278 \qquad [\text{pu}]\ , \tag{7.25}$$

$$\bar{x}_{md}(k_1) = \frac{\overline{\Psi}_{md}(k_1)}{\bar{i}_d(k_1) + \bar{i}'_{fd}(k_1)} = \frac{0.9906}{-0.0321 + 0.9960} = 1.0277 \qquad [\text{pu}]\ . \tag{7.26}$$

The estimation errors are 3.4815% and 5.7115% for the quadrature- and the direct-axis magnetizing reactance, respectively. In Table 7.4, a summarizing comparison between the data provided by manufacturer and the estimated values is presented.

Table 7.4: Comparison between the data provided by manufacturer and the estimated values for the proposed simplified approach.

Parameter	**Manufacturer value** [**pu**]	**Estimated value** [**pu**]	**Estimation error** [%]
$\bar{x}_{mq}$	0.5100	0.5278	3.4815
$\bar{x}_{md}$	1.0900	1.0277	5.7115

7.4 Bayesian approach for states estimation

Before applying the Particle Filter (PF) to the salient-pole synchronous machine models, it is important to analyze observability.

7.4.1 Observability analyses

Varying from the simplest to more complex models, different configurations were considered.

7.4.1.1 Sixth-order model

According to the procedure described in Subsection 3.2.1 to analyze observability, for the sixth-order model of Subsection 6.2.3.2,

$$
\boldsymbol{F} = \begin{bmatrix}
-\bar{r}_s \bar{x}'_{kq} & -\overline{\omega}_r(k) & 0 & -\bar{r}_s \bar{x}_{mq} & 0 & 0 \\
\overline{\omega}_r(k) & -\bar{r}_s \Xi_{11} & 0 & 0 & \bar{r}_s \Xi_{12} & \bar{r}_s \Xi_{13} \\
0 & 0 & \dfrac{\bar{r}_s}{\bar{x}_{ls}} & 0 & 0 & 0 \\
\bar{r}'_{kq} \bar{x}_{mq} & 0 & 0 & -\bar{r}'_{kq} \bar{x}_q & 0 & 0 \\
0 & -\bar{r}'_{fd} \Xi_{21} & 0 & 0 & -\bar{r}'_{fd} \Xi_{22} & -\bar{r}'_{fd} \Xi_{23} \\
0 & -\bar{r}'_{kd} \Xi_{31} & 0 & 0 & -\bar{r}'_{kd} \Xi_{32} & -\bar{r}'_{kd} \Xi_{33}
\end{bmatrix} \tag{7.27}
$$

and

$$
\boldsymbol{H} = \begin{bmatrix}
\dfrac{1}{\bar{x}_{ls}}\left(1 - \dfrac{\alpha}{\bar{x}_{ls}}\right) & 0 & 0 & -\dfrac{\alpha}{\bar{x}_{ls}\bar{x}'_{lkq}} & 0 & 0 \\
0 & \dfrac{1}{\bar{x}_{ls}}\left(1 - \dfrac{\beta}{\bar{x}_{ls}}\right) & 0 & 0 & -\dfrac{\beta}{\bar{x}_{ls}\bar{x}'_{lfd}} & -\dfrac{\beta}{\bar{x}_{ls}\bar{x}'_{lkd}} \\
0 & 0 & \dfrac{1}{\bar{x}_{ls}} & 0 & 0 & 0 \\
-\dfrac{\alpha}{\bar{x}_{ls}\bar{x}'_{lkq}} & 0 & 0 & \dfrac{1}{\bar{x}_{lkq}}\left(1 - \dfrac{\alpha}{\bar{x}_{lkq}}\right) & 0 & 0 \\
0 & -\dfrac{\beta}{\bar{x}_{ls}\bar{x}'_{lfd}} & 0 & 0 & \dfrac{1}{\bar{x}'_{lfd}}\left(1 - \dfrac{\beta}{\bar{x}'_{lfd}}\right) & \dfrac{1}{\bar{x}'_{lfd}}\left(1 - \dfrac{\beta}{\bar{x}'_{lkd}}\right) \\
0 & -\dfrac{\beta}{\bar{x}_{ls}\bar{x}'_{lkd}} & 0 & 0 & \dfrac{1}{\bar{x}'_{lkd}}\left(1 - \dfrac{\beta}{\bar{x}'_{lkd}}\right) & \dfrac{1}{\bar{x}'_{lkd}}\left(1 - \dfrac{\beta}{\bar{x}'_{lkd}}\right)
\end{bmatrix}, \tag{7.28}
$$

where

$$
\alpha = \left(\frac{1}{\bar{x}_{mq}} + \frac{1}{\bar{x}_{ls}} + \frac{1}{\bar{x}'_{lkq}}\right)^{-1}, \tag{7.29a}
$$

$$
\beta = \left(\frac{1}{\bar{x}_{md}} + \frac{1}{\bar{x}_{ls}} + \frac{1}{\bar{x}'_{lfd}} + \frac{1}{\bar{x}'_{lkd}}\right)^{-1}. \tag{7.29b}
$$

The observability matrix, $\mathcal{O}(k)$, is defined in (3.12). Numerically computing $\texttt{rank}\ \mathcal{O}(k)$, it follows that $\texttt{rank}\ \mathcal{O}(k) = 6$, $\forall k$, $\mathcal{O}(k) \in \mathbb{R}^{36\times 6}$. Therefore, the system is observable at every time step.

In practical configurations, it is not possible to measure the currents that flow through the amortisseur windings. That being the case,

$$
\boldsymbol{y} = \begin{bmatrix} \bar{i}_q & \bar{i}_d & \bar{i}_0 & \bar{i}'_{fd} \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^4 . \tag{7.30}
$$

Thus,

$$
\boldsymbol{H} = \begin{bmatrix}
\frac{1}{\bar{x}_{ls}}\left(1-\frac{\alpha}{\bar{x}_{ls}}\right) & 0 & 0 & -\frac{\alpha}{\bar{x}_{ls}\bar{x}'_{lkq}} & 0 & 0 \\
0 & \frac{1}{\bar{x}_{ls}}\left(1-\frac{\beta}{\bar{x}_{ls}}\right) & 0 & 0 & -\frac{\beta}{\bar{x}_{ls}\bar{x}'_{lfd}} & -\frac{\beta}{\bar{x}_{ls}\bar{x}'_{lkd}} \\
0 & 0 & \frac{1}{\bar{x}_{ls}} & 0 & 0 & 0 \\
0 & -\frac{\beta}{\bar{x}_{ls}\bar{x}'_{lfd}} & 0 & 0 & \frac{1}{\bar{x}'_{lfd}}\left(1-\frac{\beta}{\bar{x}'_{lfd}}\right) & \frac{1}{\bar{x}'_{lfd}}\left(1-\frac{\beta}{\bar{x}'_{lkd}}\right)
\end{bmatrix} \tag{7.31}
$$

and the numerical analysis led to $\texttt{rank}\ \mathcal{O}(k) = 6$, $\forall k$, $\mathcal{O}(k) \in \mathbb{R}^{24\times 6}$. Therefore, even when these currents are not available, the sixth-order model continues to be observable.

7.4.1.2 Seventh-order model

For the seventh-order model of Subsection 6.2.3.3,

$$
\boldsymbol{F} = \begin{bmatrix}
-\bar{r}_s\bar{x}'_{kq} & -\overline{\omega}_r & 0 & -\bar{r}_s\bar{x}_{mq} & 0 & 0 & 0 \\
\overline{\omega}_r & -\bar{r}_s\Xi_{11} & 0 & 0 & \bar{r}_s\Xi_{12} & \bar{r}_s\Xi_{13} & 0 \\
0 & 0 & \frac{\bar{r}_s}{\bar{x}_{ls}} & 0 & 0 & 0 & 0 \\
\bar{r}'_{kq}\bar{x}_{mq} & 0 & 0 & -\bar{r}'_{kq}\bar{x}_q & 0 & 0 & 0 \\
0 & -\bar{r}'_{fd}\Xi_{21} & 0 & 0 & -\bar{r}'_{fd}\Xi_{22} & -\bar{r}'_{fd}\Xi_{23} & 0 \\
0 & -\bar{r}'_{kd}\Xi_{31} & 0 & 0 & -\bar{r}'_{kd}\Xi_{32} & -\bar{r}'_{kd}\Xi_{33} & 0 \\
-\overline{\Psi}_d & \overline{\Psi}_q & 0 & 0 & 0 & 0 & 0
\end{bmatrix} \tag{7.32}
$$

and

$$
\boldsymbol{H} = \begin{bmatrix}
\frac{1}{\bar{x}_{ls}}\left(1-\frac{\alpha}{\bar{x}_{ls}}\right) & 0 & 0 & -\frac{\alpha}{\bar{x}_{ls}\bar{x}'_{lkq}} & 0 & 0 & 0 \\
0 & \frac{1}{\bar{x}_{ls}}\left(1-\frac{\beta}{\bar{x}_{ls}}\right) & 0 & 0 & -\frac{\beta}{\bar{x}_{ls}\bar{x}'_{lfd}} & -\frac{\beta}{\bar{x}_{ls}\bar{x}'_{lkd}} & 0 \\
0 & 0 & \frac{1}{\bar{x}_{ls}} & 0 & 0 & 0 & 0 \\
-\frac{\alpha}{\bar{x}_{ls}\bar{x}'_{lkq}} & 0 & 0 & \frac{1}{\bar{x}_{lkq}}\left(1-\frac{\alpha}{\bar{x}_{lkq}}\right) & 0 & 0 & 0 \\
0 & -\frac{\beta}{\bar{x}_{ls}\bar{x}'_{lfd}} & 0 & 0 & \frac{1}{\bar{x}'_{lfd}}\left(1-\frac{\beta}{\bar{x}'_{lfd}}\right) & \frac{1}{\bar{x}'_{lfd}}\left(1-\frac{\beta}{\bar{x}'_{lkd}}\right) & 0 \\
0 & -\frac{\beta}{\bar{x}_{ls}\bar{x}'_{lkd}} & 0 & 0 & \frac{1}{\bar{x}'_{lkd}}\left(1-\frac{\beta}{\bar{x}'_{lkd}}\right) & \frac{1}{\bar{x}'_{lkd}}\left(1-\frac{\beta}{\bar{x}'_{lkd}}\right) & 0
\end{bmatrix}. \tag{7.33}
$$

Numerically computing $\texttt{rank}\ \mathcal{O}(k)$, it follows that $\texttt{rank}\ \mathcal{O}(k) = 6$, $\forall k$, $\mathcal{O}(k) \in \mathbb{R}^{42\times 7}$. Therefore, the system is non-observable at any time step. To overcome this issue, rotor speed measurements must be included in the set of measurements [72]:

$$
\boldsymbol{y} = \begin{bmatrix} \dot{i}_q & \dot{i}_d & \dot{i}_0 & \dot{i}'_{kq} & \dot{i}'_{fd} & \dot{i}'_{kd} & \overline{\omega}_r \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^7 . \tag{7.34}
$$

Correspondingly, the following measurement equation is added:

$$h_7(\boldsymbol{x}, t): \quad \overline{\omega}_r = \overline{\omega}_r \qquad [\text{pu}]\ . \tag{7.35}$$

Therefore, the matrix $\boldsymbol{H}$ becomes:

$$\boldsymbol{H} = \begin{bmatrix} \frac{1}{\bar{x}_{ls}}\left(1-\frac{\alpha}{\bar{x}_{ls}}\right) & 0 & 0 & -\frac{\alpha}{\bar{x}_{ls}\bar{x}'_{lkq}} & 0 & 0 & 0 \\ 0 & \frac{1}{\bar{x}_{ls}}\left(1-\frac{\beta}{\bar{x}_{ls}}\right) & 0 & 0 & -\frac{\beta}{\bar{x}_{ls}\bar{x}'_{lfd}} & -\frac{\beta}{\bar{x}_{ls}\bar{x}'_{lkd}} & 0 \\ 0 & 0 & \frac{1}{\bar{x}_{ls}} & 0 & 0 & 0 & 0 \\ -\frac{\alpha}{\bar{x}_{ls}\bar{x}'_{lkq}} & 0 & 0 & \frac{1}{\bar{x}_{lkq}}\left(1-\frac{\alpha}{\bar{x}_{lkq}}\right) & 0 & 0 & 0 \\ 0 & -\frac{\beta}{\bar{x}_{ls}\bar{x}'_{lfd}} & 0 & 0 & \frac{1}{\bar{x}'_{lfd}}\left(1-\frac{\beta}{\bar{x}'_{lfd}}\right) & \frac{1}{\bar{x}'_{lfd}}\left(1-\frac{\beta}{\bar{x}'_{lkd}}\right) & 0 \\ 0 & -\frac{\beta}{\bar{x}_{ls}\bar{x}'_{lkd}} & 0 & 0 & \frac{1}{\bar{x}'_{lkd}}\left(1-\frac{\beta}{\bar{x}'_{lkd}}\right) & \frac{1}{\bar{x}'_{lkd}}\left(1-\frac{\beta}{\bar{x}'_{lkd}}\right) & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \tag{7.36}$$

and the system becomes observable at every time step, since $\texttt{rank}\ \mathcal{O}(k) = 7$, $\forall k$, $\mathcal{O}(k) \in \mathbb{R}^{49\times 7}$.

7.4.1.3 Eighth-order model

When the load angle δ is added to the set of states, the system becomes non-observable, as in the previous case. Once more, it is necessary to add load angle measurements to the set of measurements to make it observable again. Thus,

$$h_8(\boldsymbol{x}, t): \quad \bar{\delta} = \bar{\delta} \qquad [\text{electrical rad}] \tag{7.37}$$

and $\texttt{rank}\ \mathcal{O}(k) = 8$, $\forall k$, $\mathcal{O}(k) \in \mathbb{R}^{64\times 8}$.

■

One could wonder why measured quantities are added to the set of states. The answer is simple: when measured, they are highly contaminated with noise; therefore, after filtering, their actual value may be estimated.

7.4.2 States estimation

In Table 7.5, the running time and the total mean-squared error for the sixth-order model in the simulated scenario, as a function of the number of particles, are presented. It considered a scenario with 5, 000 samples, corresponding to 1 s of data, and all currents in the measurement vector. It is possible to conclude that the larger the number of particles, the longer the required running time, but the smaller the mean-squared error. By means of the running time per sample line, it is possible to assure that the method may be applied for real-time analyses.

Table 7.5: Running time and mean squared error for the sixth-order model in the simulated scenario for 5000 samples.

Number of particles	100	200	500	1,000	2,000	5,000
Running time [s]	23.3377	46.3551	114.7019	231.5358	459.8758	1156.2617
Running time per sample [s]	0.0047	0.0093	0.0229	0.0463	0.0920	0.2313
Mean-squared error [pu]	0.0189	0.0127	0.0083	0.0064	0.0051	0.0040

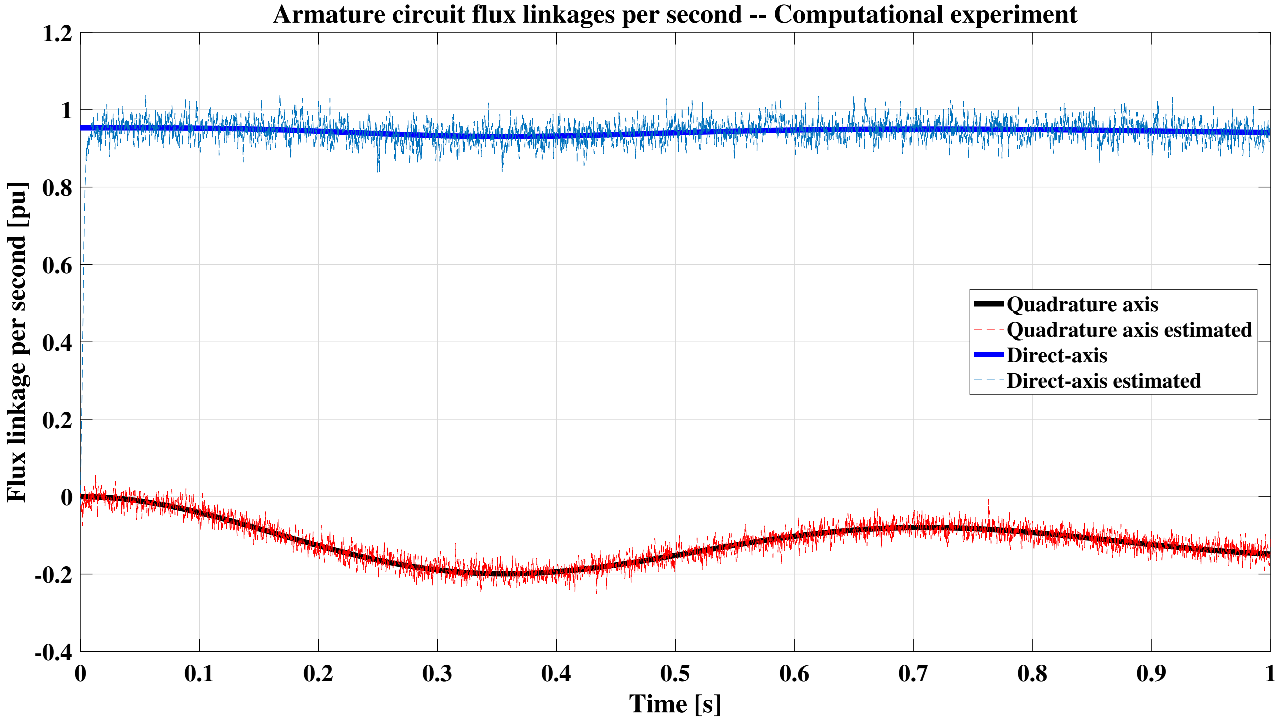

Figure 7.5: Computational experiment: Estimation of armature-circuit flux linkages per second via Particle Filter for the sixth-order model with all currents included.

For $5,000$ particles, the estimated armature circuit flux linkages per second are presented in Figure 7.5 and the rotor circuit ones, in Figure 7.6.

Although in Subsection 7.4.1.1, the observability analysis for practical configurations showed that the model with only measurable currents in the measurement vector is observable, simulation results provided a different conclusion. For $5,000$ samples and $10,000$ particles, the estimated armature circuit flux linkages per second are presented in Figure 7.7 and the rotor circuit ones, in Figure 7.8.

The reason for the difference between the theoretical and practical results will be investigated in future works.

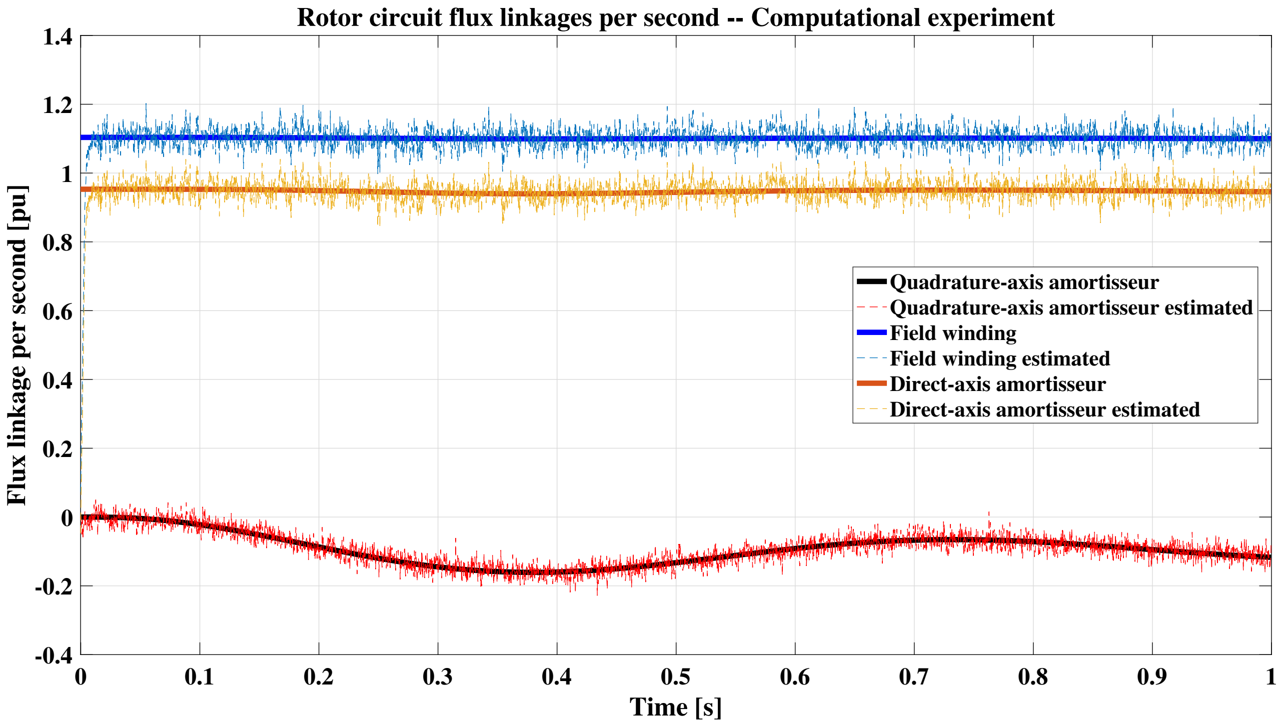

Figure 7.6: Computational experiment: Estimation of rotor-circuit flux linkages per second via Particle Filter for the sixth-order model with all currents included.

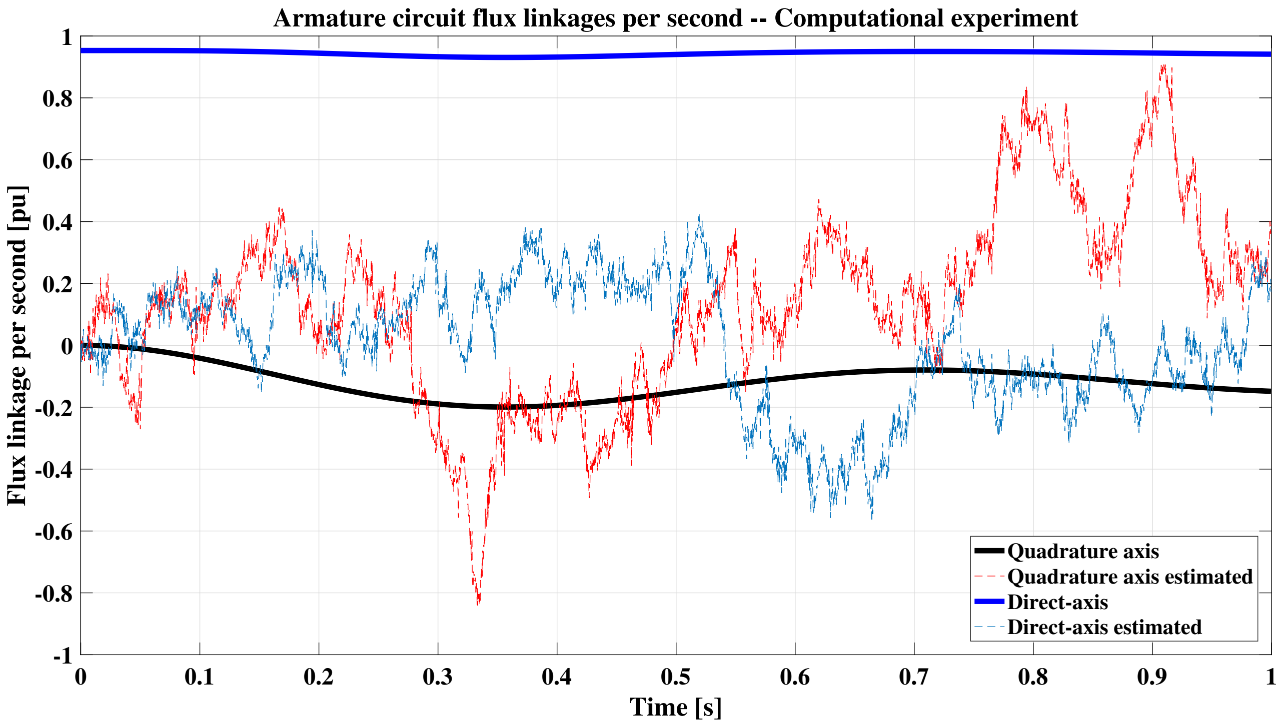

Figure 7.7: Computational experiment: Estimation of armature-circuit flux linkages per second via Particle Filter for the sixth-order model with only measurable currents included.

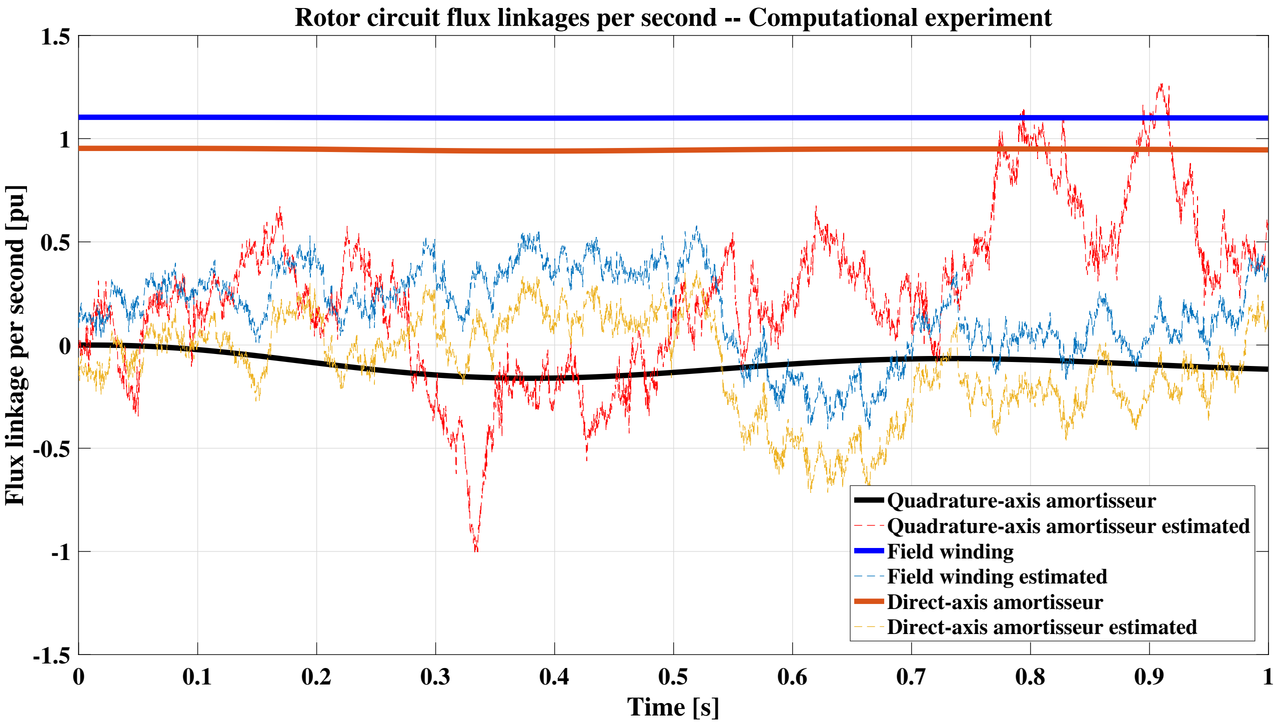

Figure 7.8: Computational experiment: Estimation of rotor-circuit flux linkages per second via Particle Filter for the sixth-order model with only measurable currents included.

7.5 Bayesian approach for states and parameters estimation

As presented in Subsection 4.5.5, the set of states is extended with the desired parameters to be estimated. In this work, the quadrature- and the direct-axis magnetizing reactances are considered.

7.5.1 Observability analysis

Once more, the procedure from Subsection 3.2.1 is used to analyze stability. At this time, the model presented in Subsection 6.2.3.5 is considered. Therefore,

$$\boldsymbol{F} = \begin{bmatrix} -\bar{r}_s\bar{x}'_{kq} & -\bar{\omega}_r(k) & 0 & -\bar{r}_s\bar{x}_{mq} & 0 & 0 & -\bar{r}_s\bar{\Psi}'_{kq} & 0 \\ \bar{\omega}_r(k) & -\bar{r}_s\Xi_{11} & 0 & 0 & \bar{r}_s\Xi_{12} & \bar{r}_s\Xi_{13} & 0 & 2\bar{r}_s\bar{x}_{md}\left(\bar{\Psi}_d+\bar{\Psi}'_{fd}+\bar{\Psi}'_{kd}\right) \\ 0 & 0 & \dfrac{\bar{r}_s}{\bar{x}_{ls}} & 0 & 0 & 0 & 0 & 0 \\ \bar{r}'_{kq}\bar{x}_{mq} & 0 & 0 & -\bar{r}'_{kq}\bar{x}_q & 0 & 0 & \bar{r}'_{kq}\bar{\Psi}_q & 0 \\ 0 & -\bar{r}'_{fd}\Xi_{21} & 0 & 0 & -\bar{r}'_{fd}\Xi_{22} & -\bar{r}'_{fd}\Xi_{23} & 0 & -2\bar{r}'_{fd}\bar{x}_{md}\left(\bar{\Psi}_d-\bar{\Psi}'_{fd}+\bar{\Psi}'_{kd}\right) \\ 0 & -\bar{r}'_{kd}\Xi_{31} & 0 & 0 & -\bar{r}'_{kd}\Xi_{32} & -\bar{r}'_{kd}\Xi_{33} & 0 & -2\bar{r}'_{kd}\bar{x}_{md}\left(\bar{\Psi}_d+\bar{\Psi}'_{fd}-\bar{\Psi}'_{kd}\right) \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{7.38}$$

and

$$
\boldsymbol{H} = \begin{bmatrix}
\frac{1}{\bar{x}_{ls}}\left(1-\frac{\alpha}{\bar{x}_{ls}}\right) & 0 & 0 & -\frac{\alpha}{\bar{x}_{ls}\bar{x}'_{lkq}} & 0 & 0 & -\frac{\kappa_1}{\bar{x}_{ls}} & 0 \\
0 & \frac{1}{\bar{x}_{ls}}\left(1-\frac{\beta}{\bar{x}_{ls}}\right) & 0 & 0 & -\frac{\beta}{\bar{x}_{ls}\bar{x}'_{lfd}} & -\frac{\beta}{\bar{x}_{ls}\bar{x}'_{lkd}} & 0 & -\frac{\kappa_2}{\bar{x}_{ls}} \\
0 & 0 & \frac{1}{\bar{x}_{ls}} & 0 & 0 & 0 & 0 & 0 \\
-\frac{\alpha}{\bar{x}_{ls}\bar{x}'_{lkq}} & 0 & 0 & \frac{1}{\bar{x}_{lkq}}\left(1-\frac{\alpha}{\bar{x}_{lkq}}\right) & 0 & 0 & -\frac{\kappa_1}{\bar{x}'_{lkq}} & 0 \\
0 & -\frac{\beta}{\bar{x}_{ls}\bar{x}'_{lfd}} & 0 & 0 & \frac{1}{\bar{x}'_{lfd}}\left(1-\frac{\beta}{\bar{x}'_{lfd}}\right) & \frac{1}{\bar{x}'_{lfd}}\left(1-\frac{\beta}{\bar{x}'_{lkd}}\right) & 0 & -\frac{\kappa_2}{\bar{x}'_{lfd}} \\
0 & -\frac{\beta}{\bar{x}_{ls}\bar{x}'_{lkd}} & 0 & 0 & \frac{1}{\bar{x}'_{lkd}}\left(1-\frac{\beta}{\bar{x}'_{lkd}}\right) & \frac{1}{\bar{x}'_{lkd}}\left(1-\frac{\beta}{\bar{x}'_{lkd}}\right) & 0 & -\frac{\kappa_2}{\bar{x}'_{lkd}}
\end{bmatrix}, \tag{7.39}
$$

where,

$$
\alpha = \left(\frac{1}{\bar{x}_{mq}} + \frac{1}{\bar{x}_{ls}} + \frac{1}{\bar{x}'_{lkq}}\right)^{-1}, \tag{7.40}
$$

$$
\beta = \left(\frac{1}{\bar{x}_{md}} + \frac{1}{\bar{x}_{ls}} + \frac{1}{\bar{x}'_{lfd}} + \frac{1}{\bar{x}'_{lkd}}\right)^{-1}, \tag{7.41}
$$

$$
\kappa_1 = \frac{\partial\overline{\Psi}_{mq}}{\partial\bar{x}_{mq}} = \frac{\bar{x}_{ls}^2\bar{x}_{lkq}^2}{\left[\bar{x}_{mq}(\bar{x}_{ls}+\bar{x}_{lkq})+\bar{x}_{ls}\bar{x}_{lkq}\right]}\left(\frac{\overline{\Psi}_q}{\bar{x}_{ls}} + \frac{\overline{\Psi}'_{kq}}{\bar{x}'_{lkq}}\right), \tag{7.42}
$$

$$
\kappa_2 = \frac{\partial\overline{\Psi}_{md}}{\partial\bar{x}_{md}} = \left[\bar{x}_{md}\left(\frac{1}{\bar{x}_{md}} + \frac{1}{\bar{x}_{ls}} + \frac{1}{\bar{x}'_{lfd}} + \frac{1}{\bar{x}'_{lkd}}\right)\right]^{-2}\left(\frac{\overline{\Psi}_d}{\bar{x}_{ls}} + \frac{\overline{\Psi}'_{fd}}{\bar{x}'_{lfd}} + \frac{\overline{\Psi}'_{kd}}{\bar{x}'_{lkd}}\right) \tag{7.43a}
$$

$$
= (\bar{x}_{md}\beta)^{-2}\left(\frac{\overline{\Psi}_d}{\bar{x}_{ls}} + \frac{\overline{\Psi}'_{fd}}{\bar{x}'_{lfd}} + \frac{\overline{\Psi}'_{kd}}{\bar{x}'_{lkd}}\right). \tag{7.43b}
$$

Numerically computing $\texttt{rank}\ \mathcal{O}(k)$, it follows that $\texttt{rank}\ \mathcal{O}(k) = 8$, $\forall k$, $\mathcal{O}(k) \in \mathbb{R}^{48\times 8}$. Therefore, the system is observable at every time step.

7.5.2 States and parameters estimation

In Subsection 7.5.1, it was shown that the eighth-order model created by the original sixth-order model augmented with the two desired parameters and with a measurement vector considering the measurements of all currents is completely observable, being able to estimate states and parameters. As in the previous cases, the theoretical results differed from the practical results: for 5, 000 samples and 10, 000 particles, the estimated armature and rotor circuits flux linkages per second are presented in Figure 7.9 and Figure 7.10, respectively, whereas the parameters are shown in Figure 7.11.

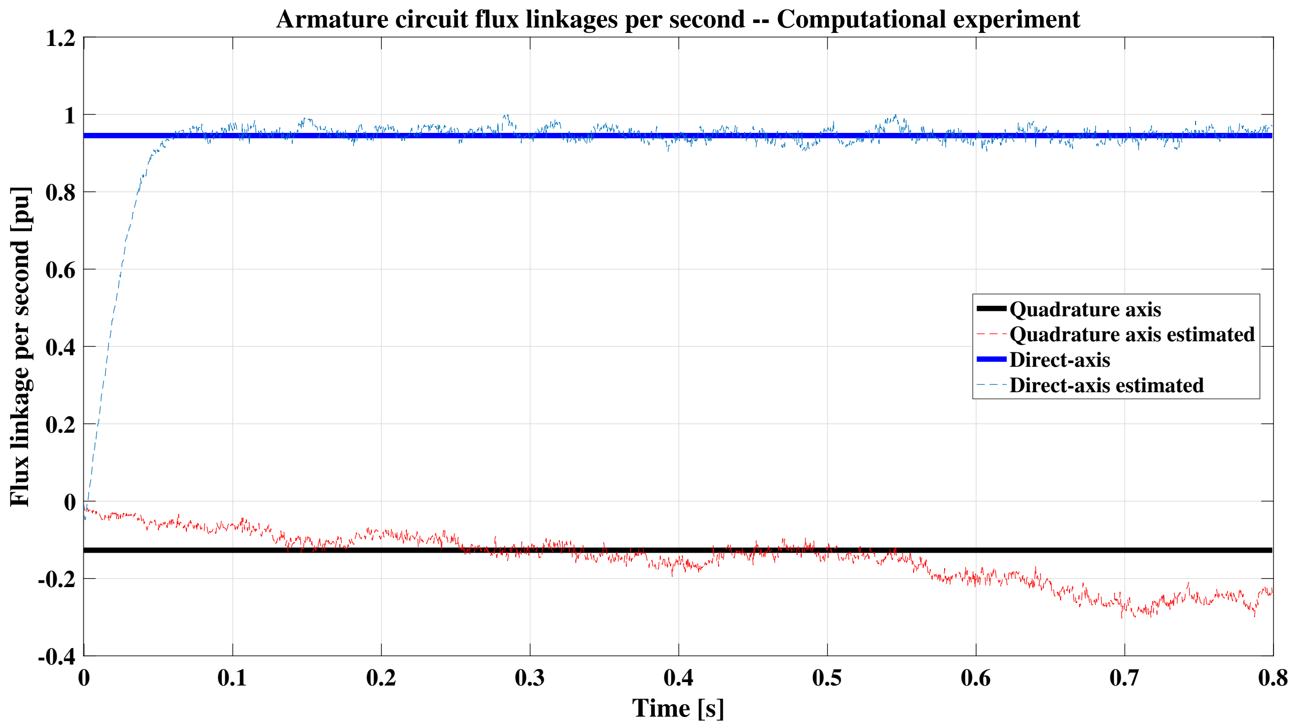

Figure 7.9: Computational experiment: Estimation of armature-circuit flux linkages per second via Particle Filter for the sixth-order model with all currents included.

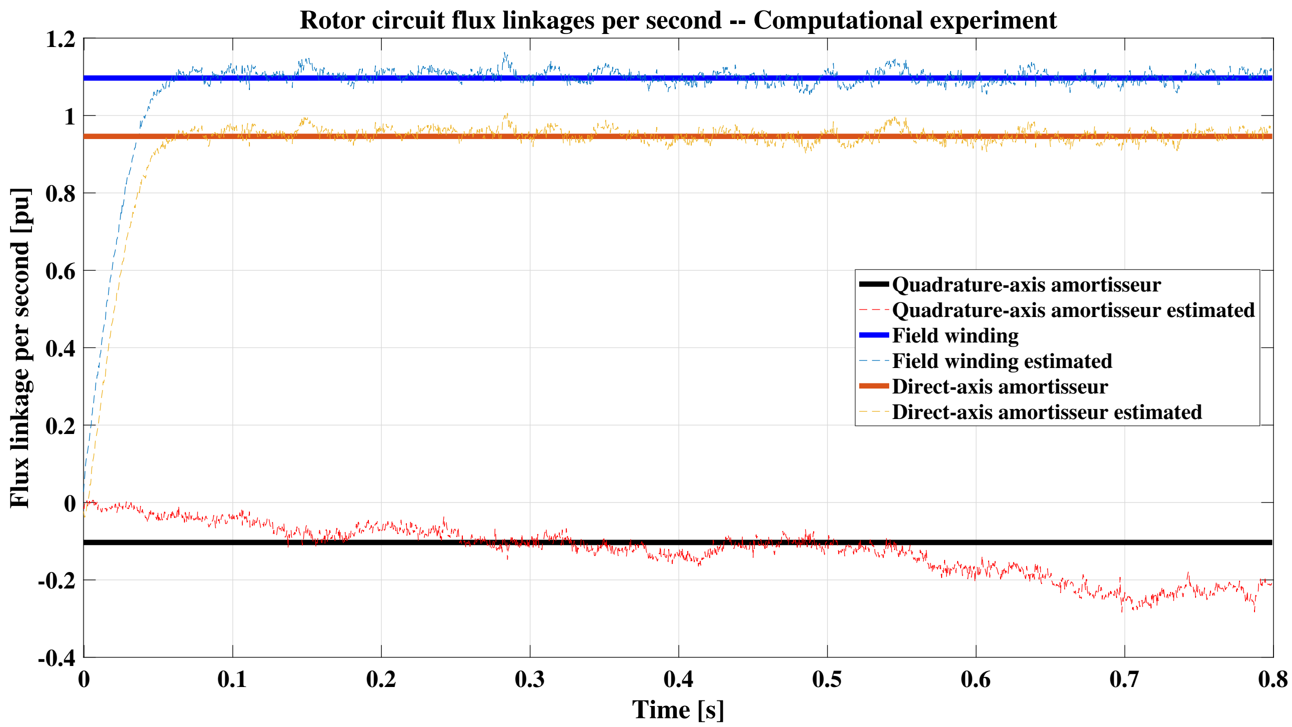

Figure 7.10: Computational experiment: Estimation of armature-circuit flux linkages per second via Particle Filter for the sixth-order model with all currents included.

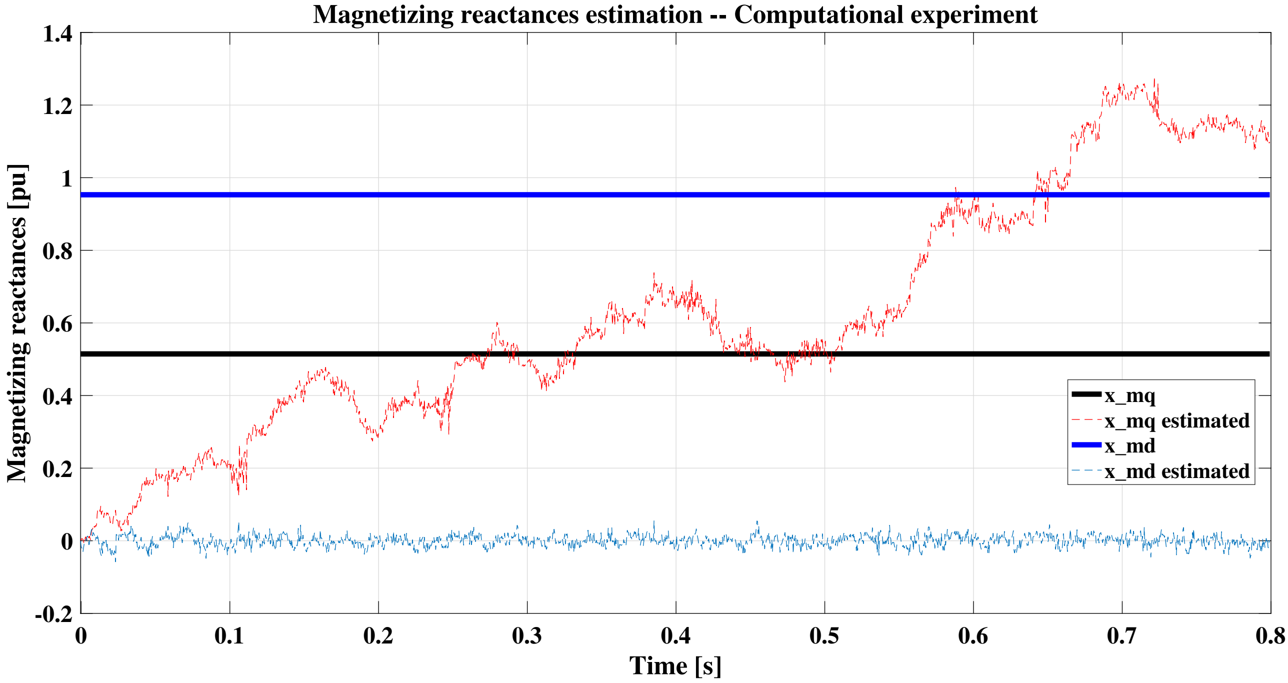

Figure 7.11: Computational experiment: Estimation of quadrature- and direct-axis magnetizing reactances via Particle Filter for the sixth-order model with all currents included.

A possible way to overcome this practical observability problem is to add rotor speed and load angle measurements to the measurement vector [1]:

$$y = \begin{bmatrix} \dot{i}_q & \dot{i}_d & \dot{i}_0 & \dot{i}'_{kq} & \dot{i}'_{fd} & \dot{i}'_{kd} & \overline{\omega}_{r,m} & \delta_m \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^8 , \tag{7.44}$$

turning it into an eighth-order vector, where $\overline{\omega}_{r,m}$ is the measured rotor speed, and δ_m is the measured load angle[3].

For 5,000 particles, the estimated armature circuit flux linkages per second are presented in Figure 7.12, the rotor circuit ones, in Figure 7.13, and the parameters, in Figure 7.14. The results are very satisfactory, with 5.0086×10^{-5} [pu] mean-squared error for the quadrature-axis magnetizing reactance and 4.0674×10^{-4} [pu], for the direct-axis magnetizing reactance; these results are quite similar to the ones presented by Monteiro, Vianna, and Giesbrecht [1].

The initial value of the particles in the positions corresponding to the magnetizing reactances comprised values randomly chosen within a range of widely established values [13]:

$$0.5 \leq \bar{x}_{mq} \leq 1 \qquad [\mathrm{pu}] \ , \tag{7.45}$$

$$0.8 \leq \bar{x}_{md} \leq 1.3 \qquad [\mathrm{pu}] \ . \tag{7.46}$$

In the work of Valverde et al. [72], it was also shown that rotor speed and load angle measurements are required for proper magnetizing reactances estimation. Due to the lack of

[3] In the lack of load angle measurements, the procedure developed in Subsection 6.2.2 may be applied, as it was done in this work.

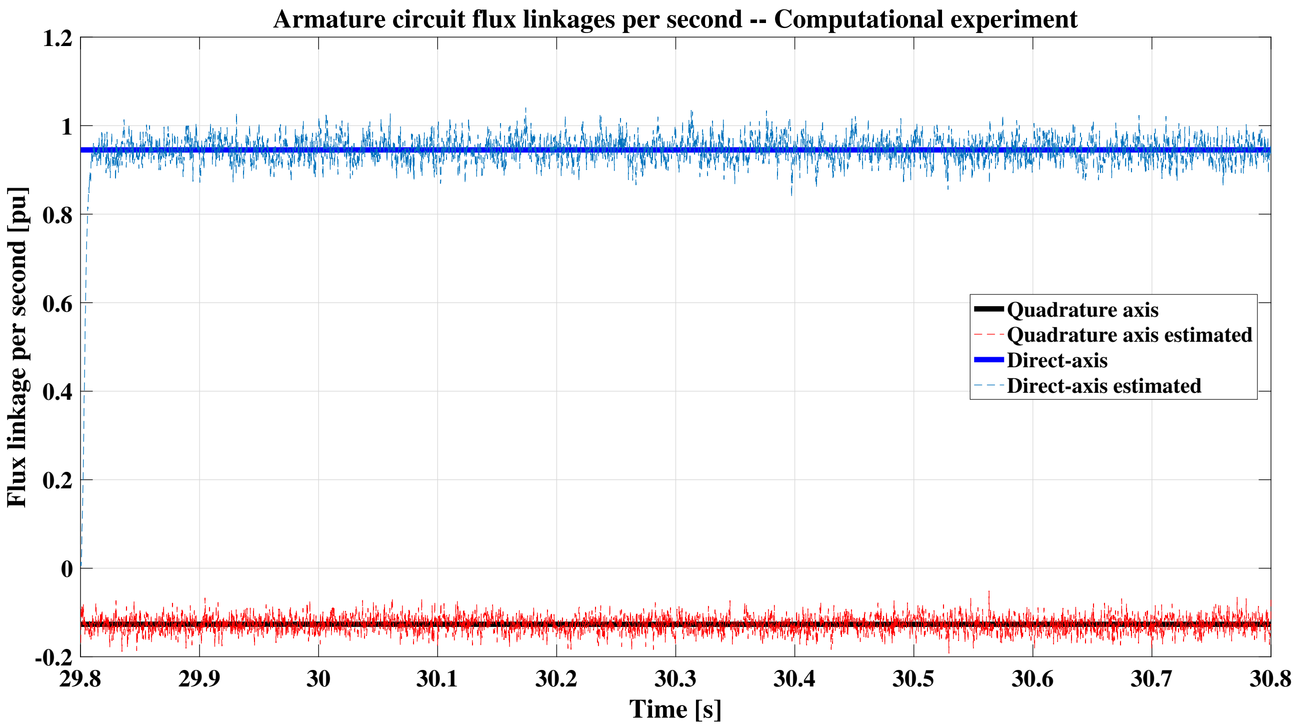

Figure 7.12: Computational experiment: Estimation of armature-circuit flux linkages per second via Particle Filter for the eighth-order model with all currents included.

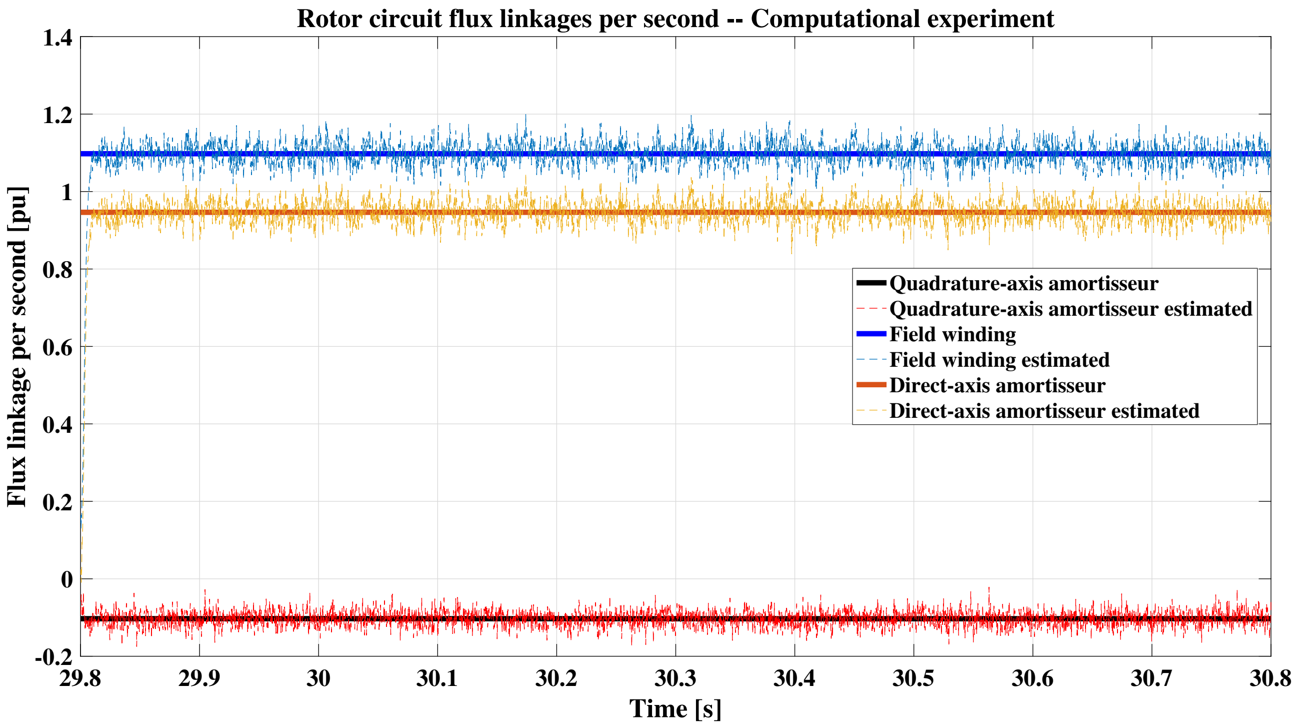

Figure 7.13: Computational experiment: Estimation of armature-circuit flux linkages per second via Particle Filter for the eighth-order model with all currents included.

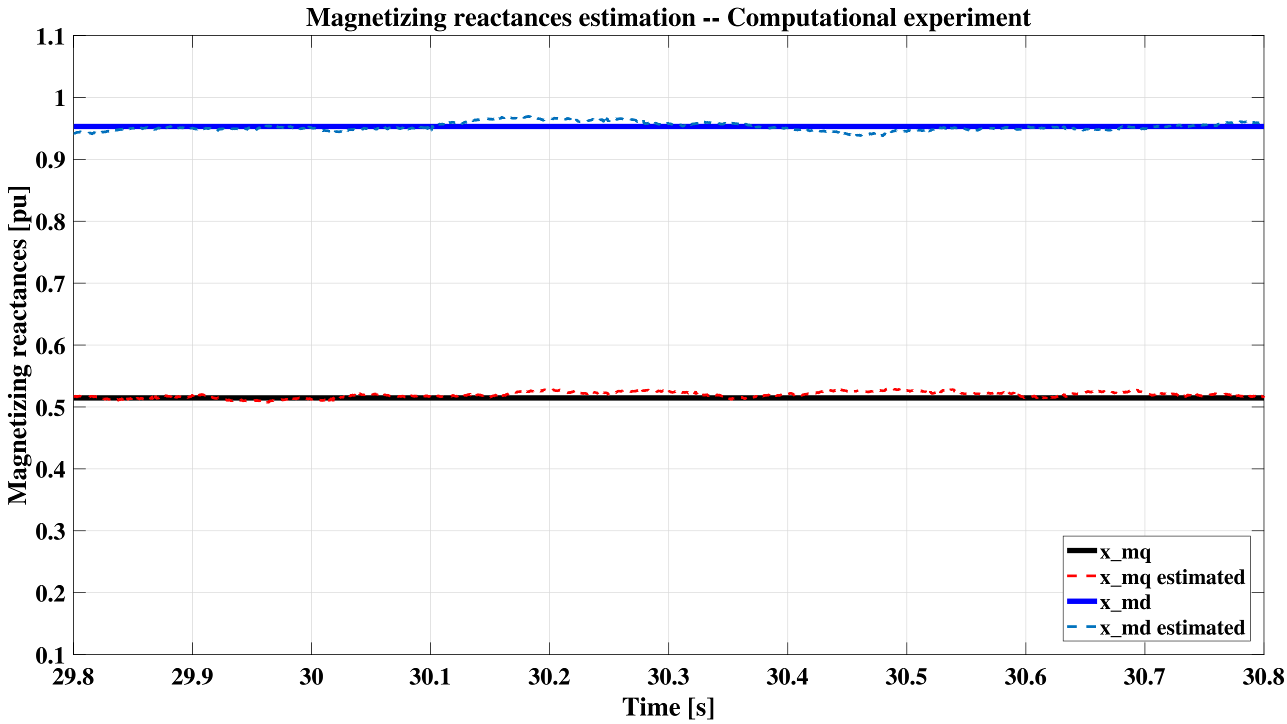

Figure 7.14: Computational experiment: Estimation of quadrature- and direct-axis magnetizing reactances via Particle Filter for the eighth-order model with all currents included.

load angle measurements, Valverde et al. [72] approximated it by using the power output at the machine terminals and the estimated value of the direct-axis reactance.

■

Since the machine used to generated the real data was not provided with equipment to measure the load angle, only the simulated data were used to estimate states and parameters by means of the Bayesian approach. In future works, the method proposed in Subsection 6.2.2 or the one from Valverde et al. [72] may be applied to overcome the absence of such load angle measurements.

Chapter 8

Conclusions and Future Directions

"Isn't it splendid to think of all the things there are to find out about? It just makes me feel glad to be alive – it's such an interesting world. It wouldn't be half so interesting if we know all about everything, would it? There would be no scope for imagination then, would there?"

— Anne Shirley-Cuthbert[1]

8.1 General conclusions

Bearing in mind the importance of salient-pole synchronous machines for the different analyses of power systems in countries where the energy generation is based on hydraulic sources, in this work, some methodologies have been proposed to estimate states and physical parameters of such machines.

The complete and most robust model of synchronous machines is based on operational equations and require more complex procedures to survey all transient and subtransient parameters. These methods are off-line – which require the machine to be taken out of operation. As an example, the load rejection test was contemplated. In result, this work proposed an automated methodology based on optimization procedures for separating every single exponential from the resulting voltage curve.

Another important challenge in the study of synchronous machines refers to the inability of measuring the load angle – that is, it is necessary to estimate it using specific procedures and algorithms. In the present work, a methodology based on the machine equations was developed to obtain load conditions that allow the calculation of the load angle by means of rotor speed and network frequency measurements and a known initial condition.

Ultimately, it is known that, in real scenarios, any measurement is corrupted with measurement noise – to a greater or a lesser extent. In this way, the analyzes become more complex, since the accuracy of the measured values is not precisely known. To get around these problems, Bayesian estimators can be applied to filter the real values amid so much noise. The present work evaluated the performance and the basis of the well-known Particle Filter (PF).

[1]Anne Shirley is a fictional character introduced in the 1908 novel Anne of Green Gables by Lucy Maud Montgomery. Montgomery wrote in her journal that the idea for Anne's story came from relatives who, planning to adopt an orphaned boy, received a girl instead.

In this sense, the application of the proposed methods proved to be satisfactory, given that the solutions obtained are consistent with the physical characteristics of the machines under analysis. The effectiveness of the methods is also evidenced by the similarity among the different solutions found by each of the approaches.

8.2 Future directions

As future works, the following research points are suggested:

1. Perform a more robust observability analysis, such as the one based on the Lie algebra, to investigate the difference found between theoretical and practical results;
2. Validate the performance of the load rejection test and the Bayesian approach methodologies on actual synchronous machines;
3. Assess the performance of Extended Kalman Filter (EKF) and the Unscented Kalman Filter (UKF) on the estimation of salient-pole synchronous machines parameters;
4. Assess some black-box identification methods on physical parameters estimation.

References

1 MONTEIRO, I. A.; VIANNA, L. M. S.; GIESBRECHT, M. Nonlinear estimation of salient-pole synchronous machines parameters via Particle Filter. In: 2019 IEEE PES Innovative Smart Grid Technologies Conference – Latin America (ISGT Latin America). Gramado, RS, BR: IEEE, Sept. 2019. P. 1–6. DOI: 10.1109/ISGT-LA.2019.8895417.

2 ________. Observador de fluxos, correntes e ângulo de carga de máquinas síncronas por meio da filtragem de partículas. In: ANAIS do XIV Simpósio Brasileiro de Automação Inteligente. Ouro Preto, MG, BR: Galoá, Oct. 2019. v. 1. DOI: 10.17648/sbai-2019-111220.

3 MONTEIRO, I. A.; MENESES, L.; GIESBRECHT, M. A novel approach on the determination of salient-pole synchronous machine magnetizing reactances from on-line measurements. In: 2020 IEEE 29th International Symposium on Industrial Electronics (ISIE). In Press: [s.n.], 2020.

4 VIANNA, L.; GONÇALVES, J.; MONTEIRO, I.; GIESBRECHT, M. Detecção de falhas de alimentação de um motor CC sem escovas via Filtro de Partículas. In: ANAIS do XIV Simpósio Brasileiro de Automação Inteligente. Ouro Preto, MG, Brazil: Galoá, Oct. 2019. v. 1. DOI: 10.17648/sbai-2019-111202.

5 GLOVER, J. D.; OVERBYE, T. J.; SARMA, M. S. **Power System Analysis & Design**. Sixth ed. Boston, MA, USA: Cengage Learning, Jan. 2017. ISBN 978-1-305-63213-4.

6 VARAIYA, P. P.; WU, F. F.; BIALEK, J. W. Smart operation of smart grid: risk-limiting dispatch. **Proceedings of the IEEE**, v. 99, n. 1, p. 40–57, Jan. 2011. ISSN 0018-9219. DOI: 10.1109/JPROC.2010.2080250.

7 BLAABJERG, F.; TEODORESCU, R.; LISERRE, M.; TIMBUS, A. V. Overview of control and grid synchronization for distributed power generation systems. **IEEE Transactions on Industrial Electronics**, v. 53, n. 5, p. 1398–1409, Oct. 2006. ISSN 1557-9948. DOI: 10.1109/TIE.2006.881997.

8 SMITH, J.; RYLANDER, M.; ROGERS, L.; DUGAN, R. It's all in the plans: maximizing the benefits and minimizing the impacts of DERs in an integrated grid.

IEEE Power and Energy Magazine, v. 13, n. 2, p. 20–29, Jan. 2015. ISSN 1558-4216. DOI: 10.1109/MPE.2014.2379855.

9 SHAHIDEHPOUR, M.; SCHWARTS, F. Don't let the sun go down on photovoltaic systems. **IEEE Power and Energy Magazine**, v. 2, n. 3, p. 40–48, June 2004. ISSN 1558-4216. DOI: 10.1109/MPAE.2004.1293599.

10 WRIGHT, S. H. Determination of synchronous machine constants by test reactances, resistances, and time constants. **Transactions of the American Institute of Electrical Engineers**, v. 50, n. 4, p. 1331–1350, Dec. 1931. ISSN 0096-3860. DOI: 10.1109/T-AIEE.1931.5055960.

11 KILGORE, L. A. Calculation of synchronous machine constants – reactances and time constants affecting transient characteristics. **Transactions of the American Institute of Electrical Engineers**, v. 50, n. 4, p. 1201–1213, Dec. 1931. ISSN 0096-3860. DOI: 10.1109/T-AIEE.1931.5055943.

12 MOMOH, J. A. **Energy Processing and Smart Grid**. Piscataway, NJ, USA: Wiley–IEEE Press, June 2018. (IEEE Press Series on Power Engineering). ISBN 978-1-119-37616-3.

13 IEEE. IEEE Guide for Test Procedures for Synchronous Machines Part I - Acceptance and Performance Testing, Part II - Test Procedures and Parameter Determination for Dynamic Analysis. **IEEE Std 115-2009 (Revision of IEEE Std 115-1995)**, p. 1–219, May 2010. DOI: 10.1109/IEEESTD.2010.5464495.

14 AGRAWAL, B. L.; DEMCKO, J. A. WSCC testing experience with generating units. In: IEEE Power Engineering Society. 1999 Winter Meeting. New York, NY, USA: IEEE, Jan. 1999. v. 1, p. 160–161. DOI: 10.1109/PESW.1999.747443.

15 DERUSSO, P. M.; ROY, R. J.; CLOSE, C. M.; DESROCHERS, A. A. **State Variables for Engineers**. Second ed. Hoboken, NJ, USA: Wiley-Interscience, Dec. 1997. ISBN 978-0-471-57795-9.

16 GHAHREMANI, E.; KAMWA, I. Dynamic state estimation in power system by applying the Extended Kalman Filter with unknown inputs to phasor measurements. **IEEE Transactions on Power Systems**, v. 26, n. 4, p. 2556–2566, Nov. 2011. ISSN 1558-0679. DOI: 10.1109/TPWRS.2011.2145396.

17 KOSTEREV, D. N.; TAYLOR, C. W.; MITTELSTADT, W. A. Model validation for the August 10, 1996 WSCC system outage. **IEEE Transactions on Power Systems**, v. 14, n. 3, p. 967–979, Aug. 1999. ISSN 0885-8950. DOI: 10.1109/59.780909.

18 GIESBRECHT, M.; MENESES, L. A. E. Detailed derivation and experimental validation of a method for obtaining load conditions for salient pole synchronous machine quadrature-axis parameters determination. **IEEE Transactions on Industrial Electronics**, v. 66, n. 7, p. 5049–5056, July 2019. ISSN 0278-0046. DOI: 10.1109/TIE.2018.2866096.

19 KRAUSE, P. C.; WASYNCZUK, O.; SUDHOFF, S. D.; PEKAREK, S. **Analysis of Electric Machinery and Drive Systems**. Third ed. Hoboken, NJ, USA: John Wiley & Sons, Inc., 2013. (IEEE Press Series on Power Engineering). ISBN 978-1-118-02429-4.

20 ZHOU, N.; MENG, D.; LU, S. Estimation of the dynamic states of synchronous machines using an extended Particle Filter. **IEEE Transactions on Power Systems**, v. 28, n. 4, p. 4152–4161, Nov. 2013. ISSN 0885-8950. DOI: 10.1109/TPWRS.2013.2262236.

21 DANDENO, P. L. Current usage suggested practices in power system stability simulations for synchronous machines. **IEEE Transactions on Energy Conversion**, EC-1, n. 1, p. 77–93, Mar. 1986. ISSN 0885-8969. DOI: 10.1109/TEC.1986.4765673.

22 ANDERSON, P. M.; FOUAD, A. A. **Power System Control and Stability**. Ed. by Mohamed E. El-Hawary. Second ed. Piscataway, NJ, USA: John Wiley & Sons, Inc., 2003. (IEEE Power Engineering Series). ISBN 0-471-23862-7.

23 KIESSLING, F.; NEFZGER, P.; NOLASCO, J. F.; KAINTZYK, U. **Overhead Power Lines: Planning, Design, Construction**. New York, NY, USA: Springer-Verlag Berlin Heidelberg, 2014. (Power Systems). ISBN 978-3-642-05556-0.

24 GRAINGER, J.; STEVENSON, W. **Power System Analysis**. [S.l.]: McGraw-Hill, 2003. (McGraw-Hill Series in Electrical and Computer Engineering: Power and Energy). ISBN 9780070585157.

25 EKANAYAKE, J. B.; JENKINS, N.; LIYANAGE, K.; WU, J.; YOKOYAMA, A. **Smart Grid: Technology and Applications**. Chichester, WS, UK: John Wiley & Sons, Ltd., Feb. 2012. ISBN 978-1-119-96909-9.

26 LIU, J. S.; CHEN, R. Blind deconvolution via sequential imputations. **Journal of the American Statistical Association**, v. 90, n. 430, p. 567–576, June 1995. ISSN 0162-1459. DOI: 10.2307/2291068.

27 HUANG, C.-T.; CHEN, Y.-T.; CHANG, C.-L.; HUANG, C.-Y.; CHIANG, H.-D.; WANG, J.-C. On-line measurement-based model parameter estimation for synchronous generators: model development and identification schemes. **IEEE Transactions on Energy Conversion**, v. 9, n. 2, p. 330–336, June 1994. ISSN 0885-8969. DOI: 10.1109/60.300140.

28 BIM, E. **Máquinas Elétricas e Acionamento**. Second ed. Rio de Janeiro, RJ, BR: Elsevier Editora Ltda., Feb. 2015. ISBN 978-85-352-5923-0.

29 WHITE, D. C.; WOODSON, H. H. **Electromechanical Energy Conversion**. New York, NY, USA: John Wiley & Sons, Inc., 1959. ISBN 978-1124129310.

30 ADKINS, B. **The General Theory of Electrical Machines**. London, UK: Chapman & Hall, 1957. ISBN 978-0-41-207840-8.

31 CONCORDIA, C. **Synchronous Machines: Theory and Performance**. New York, NY, USA: John Wiley & Sons, Inc., 1951. (General Electric Series). ISBN 978-0608100029.

32 ELGERD, O. I. **Electric Energy Systems Theory: An Introduction**. Ed. by H. H. Happ. New York, NY, USA: McGraw-Hill, 1971. ISBN 978-0-07-019230-0.

33 KUNDUR, P. **Power Systems Stability and Control**. Ed. by Neal J. Balu and Mark G. Lauby. New York, NY, USA: McGraw-Hill, Inc., 1994. (EPRI Power System Engineering Series). ISBN 0-07-035958-X.

34 KOSTENKO, M. P.; PIOTROVSKY, L. M. **Electrical Machines: Part two**. Moscow, RUS: Foreign Languages Publishing House, 1962. v. 2. ISBN 978-0714715360.

35 PADIYAR, K. R. **Power System Dynamics: Stability & Control**. Second ed. Hyderabad, India: BS Publications, 2008. ISBN 978-81-7800-186-9.

36 LIPO, T. A. **Analysis of Synchronous Machines**. Second ed. Boca Raton, FL, USA: CRC Press, May 2012. ISBN 978-1-4398-8068-5.

37 IEEE. The Authoritative Dictionary of IEEE Standards Terms. **IEEE Std 100-2000**, p. 1–1362, Dec. 2000. DOI: 10.1109/IEEESTD.2000.322230.

38 GURU, B. S.; HIZIROGLU, H. R. **Electric Machinery and Transformers**. Third ed. New York, NY, USA: Oxford University Press, Inc., 2001. (The Oxford Series in Electrical and Computer Engineering). ISBN 978-0-19-513890-0.

39 LAWRENCE, R. R. **Principles of Alternating Current Machinery**. Second ed. New York, NY, USA: McGraw-Hill Book Company, Inc., Apr. 1921. (Electrical Engineering Texts). ISBN 978-1-330-53797-8.

40 AGARWAL, P. D. Eddy-current losses in solid and laminated iron. **Transactions of the American Institute of Electrical Engineers, Part I: Communication and Electronics**, v. 78, n. 2, p. 169–181, May 1959. DOI: 10.1109/TCE.1959.6372977.

41 IEEE. Recommended phasor diagram for synchronous machines. **IEEE Transactions on Power Apparatus and Systems**, PAS-88, n. 11, p. 1593–1610, Nov. 1969. ISSN 0018-9510. DOI: 10.1109/TPAS.1969.292366.

42 ANDERSON, P. M.; AGRAWAL, B. L.; NESS, J. E. V. **Subsynchronous Resonance in Power Systems**. New York, NY, USA: Wiley–IEEE Press, 1990. ISBN 978-0-87942-258-5.

43 PARK, R. H. Two-reaction theory of synchronous machines - Generalized method of analysis - Part I. **Transactions of the American Institute of Electrical Engineers**, v. 48, n. 3, p. 716–727, July 1929. ISSN 0096-3860. DOI: 10.1109/T-AIEE.1929.5055275.

44 UMANS, S. **Fitzgerald & Kingsley's Electric Machinery**. Seventh ed. New York, NY, USA: McGraw-Hill, 2014. ISBN 978-0-07-338046-9.

45 KRON, G. **Equivalent Circuits of Electric Machinery**. New York, NY, USA: John Wiley & Sons, 1951. ISBN 978-0-84-462409-9.

46 CARTER, G. W. **The simple calculation of electrical transients: An elementary treatment of transient problems in electrical circuits by Heaviside's operational methods**. Cambridge, UK: Cambridge University Press, 1944.

47 BOYLESTAD, R. L. **Introductory Circuit Analysis**. Thirteenth ed. Essex, UK: Pearson Education, 2016. ISBN 978-1-292-09895-1.

48 STEINMETZ, C. P.; BERG, E. J. **Theory and Calculation of Alternating Current Phenomena**. Third ed. New York, NY, USA: Electrical World and Engineer Inc., 1900. ISBN 978-1-34-624243-9.

49 IEEE-1110. IEEE Guide for Synchronous Generator Modeling Practices and Applications in Power System Stability Analyses. **IEEE Std 1110-2002 Revision of IEEE Std 1110-1991**, p. 01–72, 2003. DOI: 10.1109/IEEESTD.2003.94408.

50 MACHOWSKI, J.; BIALEK, J. W.; BUMBY, J. R. **Power System Dynamics – Stability and Control**. Second ed. Chichester, WS, UK: John Wiley & Sons, Ltd, 2008. ISBN 978-0-470-72558-0.

51 MALEKPOUR, M.; AZIZIPANAH-ABARGHOOEE, R.; ZARE, M.; KIYOUMARSI, A.; TERZIJA, V. An explicit formulation for synchronous machine model in terms of the manufacturer data. **International Journal of Electrical Power & Energy Systems**, v. 108, p. 9–18, June 2019. ISSN 0142-0615. DOI: 10.1016/j.ijepes.2018.12.032.

52 STANLEY, H. C. An analysis of the induction motor. **AIEE Transactions**, v. 48, p. 716–727, 1938. ISSN 2376-7804. DOI: 10.1109/EE.1938.6431069.

53 BERETON, D. S.; LEWIS, D. G.; YOUNG, C. G. Representation of induction motor loads during power system stability studies. **AIEE Transactions**, v. 76, p. 451–461, Apr. 1957. ISSN 2379-6766. DOI: 10.1109/AIEEPAS.1957.4499587.

54 KRAUSE, P. C.; THOMAS, C. H. Simulation of symmetrical induction machinery. **IEEE Transactions on Power Apparatus Systems**, v. 84, p. 1038–1053, Nov. 1965. ISSN 0018-9510. DOI: 10.1109/TPAS.1965.4766135.

55 KRAUSE, P. C.; NOZARI, F.; SKVARENINA, T. L.; OLIVE, D. W. The theory of neglecting stator transients. **IEEE Transactions on Power Apparatus Systems**, v. 98, p. 141–148, Jan. 1979. ISSN 0018-9510. DOI: 10.1109/TPAS.1979.319542.

56 CLARKE, E. **Circuit Analysis of A-C Power Systems – Symmetrical and Related Components**. Schenectady, NY, USA: John Wiley & Sons, Inc., June 1943. One. (General Electric Advanced Engineering Program). ISBN 978-9-33-304261-1.

57 BLONDEL, A. E. **Synchronous Motors and Converters: Theory and Methods of Calculation and Testing**. Ed. by C. O. Mailloux. New York, NY, USA: McGraw-Hill Book Company, 1913. ISBN 978-1-33-061889-9.

58 DOHERTY, R. E.; NICKLE, C. A. Synchronous machines I: An extension of Blondel's two-reaction theory. **Journal of the A.I.E.E.**, v. 45, n. 10, p. 974–987, Oct. 1926. ISSN 0095-9804. DOI: 10.1109/JAIEE.1926.6537307.

59 ________. Synchronous machines III: Torque-angle characteristics under transient conditions. **Journal of the A.I.E.E.**, v. 46, n. 12, p. 1339–1339, Dec. 1927. ISSN 0095-9804. DOI: 10.1109/JAIEE.1927.6538137.

60 ________. Synchronous machines IV. **Transactions of the American Institute of Electrical Engineers**, v. 47, n. 2, p. 457–487, Apr. 1928. ISSN 0096-3860. DOI: 10.1109/T-AIEE.1928.5055006.

61 ________. Synchronous machines V: Three-phase short circuit synchronous machines. **Transactions of the American Institute of Electrical Engineers**, v. 49, n. 2, p. 700–714, Apr. 1930. ISSN 0096-3860. DOI: 10.1109/T-AIEE.1930.5055558.

62 PARK, R. H. Two-reaction theory of synchronous machines - Part II. **Transactions of the American Institute of Electrical Engineers**, v. 52, n. 2, p. 352–354, June 1933. ISSN 0096-3860. DOI: 10.1109/T-AIEE.1933.5056309.

63 FORTESCUE, C. L. Method of symmetrical co-ordinates applied to the solution of polyphase networks. **Transactions of the American Institute of Electrical Engineers**, v. XXXVII, n. 2, p. 1027–1140, July 1918. ISSN 0096-3860. DOI: 10.1109/T-AIEE.1918.4765570.

64 SHACKSHAFT, G.; PORAY, A. T. Implementation of new approach to determination of synchronous-machine parameters from tests. **Proceedings of the Institution of Electrical Engineers**, v. 124, n. 12, p. 1170–1178, Dec. 1977. ISSN 0020-3270. DOI: 10.1049/piee.1977.0246.

65 DE MELLO, F. P.; RIBEIRO, J. R. Derivation of synchronous machine parameters from tests. **IEEE Transactions on Power Apparatus and Systems**, v. 96, n. 4, p. 1211–1218, July 1977. ISSN 0018-9510. DOI: 10.1109/T-PAS.1977.32443.

66 DE MELLO, F. P.; HANNETT, L. H. Validation of synchronous machine models and derivation of model parameters from tests. **IEEE Transactions on Power Apparatus and Systems**, PAS-100, n. 2, p. 662–672, Feb. 1981. ISSN 0018-9510. DOI: 10.1109/TPAS.1981.316963.

67 COULTES, M. E.; WATSON, W. Synchronous machine models by standstill frequency response tests. **IEEE Transactions on Power Apparatus and Systems**, PAS-100, n. 4, p. 1480–1489, Apr. 1981. ISSN 0018-9510. DOI: 10.1109/TPAS.1981.316568.

68 DANDENO, P. L.; KUNDUR, P.; PORAY, A. T.; Z. EL-DIN, H. m. Adaptation and validation of turbogenerator model parameters through on-Line frequency response measurements. **IEEE Transactions on Power Apparatus and Systems**, PAS-100, n. 4, p. 1656–1664, Apr. 1981. ISSN 0018-9510. DOI: 10.1109/TPAS.1981.316560.

69 LEWIS, W. A. A basic analysis of synchronous machines – Part I. **Transactions of the American Institute of Electrical Engineers. Part III: Power Apparatus and Systems**, v. 77, n. 3, p. 436–453, Apr. 1958. ISSN 2379-6766. DOI: 10.1109/AIEEPAS.1958.4499954.

70 HARRIS, M. R.; LAWRENSON, P. J.; STEPHENSON, J. M. **Per-unit Systems with Special Reference to Electrical Machines**. Cambridge, UK: University Press, 1970. (IEE Monograph Series). ISBN 978-0-521-07857-3.

71 KIRTLEY, J. L. **Electric Power Principles: Sources, Conversion, Distribution and Use**. West Sussex, UK: John Wiley & Sons, Ltd., 2010. (Massachusetts Institute of Technology). ISBN 978-0-470-68636-2.

72 VALVERDE, G.; KYRIAKIDES, E.; HEYDT, G. T.; TERZIJA, V. Nonlinear estimation of synchronous machine parameters using operating data. **IEEE Transactions on Energy Conversion**, v. 26, n. 3, p. 831–839, Sept. 2011. ISSN 0885-8969. DOI: 10.1109/TEC.2011.2141136.

73 HOSSEINI, S. M.; ABDOLLAHI, R.; KARRARI, M. Inclusive design and implementation of online load angle measurement for real-time transient stability improvement of a synchronous generator in a smart grid. **IEEE Transactions on Industrial Electronics**, v. 65, n. 11, p. 8966–8972, Nov. 2018. ISSN 0278-0046. DOI: 10.1109/TIE.2018.2811394.

74 LEVI, E.; LEVI, V. A. Impact of dynamic cross-saturation on accuracy of saturated synchronous machine models. **IEEE Transactions on Energy Conversion**, v. 15, n. 2, p. 224–230, June 2000. ISSN 0885-8969. DOI: 10.1109/60.867004.

75 FILHO, E. R.; NUNES JUNIOR, F. L.; NUNES, S. O. Synchronous machine field current calculation taking into account the magnetic saturation. **Revista Controle & Automação**, v. 13, n. 2, p. 165–170, May 2002.

76 JING, W.; TAN, G. Modeling of salient-pole synchronous motor considering saturation effect. In: 2010 International Conference on Computer Application and System Modeling (ICCASM 2010). Taiyuan, China: IEEE, Oct. 2010. v. 3, p. 277–281. DOI: 10.1109/ICCASM.2010.5620132.

77 ZAKER, B.; GHAREHPETIAN, G. B.; KARRARI, M. Improving synchronous generator parameters estimation using d-q axes tests and considering saturation effect. **IEEE Transactions on Industrial Informatics**, 2017. ISSN 1551-3203. DOI: 10.1109/TII.2017.2759502.

78 THERRIEN, F.; CHAPARIHA, M.; JATSKEVICH, J. Constant-parameter synchronous machine model including main flux saturation. **IET Electric Power Applications**, v. 10, n. 6, p. 477–487, June 2016. ISSN 1751-8679. DOI: 10.1049/iet-epa.2015.0447.

79 ZAKER, B.; GHAREHPETIAN, G. B.; KARRARI, M. Improving synchronous generator parameters estimation using d and q axes tests and considering saturation effect. **IEEE Transactions on Industrial Informatics**, v. 14, n. 5, p. 1898–1908, May 2018. ISSN 1941-0050. DOI: 10.1109/TII.2017.2759502.

80 SHACKSHAFT, G. New approach to the determination of synchronous-machine parameters from tests. **Proceedings of the Institution of Electrical Engineers**, v. 121, n. 11, p. 1385–1392, Nov. 1974. ISSN 0020-3270. DOI: 10.1049/piee.1974.0289.

81 GIESBRECHT, M. A method to determine load conditions for salient pole synchronous machine quadrature axis parameters determination. In: 2017 IEEE 26th International Symposium on Industrial Electronics (ISIE). Edinburgh, UK: IEEE, June 2017. P. 202–208. DOI: 10.1109/ISIE.2017.8001248.

82 ZADEH, L. A. From circuit theory to system theory. **Proceedings of the IRE**, v. 50, n. 5, p. 856–865, May 1962. ISSN 2162-6634. DOI: 10.1109/JRPROC.1962.288302.

83 SARACHIK, P. E. **Principles of Linear Systems**. Cambridge, UK: Cambridge University Press, 1997. ISBN 978-0-521-57057-2.

84 BARRETO, G. **Modelagem Computacional Distribuída e Paralela de Sistemas e de Séries Temporais Multivariáveis no Espaço de Estado**. Feb. 2002. PhD thesis – University of Campinas, Campinas, SP, BR. Programa de Pós-Graduação em Engenharia Elétrica.

85 AXLER, S. **Linear Algebra Done Right**. San Francisco, CA, USA: Springer International Publishing, 2015. (Undergraduate Texts in Mathematics). ISBN 978-3-319-11079-0. DOI: 10.1007/978-3-319-11080-6.

86 LUENBERGER, D. G. **Introduction to Dynamic Systems – Theory, Models, and Applications**. Palo Alto, CA, USA: John Wiley & Sons, Inc., May 1979. ISBN 0-471-02594-1.

87 ISIDORI, A. **Nonlinear control systems**. Third ed. London, UK: Springer-Verlag London, 1995. (Communications and control engineering series). ISBN 978-1-4471-3909-6. DOI: 10.1007/978-1-84628-615-5.

88 JAZWINSKI, A. H. **Stochastic Process and Filtering Theory**. Ed. by Richard Bellman. New York, NY, USA: Academic Press, Inc., Jan. 1970. v. 64. (Mathematics in science and engineering). ISBN 978-0-08-096090-6.

89 KALMAN, R. Contributions to the theory of optimal control. **Boletin de la Sociedad Matematica Mexicana**, v. 5, p. 102–119, 1960. Research Institute for Advanced Studies (Martin Marietta Corporation). ISSN 2296-4495.

90 ________. On the general theory of control systems. **International Federation of Automatic Control Proceedings**, v. 1, n. 1, p. 481–491, Aug. 1960. ISSN 1474-6670. DOI: doi.org/10.1016/S1474-6670(17)70094-8.

91 LEE, E. B.; MARKUS, L. **Foundations of Optimal Control Theory**. New York, NY, USA: John Wiley & Sons, Inc., 1967. ISBN 978-0-4715-2263-8.

92 GRIFFITH, E. W.; KUMAR, K. S. P. On the observability of nonlinear systems: I. **Journal of Mathematical Analysis and Applications**, v. 35, n. 1, p. 135–147, July 1971. ISSN 0022-247X. DOI: https://doi.org/10.1016/0022-247X(71)90241-1.

93 KRENER, A. J. **A generalization of the Pontryagin maximal principle and the bang-bang principle**. 1971. PhD thesis – University of California, Berkeley, CA, USA.

94 ________. A generalization of Chow's theorem and the bang-bang theorem to nonlinear control problems. **SIAM Journal on Control**, v. 12, n. 1, p. 43–52, Mar. 1974. ISSN 0036-1402. DOI: https://doi.org/10.1137/0312005.

95 SUSSMANN, H. J.; JURDJEVIC, V. Controllability of nonlinear systems. **Journal of Differential Equations**, v. 12, n. 1, p. 95–116, July 1972. ISSN 0022-0396. DOI: 10.1016/0022-0396(72)90007-1.

96 HERMANN, R.; KRENER, A. Nonlinear controllability and observability. **IEEE Transactions on Automatic Control**, v. 22, n. 5, p. 728–740, Oct. 1977. ISSN 1558-2523. DOI: 10.1109/TAC.1977.1101601.

97 AGUIRRE, L. A.; PORTES, L. L.; LETELLIER, C. Structural, dynamical and symbolic observability: from dynamical systems to networks. **Plos One**, v. 13, n. 10, p. 1–21, Oct. 2018. ISSN 1932-6203. DOI: 10.1371/journal.pone.0206180.

98 NAHAR, J.; LIU, J.; SHAH, S. L. Parameter and state estimation of an agro-hydrological system based on system observability analysis. **Computer and Chemical Engineering**, v. 121, p. 450–464, Feb. 2019. ISSN 0098-1354. DOI: 10.1016/j.compchemeng.2018.11.015.

99 CANDY, J. V. **Bayesian Signal Processing: Classical, Modern, and Particle Filtering Methods**. Second ed. Hoboken, NJ, USA: John Wiley & Sons, Inc, 2016. ISBN 978-1-11-912545-7.

100 NELLES, O. **Nonlinear System Identification – From Classical Approaches to Neural Networks and Fuzzy Models**. Germany: Springer – Verlag Berlin Heidelberg, 2001. ISBN 978-3-662-04323-3.

101 THOMAS, C. H. Discussion of 'Analogue computer representations of synchronous generators in voltage-regulation studies'. **Transactions of the American Institute of Electrical Engineers. Part III: Power Apparatus and Systems**, v. 75, n. 3, p. 1182–1184, Jan. 1956. ISSN 0097-2460. DOI: 10.1109/AIEEPAS.1956.4499421.

102 RUDENBERG, R. Saturated synchronous machines under transient conditions in the pole axis. **Transactions of the American Institute of Electrical Engineers**, v. 61, n. 6, p. 297–306, June 1942. ISSN 2330-9431. DOI: 10.1109/T-AIEE.1942.5058532.

103 ROBERT, C. P.; CASELLA, G. **Monte Carlo Statistical Methods**. Second ed. New York, NY, USA: Springer Science+Business Media, 2004. (Springer texts in statistics). ISBN 978-1-4757-4145-2. DOI: 10.1007/978-1-4757-4145-2.

104 FISHER, J. R. Optimal nonlinear filtering. **Advances in Control Systems**, v. 5, p. 197–300, 1967. ISSN 0065-2466. DOI: https://doi.org/10.1016/B978-1-4831-6718-3.50010-4.

105 SIMON, D. **Optimal State Estimation: Kalman, $\mathcal{H}_\infty$, and Nonlinear Approaches**. Hoboken, NJ, USA: John Wiley & Sons, Inc., June 2006. ISBN 978-0-471-70858-2.

106 CAPPÉ, O.; MOULINES, E.; RYDEN, T. **Inference in Hidden Markov Models**. New York, NY, USA: Springer-Verlag New York, 2005. (Springer Series in Statistics). ISBN 978-0-387-28982-3.

107 DOUCET, A.; FREITAS, N. de; GORDON, N. **Sequential Monte Carlo Methods in Practice**. New York, NY, USA: Springer Science+Business Media Inc., 2001. (Statistics for engineering and information science). ISBN 978-1-4757-3437-9. DOI: 10.1007/978-1-4757-3437-9.

108 HOGG, R. V.; MCKEAN, J. W.; CRAIG, A. T. **Introduction to Mathematical Statistics**. Eighth ed. Boston, MA, USA: Pearson Education, Inc., 2019. ISBN 978-0-13-468699-8.

109 MAYBECK, P. S. **Stochastics Models, Estimation, and Control: Introduction**. First ed. New York, NY, USA: Academic Press, May 1979. 141A. (Mathematics in Science and Engineering). ISBN 978-0-08-095650-3.

110 GIESBRECHT, M. **Propostas imuno-inspiradas para identificação de sistemas e realização de séries temporais multivariáveis (Immuno-inspired approaches for state-space multivariable system identification and time series realization)**. June 2013. PhD thesis – University of Campinas, Campinas, SP, BR.

111 PAPOULIS, A.; PILLAI, S. U. **Probability, Random Variables and Stochastic Processes**. Fourth ed. New York, NY, USA: McGraw-Hill, 2002. ISBN 978-0-07-122661-5.

112 HOGG, R. V.; MCKEAN, J. W.; CRAIG, A. T. **Introduction to Mathematical Statistics**. Eighth ed. New York, NY, USA: Pearson Education, Inc., 2019. ISBN 978-0-13-468699-8.

113 LIU, J. S. **Monte Carlo Strategies in Scientific Computing**. New York, NY, USA: Springer-Verlag New York, 2004. ISBN 978-0-387-76371-2. DOI: 10.1007/978-0-387-76371-2.

114 FITZGERALD, W. J. Markov chain Monte Carlo methods with applications to signal processing. **Signal Processing**, v. 81, n. 1, p. 3–18, Jan. 2001. ISSN 0165-1684. DOI: 10.1016/S0165-1684(00)00187-0.

115 ANDRIEU, C.; FREITAS, N. de; DOUCET, A.; JORDAN, M. I. An introduction to Monte Carlo Markov chain for machine learning. **Machine Learning**, v. 50, n. 1–2, p. 5–43, Jan. 2003. ISSN 1573-0565. DOI: 10.1023/A:1020281327116.

116 EFRON, B. Breakthroughs in Statistics. In: ed. by S. Kotz and N. L. Johnson. New York, NY, USA: Springer-Verlag New York, Inc., 1992. Bootstrap Methods: Another Look at the Jackknife, p. 569–593. (Springer Series in Statistics book series). ISBN 978-1-4612-4380-9. DOI: 10.1007/978-1-4612-4380-9_41.

117 EFRON, B.; TIBSHIRANI, R. J. **An Introduction to the Bootstrap**. First ed. Boca Raton, FL, USA: CRC Press, May 1993. (Chapman & Hall/CRC Monographs on Statistics and Applied Probability). ISBN 978-0-41-204231-7.

118 DAVISON, A. C.; HINKLEY, D. V. **Bootstrap Methods and their Application**. Cambridge, UK: Cambridge University Press, June 1997. (Cambridge Series in Statistical and Probabilistic Mathematics). ISBN 978-0-511-80284-3. DOI: 10.1017/CBO9780511802843.

119 DUDA, R. O.; HART, P. E.; STORK, D. G. **Pattern Classification**. Second ed. New York, NY, USA: Wiley-Interscience, Nov. 2000. ISBN 978-0-471-05669-0.

120 VAN TREES, H. J. **Detection, Estimation, and Modulation Theory – Part One**. New York, NY, USA: John Wiley & Sons, 2001. ISBN 978-0-471-22108-1.

121 KALMAN, R. E. **Analysis and Synthesis of Linear Systems Operating on Randomly Sampled Data**. Aug. 1957. PhD thesis – Columbia University, New York, NY, USA.

122 BALCHEN, J. G. How have we arrived at the present state of knowledge in process control? Is there a lesson to be learned? **Modeling, Identification and Control**, v. 20, n. 2, p. 63–73, 1999. ISSN 1890-1328. DOI: 10.4173/mic.1999.2.1.

123 GEVERS, M. A personal view of the development of system identification: a 30-year journey through an exciting field. **IEEE Control Systems Magazine**, v. 26, n. 6, p. 93–105, Dec. 2006. ISSN 1066-033X. DOI: 10.1109/MCS.2006.252834.

124 SCHMIDT, S. F. Advances in Control Systems. In: ed. by C. T. Leondes. New York, NY, USA: Academic Press, Inc., 1966. v. 3 Application of State-Space Methods to Navigation Problems, p. 293–340. ISBN 978-1-4831-6716-9.

125 GREWAL, M. S.; ANDREWS, A. P. **Kalman Filtering: Theory and Practice using MATLAB**. Fourth ed. Hoboken, NJ, USA: John Wiley & Sons, Inc., 2015. ISBN 978-0-47-139254-5.

126 CRASSIDIS, J. L.; JUNKINS, J. L. **Optimal Estimation of Dynamic Systems**. Ed. by Goong Chen and Thomas J. Bridges. Boca Raton, FL, USA: Chapman & Hall, 2004. (CRC Applied Mathematics and Nonlinear Science Series). ISBN 978-1-43-983985-0.

127 ANALYTIC SCIENCES CORPORATION, T. **Applied Optimal Estimation**. Ed. by Arthur Gelb. Cambridge, MA, USA: The MIT Press, 1974. ISBN 978-0-262-57048-0.

128 JULIER, S. J.; UHLMANN, J. K. Unscented filtering and nonlinear estimation. **Proceedings of the IEEE**, v. 92, n. 3, p. 401–422, Nov. 2004. ISSN 1558-2256. DOI: 10.1109/JPROC.2003.823141.

129 JULIER, S.; UHLMANN, J.; DURRANT-WHYTE, H. F. A new method for the nonlinear transformation of means and covariances in filters and estimators. **IEEE Transactions on Automatic Control**, v. 45, n. 3, p. 477–482, Mar. 2000. ISSN 0018-9286. DOI: 10.1109/9.847726.

130 UHLMANN, J. **First-Hand: The Unscented Transform – An interview with Jeffrey Uhlmann**. Columbia, MO, USA: Engineering and Technology History Wiki, Nov. 2012. Online at ethw.org/First-Hand:The-Unscented-Transform.

131 ITO, K.; XIONG, K. Gaussian filters for nonlinear filtering problems. **IEEE Transactions on Automatic Control**, v. 45, n. 5, p. 910–927, May 2000. ISSN 0018-9286. DOI: 10.1109/9.855552.

132 ARASARATNAM, I.; HAYKIN, S.; ELLIOTT, R. J. Discrete-time nonlinear filtering algorithms using Gauss–Hermite quadrature. **Proceedings of the IEEE**, v. 95, n. 5, p. 953–977, May 2007. ISSN 0018-9219. DOI: 10.1109/JPROC.2007.894705.

133 BUCY, R. S.; SENNE, K. D. Digital synthesis of nonlinear filters. **Automatica**, v. 7, n. 3, p. 287–298, May 1971. ISSN 0005-1098. DOI: 10.1016/0005-1098(71)90121-X.

134 ALSPACH, D. L.; SORENSON, H. W. Nonlinear Bayesian estimation using Gaussian-Sum approximations. **IEEE Transactions on Automatic Control**, v. 17, n. 4, p. 439–448, Aug. 1972. ISSN 0018-9286.

135 SORENSON, H. W.; ALSPACH, D. L. Recursive Bayesian estimation using Gaussian sums. **Automatica**, v. 7, n. 4, p. 465–479, July 1971. ISSN 0005-1098.

136 EVENSEN, G. Sequential data assimilation with nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. **Journal of Geophysical Research: Oceans**, v. 99, n. C5, p. 143–162, May 1994. ISSN 2169-9291. DOI: doi:10.1029/94JC00572.

137 ________. The ensemble Kalman Filter for combined state and parameter estimation. **IEEE Control Systems Magazine**, v. 29, n. 3, p. 83–104, June 2009. ISSN 1066-033X. DOI: 10.1109/MCS.2009.932223.

138 METROPOLIS, N.; ULAM, S. The Monte Carlo method. **Journal of the American Statistical Association**, v. 44, n. 247, p. 335–341, Sept. 1949. ISSN 0162-1459. DOI: 10.2307/2280232.

139 WIENER, N. **I am a Mathematician – The later life of a prodigy**. Cambridge, MA, USA: The MIT Press, Aug. 1964. (Autobiography-Science). ISBN 978-0-26-273007-5.

140 HAMMERSLEY, J. M.; MORTON, K. W. Poor man's Monte Carlo. **Journal of the Royal Statistical Society**, v. 16, n. 1, p. 23–38, Jan. 1954. ISSN 1369-7412.

141 KITAGAWA, G. Monte Carlo filter and smoother for non-Gaussian nonlinear state space models. **Journal of Computational and Graphical Statistics**, v. 5, n. 1, p. 1–25, Dec. 1996. ISSN 1537-2715. DOI: 10.1080/10618600.1996.10474692.

142 CARPENTER, J.; CLIFFORD, P.; FEARNHEAD, P. Improved particle filter for nonlinear problems. **IEE Proceedings - Radar, Sonar and Navigation**, v. 146, n. 1, p. 2–7, 1999. ISSN 1350-2395. DOI: 10.1049/ip-rsn:19990255.

143 HANDSCHIN, J. E.; MAYNE, D. Q. Monte Carlo techniques to estimate the conditional expectation in multi-stage non-linear filtering. **International Journal of Control**, v. 9, n. 5, p. 547–559, July 1969. ISSN 0020-7179. DOI: 10.1080/00207176908905777.

144 DJURIC, P. M.; KOTECHA, J. H.; ZHANG, J.; HUANG, Y.; GHIRMAI, T.; BUGALLO, M. F.; MIGUEZ, J. Particle filtering. **IEEE Signal Processing Magazine**, v. 20, n. 5, p. 19–38, Sept. 2003. ISSN 1558-0792. DOI: 10.1109/MSP.2003.1236770.

145 ARULAMPALAM, M. S.; MASKELL, S.; GORDON, N.; CLAPP, T. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. **IEEE Transactions on Signal Processing**, v. 50, n. 2, p. 174–188, Feb. 2002. ISSN 1941-0476. DOI: 10.1109/78.978374.

146 GORDON, N. J.; SALMOND, D. J.; SMITH, A. F. M. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. **IEEE Proceedings F - Radar and Signal Processing**, v. 140, n. 2, p. 107–113, Apr. 1993. ISSN 0956-375X. DOI: 10.1049/ip-f-2.1993.0015.

147 ISAARD, M.; BLAKE, A. Contour tracking by stochastic propagation of conditional density. In: ________. **Computer Vision – ECCV '96**. Berlin, GER: Springer Berlin Heidelberg, 1996. v. 1064. (Lecture Notes in Computer Science), p. 343–356. ISBN 978-3-540-49949-7. DOI: 10.1007/BFb0015549.

148 MACCORMICK, J.; BLAKE, A. A probabilistic exclusion principle for tracking multiple objects. In: PROCEEDINGS of the Seventh IEEE International Conference on Computer Vision. Kerkyra, GR: IEEE, Sept. 1999. v. 1, p. 572–578. DOI: 10.1109/ICCV.1999.791275.

149 MORAL, P. D. Measure valued processes and interacting particle systems: Application to non-linear filtering problems. **The Annals of Applied Probability**, v. 8, n. 2, p. 438–495, Aug. 1998. ISSN 2168-8737. DOI: 10.1214/aoap/1028903535.

150 ANDRIEU, C.; DOUCET, A.; SINGH, S. S.; TADIC, V. B. Particle methods for change detection, system identification, and control. **Proceedings of the IEEE**, v. 92, n. 3, p. 423–438, Mar. 2004. ISSN 1558-2256. DOI: 10.1109/JPROC.2003.823142.

151 CRISAN, D.; DOUCET, A. A survey of convergence results on particle filtering methods for practitioners. **IEEE Transactions on Signal Processing**, v. 50, n. 3, p. 736–746, Mar. 2002. ISSN 1941-0476. DOI: 10.1109/78.984773.

152 KANAZAWA, K.; KOLLER, D.; RUSSELL, S. Stochastic Simulation Algorithms for Dynamic Probabilistic Networks. In: PROCEEDINGS of the Eleventh Conference on Uncertainty in Artificial Intelligence (UAI1995). Montréal, Qué, CAN: Morgan Kaufmann Publishers Inc., Aug. 1995. P. 346–351.

153 HIGUCHI, T. Monte Carlo filter using the genetic algorithm operators. **Journal of Statistical Computation and Simulation**, Taylor & Francis, v. 59, n. 1, p. 1–23, Feb. 1997. ISSN 1563-5163. DOI: 10.1080/00949659708811843. eprint: https://doi.org/10.1080/00949659708811843.

154 GEWEKE, J. Bayesian inference in econometric models using Monte Carlo integration. **Econometrica**, v. 57, n. 6, p. 1317–1339, Nov. 1989. ISSN 1468-0262. DOI: 10.2307/1913710.

155 HENDRY, D. F.; RICHARD, J.-F. Likelihood evaluation for dynamic latent variables models. In: ed. by Hans M. Amman, David A. Belsey and Louis F. Pau. Dordrecht, NLD: Springer Science+Business Media, 1992. v. 22 One, p. 3–18. (Computational Economics and Econometrics. Advanced Studies in Theoretical and Applied Econometrics). ISBN 978-94-011-3162-9.

156 CAPPÉ, O.; GODSILL, S.; MOULINES, eric. An overview of existing methods and recent advances in sequential Monte Carlo. **Proceedings of the IEEE**, v. 95, n. 5, p. 899–924, May 2007. ISSN 1558-2256. DOI: 10.1109/JPROC.2007.893250.

157 WANG, D.; ZHAO, J.; WANG, W. Particle filter based robust mobile robot localization. In: IEEE. INTERNATIONAL Conference on Mechatronics and Automation. [S.l.: s.n.], 2009. P. 817–821.

158 RUBIN, D. B. Using the SIR algorithm to simulate posterior distributions. **Bayesian Statistics**, v. 3, n. 1, p. 395–402, Jan. 1988. ISSN 1931-6690.

159 KITAGAWA, G. Non-Gaussian state-space modeling of nonstationary time series. **Journal of the American Statistical Association**, v. 82, n. 400, p. 1032–1041, Dec. 1987. ISSN 1537-274X. DOI: 10.2307/2289375.

160 HÜRZELER, M.; KÜNSCH, H. R. Monte Carlo approximations for general states-space models. **Journal of Computational and Graphical Statistics**, v. 7, n. 2, p. 175–193, May 1998. ISSN 1537-2715. DOI: 10.1080/10618600.1998.10474769.

161 SILVERMAN, B. W. **Density Estimation for Statistics and Data Analysis**. London, UK: Chapman and Hall/CRC, 1986. (Chapman & Hall/CRC Monographs on Statistics and Applied Probability). ISBN 978-0-412-24620-3.

162 GILKS, W. R.; BERZUINI, C. Following a moving target – Monte Carlo inference for dynamic Bayesian models. **Journal of the Royal Statistical Society: Series B (Statistical Methodology)**, v. 63, n. 1, p. 127–146, Feb. 2001. ISSN 1467-9868. DOI: 10.1111/1467-9868.00280.

163 PITT, M. K.; SHEPHARD, N. Filtering via simulation: Auxiliary Particle Filters. **Journal of the American Statistical Association**, v. 94, n. 466, p. 590–599, June 1999. ISSN 1537-274X. DOI: 10.2307/2670179.

164 METROPOLIS, N.; ROSENBLUTH, A. W.; ROSENBLUTH, M. N.; TELLER, A. H.; TELLE, E. Equations of state calculations by fast computing machines. **The Journal of Chemical Physics**, v. 21, n. 6, p. 1087–1092, June 1953. ISSN 0021-9606. DOI: 10.1063/1.1699114.

165 HASTINGS, W. K. Monte Carlo sampling methods using Markov chains and their applications. **Biometrika**, v. 57, n. 1, p. 97–109, Apr. 1970. ISSN 1464-3510. DOI: 10.2307/2334940.

166 GELFAND, A. E.; SMITH, A. F. M. Sampling-based approaches to calculating marginal densities. **Journal of the American Statistical Association**, v. 85, n. 410, p. 398–409, June 1990. ISSN 1537-274X. DOI: 10.2307/2289776.

167 KITAGAWA, G. A self-organizing state-space model. **Journal of the American Statistical Association**, v. 93, n. 443, p. 1203–1215, Sept. 1998. ISSN 1537-274X. DOI: 10.2307/2669862.

168 SOLO, V. Some aspects of recursive parameter estimation. **International Journal of Control**, v. 32, n. 3, p. 395–410, Jan. 1980. ISSN 1366-5820. DOI: 10.1080/00207178008922864.

169 ANDERSON, B. D. O.; MOORE, J. B. **Optimal Filtering**. Ed. by Thomas Kailath. Englewood Cliffs, NJ, USA: Prentice-Hall, Inc., 1979. (Prentice-Hall Information and System Sciences Series). ISBN 978-0-13-638122-8.

170 LJUNG, L.; SODERSTROM, T. **Theory and Practice of Recursive Identification**. Boston, MA, USA: The MIT Press, Feb. 1987. (MIT Press Series in Signal Processing, Optimization, and Control). ISBN 978-0-26-262058-1.

171 AGUIRRE, L. A. **Introdução à Identificação de Sistemas: Técnicas lineares e não lineares – Teoria e Aplicação**. Fourth ed. Belo Horizonte, MG, BR: Editora UFMG, Jan. 2015. ISBN 978-8-54-230079-6.

172 LJUNG, L. On consistency and identifiability. In: **Stochastic Systems: Modeling, Identification and Optimization, I**. Ed. by Roger J. B. Wets. Berlin, Heidelberg: Springer Berlin Heidelberg, 1976. P. 169–190. ISBN 978-3-642-00784-2. DOI: 10.1007/BFb0120772.

173 ________. **System Identification: Theory for the User**. Second ed. Upper Saddle River, NJ, USA: Prentice Hall PTR, 1999. (Prentice Hall Information and System Sciences Series). ISBN 978-0-13-656695-3.

174 BOHLIN, T. P. **Practical grey-box process identification: theory and applications**. London, UK: Springer-Verlag London Limited, 2006. (Advances in Industrial Control). ISBN 978-1-84628-402-1.

175 NIELSEN, H. A.; MADSEN, H. Modelling the heat consumption in district heating systems using a grey-box approach. **Energy and Buildings**, v. 38, n. 1, p. 63–71, 2006. ISSN 0378-7788. DOI: https://doi.org/10.1016/j.enbuild.2005.05.002.

176 KARRARI, M.; MALIK, O. P. Identification of Heffron-Phillips model parameters for synchronous generators using online measurements. **IEEE Proceedings - Generation, Transmission and Distribution**, v. 151, n. 3, p. 313–320, 2004. ISSN 1350-2360. DOI: 10.1049/ip-gtd:20040275.

177 ________. Identification of physical parameters of a synchronous generator from online measurements. **IEEE Transactions on Energy Conversion**, v. 19, n. 2, p. 407–415, 2004. ISSN 0885-8969. DOI: 10.1109/TEC.2003.822296.

178 KIM, D.; MOON, Y.; LEE, J.; RYU, H.; KIM, T. A new method of recording generator dynamics and its application to the derivation of synchronous machine parameters for power system stability studies. **IEEE Transactions on Energy Conversion**, v. 33, n. 2, p. 605–616, June 2018. ISSN 0885-8969. DOI: 10.1109/TEC.2017.2772234.

179 DEL ANGEL, A.; GEURTS, P.; ERNST, D.; GLAVIC, M.; WEHENKEL, L. Estimation of rotor angles of synchronous machines using artificial neural networks and local PMU-based quantities. **Neurocomputing**, v. 70, n. 16–18, p. 2668–2678, Oct. 2007. ISSN 0925-2312. DOI: 10.1016/j.neucom.2006.12.017.

180 GHAHREMANI, E.; KAMWA, I. Simultaneous state and input estimation of a synchronous machine using the Extended Kalman Filter with unknown inputs. In: 2011 IEEE International Electric Machines Drives Conference (IEMDC). Niagara Falls, ON, Canada: IEEE, May 2011. P. 1468–1473. DOI: 10.1109/IEMDC.2011.5994825.

181 FAN, L.; WEHBE, Y. Extended Kalman filtering based real-time dynamic state and parameter estimation using PMU data. **Electric Power System Research**, v. 103, p. 168–177, Oct. 2013. ISSN 0378-7796. DOI: 10.1016/j.epsr.2013.05.016.

182 LI, Y.; LI, J.; QI, J.; CHEN, L. Robust cubature Kalman Filter for dynamic state estimation of synchronous machines under unknown measurement noise statistics. **IEEE Access**, v. 7, p. 29139–29148, Feb. 2019. ISSN 2169-3536. DOI: 10.1109/ACCESS.2019.2900228.

183 JOSEPH, T.; TYAGI, B.; KUMAR, V. Unbiased minimum variance filter-based generator state estimation using PMU measurements for unknown generator input. **IEEE Systems Journal**, v. 13, n. 3, p. 3176–3184, Sept. 2019. ISSN 1937-9234. DOI: 10.1109/JSYST.2018.2878413.

184 ZHAO, J.; MILI, L. A decentralized H_∞ Unscented Kalman Filter for dynamic state estimation against uncertainties. **IEEE Transactions on Smart Grid**, v. 10, n. 5, p. 4870–4880, Sept. 2019. ISSN 1949-3061. DOI: 10.1109/TSG.2018.2870327.

185 PHADKE, A. G.; THORP, J. S.; ADAMIAK, M. G. A new measurement technique for tracking voltage phasors, local system frequency, and rate of change of frequency. **IEEE Transactions on Power Apparatus and Systems**, PAS-102, n. 5, p. 1025–1038, May 1983. ISSN 0018-9510. DOI: 10.1109/TPAS.1983.318043.

186 MA, J.; MAKAROV, Y.; DONG, Z. Emerging Techniques in Power System Analysis. In: ed. by Zhaoyang Dong and Pei Zhang. Berlin, GER: Springer, 2010. Phasor Measurement Unit and Its Application in Modern Power Systems, p. 147–184. ISBN 978-3-642-04282-9. DOI: 10.1007/978-3-642-04282-9-6.

187 LIU, Y.; LIN, F.; CHU, X. Transient stability prediction based on PMU and FCRBFN. In: PROCEEDINGS of International Conference on Power System Management and Control. London, UK: IET, Apr. 2002. P. 281–284. DOI: 10.1049/cp:20020048.

188 BORTONI, E. C.; JARDINI, J. A. Identification of synchronous machine parameters using load rejection test data. **IEEE Transactions on Energy Conversion**, v. 17, n. 2, p. 242–247, June 2002. ISSN 0885-8969. DOI: 10.1109/TEC.2002.1009475.

189 SILVA, P. V. V.; BORTONI, E. C.; ROCHA, E. J. J. Identification of synchronous machines parameters using genetic algorithm and load rejection test. In: 2017 IEEE Power Energy Society General Meeting. Chicago, IL, USA: IEEE, July 2017. P. 1–5. DOI: 10.1109/PESGM.2017.8274095.

190 RADJEAI, H.; BARAKAT, A.; TNANI, S.; CHAMPENOIS, G. Identification of synchronous machine by Standstill Frequency Response (SSFR) method - influence of the stator resistance. In: THE XIX International Conference on Electrical Machines - ICEM 2010. Rome, ITL: IEEE, Sept. 2010. P. 1–5. DOI: 10.1109/ICELMACH.2010.5608146.

191 COULTES, M. E.; WATSON, W. Synchronous machine models by standstill frequency response tests. **IEEE Transactions on Power Apparatus and Systems**, PAS-100, n. 4, p. 1480–1489, Apr. 1981. ISSN 0018-9510. DOI: 10.1109/TPAS.1981.316568.

192 SUGIURA, O.; AKIYAMA, Y. Precise method for measuring Xd and Xq based on slip test of synchronous machines. In: CONFERENCE Record of the 1993 IEEE Industry Applications Conference Twenty-Eighth IAS Annual Meeting. Toronto, ON, CA: IEEE, Oct. 1993. v. 1, p. 155–162. DOI: 10.1109/IAS.1993.298918.

193 JACK, A. G.; BEDFORD, T. J. An analysis of the results from the computation of transients in synchronous generators using frequency domain methods. **IEEE Transactions on Energy Conversion**, v. 3, n. 2, p. 375–383, June 1988. ISSN 0885-8969. DOI: 10.1109/60.4744.

194 KAMWA, I.; VIAROUGE, P.; DICKINSON, E. J. Optimal estimation of the generalized operational impedances of synchronous machines from short-circuit tests. **IEEE Transactions on Energy Conversion**, v. 5, n. 2, p. 401–407, June 1990. ISSN 0885-8969. DOI: 10.1109/60.107239.

195 TESSAROLO, A.; MOHAMADIAN, S.; BORTOLOZZI, M. A new method for determining the leakage inductances of a nine-phase synchronous machine from no-load and short-circuit tests. **IEEE Transaction on Energy Conversion**, v. 30, n. 4, p. 1515–1527, Dec. 2015. ISSN 0885-8969. DOI: 10.1109/TEC.2015.2458182.

196 GIESBRECHT, M. A phasor diagram based method to determine load conditions for load rejection tests for round rotor synchronous machine quadrature axis parameters determination. In: 2016 IEEE 25th International Symposium on Industrial Electronics (ISIE). Santa Clara, CA, USA: IEEE, June 2016. P. 122–127. DOI: 10.1109/ISIE.2016.7744877.

197 DA COSTA BORTONI, E.; JARDINI, J. A. Synchronous machines parameters identification using load rejection test data. In: 1997 IEEE International Electric Machines and Drives Conference Record. [S.l.: s.n.], May 1997. wb1/1.1–wb1/1.3. DOI: 10.1109/IEMDC.1997.604292.

198 DE OLIVEIRA, S. E. M.; DE SOUZA, J. A. Effect of field-voltage source impedance on load-rejection test results of large-rating synchronous generators. **IEEE Transactions on Energy Conversion**, v. 26, n. 1, p. 30–35, Mar. 2011. ISSN 0885-8969. DOI: 10.1109/TEC.2010.2066277.

199 TIOMO, D.; KENFACK, E. Y.; WAMKEUE, R. Dynamic study and analysis of synchronous generator under sudden short-circuit and load rejection tests. In: 2019 IEEE/IAS 55th Industrial and Commercial Power Systems Technical Conference. Calgary, CAN: IEEE, May 2019. DOI: 10.1109/ICPS.2019.8733379.

200 WAMKEUE, R.; KAMWA, I. Detailed analysis of load rejection test of autonomous synchronous generator. In: ICEMS'2001. Proceedings of the Fifth International Conference on Electrical Machines and Systems. Shenyang, CHI: IEEE, Aug. 2001. DOI: 10.1109/ICEMS.2001.970702.

201 WAMKEUE, R.; CHRISTIAN, J.; KAMWA, I. New Approach for Partial and Full Load Rejection Analysis of Synchronous Generator. In: CANADIAN Conference on Electrical and Computer Engineering. Ottawa, ON, CA: IEEE, May 2006. P. 1240–1245. DOI: 10.1109/CCECE.2006.277528.

202 WAMKEUE, R.; JOLETTE, C.; KAMWA, I. Analytical response of synchronous generators during load rejection and field short-circuit tests. **Electric Power Components and Systems**, v. 35, n. 7, p. 803–821, Apr. 2007. ISSN 1532-5016. DOI: 10.1080/15325000601175199.

203 WAMKEUE, R.; BAETSCHER, F.; KAMWA, I. Hybrid-state-model-based time-domain identification of synchronous machine parameters from saturated load rejection test records. **IEEE Transactions on Energy Conversion**, v. 23, n. 1, p. 68–77, 2008. ISSN 0885-8969. DOI: 10.1109/TEC.2007.914663.

204 WAMKEUE, R.; JOLETTE, C.; KAMWA, I. Advanced modeling of a synchronous generator under line-switching and load-rejection tests for isolated grid applications. **IEEE Transactions on Energy Conversion**, v. 25, n. 3, p. 680–689, Sept. 2010. ISSN 0885-8969.

205 WAMKEUE, R.; JOLETTE, C.; KAIROUS, D. Estimation of synchronous generator parameters from load rejection analytical responses. In: 2010 9th International Conference on Environment and Electrical Engineering. Prague, CZE: IEEE, May 2010. DOI: 10.1109/EEEIC.2010.5489967.

206 DANDENO, P. L. Supplementary definitions associated test methods for obtaining parameters for synchronous machine stability study simulations. **IEEE Transactions on Power Apparatus and Systems**, PAS-99, n. 4, p. 1625–1633, July 1980. ISSN 0018-9510. DOI: 10.1109/TPAS.1980.319588.

207 TSAI, H.; KEYHANI, A.; DEMCKO, J.; FARMER, R. G. On-line synchronous machine parameter estimation from small disturbance operating data. **IEEE Transactions on Energy Conversion**, v. 10, n. 1, p. 25–36, Mar. 1995. ISSN 0885-8969. DOI: 10.1109/60.372565.

208 DIGGLE, R.; DINELEY, J. L. Generator works testing sudden-short-circuit or standstill variable-frequency-response method. **IEE Proceedings C - Generation, Transmission and Distribution**, v. 128, n. 4, p. 177–182, July 1981. ISSN 0143-7046. DOI: 10.1049/ip-c.1981.0031.

209 DANDENO, P. L.; PORAY, A. T. Development of detailed turbogenerator equivalent circuits from standstill frequency response measurements. **IEEE Transactions on Power Apparatus and Systems**, PAS-100, n. 4, p. 1646–1655, Apr. 1981. ISSN 0018-9510. DOI: 10.1109/TPAS.1981.316559.

210 MALLICK, J. A.; WILSON, G. L.; UMANS, S. D. Modeling of solid-rotor turbogenerators Part II: Example of model derivation and use in digital simulation. **IEEE Transactions on Power Apparatus and Systems**, PAS-97, n. 1, p. 278–291, Jan. 1978. ISSN 0018-9510. DOI: 10.1109/TPAS.1978.354479.

211 ZAFARABADI, G.; BOROUJENI, E. A. Identification of generator parameters from SSFR test for Montazer-Qaem powerplant. In: 2008 43rd International Universities Power Engineering Conference. Padova, ITA: IEEE, July 2008. P. 1–6. DOI: 10.1109/UPEC.2008.4651648.

212 BELQORCHI, A.; KARAAGAC, U.; MAHSEREDJIAN, J.; KAMWA, I. Standstill frequency response test and validation of a large hydrogenerator. **IEEE Transactions on Power Systems**, v. 34, n. 3, p. 2261–2269, May 2019. ISSN 1558-0679. DOI: 10.1109/TPWRS.2018.2889510.

213 PARK, R. H.; ROBERTSON, B. L. The reactances of synchronous machines. **Transactions of the American Institute of Electrical Engineers**, v. 47, n. 0, p. 514–535, Apr. 1928. ISSN 2330-9431. DOI: 10.1109/T-AIEE.1928.5055010.

214 UENOSONO, C.; OKADA, T.; KOH, H. Extended slip test for measuring synchronous machine constants. **Electrical Engineering in Japan**, v. 97, n. 3, p. 287–293, May 1977. ISSN 0036-9691. DOI: 10.1002/eej.4390970307.

215 WAMKEUE, R.; KAMWA, I.; DAI-DO, X. Short-circuit test based maximum likelihood estimation of stability model of large generators. **IEEE Transactions on Energy Conversion**, v. 14, n. 2, p. 167–174, June 1999. ISSN 0885-8969. DOI: 10.1109/60.766977.

216 SHAMSOLLAHI, P.; MALIK, O. P. On-line identification of synchronous generator using neural networks. In: PROCEEDINGS of 1996 Canadian Conference on Electrical and Computer Engineering. Calgary, AB, CA: IEEE, May 1996. v. 2, p. 595–598. DOI: 10.1109/CCECE.1996.548223.

217 WAMKEUE, R.; KAMWA, I.; DAI-DO, X.; KEYHANI, A. Iteratively reweighted least-squares for maximum likelihood identification of synchronous machine parameters from on-line tests. **IEEE Transactions on Energy Conversion**, v. 14, n. 2, p. 159–166, June 1999. ISSN 0885-8969. DOI: 10.1109/60.766971.

218 VERMEULEN, H. J.; STRAUSS, J. M.; SHIKOANA, V. Online estimation of synchronous generator parameters using PRBS perturbations. **IEEE Transactions on Power Systems**, v. 17, n. 3, p. 694–700, Aug. 2002. ISSN 0885-8950. DOI: 10.1109/TPWRS.2002.800915.

219 HEFFRON, W. G.; PHILLIPS, R. A. Effect of a modern amplidyne voltage regulator on underexcited operation of large turbine generators. **Transactions of the American Institute of Electrical Engineers. Part III: Power Apparatus and Systems**, v. 71, n. 3, p. 692–697, Aug. 1952. ISSN 0097-2460. DOI: 10.1109/AIEEPAS.1952.4498530.

220 OVERSCHEE, P. van; MOOR, B. de. N4SID: Subspace algorithms for the identification of combined deterministic stochastic systems. **Automatica**, v. 30, n. 1, p. 75–93, Jan. 1994. ISSN 0005-1098. DOI: 10.1016/0005-1098(94)90230-5.

221 PILLUTLA, S.; KEYHANI, A.; KAMWA, I. Neural network observers for online tracking of synchronous generator parameters. **IEEE Transactions on Energy Conversion**, v. 14, n. 1, p. 23–30, Mar. 1999. ISSN 0885-8969. DOI: 10.1109/60.749143.

222 MELGOZA, J. J. R.; HEYDT, G. T.; KEYHANI, A.; AGRAWAL, B. L.; SELIN, D. Synchronous machine parameter estimation using the Hartley series. **IEEE Transactions on Energy Conversion**, v. 16, n. 1, p. 49–54, Mar. 2001. ISSN 0885-8969. DOI: 10.1109/60.911403.

223 ________. An algebraic approach for identifying operating point dependent parameters of synchronous machines using orthogonal series expansions. **IEEE Transactions on Energy Conversion**, v. 16, n. 1, p. 92–98, Mar. 2001. ISSN 0885-8969. DOI: 10.1109/60.911410.

224 SADABADI, M. S.; KARRARI, M.; MALIK, O. P. Nonlinear Identification of Synchronous Generator Using Hammerstein Model with Piecewise Linear Static Maps. In: 2007 IEEE Lausanne Power Tech. Lausanne, CH: IEEE, July 2007. P. 1067–1071. DOI: 10.1109/PCT.2007.4538463.

225 ICHIKAWA, S.; TOMITA, M.; DOKI, S.; OKUMA, S. Sensorless control of permanent-magnet synchronous motors using online parameter identification based on system identification theory. **IEEE Transactions on Industrial Electronics**, v. 53, n. 2, p. 363–372, Apr. 2006. ISSN 0278-0046. DOI: 10.1109/TIE.2006.870875.

226 GOLUB, G. H.; PEREYA, V. The differentiation of pseudo-inverses and nonlinear least-squares problems whose variables separate. **SIAM Journal on Numerical Analysis**, v. 10, n. 2, p. 413–432, Apr. 1973. ISSN 1095-7170. DOI: 10.1137/0710036.

227 O'LEARY, D. P.; RUST, B. W. Variable projection for nonlinear least-squares problems. **Computational Optimization and Applications**, v. 54, n. 3, p. 579–593, Aug. 2012. ISSN 1573-2894. DOI: 10.1007/s10589-012-9492-9.

228 IEEE-SA STANDARDS BOARD. **IEEE Recommended Practice for Preferred Metric Units for Use in Electrical and Electronics Science and Technology**. New York, NY, USA: The Institute of Electrical and Electronics Engineers, Inc., May 2019. P. 1–24. ISBN 978-1-5044-5551-0. DOI: 10.1109/IEEESTD.2019.8715835.

229 BUTCHER, J. C. **Numerical Methods for Ordinary Differential Equations**. Second ed. West Sussex, England, UK: John Wiley & Sons, Ltd., 2008. ISBN 978-0-470-72335-7.

230 EULER, L. De integratione aequationum differentialium per approximationem. **Opera Omnia**, v. 11, p. 424–434, 1913.

231 RUNGE, C. Über die numerische auflösung von differentialgleichungen. **Mathematische Annalen**, v. 46, n. 2, p. 167–178, June 1895. ISSN 1432-1807. DOI: 10.1007/BF01446807.

232 HEUN, K. Neue methode zur approximativen integration der differentialgleichungen einer unabhängigen veränderlichen. **Journal of Applied Mathematics and Physics**, v. 46, p. 435–453, 1900. ISSN 1420-9039.

233 KUTTA, W. Beitrag zur näherungsweisen integration totaler differentialgleichungen. **Journal of Applied Mathematics and Physics**, v. 45, p. 23–28, 1901. ISSN 1420-9039.

Appendix I – The Operational Impedances

Since in many important problems, the primary interest is in the results as viewed from the machine armature terminals, as in computing short-circuit currents, it is convenient to write the machine equations in a more compact form by eliminating the rotor currents. This may be done by:

(*i*) substituting the rotor flux-linkage relations into the rotor-circuit voltage equations;

(*ii*) solving these for the rotor currents in terms of the field voltage and the armature currents;

(*iii*) and substituting the resulting relations in the armature flux-linkage relations.

That may be a more or less difficult job of solving several simultaneous equations depending on the complexity of the amortisseur, but it is evident that if the derivative operator p is treated algebraically, it becomes a much simpler task [31]. This method arrives at a result in the form shown in (8.2).

R. H. Park, in his original paper [43], did not specify the number of rotor circuits. Instead, he expressed the stator flux linkages in terms of operational impedances and a transfer function relating stator flux linkages to field voltage.

I.1 Park's equations in operational form

Park [43] published the original $qd0$-voltages equations in the form:

$$\bar{v}_q = -\bar{r}_s \dot{i}_q + \overline{\omega}_r \overline{\Psi}_d + \frac{p}{\omega_b} \overline{\Psi}_q \qquad [\mathrm{pu}] \ , \tag{8.1a}$$

$$\bar{v}_d = -\bar{r}_s \dot{i}_d - \overline{\omega}_r \overline{\Psi}_q + \frac{p}{\omega_b} \overline{\Psi}_d \qquad [\mathrm{pu}] \ , \tag{8.1b}$$

$$\bar{v}_0 = -\bar{r}_s \dot{i}_0 + \frac{p}{\omega_b} \overline{\Psi}_0 \qquad [\mathrm{pu}] \ , \tag{8.1c}$$

where

$$\overline{\Psi}_q = -\bar{x}_q(p) \dot{i}_q \qquad [\mathrm{pu}] \ , \tag{8.2a}$$

$$\overline{\Psi}_d = -\bar{x}_d(p)\dot{i}_d + G(p)\bar{v}'_{fd} \qquad [\text{pu}] \;, \tag{8.2b}$$

$$\overline{\Psi}_0 = -\bar{x}_{ls}\dot{i}_0 \qquad [\text{pu}] \;. \tag{8.2c}$$

In these equations, positive stator current is assumed out of the machine. The operator $\bar{x}_q(p)$ is referred to as the quadrature-axis operational impedance, $\bar{x}_d(p)$ is the direct-axis operational impedance, and $G(p)$ is the dimensionless transfer function relating stator flux linkages per second to field voltage.

I.2 Direct-axis operational impedance

Let the network shown in Figure I.1. It represents an equivalent circuit with one damper winding in the direct-axis when $\bar{v}'_{fd}$ is set to zero[2].

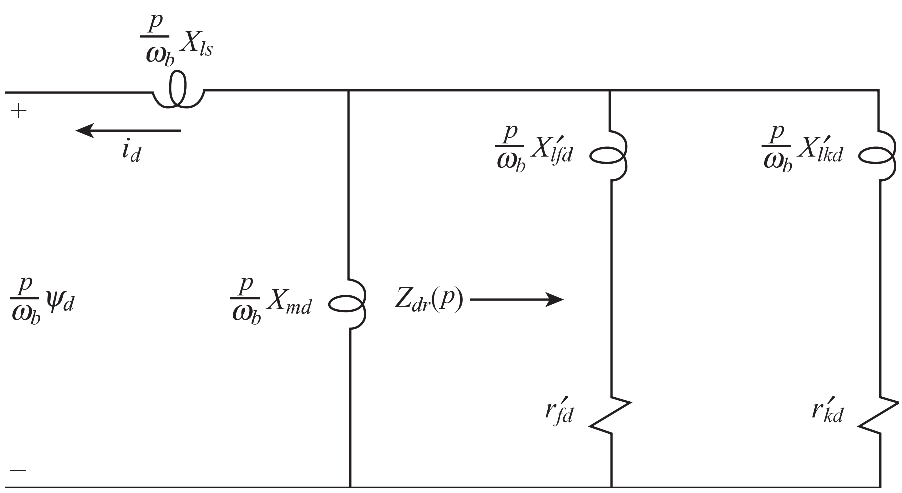

Figure I.1: Equivalent circuit with one damper winding in the direct-axis for the calculation of $\bar{x}_d(p)$. Adapted from Krause et al. [19, p. 275].

It is helpful in this and in the following derivations to express the input impedance of the rotor circuits in the form

$$Z_{dr}(p) = R_{ed}\frac{(1+\tau_{da}p)(1+\tau_{db}p)}{(1+\tau_{Da}p)} \qquad [\text{pu}] \;, \tag{8.3}$$

where

$$R_{ed} = \frac{\bar{r}'_{fd}\bar{r}'_{kd}}{\bar{r}'_{fd}+\bar{r}'_{kd}} \qquad [\text{pu}] \;, \tag{8.4a}$$

$$\tau_{da} = \frac{\bar{x}'_{lfd}}{\omega_b\bar{r}'_{fd}} \qquad [\text{s}] \;, \tag{8.4b}$$

[2]Although it is customary to use the Laplace operator s rather than the operator p, this work will keep its initial notation – the one adopted by Park [43] and Carter [46] – and remain using p to denote the Laplace operator.

$$\tau_{db} = \frac{\bar{x}'_{lkd}}{\omega_b \bar{r}'_{kd}} \qquad [\mathrm{s}] \ , \tag{8.4c}$$

$$\tau_{Da} = \frac{\bar{x}'_{lfd} + \bar{x}'_{lkd}}{\omega_b(\bar{r}'_{fd} + \bar{r}'_{kd})} = R_{eq}\left(\frac{\tau_{da}}{\bar{r}'_{kd}} + \frac{\tau_{db}}{\bar{r}'_{fd}}\right) \qquad [\mathrm{s}] \ . \tag{8.4d}$$

The operational impedance for a field and a damper winding in the d-axis can be obtained by setting $\bar{v}'_{fd}$ to zero and determining the equivalent impedance relative to the terminal in Figure I.1, which is

$$\frac{p\bar{x}_d(p)}{\omega_b} = \frac{p\bar{x}_{ls}}{\omega_b} + \frac{(p\bar{x}_{md}/\omega_b)Z_{dr}(p)}{Z_{dr}(p) + (p\bar{x}_{md}/\omega_b)} \ . \tag{8.5}$$

Solving the equation above for $\bar{x}_d(p)$ yields the operational impedance in the d-axis

$$\bar{x}_d(p) = \bar{x}_d \frac{(1+\tau'_d p)(1+\tau''_d p)}{(1+\tau'_{d0}p)(1+\tau''_{d0}p)} \ , \tag{8.6}$$

where

$$\tau'_d = \frac{1}{\omega_b \bar{r}'_{fd}}\left(\bar{x}'_{lfd} + \frac{\bar{x}_{md}\bar{x}_{ls}}{\bar{x}_{md} + \bar{x}_{ls}}\right) \qquad [\mathrm{s}] \ , \tag{8.7a}$$

$$\tau''_d = \frac{1}{\omega_b \bar{r}'_{kd}}\left(\bar{x}'_{lkd} + \frac{\bar{x}_{md}\bar{x}_{ls}\bar{x}'_{lfd}}{\bar{x}_{md}\bar{x}_{ls} + \bar{x}_{md}\bar{x}'_{lfd} + \bar{x}_{ls}\bar{x}'_{lfd}}\right) \qquad [\mathrm{s}] \tag{8.7b}$$

are the direct-axis transient and sub-transient short-circuit time constants, respectively, and

$$\tau'_{d0} = \frac{1}{\omega_b \bar{r}'_{fd}}(\bar{x}'_{lfd} + \bar{x}_{md}) \qquad [\mathrm{s}] \ , \tag{8.7c}$$

$$\tau''_{d0} = \frac{1}{\omega_b \bar{r}'_{kd}}\left(\bar{x}'_{lkd} + \frac{\bar{x}_{md}\bar{x}'_{lfd}}{\bar{x}_{md} + \bar{x}'_{lfd}}\right) \qquad [\mathrm{s}] \tag{8.7d}$$

are the direct-axis transient and sub-transient open-circuit time constants, respectively.

The transfer function $G(p)$ may be evaluated by expressing the relationship between stator flux linkages per second to field voltage, $\bar{r}'_{fd}$, with $\bar{i}_d$ equal to zero. Hence, from (8.2),

$$G(p) = \left.\frac{\Psi_d}{\bar{v}'_{fd}}\right|_{\bar{i}_d=0} \qquad [1] \ . \tag{8.8}$$

Although [1] is generally omitted in specifying the values of dimensionless quantities, it is presented in this section for an improved representation.

From Figure I.2,

$$G(p) = \frac{\bar{x}_{md}}{\bar{r}'_{fd}} \frac{(1+\tau_{kd}p)}{(1+\tau'_{d0}p)(1+\tau''_{d0}p)} \qquad [1] \ , \tag{8.9}$$

where

$$\tau_{kd} = \frac{\bar{x}'_{lkd}}{\omega_b \bar{r}'_{kd}} \qquad [\mathrm{s}] \tag{8.10}$$

is the direct-axis damper leakage time constant.

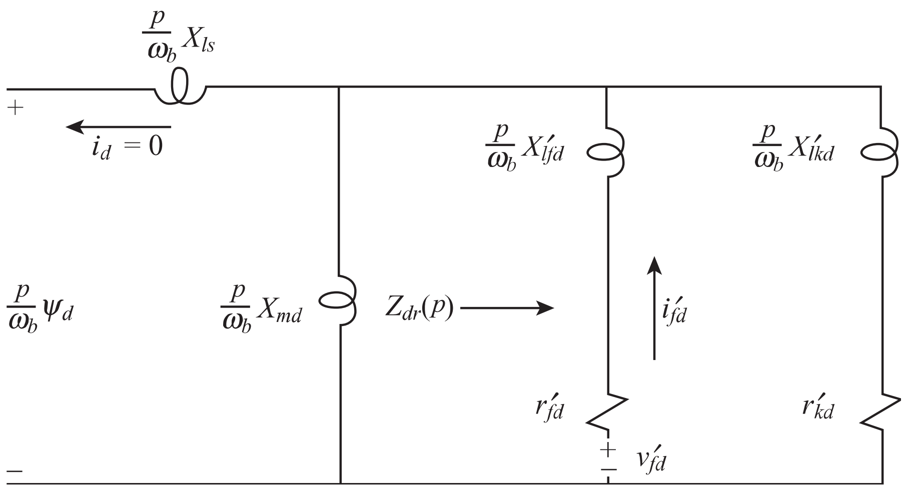

Figure I.2: Equivalent circuit with one damper winding in the direct-axis for calculation of $G(p)$. Adapted from Krause et al. [19, p. 275].

I.3 Quadrature-axis operational impedance

Let the network shown in Figure I.3. It represents an equivalent circuit with one damper winding in the quadrature-axis.

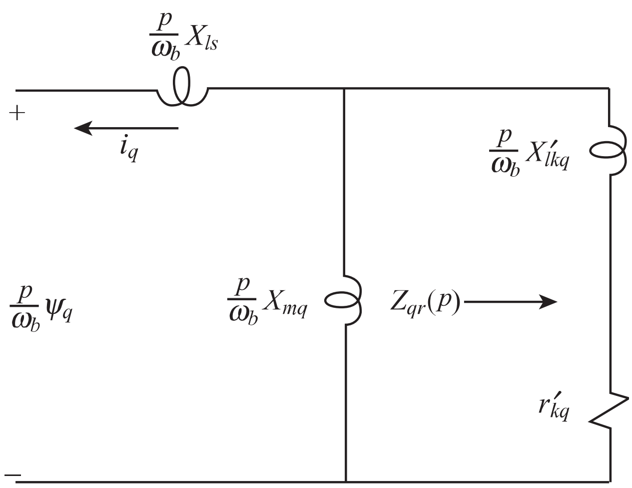

Figure I.3: Equivalent circuit with one damper winding in the quadrature-axis. Adapted from Krause et al. [19, p. 273].

As in the previous case, let the input impedance of the rotor circuits in the form of

$$Z_{qr}(p) = \bar{r}'_{kq} + \frac{p}{\omega_b}\bar{x}'_{lkq} \qquad [\mathrm{pu}]\ . \tag{8.11}$$

From Figure I.3 and (8.2), the equivalent impedance relative to the terminal is

$$\frac{p}{\omega_b}\bar{x}_q(p) = \frac{p}{\omega_b}\bar{x}_{ls} + \frac{(p/\omega_b)\bar{x}_{mq}Z_{qr}(p)}{Z_{qr}(p) + (p/\omega_b)\bar{x}_{mq}} \qquad [\mathrm{pu}]\ , \tag{8.12}$$

which solved for $\bar{x}_q(p)$ yields the operational impedance for one damper winding in the q-axis

$$\bar{x}_q(p) = \bar{x}_q \frac{(1 + \tau''_q p)}{(1 + \tau''_{q0} p)} \qquad [\text{pu}] \ , \tag{8.13}$$

where

$$\tau''_q = \frac{1}{\omega_b \bar{r}'_{kq}} \left(\bar{x}_{lkq} + \frac{\bar{x}_{mq}\bar{x}_{ls}}{\bar{x}_{mq} + \bar{x}_{ls}} \right) \qquad [\text{s}] \tag{8.14a}$$

and

$$\tau''_{q0} = \frac{1}{\omega_b \bar{r}'_{kq}} \left(\bar{x}_{lkq} + \bar{x}_{mq} \right) \qquad [\text{s}] \tag{8.14b}$$

are the quadrature-axis sub-transient short-circuit time constant and the quadrature-axis sub-transient open-circuit time constant, respectively [30].

In the above definitions, open- and short-circuit refers to the conditions of the stator circuits. All of these time constants are approximations of the actual time constants, and when used to determine machine parameters, they can lead to substantial errors in predicting the dynamic behavior of a synchronous machine [19].

Appendix II – Results

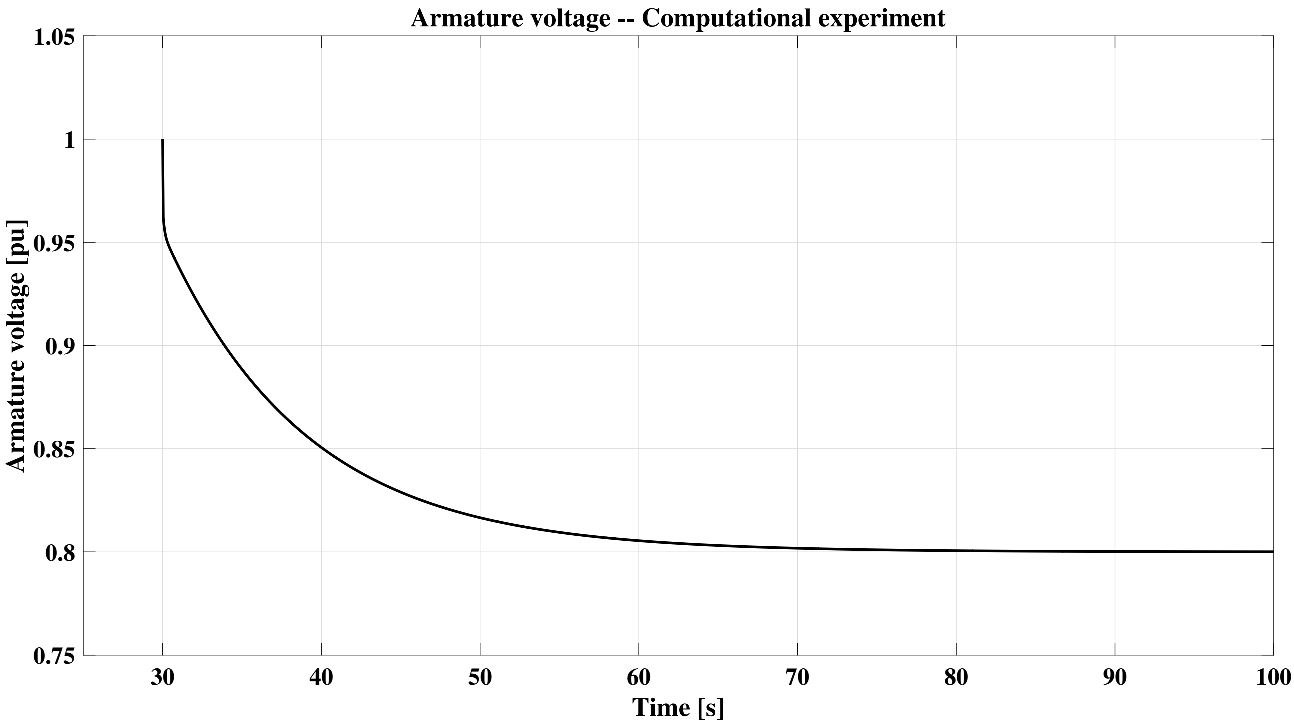

Figure II.1: Computational data: Armature voltage after the direct-axis load rejection.

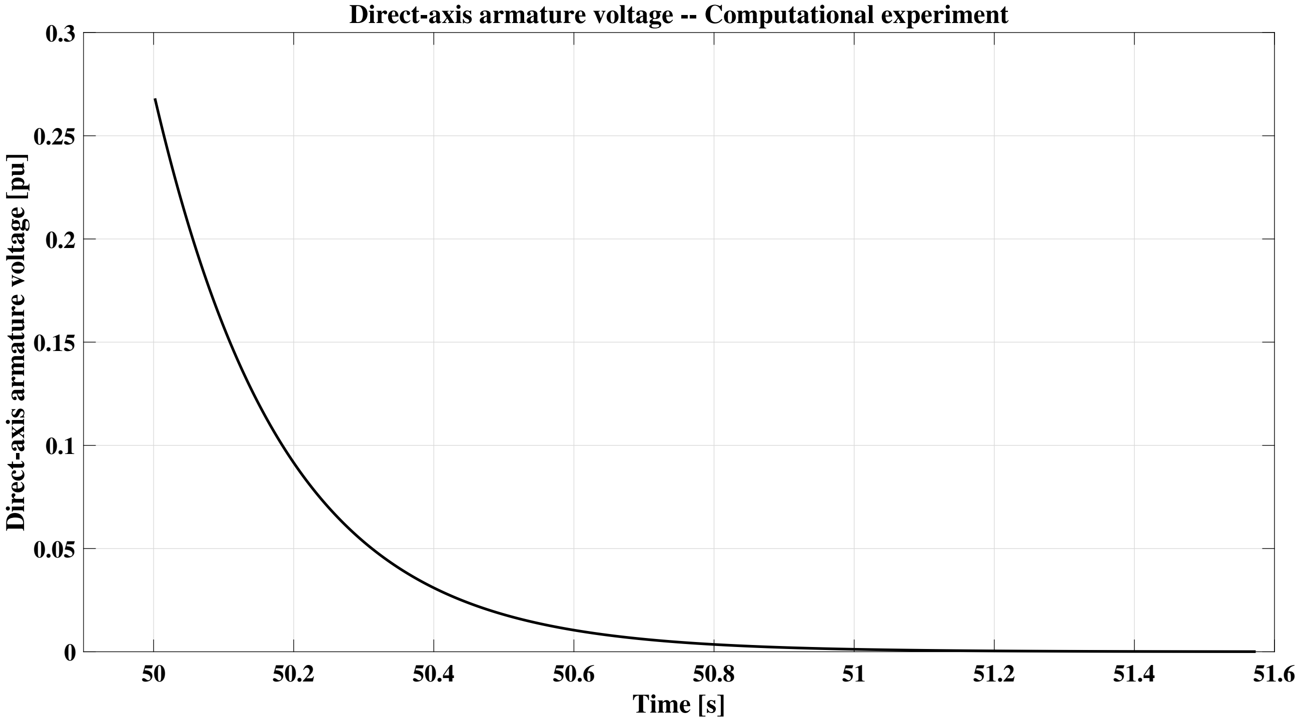

Figure II.2: Computational data: Armature-voltage direct-axis component after the quadrature-axis load rejection.

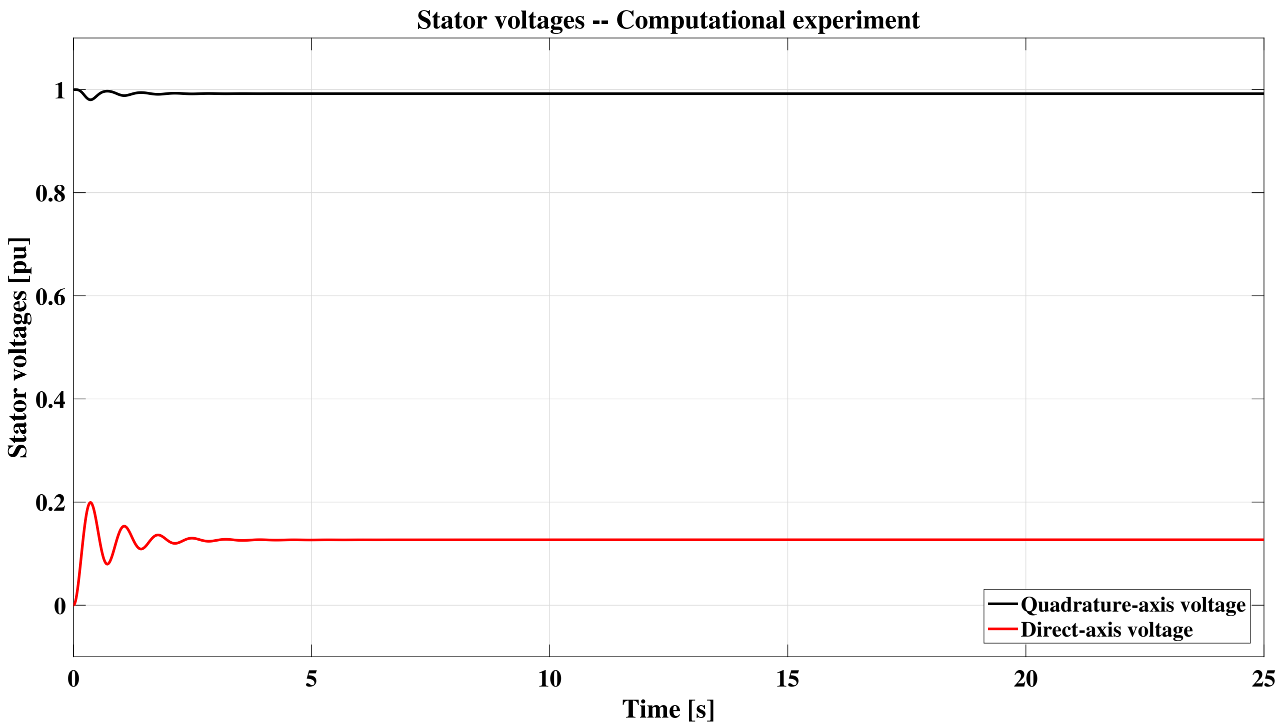

Figure II.3: Computational data: Quadrature- and direct-axis voltages measurement.

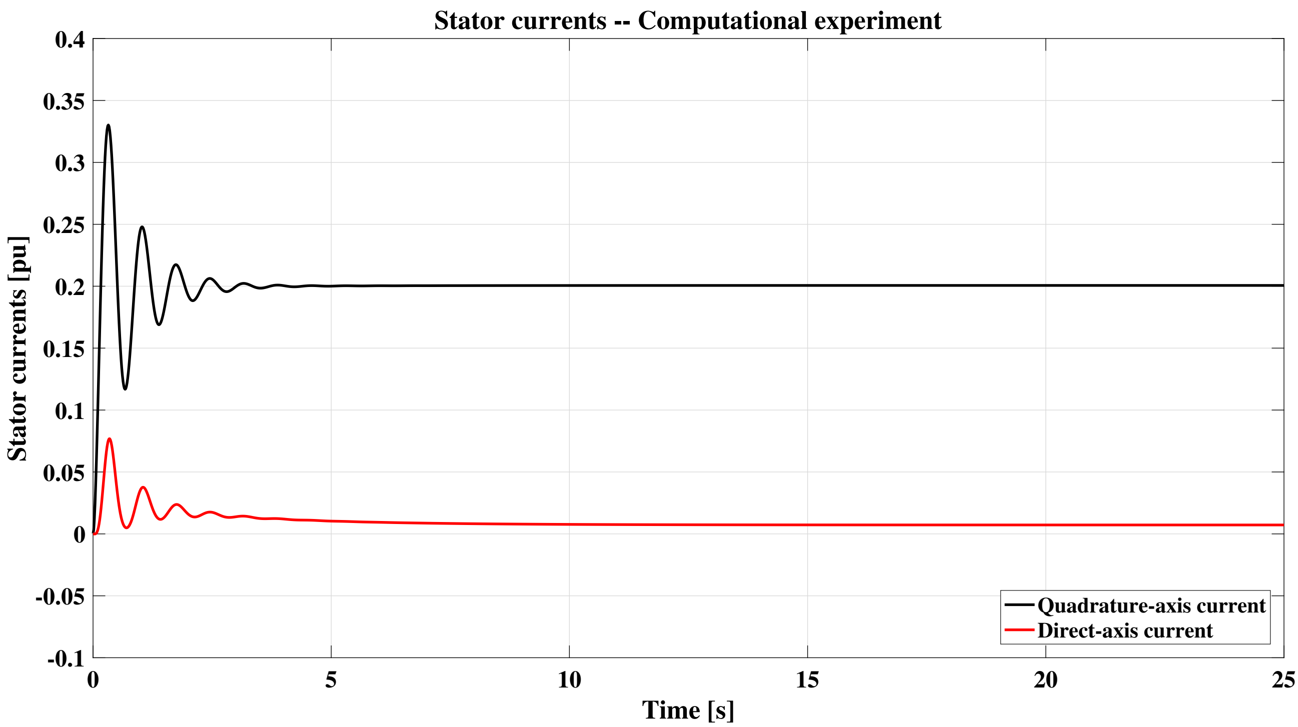

Figure II.4: Computational data: Quadrature- and direct-axis currents measurement.

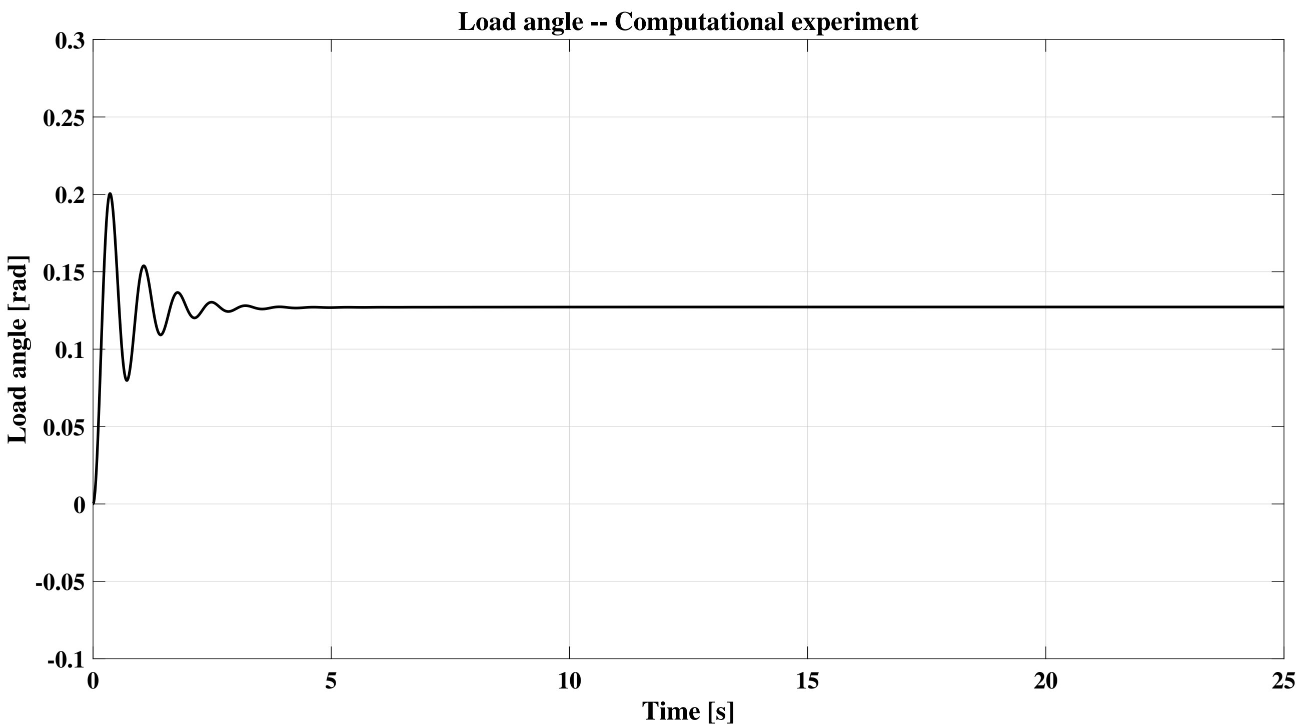

Figure II.5: Computational data: Load angle measurement.

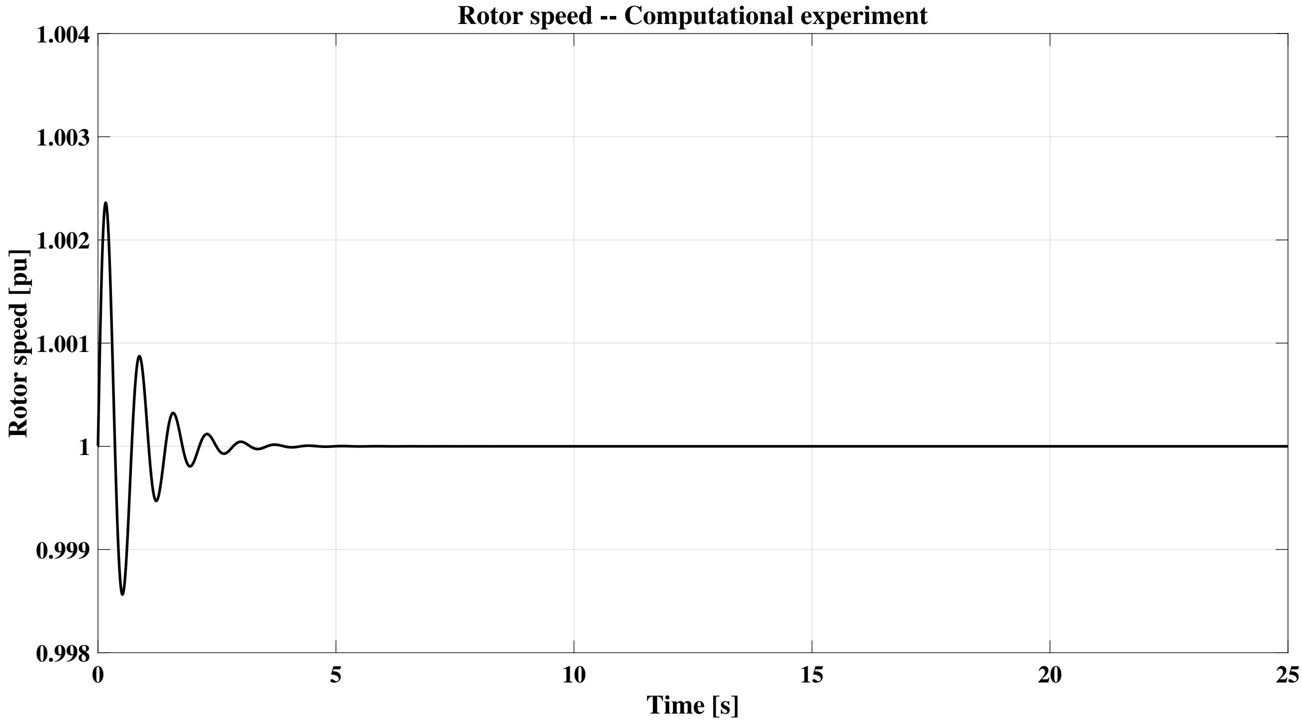

Figure II.6: Computational data: Rotor speed measurement.

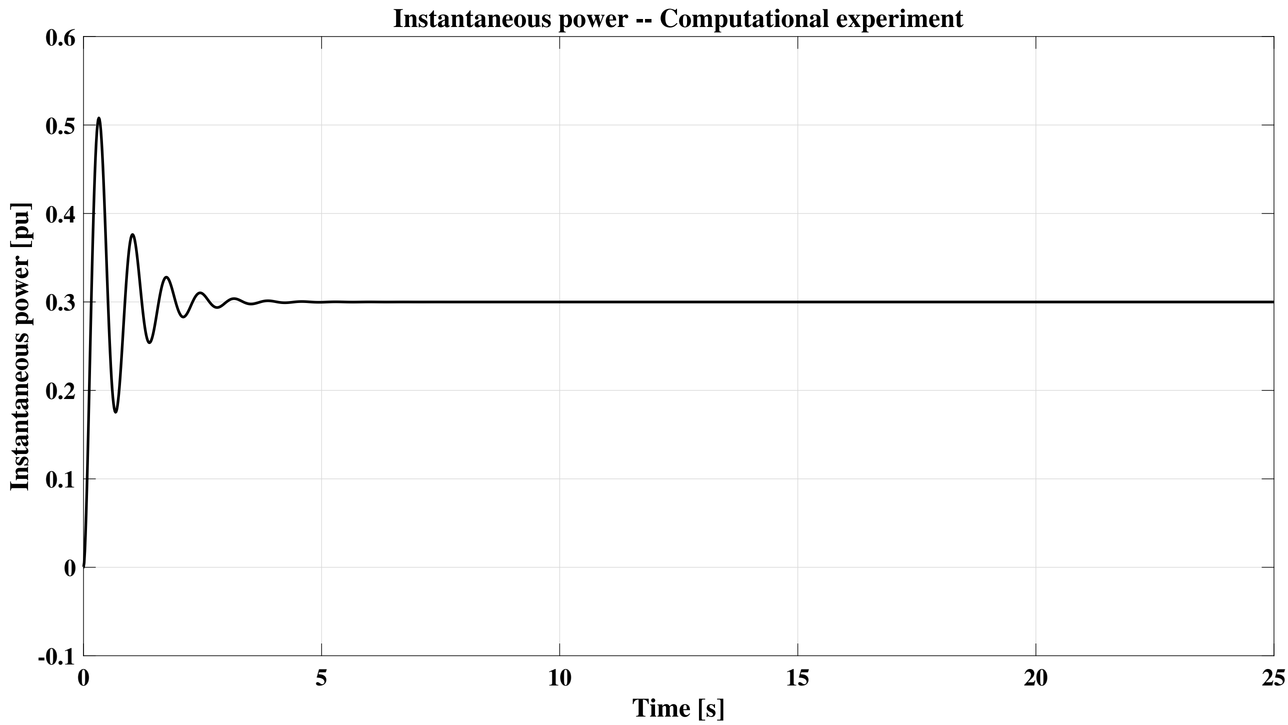

Figure II.7: Computational data: Instantaneous power measurement.

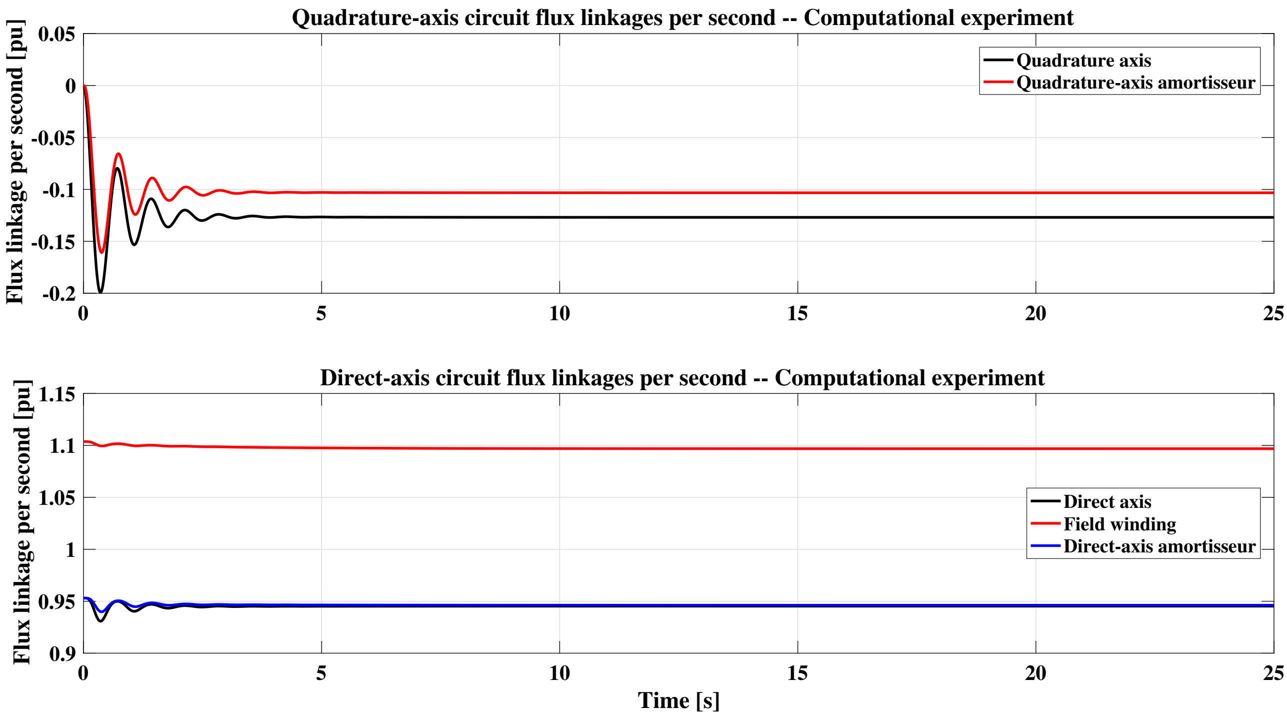

Figure II.8: Computational data: Calculated flux linkages per second.

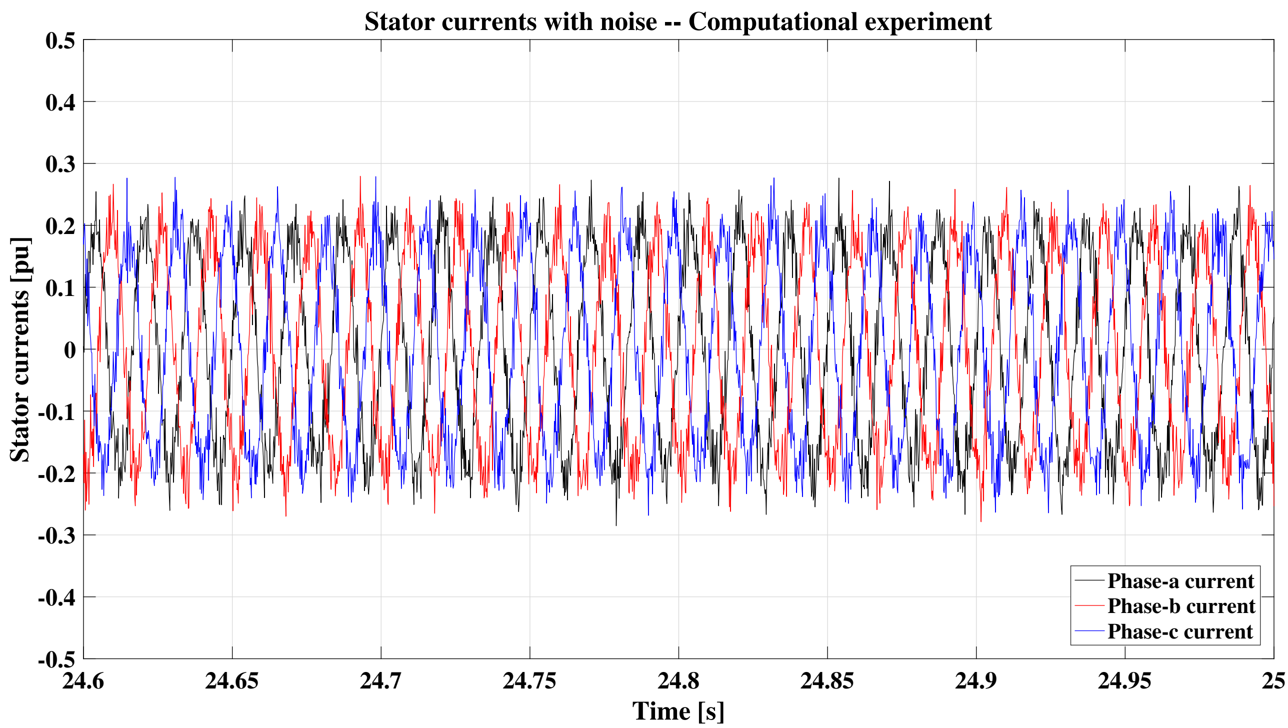

Figure II.9: Computational data: Stator currents measurements with noise added.

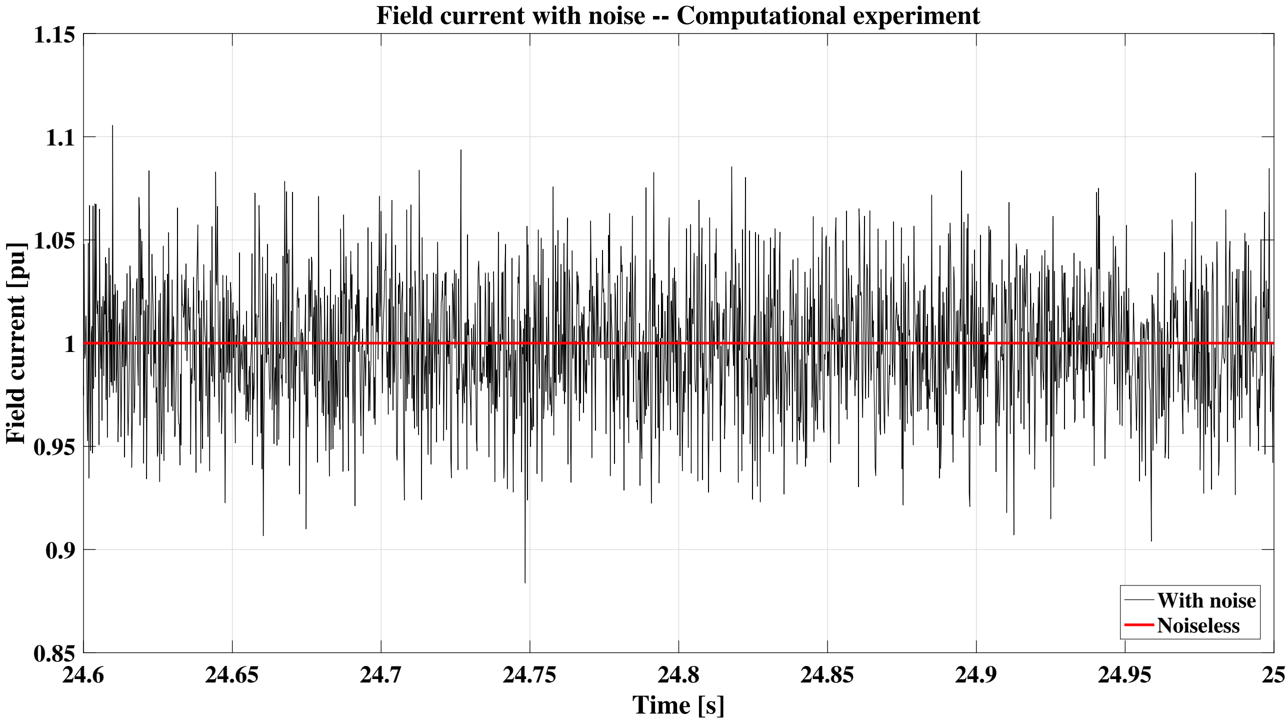

Figure II.10: Computational data: Field current measurement with noise added.

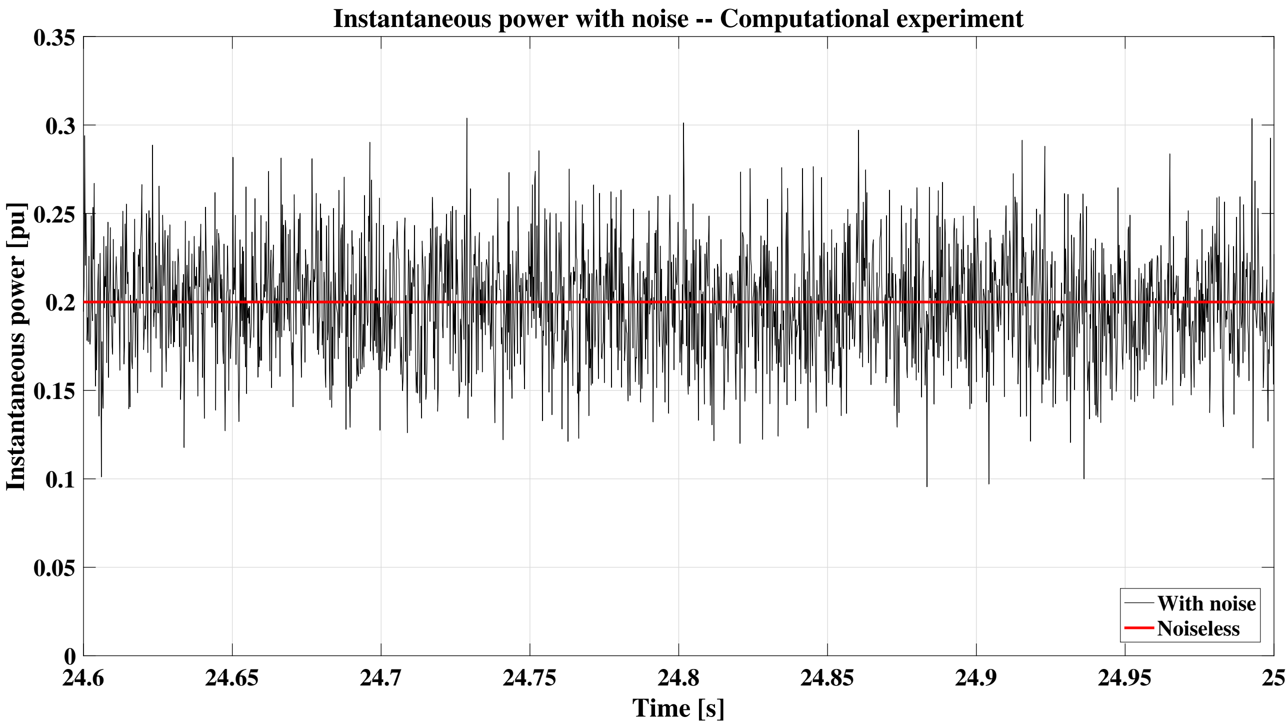

Figure II.11: Computational data: Instantaneous power measurement with noise added.

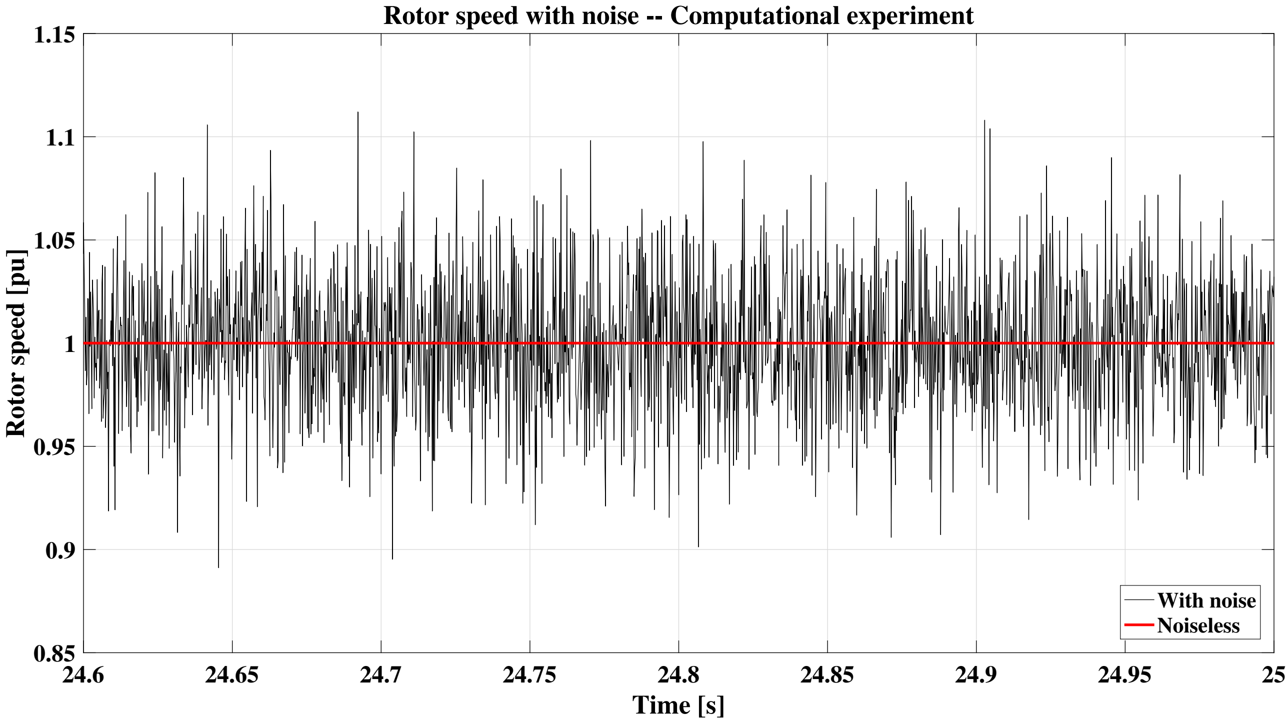

Figure II.12: Computational data: Rotor speed measurement with noise added.

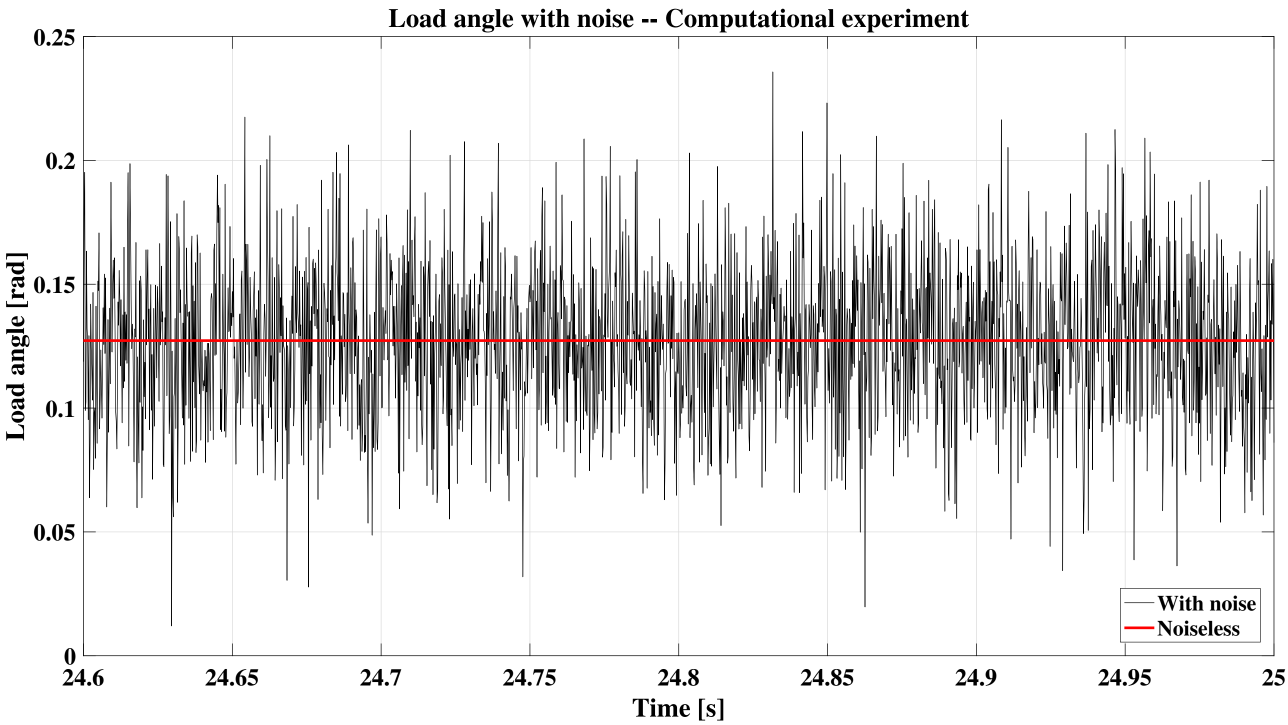

Figure II.13: Computational data: Load angle measurement with noise added.

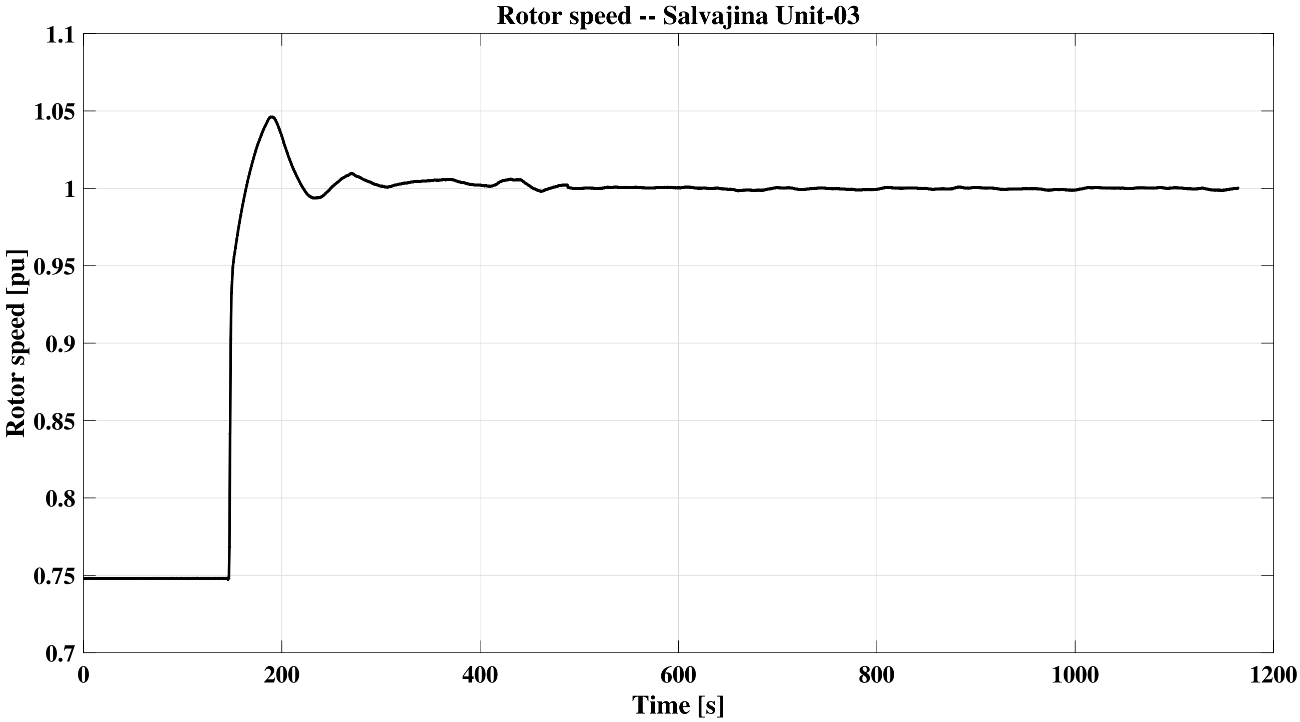

Figure II.14: Salvajina Unit-03 data: Rotor speed measurement.

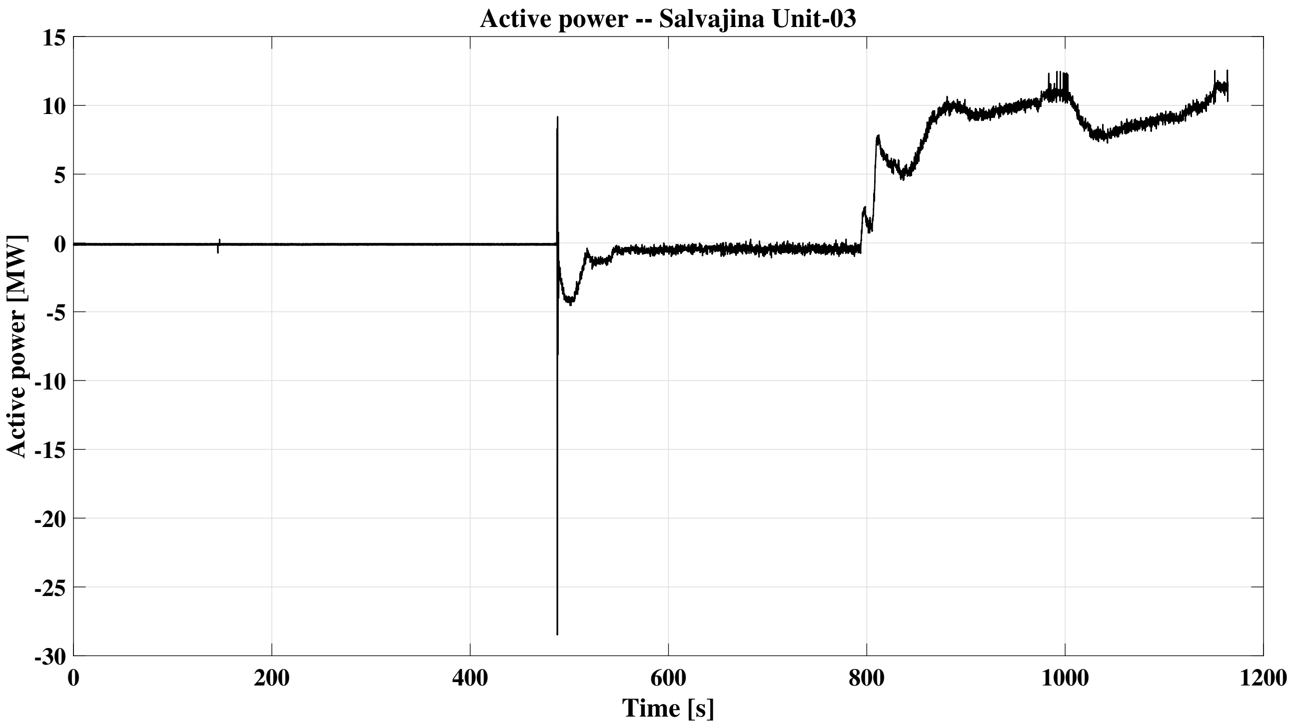

Figure II.15: Salvajina Unit-03 data: Active power measurement.

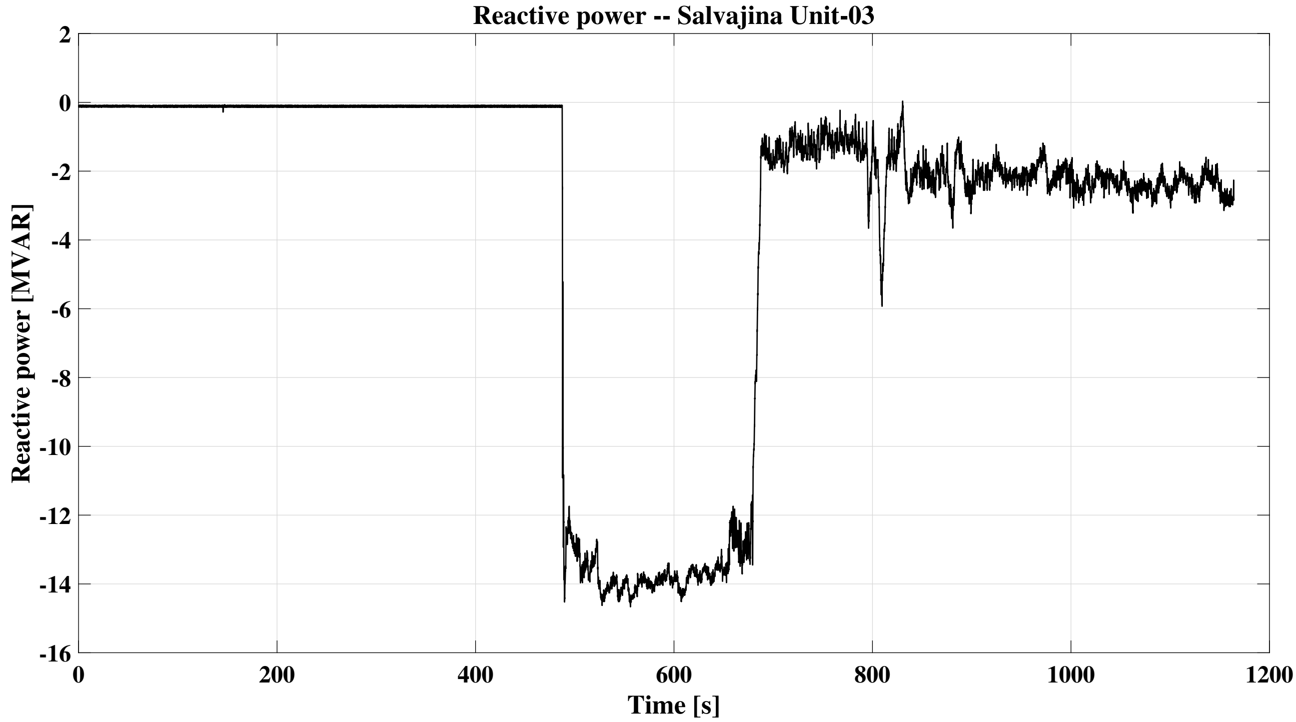

Figure II.16: Salvajina Unit-03 data: Reactive power measurement.

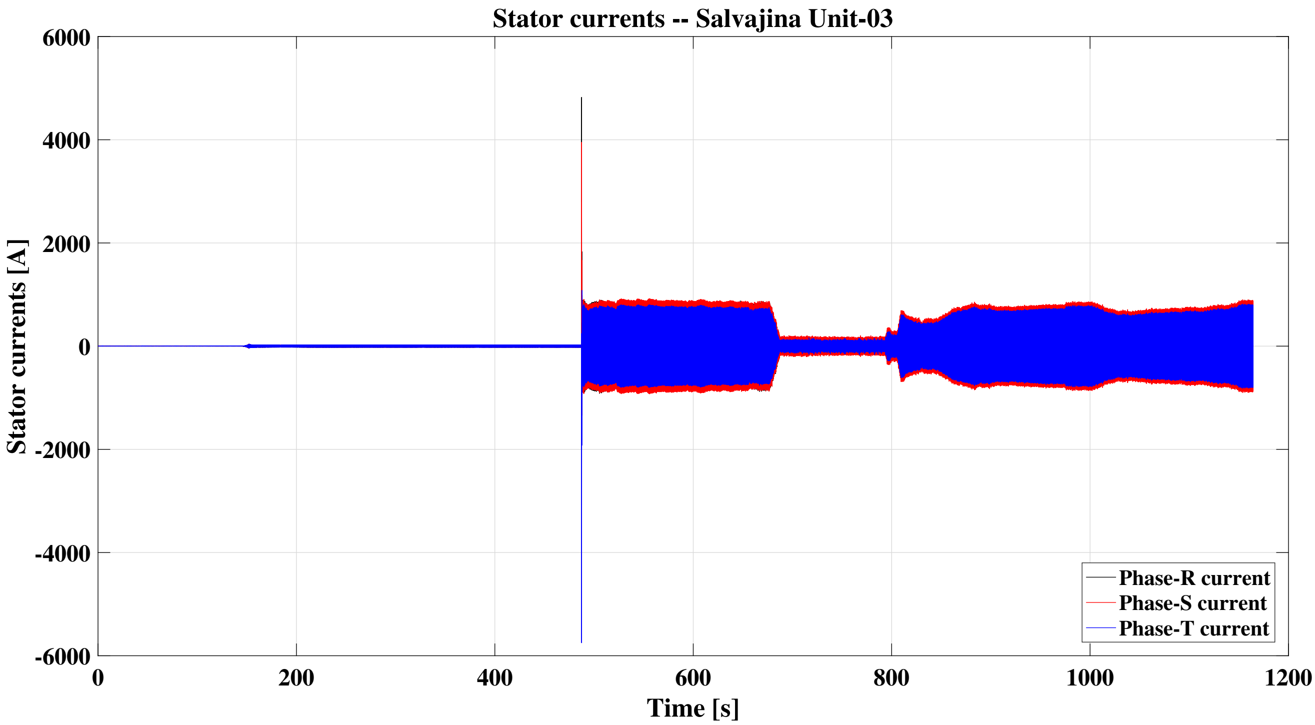

Figure II.17: Salvajina Unit-03 data: Stator currents measurements.

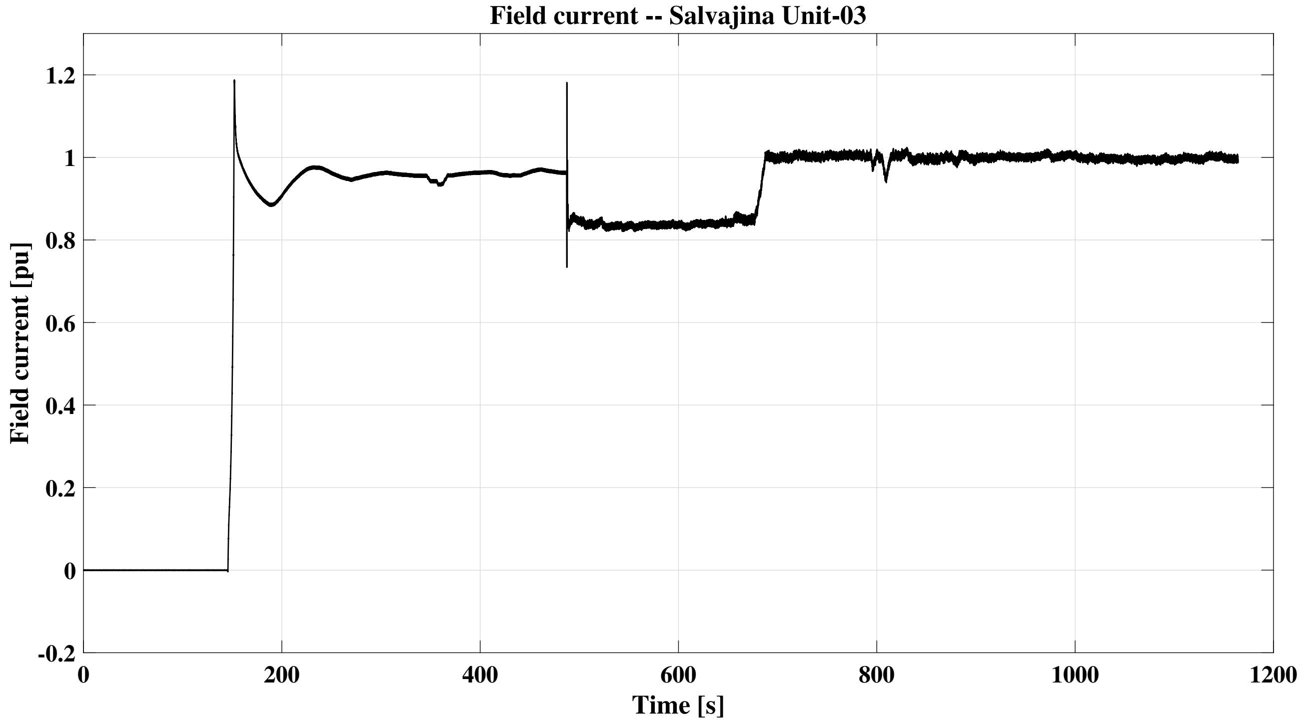

Figure II.18: Salvajina Unit-03 data: Field current measurement.

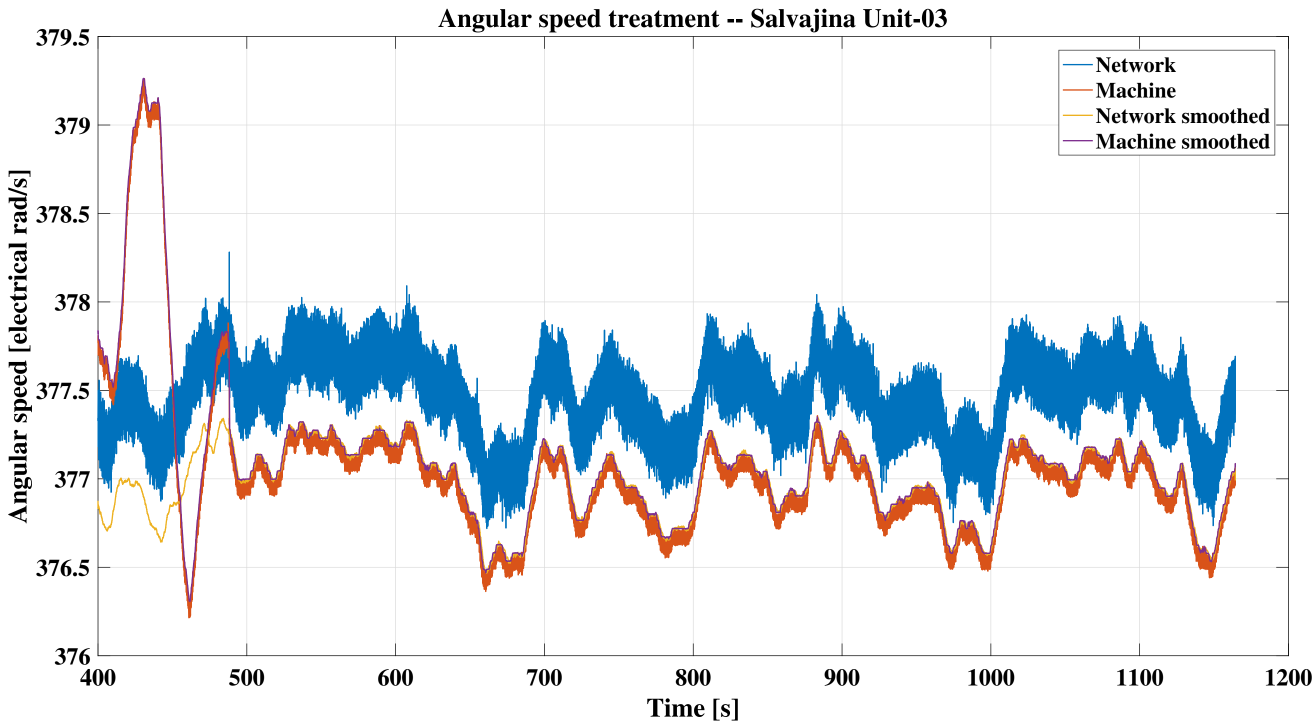

Figure II.19: Salvajina Unit-03 data: Angular speed treatment.

Annex A – Trigonometric Relationships

$$\sin x \cos y = 0.5 \sin (x+y) + 0.5 \cos (x-y) \tag{A.1}$$

$$\sin x \sin y = 0.5 \cos (x-y) - 0.5 \cos (x+y) \tag{A.2}$$

$$\cos x \cos y = 0.5 \cos (x+y) + 0.5 \cos (x-y) \tag{A.3}$$

$$\cos x + \cos (x - 2\pi/3) + \cos (x - 4\pi/3) = 0 \tag{A.4}$$

$$\sin x + \sin (x - 2\pi/3) + \sin (x - 4\pi/3) = 0 \tag{A.5}$$

$$\cos^2 x + \cos^2 (x - 2\pi/3) + \cos^2 (x - 4\pi/3) = 3/2 \tag{A.6}$$

$$\sin^2 x + \sin^2 (x - 2\pi/3) + \sin^2 (x - 4\pi/3) = 3/2 \tag{A.7}$$

Annex B – General Figures and Tables

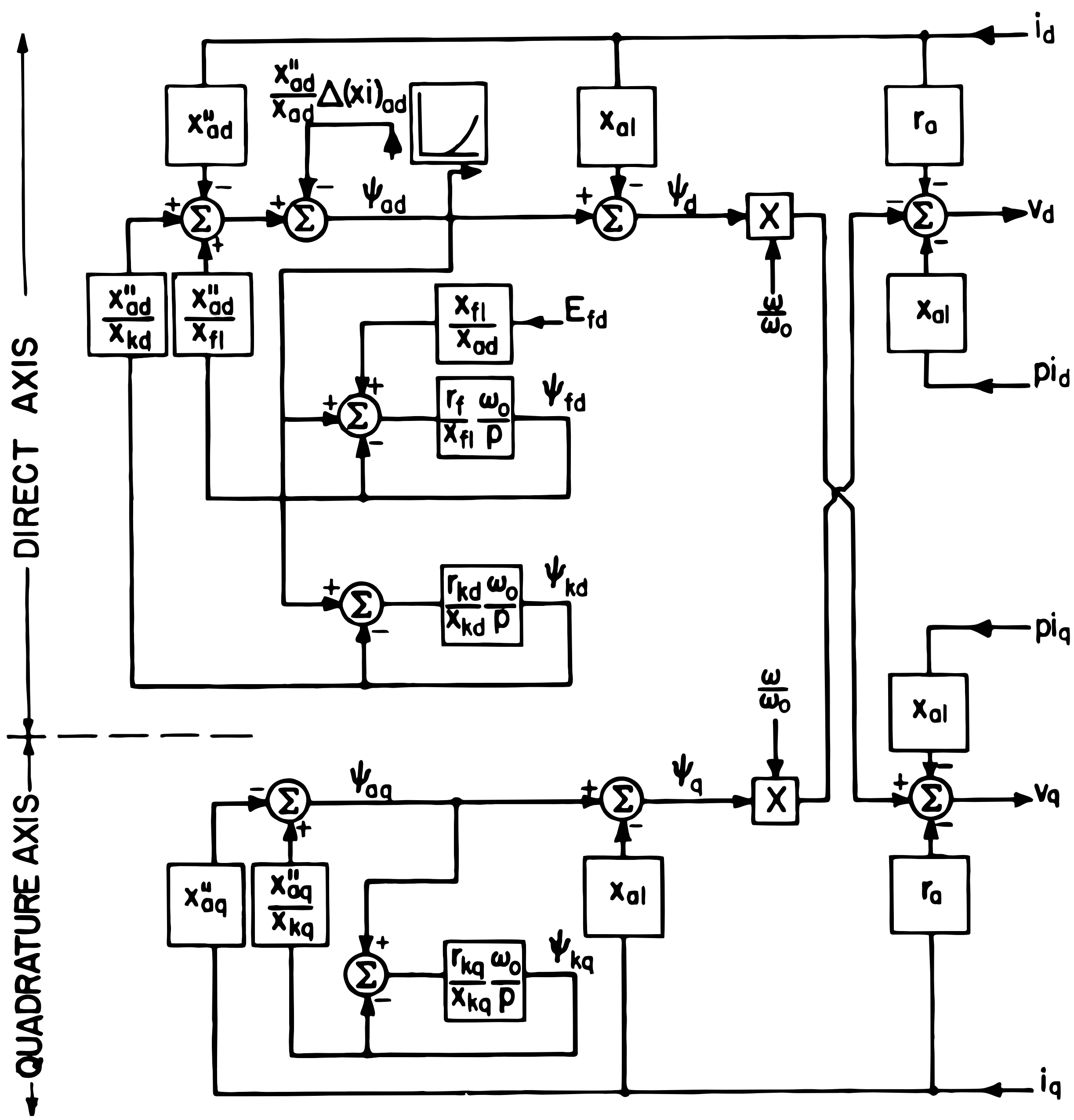

Figure B.1: Block diagram of the synchronous machine – Version 01.
From Thomas [101].

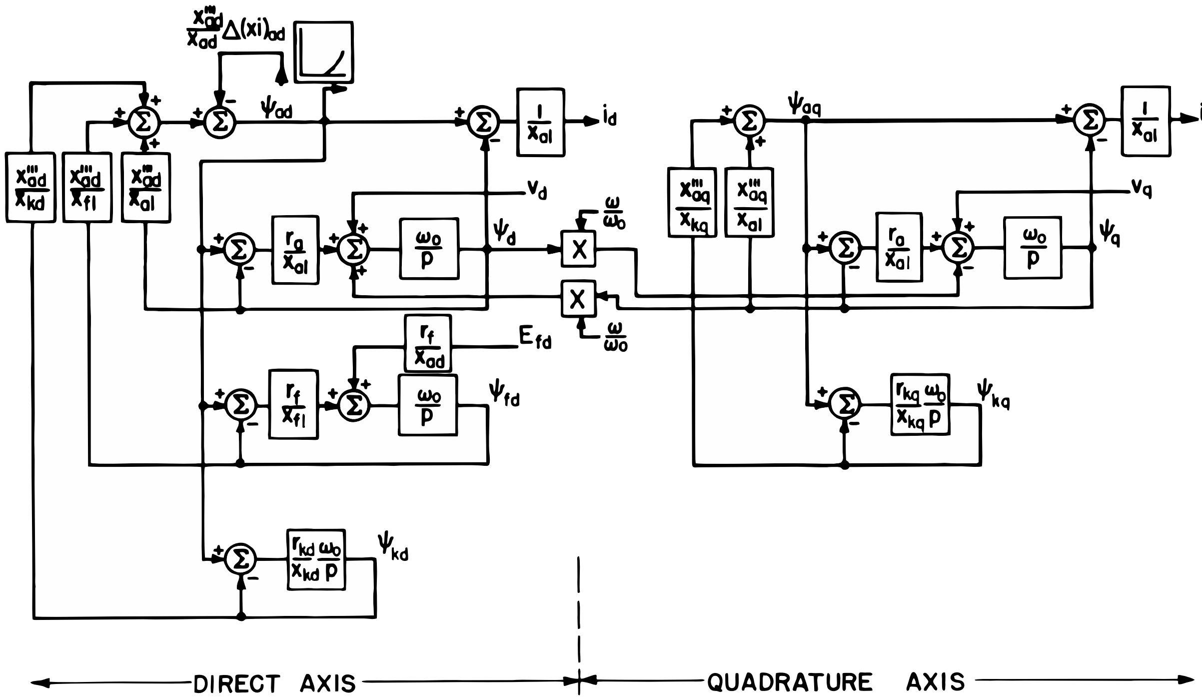

Figure B.2: Block diagram of the synchronous machine – Version 02.
From Thomas [101].

Table B.1: Salient-pole synchronous generator parameters provided by the manufacturer.

Parameter	Value
Rated power	126 MVA
Line-to-line voltage	13.8 kV
Frequency	60 Hz
Inertia coefficient	4.055 s
Friction factor	0 pu
Pole pairs	8
Stator winding resistance	0.00120 pu
Quadrature-axis amortisseur winding resistance	0.02993 pu
Field-winding resistance	0.00027 pu
Direct-axis amortisseur winding resistance	0.01995 pu
Stator winding leakage inductance	0.11800 pu
Quadrature-axis amortisseur winding leakage inductance	0.16450 pu
Field-winding leakage inductance	0.15070 pu
Direct-axis amortisseur winding leakage inductance	0.10970 pu
Quadrature-axis magnetizing inductance	0.51460 pu
Direct-axis magnetizing inductance	0.95300 pu

Annex C – The International System of Units

Quantity	Unit Name	Unit symbol
Acceleration	Meter per second squared	m/s^2
Active power	Watt	W
Angular velocity	Radian per second	rad/s
Area	Square meter	m^2
Capacitance	Farad	F
Complex power, apparent power	Volt-Ampere	VA
Density, mass density	Kilogram per cubic meter	kg/m^3
Celsius temperature	Degree Celsius	ºC
Electric charge, amount of electricity	Coulomb	C
Electric current	Ampere	A
Electric potential difference, electromotive force	Volt	V
Electric resistance	Ohm	Ω
Energy, work, amount of heat	Joule	J
Force	Newton	N
Frequency	Hertz	Hz
Inductance	Henry	H
Inertia	Kilogram meter squared	$kg.m^2$
Length	Meter	m
Magnetic field strength	Ampere per meter	A/m
Magnetic flux	Weber	Wb
Magnetic flux linkage	Weber turn	Wb-t
Magnetomotive force	Ampere	A
Mass	Kilogram	kg
Permeability	Henry per meter	H/m
Permittivity	Farad per meter	F/m

(continue in the next page)

(continued)

Quantity	**Unit Name**	**Unit symbol**
Plane angle	Radian[3]	1
Reactive power	Volt-Ampere-Reactive	VAR
Relative permeability	One	1
Speed, velocity	Meter per second	m/s
Torque	Newton meter	N.m
Time	Second	s
Volume	Cubic meter	m^3

[3]The radian is a special name for the number one that may be used to convey information about the quantity concerned. In practice, the symbol rad is used where appropriate, but the symbol for the derived unit one is generally omitted in specifying the values of dimensionless quantities [228].

Annex D – Numerical Differential Equation Methods

The following sections are based on the work of Butcher [229] and Kundur [33]. One should refer to them for further details on convergence, truncation error, and stability, and for numerical examples.

D.1 The Euler method

The famous method of Euler was published in his three-volume work *Institutiones Calculi Integralis* from 1768–1770, republished in his collected works [230]. This fundamental idea is based on a very simple principle. Suppose that a particle is moving in such a way that, at x_0, its position is equal to y_0 and that, at this time, the velocity is known to be v_0. The simple principle is that, in a short period of time, so short that there has not been time for the velocity to change significantly from v_0, the change in position will be approximately equal to the change in time multiplied by v_0.

If the motion of the particle is governed by a differential equation, the value of v_0 will be known as a function of x_0 and y_0. Hence, given x_0 and y_0, the solution at x_1, assumed to be close to x_0, can be calculated as:

$$y_1 = y_0 + (x_1 - x_0)v_0 \,, \tag{D.1}$$

which can be found from known values only of x_0, x_1, and y_0. Assuming that v_1, found using the differential equation from the values x_1 and y_1, is sufficiently accurate, a second step can be taken to find y_2, an approximate solution at x_2, using the formula

$$y_2 = y_1 + (x_2 - x_1)v_1 \,. \tag{D.2}$$

A sequence of approximations $y_1, y_2, y_3, \cdots$ to the solution of the differential equation at $x_1, x_2, x_3, \cdots$ is intended to lead eventually to acceptable approximations, at increasingly distant times from where the initial data was given.

D.1.1 The method itself

Consider the first-order differential equation

$$\frac{dx}{dt} = f(x,t)\,, \tag{D.3}$$

with $x = x_0$ at $t = t_0$. Figure D.1 illustrates the principle of applying the Euler method.

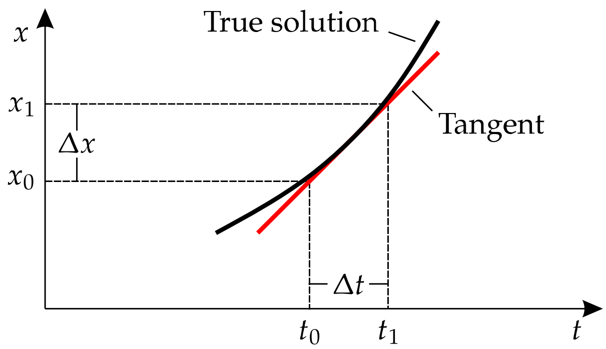

Figure D.1: The principle of applying the Euler method.
Adapted from Kundur [33, p. 836].

At $x = x_0$, $t = t_0$, the curve representing the true solution can be approximated by its tangent having a slope

$$\frac{dx}{dt}\Big|_{x=x_0} = f(x_0, t_0)\,. \tag{D.4}$$

Therefore,

$$\Delta x = \frac{dx}{dt}\Big|_{x=x_0}\Delta t\,. \tag{D.5}$$

The value of x at $t = t_1 = t_0 + \Delta t$ is given by

$$x_1 = x_0 + \Delta x = x_0 + \frac{dx}{dt}\Big|_{x=x_0}\Delta t\,. \tag{D.6}$$

The Euler method is equivalent to using the first two terms of the Taylor series expansion for x around the point (x_0, t_0):

$$x_1 = x_0 + \Delta t(\dot{x}_0) + \frac{\Delta t^2}{2!}(\ddot{x}_0) + \frac{\Delta t^3}{3!}(\dddot{x}_0) + \cdots\,. \tag{D.7}$$

After using the Euler technique for determining $x = x_1$ corresponding to $t = t_1$, it is possible to take another short time step Δt and determine x_2 corresponding to $t_2 = t_1 + \Delta t$ as follows:

$$x_2 = x_1 + \frac{dx}{dt}\Big|_{x=x_1}\Delta t\,. \tag{D.8}$$

By applying the technique successively, values of x can be determined corresponding to different values of t.

The method considers only the first derivative of x and is, therefore, referred to as a *first-order* method. To give sufficient accuracy for each step, Δt has to be small. This will increase round-off errors, and the computational effort required will be very high.

D.2 Runge-Kutta methods

The idea of generalizing the Euler method, by allowing for a number of evaluations of the derivative to take place in a step, is generally attributed to Runge [231]. Further contributions were made by Heun [232] and Kutta [233]. The latter completely characterized the set of Runge–Kutta methods of order four and proposed the first methods of order five.

Since the advent of digital computers, fresh interest has been focused on Runge–Kutta methods, and a large number of research workers have contributed to recent extensions to the theory, and to the development of particular methods.

The Runge–Kutta methods approximate the Taylor series solution; however, unlike the formal Taylor series solution, the Runge–Kutta methods do not require explicit evaluation of derivatives higher than the first. The effects of higher derivatives are included by several evaluations of the first derivative. Depending on the number of terms effectively retained in the Taylor series, there are Runge–Kutta methods of different orders.

D.2.1 Fourth-order Runge–Kutta method

Referring to the differential equation (D.3), the general formula giving the value of x for the $(n+1)^{st}$ step is

$$x_{n+1} = x_n + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4)\,, \tag{D.9}$$

where

$$k_1 = f(x_n, t_n)\Delta t\,, \tag{D.10a}$$

$$k_2 = f(x_n + \frac{k_1}{2}, t_n + \frac{\Delta t}{2})\Delta t\,, \tag{D.10b}$$

$$k_3 = f(x_n + \frac{k_2}{2}, t_n + \frac{\Delta t}{2})\Delta t\,, \tag{D.10c}$$

$$k_4 = f(x_n + k_3, t_n + \Delta t)\Delta t\,. \tag{D.10d}$$

The physical interpretation of the above solution is as follows:

$$k_1 = (\text{slope at the beginning of time step})\Delta t\,,$$

$$k_2 = (\text{first approximation to slope at midstep})\Delta t\,,$$

$$k_3 = (\text{second approximation to slope at midstep})\Delta t\,,$$

$$k_4 = (\text{slope at the end of step})\Delta t\,,$$

$$\Delta x = \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4)\,.$$

Thus, Δx is the incremental value of x given by the weighted average of estimates based on slopes at the beginning, midpoint, and end of the time step.

This method is equivalent to considering up to the fourth derivative terms in the Taylor series expansion; it has an error of the order of Δt^5.