Università degli Studi di Padova

Leuphana Universität Lüneburg

Department of Management and Engineering

Ph.D. program in Mechatronics and Product Innovation Engineering

Cycle: XXXVII

Institute for Production Technology and Systems

Ph.D. program in Technical Sciences

Advanced rare earth-free motor drives for energy-efficient pumps in urban applications

Coordinator: Prof. Giulio Timelli

Supervisor: Prof. Mauro Zigliotto

Dean: Prof. Dr. Paul Drews

Supervisor: Prof. Dr.-Ing. Paolo Mercorelli

Ph.D. Candidate: Saverio Rigon

LEUPHANA
UNIVERSITÄT LÜNEBURG

Advanced rare earth-free motor drives for energy-efficient pumps in urban applications

From the Faculty of Management and Technology of
Leuphana Universität Lüneburg for the conferral of the degree
Doktor der Ingenieurwissenschaften
– Dr.-Ing. –
approved dissertation submitted by Saverio Rigon
born on 14.03.1996 in Vicenza

Submitted on: 12.05.2025

Oral defense (Disputation) on: 12.05.2025

Supervisor (Università degli Studi di Padova): Prof. Mauro Zigliotto
Supervisor (Leuphana Universität Lüneburg): Prof. Dr.-Ing. Paolo Mercorelli

Published under the title:
Advanced rare earth-free motor drives for energy-efficient pumps in urban applications

Year of publication: 2025

Acknowledgements

I would like to thank my family and friends for supporting me throughout this wonderful journey, for their love and encouragement.

I would like to express my deepest gratitude to my *Italian* supervisor, Prof. Mauro Zigliotto, for his guidance, support, and encouragement throughout my studies. I am grateful for his patience, enthusiasm, and willingness to share his knowledge and experience. I am also deeply thankful to my *German* supervisor, Prof. Paolo Mercorelli, for his support, his valuable advice, and for kindly hosting me in his research group. I will always remember the long walks after work, during which we discussed our ideas.

I have met so many friends and colleagues during this PhD journey, and I would like to thank all of them for the great times we shared. In Vicenza, I want to express my gratitude to Prof. Riccardo Antonello for his invaluable help and for the memorable times we had together. His contributions were essential to completing this work, and I have learned so much from him. I also wish to thank Prof. Fabio Tinazzi, an exceptional mentor and person whom I will always remember fondly. Furthermore, I want to thank all the PhD colleagues, students, and technicians who shared their working days in the mechatronics laboratory. Those moments are the best part of this journey and hold a special place in my heart. In particular, I want to thank Andrea and Ismaele for their strong friendship and the wonderful times we spent together.

I also want to express my gratitude to the staff I met at Tekneidos, the Italian company where I had the opportunity to work. In particular, I want to thank Eng. Ivan Tocca, from whom I learned so much and shared countless laughs.

In Lüneburg, I have many people to thank. First and foremost, I want to thank Dr. Eng. Benedikt Haus, a true friend and colleague, who helped me many times in various contexts, I will always owe you one. I also want to thank Jannis and Felix for the wonderful moments we shared and the great times we had in the lab.

Abstract

The growing emphasis on environmental sustainability has elevated the importance of energy-efficient technologies, particularly in electric motors, which account for over half of global electricity consumption. The adoption of high-efficiency standards for electric motors, along with innovations in control strategies and hardware design, is crucial for reducing energy consumption and mitigating environmental impact.

Although induction motors dominate the electric motor market because of their robustness and cost-effectiveness, there is a growing interest in synchronous reluctance motors (SynRMs) and other high-efficiency alternatives that avoid the use of rare-earth materials, which pose both geopolitical and environmental challenges.

This research focuses on the development of capacitorless drives and sensorless control techniques for SynRMs, particularly in applications such as pumps, where cost, efficiency, and reliability are essential.

The study explores the design and implementation of advanced control methods, eliminating bulky electrolytic capacitors and position sensors to enhance drive durability, reduce costs, and improve remote operability without requiring additional hardware. By collaborating with industry partners, this project develops practical solutions for adjustable-speed drives in pump systems.

For several decades, Kalman filter (KF)-based sensorless control techniques have been studied due to their numerous advantages. However, their application in real-world products has been limited by the cumbersome manual tuning required. Key advancements in this PhD project include the use of Adaptive Kalman Filters (AKFs), which feature much easier and nearly automatic tuning procedures. The nonlinear model of SynRMs necessitated the use of nonlinear KF algorithms. Among the various alternatives, determining the most suitable option for synchronous motor drive applications was crucial. To this end, the research initially focused on evaluating different algorithms, particularly comparing the extended Kalman filter (EKF) and unscented Kalman filter (UKF) approaches, assessing their performance in sensorless SynRM control. Throughout the project, the selection of the best algorithm and its straightforward tuning were considered pivotal in promoting the use of KF-based sensorless SynRM drives in pump applications.

The findings contribute to the development of reliable, sustainable, and cost-effective electric drive solutions, positioning SynRMs as a key component in the transition to environmentally responsible HVAC and pump systems.

Sommario

La crescente attenzione alla sostenibilità ambientale ha aumentato l'importanza delle tecnologie a basso consumo energetico, in particolare nei motori elettrici, che rappresentano oltre la metà del consumo globale di elettricità. L'adozione di standard ad alta efficienza per i motori elettrici, insieme a innovazioni nelle strategie di controllo e nel design dell'hardware, sono passi cruciali per ridurre il consumo energetico e mitigare gli impatti ambientali.

Sebbene i motori a induzione dominino il mercato dei motori elettrici grazie alla loro robustezza e convenienza economica, c'è un interesse crescente per i motori sincroni a riluttanza (SynRM) e altre alternative ad alta efficienza che evitano l'uso di terre rare, le quali pongono sfide sia geopolitiche che ambientali.

Questa ricerca si concentra sullo sviluppo di azionamenti senza condensatori e tecniche di controllo sensorless per SynRM, in particolare in applicazioni come le pompe, dove costo, efficienza e affidabilità sono essenziali.

Questa tesi esplora la progettazione e l'implementazione di metodi di controllo avanzati, eliminando ingombranti condensatori elettrolitici e sensori di posizione per migliorare la durabilità degli azionamenti, ridurre i costi e aumentare la possibilità di operare da remoto senza la necessità di hardware aggiuntivo. Collaborando con partner industriali, il progetto sviluppa soluzioni pratiche per azionamenti a velocità variabile nei sistemi di pompaggio.

Da diversi decenni, il controllo sensorless basato su filtri di Kalman (KF) è stato studiato per i suoi numerosi vantaggi. Tuttavia, la sua applicazione in prodotti reali è stata limitata dalla complessa taratura manuale. Tra i principali temi di questo progetto di dottorato, figura l'uso di filtri di Kalman adattivi (AKF), che presentano procedure di taratura molto più semplici e quasi automatiche. Il modello non lineare, come quello del motore sincrono a riluttanza, richiede l'uso di filtri di Kalman non lineari. Tra le varie alternative, è stato fondamentale determinare la tecnica adattativa più appropriata per le applicazioni con motori sincroni. In tal senso, la ricerca si è inizialmente concentrata sulla valutazione di diversi algoritmi, confrontando in particolare gli approcci basati sull'Extended Kalman Filter (EKF) e sull'Unscented Kalman Filter (UKF), valutandone le prestazioni nel controllo sensorless per SynRM. Nello svolgimento del progetto, la scelta del miglior algoritmo e la sua immediata taratura sono stati considerati elementi chiave per promuovere l'uso degli azionamenti sensorless per SynRM basati su KF, in un ambito come quello dei sistemi di pompaggio.

I risultati contribuiscono allo sviluppo di soluzioni di azionamenti elettrici affidabili, sostenibili ed economici, posizionando i motori sincroni a riluttanza come una parte essenziale nella transizione verso sistemi HVAC e di pompaggio più rispettosi dell'ambiente.

Zusammenfassung

Energieeffiziente Technologien haben im Rahmen der Nachhaltigen Entwicklung insbesondere im Feld der elektrischen Antriebstechnik stark an Bedeutung gewonnen, denn Elektromotoren sind heute für über die Hälfte des globalen Stromverbrauchs verantwortlich. Die Einführung von neuen Standards für hocheffiziente Elektromotoren sowie Innovationen in Regelungsalgorithmen und Hardware-Designs sind entscheidend, um den Energieverbrauch zu senken und Umweltbelastungen zu reduzieren.

Obwohl Asynchronmotoren aufgrund ihrer Robustheit und Kosteneffizienz den Markt für Elektromotoren dominieren, wächst das Interesse an Synchron Reluktanzmotoren (SynRMs) und anderen hocheffizienten Alternativen, die auf seltene Erden verzichten. Diese Materialien stellen sowohl geopolitische als auch ökologische Herausforderungen dar.

Diese Forschung konzentriert sich auf die Verbesserung von kondensatorlosen SynRM-Antriebssystemen und ihrer sensorlosen Regelung, insbesondere in Pumpenanwendungen, wo Kosten, Effizienz und Zuverlässigkeit im Mittelpunkt stehen. Konkret wird die Herleitung und Implementierung fortgeschrittener Regelungsalgorithmen für Antriebssysteme behandelt, bei denen voluminöse Elektrolytkondensatoren und Absolutwinkelgeber vermieden werden, um die Haltbarkeit des Systems zu erhöhen, die Kosten zu senken und die Fernbedienbarkeit ohne zusätzliche Hardware zu verbessern. Durch die Zusammenarbeit mit Industriepartnern werden praktische Lösungen für drehzahlverstellbare Antriebe in Pumpemsystemen entwickelt.

Seit mehreren Jahrzehnten wird die sensorlose Regelung von Synchronmotoren auf Basis von Kalmanfiltern (KF) untersucht, da sie zahlreiche praktische Vorteile mit sich bringt. Ihre Anwendung in realen Produkten wurde bisher jedoch durch den Aufwand der manuellen Abstimmung und Feineinstellung gebremst. Wichtige Fortschritte in diesem Promotionsprojekt umfassen die Nutzung von adaptiven Kalmanfiltern (AKFs), die deutlich einfacher, nahezu automatisch und adaptiv abgestimmt werden können. Das nichtlineare magnetische dynamische Verhalten von SynRMs macht den Einsatz nichtlinearer KF-Algorithmen mit entsprechenden nichtlinearen Modellen erforderlich. Um die am besten geeignete KF-Variante für diese konkrekte sensorlose SynRM-Anwendung auszumachen, konzentriert sich die Forschung zunächst auf einen eingehenden Vergleich mit quantitativer Bewertung verschiedener Algorithmen, insbesondere des Extended Kalmanfilters (EKF) und des Unscented Kalmanfilters (UKF). Die Auswahl des Algorithmus und dessen einfache

Abstimmung sind in der Folge ein zentraler Punkt dieses Projekts, mit dem Ziel die Nutzung von KF-basierten sensorlosen SynRM-Antrieben zu fördern.

Die erlangten Erkenntnisse tragen zur Entwicklung zuverlässiger, nachhaltiger und kosteneffizienter elektrischer Antriebslösungen bei und positionieren SynRMs als einen Schlüsselbestandteil des Übergangs zu umweltfreundlichen Klima- und Pumpensystemen.

Acronyms

AKF Adaptive Kalman Filter

ASD Adjustable Speed Drive

CPL Constant Power Load

EKF Extended Kalman Filter

FCP Fast Control Prototype

HVAC Heating Ventilation and Air Conditioning

IAE Innovation-based Adaptive Estimation

IEC International Electrotechnical Commission

IM Induction Motor

IPMSM Interior Permanent Magnet Synchronous Motor

LUT Look-Up Table

MCS Monte Carlo Simulation

MLE Maximum-Likelihood Estimation

MLEKF Maximum-Likelihood Estimation Kalman Filter

MS Master-Slave

MSE Mean Squared Error

MSKF Master-Slave Kalman Filter

MTPA Maximum-Torque-Per-Ampere

PDF Probability Density Function

PLL Phase-Locked Loop

PMASynRM Permanent Magnet Assisted SynRM

PMSM Permanent Magnet Synchronous Motor

PWHD Partially Weighted Harmonic Distortion

RMS Root Mean Square

RRE Rare Earth Elements

SVM Space Vector Modulation

SynRM Synchronous Reluctance Motor

THD Total Harmonic Distortion

UKF Unscented Kalman Filter

VSD Variable Speed Drive

VSI Voltage Source Inverter

Nomenclature

$\boldsymbol{\lambda}_{\mathrm{dq}}$	Flux linkages in the synchronous reference frame
Ω_E	Steady state electrical speed of the motor
ω_E	Electrical speed of the motor
Ω_M	Steady state mechanical speed of the motor
ω_M	Mechanical speed of the motor
ϑ_E	Electrical position of the motor
ϑ_M	Mechanical position of the motor
$\boldsymbol{L}_{d,q}^{\mathrm{app}}$	Apparent inductance
$\boldsymbol{L}_{d,q}^{\mathrm{diff}}$	Differential inductance
C	Capacitor
C_d	Damping capacitor
f_g	Grid frequency
f_{PWM}	PWM frequency
f_s	Sampling frequency
i_C	Capacitor current
i_G	Grid current
I_L	Steady state inverter input current
i_L	Inverter input current
i_l	Small signal inverter input current
i_M	General motor current
I_S	Steady state rectifier output current

i_S	Rectifier output current
i_s	Small signal rectifier output current
$i_{\alpha,\beta}$	Currents in the stationary reference frame
i_{abc}	Three phase currents
$I_{D,Q}$	Steady state currents in the synchronous reference frame
$i_{D,Q}$	Currents in the synchronous reference frame
$i_{d,q}$	Small signal currents in the synchronous reference frame
L	Equivalent grid inductance
L_g	Grid inductance
P_L	Steady state inveter input power
p_L	Inverter input power
p_l	Small signal inveter input power
R	Equivalent grid resistance
R_d	Damping resistance
R_g	Grid resistance
R_L	Negative load resistance
r_s	Stator resistance
T_c	Control period
T_s	Sampling time
U_C	Steady state DC-link voltage
u_C	DC-link voltage
u_c	Small signal DC-link voltage
u_G	Grid voltage
U_S	Steady state rectified voltage
u_S	Rectified voltage
u_s	Small signal rectified voltage
$u_{\alpha,\beta}$	Voltages in the stationary reference frame
u_{abc}	Three phase voltages

$U_{D,Q}$ Steady state voltages in the synchronous reference frame

$u_{D,Q}$ Voltages in the synchronous reference frame

$u_{d,q}$ Small signal voltages in the synchronous reference frame

Contents

CHAPTER 1

General framework

In recent years, environmental issues have garnered increasing attention. To mitigate the ever-growing human environmental footprint, the efficiency of electric motors has become significantly more important. The global energy economy is increasingly electrifying. Since 2010, electricity demand has grown on average by 2.7% per year, while overall energy demand has risen by 1.4% per year [1]. Electricity is replacing fossil fuels for providing heat, mobility, and industrial energy demand. Innovations such as smart grids and advancements in the efficiency of electric motors and appliances have further increased electricity's appeal.

According to [2], 53% of global electric energy consumption is used to power electric motors. The International Electrotechnical Commission (IEC) introduced efficiency classes and requirements for AC electric motors in 2009. In 2014, they expanded the efficiency range covered by the standard `IEC/EN 60034-30-1`. The IEC efficiency classes are: `IE1` (standard efficiency), `IE2` (high efficiency), `IE3` (premium efficiency), and `IE4` (super-premium efficiency). In 2019, they upgraded the efficiency requirements for 2-, 4-, 6-, and 8-pole motors. Since 1 July 2021, motors between 0.75 kW and 100 kW are required to meet a minimum efficiency class of `IE3`, while motors between 0.12 and 0.75 kW must meet `IE2`. As of 1 July 2023, motors between 75 kW and 200 kW are required to meet the even higher efficiency class of `IE4`. The minimum efficiency requirements for 2-, 4-, 6-, and 8-pole motors for a given `IE` class are shown in Fig. 1.1.

Despite their lower efficiency, approximately 80% of electric motors are Induction Motors (IMs) due to their low cost and robustness. The remaining 20% of electric drives belong to the synchronous motor family, which is divided into surface Permanent Magnet Synchronous Motors (PMSMs), Interior Permanent Magnet Synchronous Motors (IPMSMs), and Synchronous Reluctance Motors (SynRMs). Permanent magnet electric motors offer the highest torque-to-volume ratio, making them highly attractive in automotive and electric mobility applications, where drive volume is a critical factor.

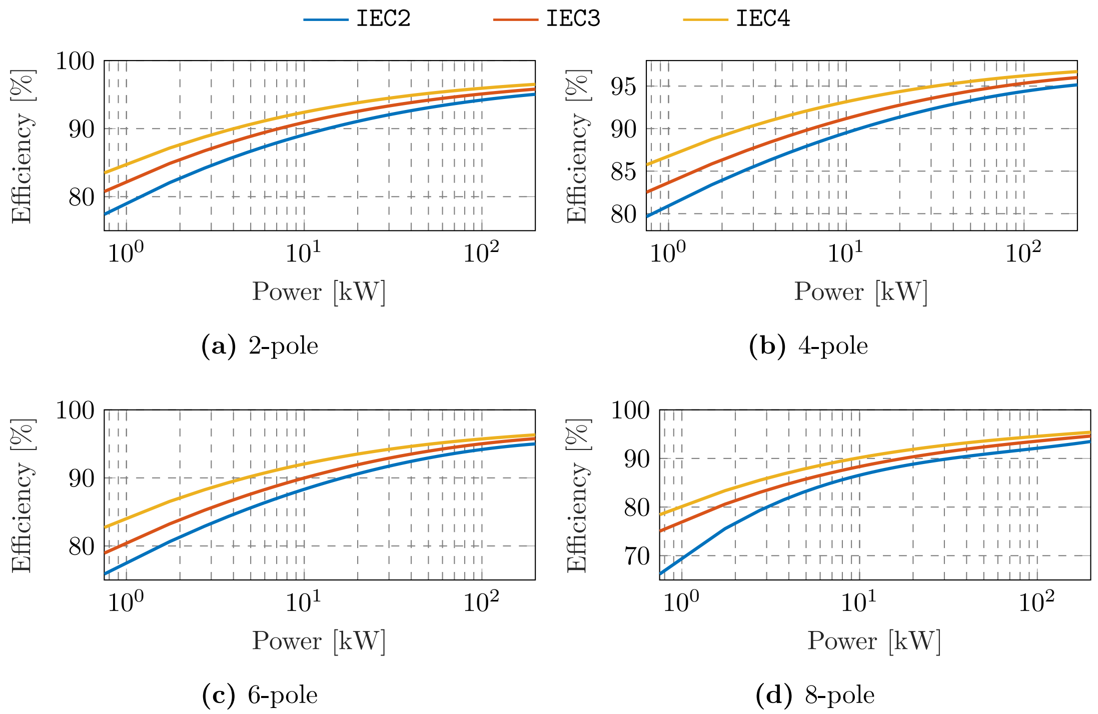


Figure 1.1 IE efficiency classes for electric motors from 75 kW to 200 kW at 50 Hz, according to the standard `IEC/EN 60034-30-1` [2]. (a) 2-pole machine; (b) 4-pole machine; (c) 6-pole machine; (d) 8-pole machine;

In addition to mobility applications, the Heating Ventilation and Air Conditioning (HVAC) sector also requires more efficient solutions. Recently, building energy consumption in the European Union accounted for 37% of final energy use, surpassing the industrial sector (28%) and transportation (32%). There are considerable opportunities to reduce energy consumption in buildings. In large buildings, computerized central stations often monitor and control services such as heating, air conditioning, and water supply. Electronically controlled pumps, integrated into such building automation systems, increase convenience while saving energy.

Improving efficiency reduces waste while maintaining the same output, resulting in a positive impact on energy demand. This leads to a decreased reliance on fossil fuel sources, thereby mitigating environmental impacts. The demand for increasingly efficient solutions in the field of electric drives translates not only into the development of higher-performance electric machines but also into the adoption of innovative and more complex control strategies. PMSMs and IPMSMs motors are currently among the most efficient electric machines available. However, their application involves the use of Rare Earth Elements (RRE) for permanent magnets.

1.1 Rare Earth Elements

RRE are often considered the metaphorical *gold* of the twenty-first century. Their extensive use spans a wide range of technological applications, including microchips, magnets, superconductors, home appliances, wind generators, automotive drives, defense, and aerospace industries.

The main producer is China, whose trade monopoly enables the Beijing government to use rare earth embargos as a geopolitical tool against countries that do not align with Chinese interests [3]. For instance, in 2010, a diplomatic issue between Japan and China, involving a Chinese fishing boat in Japanese waters, was managed by the latter by reducing rare earth exports to Japan. From this point onwards, finding alternatives to China's dominance has become a critical priority [4].

According to [5], China's mining production was 240 t in 2023 (14.29% more than in 2022), while the United States produced 43 t (2.3% more than in 2022). China and the USA are responsible for 69% and 12% of global production, respectively. The three largest reserves [1] of RRE are located in China (44 Mt), Vietnam (22 Mt), and Brazil (21 Mt).

In addition to political concerns, the environmental impact of rare earth mining is a significant issue. Pressures to extract and process REEs are accelerating globally, as the International Energy Agency (IEA) suggests that meeting *Net Zero Emissions* goals would require REE extraction to increase by a factor of 10 by 2030.

However, REE mining has been linked to greater environmental impacts compared to other minerals and metals. REEs are usually present in very low concentrations and are often combined, meaning that their extraction and separation are expensive, require large amounts of energy and water, and generate substantial waste. Furthermore, they are frequently associated with radioactive and hazardous elements such as uranium, thorium, arsenic, and other heavy metals, posing significant health and environmental risks. Extraction methods include open-pit mining (generally involving intense water usage), underground mining, and in-situ leaching [7], [8].

While there are high expectations for REE recycling, this remains a marginal source (less than 1%). Numerous obstacles hinder REE recycling, including the low concentration of REEs in end-products and the difficulties in separating individual REEs from one another. Recycling is also far from being a clean industry, requiring substantial energy and generating hazardous waste.

Given these environmental and geopolitical challenges, research is increasingly focused on developing electric drives that do not depend on rare earth materials. This has been a driving force behind the advancement of the SynRMs.

1.2 Synchronous Reluctance Motors

The SynRM was developed in the twentieth century, but its use in past years was minimal due to its lower efficiency compared to other available solutions. However, recent advancements in rotor design have significantly improved motor efficiency, positioning the SynRM as a viable alternative to the IM [9]. The stator design is similar to that used in conventional induction and synchronous machines, while the

[1] Reserves may be considered a working inventory of mining companies' supplies of an economically extractable mineral commodity. As such, the magnitude of that inventory is necessarily limited by many considerations, including cost of drilling, taxes, price of the mineral commodity being mined, and the demand for it. Reserves will be developed to the point of business needs and geologic limitations of economic ore grade and tonnage [6, Appendix C].

rotor is constructed from punched laminates [10]. The absence of a rotor cage or field winding reduces the manufacturing cost of the synchronous reluctance machine, making it more affordable than conventional machines.

The operating principle of the SynRM is based on magnetic asymmetry. The rotor aligns itself with the rotating magnetic field generated by the stator windings to minimize the machine's reluctance.

The SynRM offers several advantages, including a simple and robust mechanical structure, low-cost manufacturing, and independence from the fluctuating prices of permanent magnets. It is increasingly considered a strong candidate to replace induction motors in many Variable Speed Drives (VSDs) applications. However, one significant drawback is the highly non-linear relationship between currents and fluxes, which complicates current control design [11].

One fundamental difference between SynRM and IM applications is the requirement for an inverter to control the SynRM. While induction motors can be directly connected to the grid, the SynRM requires a power electronics interface. Although this adds complexity and increases initial costs, it leads to lower running costs due to the SynRM's superior efficiency. Additionally, the inverter facilitates variable-speed operation and enables the implementation of soft-starting and smoother emergency stopping procedures. These features reduce mechanical stress on the motor and transmission systems, enhancing the overall life and reliability of the system.

For applications such as pumps, which prioritize low costs, minimal maintenance, and high reliability, SynRMs present a compelling alternative to IMs. However, these motors typically require rotor position measurement. To maintain competitive drive costs, SynRMs are almost always paired with sensorless control algorithms.

1.3 Sensorless drives

Sensorless algorithms replace rotor position measurements with an observer that relies solely on current—and, in some cases, voltage—measurements. By eliminating the need for encoders or resolvers, these algorithms remove associated signal conditioning circuits, cabling, and connectors. This not only reduces costs but also improves system reliability.

In automotive and traction applications where position sensors are not entirely removed, sensorless observers can serve as redundant components, compensating for potential sensor failures in real-time. This redundancy further enhances system reliability while maintaining cost-effectiveness.

The literature offers various techniques for sensorless position estimation in alternating current motors, broadly categorized into saliency-based and model-based methods. Saliency-based methods are particularly effective under stationary or low-speed conditions, whereas model-based approaches perform better at high speeds.

Among model-based techniques, the most prevalent method estimates the rotor position of synchronous motors by reconstructing the back-electromotive force (BEMF). These methods vary in the complexity of their model approximations. Extending sensorless techniques already developed for PMSMs and IPMSMs to SynRMs poses additional challenges due to the pronounced magnetic nonlinearity of reluctance motors [12].

Applications where reluctance motors can compete effectively with induction motors often do not require high starting torque. Examples include pumps, HVAC systems, and wind turbines [13]. In such cases, robust sensorless solutions can integrate model-based techniques with simpler start-up algorithms. These hybrid approaches enable low-speed operation without needing initial position estimation or extensive machine parameter identification [14]. To minimize open-loop operation time, it is beneficial to extend the applicability of model-based sensorless techniques as far as possible into the low-speed range [15].

The removal of the position sensor significantly enhances system reliability, reducing both fixed and operational costs. Another critical component in many HVAC drives, particularly in the pump sector, is the electrolytic capacitor bank, present in the Voltage Source Inverter (VSI) that feed the SynRM. Its role and optimization also contribute to improving the overall system performance and longevity.

1.4 Capacitorless drives

Electrolytic capacitors are widely used to decouple the grid side from the inverter side in electric drives. However, they are also the most failure-prone component of the drive, often responsible for reduced system reliability and shorter operational lifespans [16]. Replacing this fragile component with a more robust and long-lasting alternative, such as a film capacitor, is a critical step in advancing the reliability of drives in the HVAC sector, and in helping the transition from IM to SynRM drives.

While film capacitors offer greater durability and reliability, their adoption leads to reduced DC-link capacitance. This reduction can destabilize the drive as the power demand from the load increases. If instability occurs, DC-link voltage oscillations may arise, potentially causing failure in critical electronic components like IGBTs. Additionally, grid current distortion can occur, preventing the drive from meeting `IEC` standards.

Developing control strategies to mitigate these drawbacks without introducing additional hardware is crucial for enabling reliable capacitorless drives.

1.5 Thesis Contributions

Given the aforementioned open challenges, this thesis aims to advance the field of electric drives in the HVAC sector by addressing two key areas: the development of capacitorless drives and the sensorless control of SynRM.

The first part of the thesis focuses on improving the stability of capacitorless drives and addressing issues related to replacing large electrolytic capacitors with slim film capacitors. The primary objective was to design a control strategy capable of ensuring stability and compliance with standards without the need for extra hardware components.

The second part of the thesis delves into sensorless control of SynRM drives, specifically through the application of Kalman Filter observers. Particular emphasis was placed on the unsolved issue of filter tuning, which remains a significant barrier to widespread adoption. This research explores innovative solutions on this side.

Through these contributions, this thesis seeks to address critical reliability and control challenges, paving the way for the adoption of SynRM drives as a more efficient and sustainable solution in the HVAC industry.

Part I

Electrolytic capacitorless drives

CHAPTER 2

Introduction

Heating, ventilation, and air conditioning (HVAC) systems are among the largest consumers of electricity in buildings, serving both industrial and domestic sectors.

According to [17], approximately 30% of faults in converters are caused by capacitor degradation, making them the weakest link in Adjustable Speed Drives (ASDs). This highlights the need for monitoring and maintenance scheduling to prevent serious degradation or breakdowns [16]. As a cost-effective alternative, electrolytic capacitors can be replaced with film capacitors, which have a much longer lifetime [18], lower cost, but reduced capacitance for the same physical volume [19].

ASDs implementing low-capacitance film capacitors are referred to as *capacitorless drives*. However, this does not come without consequences. Reducing the overall capacitance can cause stability issues in the DC link due to resonances, especially when there are high line impedances (e.g., when the medium voltage cabinet is located far from the point of use) or when a line filter is used to mitigate noise conducted to the mains. At the same time, the downward shift of the DC-link capacitance may negatively impact both the line Total Harmonic Distortion (THD) and the Partially Weighted Harmonic Distortion (PWHD) [20]. Another significant drawback of capacitorless ASDs is the large DC-link voltage fluctuation, which can cause severe control issues in the presence of Constant Power Loads (CPLs), as frequently happens in HVAC applications. Intuitively, in the presence of DC bus variations, CPLs demand increasing current as the DC bus voltage decreases, resulting in a so-called negative input impedance. This can lead to ASD input current oscillations or instability [21].

2.1 State of the art

Recently, there has been growing interest in capacitorless frequency converters, as evidenced by an increasing number of publications. The focus has been primarily on flux weakening control in single-phase AC input applications [21], [22], [23], DC-link voltage damping [20], [24], [25], [26], and line current quality [27], [28]. The latter two issues are interrelated, as reducing DC-link voltage fluctuations improves the harmonic content of the line current. However, flux weakening control is not considered within this work, as its effect is less pronounced in three-phase AC input applications. To reduce costs, solutions requiring additional hardware, such as the one proposed in [28], were not considered.

Additionally, solutions involving extra hardware [28] or addressing overvoltage during braking operations [29], [30] are not relevant for this discussion. These solutions either fail to meet cost-effectiveness requirements or are not applicable to electrical pumps, which are the specific focus of this PhD work.

It is important to note that damping DC-link oscillations always requires injecting additional current or voltage signals, which may result in additional load torque ripple [31]. Furthermore, any countermeasure relying on current injection will only be effective if the reference signal falls within the bandwidth of the current controller. Finally, it remains unclear how injection strategies aimed at damping DC-link voltage oscillations affect the quality of the line current. These aspects merit further investigation.

2.2 Research objectives and structure

Most research contributions focus on synchronous permanent magnet and induction motors, with only a few addressing SynRM [20], [26], [32]. This thesis aims to provide an analytical foundation for the rational design of a capacitorless SynRM-based pump drive. While several papers address capacitorless AC drives, very few focus on SynRM, and even fewer adopt a rigorous approach with clear guidelines for control parameter tuning. This constitutes the focus of the first part of this PhD thesis.

The rigorous analytical treatment presented in this thesis enables a deeper understanding of the complex interactions between DC-link oscillations, line current quality, and control signals in terms of power reference generation. During this research, practical implementation has also been a priority, and useful guidelines for tuning control parameters are provided.

An improved technique was developed, combining voltage and current reference signals, in contrast to conventional methods that typically rely on only one.

The proposed dual-way damping power generation allows current loops to operate within their bandwidth limits, while the voltage feedforward automatically handles the remainder. The mathematical analysis also elucidates the relationship between DC bus voltage damping and line current quality. As a third major contribution, this research achieves a significant reduction in line current harmonic content compared to existing strategies.

The first part is structured as follows: Ch. 3 introduces the mathematical foundations; in particular, Sect. 3.1 lays the groundwork for the system stability analysis, Sect. 3.2 computes the required damping power, and Sect. 3.3 proposes an improved technique for its generation using combined voltage and current injections. Ch. 4 describes the experimental setup used as a proof of concept and provides detailed implementation information for the proposed techniques. The experimental results, which align with expectations, are thoroughly discussed. Finally, conclusions are presented in Ch. 5.

CHAPTER 3

Capacitorless AC Drive Model

3.1 Mathematical foundations

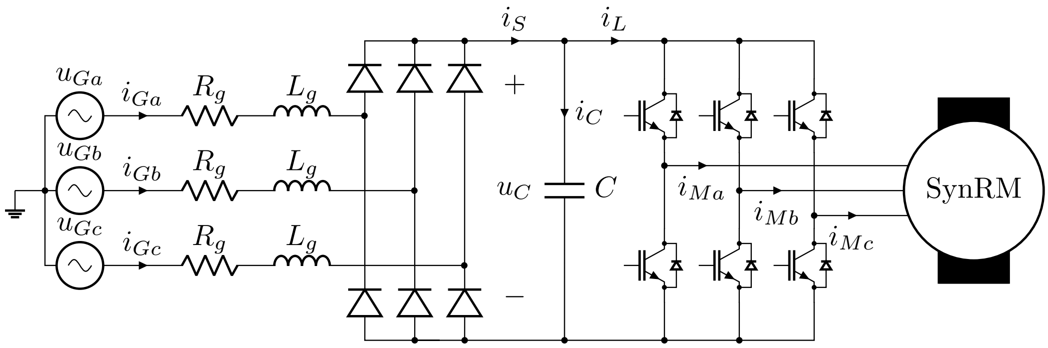


Figure 3.1 Overall circuit diagram of a three-phase adjustable speed AC drive.

The overall circuit diagram of a conventional three-phase adjustable AC drive is shown in Fig. 3.1. It consists of six-pulse diode-bridge rectifier fed by a three-phase grid, followed by a two-level inverter driving a Synchronous Reluctance Motor (SynRM). In the figure, R_g and L_g are the grid resistance and inductance, while u_{Ga}, u_{Gb} and u_{Gc} are the phase-to-neutral grid voltages, with peak value U_g and frequency ω_g. The currents i_S and i_L are, respectively, those at the output of the three-phase rectifier, and at the input of the of the inverter; u_C and i_C are instead the voltage and current of the DC-link capacitor. With regards to the machine, i_{Ma}, i_{Mb} and i_{Mc} are the phase motor currents.

The system of Fig. 3.1 can be reduced to the equivalent model of Fig. 3.2, thanks to the fact that only two of the six diodes of the rectifier are conducting while the others are blocked at any time. The ideal voltage source represents the rectified

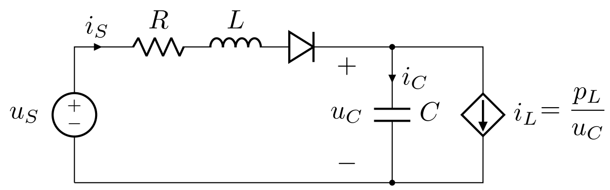


Figure 3.2 Simplified circuit diagram of a three-phase adjustable speed AC drive.

voltage, which is equal to (assuming ideal diodes):

$$u_S = \max\{u_{Ga}, u_{Gb}, u_{Gc}\} - \min\{u_{Ga}, u_{Gb}, u_{Gc}\} \tag{3.1}$$

It can be alternatively expressed in terms of its Fourier series expansion, namely:

$$u_S = U_S \left[1 - \sum_{n=1}^{\infty} \frac{2}{(6n)^2 - 1} \cos(6n\omega_g t) \right] \tag{3.2}$$

where

$$U_s = \frac{3\sqrt{3}}{\pi} U_g \tag{3.3}$$

The amplitude $U_g = \sqrt{2} U_{g,rms}$ and the frequency $\omega_g = 2\pi f_g$ are grid parameters that depend on the local electric energy provider. The resistance and inductance of the equivalent series impedance of the voltage source are equal to:

$$R = 2R_g + \frac{3\omega_g L_g}{\pi}, \qquad L = 2L_g \tag{3.4}$$

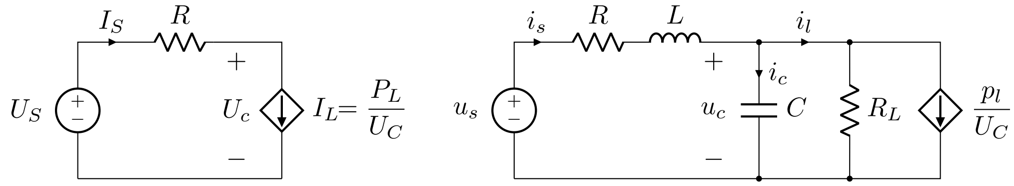
where the term $3\omega_g L_g/\pi$ is the *nonohmic* voltage drop due to the non-ideal commutation of the diodes [33, p. 451]. The current source i_L models the average current absorption on a PWM period of the voltage-source inverter.

The dynamics of the simplified model depicted in Fig. 3.2 is described by the following equations:

$$\begin{cases} L \dfrac{di_S}{dt} = u_S - R\, i_S - u_C \\ C \dfrac{du_C}{dt} = i_S - i_L\,, \quad i_S \geq 0 \end{cases} \tag{3.5}$$

where i_S and u_C are regarded as state variables, while u_S and i_L as inputs. Typically the current i_L is indirectly controlled by adjusting the power $p_L = u_C i_L$ absorbed by the inverter (indeed the inverter load, if the inverter losses are negligible). In this new input variable, the dynamical system (3.5) is nonlinear.

The simplified system (3.5) shown in Fig. 3.2 can be linearized around a working point, resulting in two subsystems, the average constant model and the small-signal model, depicted in Fig. 3.3a and Fig. 3.3b, respectively. In the followings, the quantities i_s, u_c, p_l and u_s represent small variations around the average values assumed at the selected working point, denoted by I_S, U_C, P_L and U_S, so that, for example, $u_C = U_C + u_c$.



(a) average signal model **(b)** small signal model

Figure 3.3 Average and small signal representations of the overall AC drive circuit.

3.1.1 Average signal model

The average signal model is obtained by considering the average values of the state variables i_S and u_C from system (3.5) represented in Fig. 3.3a and denoted by I_S and U_C, respectively. I_S is the average current after the rectifier, while U_C is the average voltage across the DC-link capacitor.

$$\begin{cases} U_S = R\,I_S + U_C \\ I_S = I_L = \dfrac{P_L}{U_C} \end{cases} \tag{3.6}$$

3.1.2 Small signal Model

For the analysis of system stability, the small-signal model is particularly relevant, as it describes the behavior of the system in response to small perturbations around a specific operating point (U_C, P_L). The dynamic equations governing the system are given by:

$$\begin{cases} L\,\dfrac{di_s}{dt} = u_s - R\,i_s - u_c \\ C\,\dfrac{du_c}{dt} = i_s - i_l \end{cases} \tag{3.7}$$

To obtain the linearized model, the inverter current i_l is expressed as the small signal part of a first-order Taylor expansion of the total current i_L around the operating point (U_C, P_L). The total current i_L can be written as:

$$i_L = f(p_l,\, u_c) = \frac{p_l}{u_c} = I_L + i_l \tag{3.8}$$

Applying the Taylor expansion results in:

$$i_L \approx f(P_L,\, U_C) + \left.\frac{\partial f}{\partial p_l}\right|_{(U_C,P_L)} (p_l - P_L) + \left.\frac{\partial f}{\partial u_c}\right|_{(U_C,P_L)} (u_c - U_C) + \dots \tag{3.9}$$

$$\approx \underbrace{\frac{P_L}{U_C}}_{I_L} + \underbrace{\frac{1}{U_C}\,p_l - \frac{P_L}{U_C^2}\,u_c}_{i_l} \tag{3.10}$$

In this formulation, $I_L = \dfrac{P_L}{U_C}$ represents the steady-state current, while i_l is the small-signal perturbation. Substituting i_l into the original system equations (3.7),

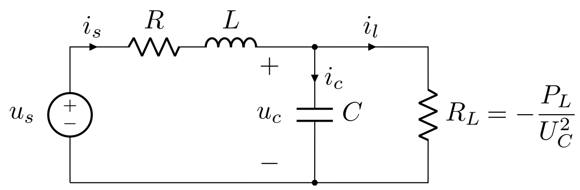


Figure 3.4 Small signal circuit in the case of constant power load, i.e. with $p_l = 0$.

the small-signal model is obtained as:

$$\begin{cases} L\,\dfrac{di_s}{dt} = u_s - R\,i_s - u_c \\ C\,\dfrac{du_c}{dt} = i_s - \dfrac{1}{U_C}\,p_l + \dfrac{P_L}{U_C^2}\,u_c \end{cases} \tag{3.11}$$

or, in state-space form:

$$\begin{bmatrix} \dfrac{di_s}{dt} \\ \dfrac{du_c}{dt} \end{bmatrix} = \begin{bmatrix} -\dfrac{R}{L} & -\dfrac{1}{L} \\ \dfrac{1}{C} & \dfrac{P_L}{CU_C^2} \end{bmatrix} \begin{bmatrix} i_s \\ u_c \end{bmatrix} + \begin{bmatrix} 0 \\ -\dfrac{1}{CU_C} \end{bmatrix} p_l + \begin{bmatrix} \dfrac{1}{L} \\ 0 \end{bmatrix} u_s \tag{3.12}$$

provided that the rectifier output current i_S is always positive. The electric representation of the small signal model is shown in Fig. 3.3b.

In case of a CPL, i.e. $p_l = 0$ and $p_L = P_L$, the small-signal linearized model (3.6)-(3.12) is equivalent to the electric circuit of Fig. 3.4, where the inverter behaves as a resistive load with *negative resistance* $R_L = -U_C^2/P_L$, connected in parallel to the DC-link capacitor. In the circuit, the fluctuations on the DC-link voltage u_c and current i_s are induced by the rectified voltage u_s, that can be considered as an exogenous disturbance for the system under consideration.

The influence of u_s on the state variables is described by the following transfer functions:

$$G_i(s) = \frac{i_s(s)}{u_s(s)} = \frac{Cs + \frac{1}{R_L}}{\Delta(s)}, \qquad G_u(s) = \frac{u_c(s)}{u_s(s)} = \frac{1}{\Delta(s)} \tag{3.13}$$

with

$$\Delta(s) = LCs^2 + \left(\frac{L}{R_L} + RC\right) s + \left(\frac{R}{R_L} + 1\right) \tag{3.14}$$

3.1.3 Stability of the small signal model

Typically the DC-link capacitor C is dimensioned to filter out the harmonics at multiples of sixfold the grid frequency ω_g contained in u_s, provided that (3.13) are BIBO stable.

According to the Routh-Hurwitz criterion, the stability is guaranteed when all the coefficients of the characteristic polynomial (3.14) are positive. Since $R/|R_L| \ll$

1 in almost all practical situations, the stability is guaranteed when the second coefficient is positive, namely:

$$\frac{L}{R_L} + RC > 0 \qquad \Rightarrow \qquad C > \frac{L}{R\,|R_L|} \tag{3.15}$$

It is important to note that the stability condition expressed by (3.15) is influenced by the equivalent line impedance, specifically the grid inductance L_g and resistance R_g, as given by (3.4).

In particular, the higher the line inductance L_g, the larger the capacitance bank required to satisfy the stability condition (3.15). On the other hand, the line resistance R_g counteracts the effect of the inductance, thereby reducing the required capacitance value.

For an effective filtering of the rectified voltage harmonics, the value of the DC-link capacitor is selected so that the natural frequency $\omega_n \approx 1/\sqrt{LC}$ of the poles of (3.13) is sufficiently smaller than ω_g. This usually requires to select large electrolytic DC-link capacitors, that are expensive and with limited life-time.

A better option for cost reduction would consist of using smaller thin-film capacitors, at the expense however of reduced stability margins and filtering capabilities. A way to overcome these issues consists of manipulating the inverter power input p_l in (3.6)-(3.12) according to the measured fluctuations of the DC-link voltage, in order to increase the overall system damping. In doing so, p_l can be regarded as a control input for the system under consideration.

Suitable control strategies (active damping strategies) to improve both stability and damping in systems with small DC-link capacitors are presented in the next chapter.

3.2 Computation of the damping power

Throughout the following analysis, the quantities i_s, u_c, p_l, and u_s denote small deviations from the average values at the chosen operating point, represented by I_S, U_C, P_L, and U_S, respectively. Suppose that the inverter load can be controlled so that its small-signal power absorption is equal to:

$$p_l = k_p(u_c - u_s) \; - \; k_d \frac{du_s}{dt} \tag{3.16}$$

in which the inverter losses are neglected and where u_c is the measured small-signal DC-link voltage, while u_s is an estimation of the small-signal rectified voltage, namely the second term in (3.2), obtained with a signal generator suitably synchronized with the grid voltage, as it will be explain in details in Sect. 4.1.1. After rewriting (3.16) as

$$p_l = k_p(u_c - u_s) \; + \; k_d \frac{d(u_c - u_s)}{dt} \; - \; k_d \frac{du_c}{dt} \tag{3.17}$$

and by replacing in (3.6)-(3.12) it is immediate to notice that the overall system behaves as the electric circuit of Fig. 3.5, where:

$$R_d = \frac{U_C}{k_p}\,, \qquad C_d = \frac{k_d}{U_C} \tag{3.18}$$

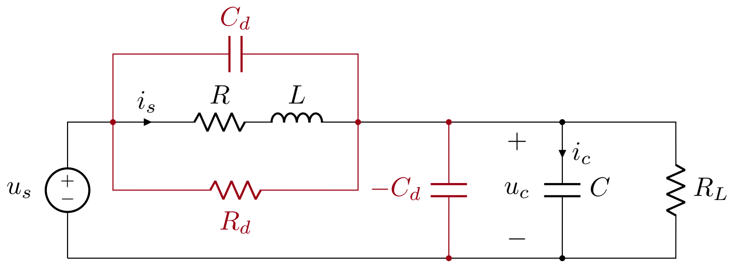


Figure 3.5 Small signal circuit in the case of $p_l \neq 0$, according to its definition in (3.16).

Table 3.1 Equivalent circuit parameters.

Symbol	Parameter	Value
$P_{L,1}$	Load Power @ (3500 rpm, 2.5 N m)	1055 W
$P_{L,2}$	Load Power @ (3000 rpm, 1.84 N m)	689 W
R_g	Grid resistance	100 mΩ
L_g	Grid inductance	1.15 mH
R	Equivalent resistance	3.04 Ω
L	Equivalent inductance	2.3 mH
C	Film Capacitor	12 µF
U_S	Constant source voltage	538 V
U_C	Constant capacitor voltage	537 V

Accordingly, the transfer functions (3.13) modify as follows:

$$G_i(s) = \frac{i_s(s)}{u_s(s)} = \frac{(C - C_d)\, s + \frac{1}{R_L}}{\Delta_d(s)} \tag{3.19}$$

$$G_u(s) = \frac{u_c(s)}{u_s(s)} = \frac{LC_d s^2 + \left(\frac{L}{R_d} + RC_d\right) s + \left(\frac{R}{R_d} + 1\right)}{\Delta_d(s)} \tag{3.20}$$

with

$$\Delta_d(s) = LCs^2 + \left(\frac{L}{R_d} + \frac{L}{R_L} + RC\right) s + \left(\frac{R}{R_d} + \frac{R}{R_L} + 1\right) \tag{3.21}$$

or, after comparing with (3.14)

$$\Delta_d(s) = \Delta(s) + \frac{1}{R_d}(Ls + R) \tag{3.22}$$

Consider the nominal parameters reported in Tab. 3.1, with P_L as the inverter power absorption.

By sketching the root-locus of (3.22) with respect to variations of the parameter $1/R_d$ (see Fig. 3.6), it can be easily noticed that as R_d decreases, i.e. the proportional gain k_p in (3.16) increases, the poles of (3.20) become more damped, being "attracted" by the zero $-R/L$ in the open left-half of the complex plane.

If the original undamped system is unstable, stability is also recovered by decreasing R_d.

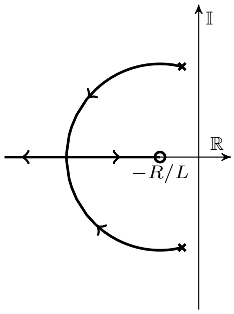
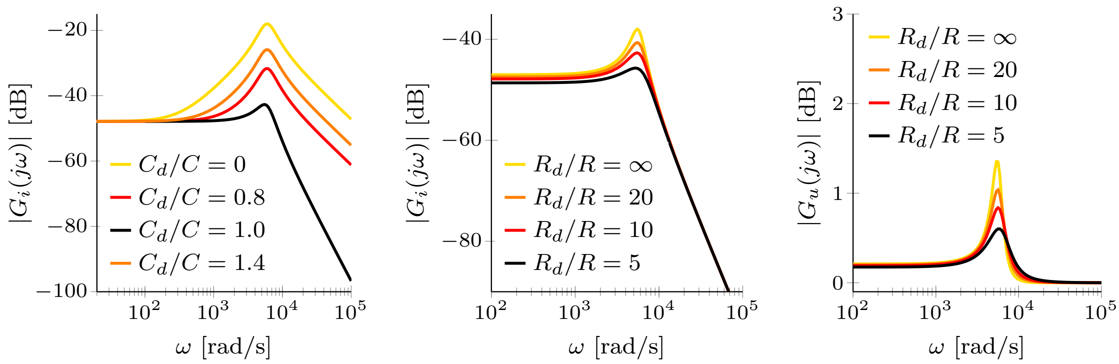


Figure 3.6 Root locus of closed-loop poles for increasing $1/R_d$ (i.e. increasing k_p).



(a) $G_i(s)$ with $R/R_d = 10$ **(b)** $G_i(s)$ with $C/C_d = 1$ **(c)** $G_u(s)$ with $C/C_d = 1$

Figure 3.7 Relevant dynamic of the actively damped system: frequency response magnitude, obtained through the circuit parameters in Tab. 3.1.

As $R_d \to 0$, both poles become real, with one pole moving toward the zero $-R/L$, while the other going to $-\infty$. Although a small value of R_d (large value of k_p) is advisable to increase both the damping and the frequency of the poles of (3.20), a too small value would require an impractical large power absorption in (3.16); hence, a trade-off must be found between the two conflicting requirements.

The benefit of the "virtual" capacitance C_d, i.e. the derivative compensation term in (3.16), can be understood by referring to the first transfer function in (3.20), and the related Bode plot of Fig. 3.7a.

If $C_d = 0$, then the zero remains unchanged with respect to the undamped case in (3.13), causing an amplification of the current harmonics beyond the frequency $1/(|R_L|C)$ of the zero. By increasing C_d, the frequency of the zero increases, and less amplification results for the current harmonics.

By setting $C_d = C$, the zero is completely removed from the transfer function, thus avoiding any high frequency amplification of the current harmonics. As evident from Fig. 3.7a, this is indeed the optimal choice, because it implies the minimization of the $\mathcal{H}_\infty$ norm:

$$\|G_i(s)\|_\infty = \max_{u_s(t)} \frac{\|i_s(t)\|_{\mathrm{rms}}}{\|u_s(t)\|_{\mathrm{rms}}} = \max_{\omega \in \mathbb{R}} |G_i(j\omega)| \tag{3.23}$$

that quantifies how much the Root Mean Square (RMS)-norm (power-norm)

$$\|i_s(t)\|_{\text{rms}} = \lim_{T\to\infty}\sqrt{\frac{1}{2T}\int_{-T}^{+T}|i_s(\tau)|^2 d\tau} \tag{3.24}$$

of the current i_s is affected by the rectified voltage u_s – a smaller value of the $\mathcal{H}_\infty$ norm implies a smaller current rms value (i.e less current oscillations) for a given voltage RMS value [34].
Given $C_d = C$, the Bode plots of Fig. 3.7b and Fig. 3.7c show that the overall system damping increases as the "virtual" resistance R_d decreases. However, this trend is barely noticeable on the DC-link voltage u_c (see Fig. 3.7c), since the second of (3.16) almost resembles the transfer function of an all-pass system (seeFig. 3.7c) when $C = C_d$. In this situation, u_c closely follows the profile of the rectified voltage u_s.

3.3 Generation of the damping power

The small-signal power absorption p_l required to implement the active damping strategy of Sec. 3.2 has to be produced by the motor connected to the inverter. This is assumed as a SynRM with electrical equations

$$\begin{aligned} u_D &= r_s i_D + L_d^{\text{diff}}\frac{di_D}{dt} - \omega_E L_q^{\text{app}} i_Q \\ u_Q &= r_s i_Q + L_q^{\text{diff}}\frac{di_Q}{dt} + \omega_E L_d^{\text{app}} i_D \end{aligned} \tag{3.25}$$

when referred to the dq reference frame, synchronous with the rotor. If the inverter power losses are negligible, the total power p_L drawn from the DC bus is that absorbed by the motor, namely

$$\begin{aligned} p_L = \frac{3}{2}(u_D i_D + u_Q i_Q) &= \frac{3}{2}(U_D I_D + U_Q I_Q) + \cdots \\ \cdots + \frac{3}{2}(U_D i_d + U_Q i_q) &+ \frac{3}{2}(u_d I_D + u_q I_Q) + \frac{3}{2}(u_d i_d + u_q i_q) \end{aligned} \tag{3.26}$$

The first term in the previous expression (3.26) is the average power P_L absorbed at the selected operating point, while the remaining terms are related to the small-signal power absorption p_l. In the following, the latter term in (3.26) will be neglected, being an higher-order infinitesimal when compared to the remaining terms. The motor is assumed to operate at constant speed under Maximum-Torque-Per-Ampere (MTPA) conditions, with current vector components

$$I_D = I\cos\vartheta\,, \quad I_Q = I\sin\vartheta \tag{3.27}$$

where ϑ is the MTPA angle ϑ^{MTPA} and corresponding voltage vector components

$$\begin{aligned} U_D &= r_s I_D - \Omega_E L_q^{\text{app}} I_Q = (r_s\cos\vartheta - \Omega_E L_q^{\text{app}}\sin\vartheta)\, I \\ U_Q &= r_s I_Q + \Omega_E L_d^{\text{app}} I_D = (r_s\sin\vartheta + \Omega_E L_d^{\text{app}}\cos\vartheta)\, I \end{aligned} \tag{3.28}$$

The latter are obtained from (3.25), by assuming a constant electromechanical speed Ω_E.

The small-signal power absorption p_l required to increase the system damping is produced by generating a current perturbation vector directed orthogonally to the current vector (3.27) (i.e. tangent to a constant torque curve in the $i_D - i_Q$ plane) in order to avoid the generation of any torque ripple. Its components are

$$i_d = I_p \sin\vartheta\,, \quad i_q = -I_p \cos\vartheta \tag{3.29}$$

where I_p is the magnitude of the current perturbation vector. They can be generated by applying the voltage perturbations

$$\begin{aligned} u_d &= \left(r_s I_p + L_d^{\text{diff}} \frac{dI_p}{dt}\right) \sin\vartheta + \Omega_E L_q^{\text{app}} I_p \cos\vartheta \\ u_q &= -\left(r_s I_p + L_q^{\text{diff}} \frac{dI_p}{dt}\right) \cos\vartheta + \Omega_E L_d^{\text{app}} I_p \sin\vartheta \end{aligned} \tag{3.30}$$

which are again obtained by resorting to (3.25).

In a conventional Field Oriented Control (FOC) scheme, these current and voltage perturbations can be induced by adding (3.29) and (3.30) to, respectively, the reference and output signals of the current regulators. Note that:

- in a FOC with axes decoupling, the second terms of the right hand side in (3.30) are part of the decoupling voltages, so that only the first terms need to be explicitly injected.

- both (3.29) and (3.30) must be simultaneously injected in the FOC to produce the desired power absorption p_l. In fact, if only (3.29) were present, it would be impossible to generate a perturbation at a frequency beyond the current control bandwidth. Conversely, if only (3.30) were present, then the current controllers would neutralize the voltage harmonics within their control bandwidth (in fact, the voltage injections (3.30) appear as disturbances for the current regulation loops).

The generation of both (3.29) and (3.30) requires to relate the current perturbation magnitude I_p to the desired power absorption p_l.

By combining (3.29)-(3.30) with (3.27)-(3.28) it holds that:

$$
\begin{aligned}
p_l &= \tfrac{3}{2}(U_D i_d + U_Q i_q) + \tfrac{3}{2}(u_d I_D + u_q I_Q) = \cdots \\
&= \tfrac{3}{2}\left(\cancel{r_s \cos\vartheta} - \Omega_E L_q^{\text{app}} \sin\vartheta\right) I I_p \sin\vartheta + \cdots \\
&\quad \cdots - \tfrac{3}{2}\left(\cancel{r_s \sin\vartheta} + \Omega_E L_d^{\text{app}} \cos\vartheta\right) I I_p \cos\vartheta + \cdots \\
&\qquad \cdots + \tfrac{3}{2}\left[\left(\cancel{r_s I_p} + L_d^{\text{diff}} \tfrac{dI_p}{dt}\right) \sin\vartheta + \Omega_E L_q^{\text{app}} I_p \cos\vartheta\right] I \cos\vartheta + \cdots \\
&\qquad\quad \cdots + \tfrac{3}{2}\left[-\left(\cancel{r_s I_p} + L_q^{\text{diff}} \tfrac{dI_p}{dt}\right) \cos\vartheta + \Omega_E L_d^{\text{app}} I_p \sin\vartheta\right] I \sin\vartheta \\
&= -\frac{3}{2}(L_d^{\text{app}} \cos^2\vartheta + L_q^{\text{app}} \sin^2\vartheta)\, \Omega_E I I_p + \frac{3}{2}(L_d^{\text{diff}} - L_q^{\text{diff}}) I \sin\vartheta \cos\vartheta \frac{dI_p}{dt} + \cdots \\
&\quad \cdots + \frac{3}{2}(L_d^{\text{app}} \sin^2\vartheta + L_q^{\text{app}} \cos^2\vartheta)\, \Omega_E I I_p \\
&= -\frac{3}{2}(L_d^{\text{app}} - L_q^{\text{app}})(\cos^2\vartheta - \sin^2\vartheta)\, \Omega_E I I_p + \frac{3}{2}(L_d^{\text{diff}} - L_q^{\text{diff}}) I \sin\vartheta \cos\vartheta \frac{dI_p}{dt} \\
&= \frac{3}{2}(L_d^{\text{diff}} - L_q^{\text{diff}}) \frac{\sin(2\vartheta)}{2} I \frac{dI_p}{dt} - \frac{3}{2}(L_d^{\text{app}} - L_q^{\text{app}}) \cos(2\vartheta) \Omega_E I I_p
\end{aligned}
\tag{3.31}
$$

from which the following transfer function from p_l to I_p can be obtained:

$$
\frac{I_p(s)}{p_l(s)} = \frac{1}{\kappa_1 \sin(2\vartheta)\, s - \kappa_2 \Omega_E \cos(2\vartheta)} \tag{3.32}
$$

with

$$
\kappa_1 = \frac{3(L_d^{\text{diff}} - L_q^{\text{diff}}) I}{4} \qquad \kappa_2 = \frac{3(L_d^{\text{app}} - L_q^{\text{app}}) I}{2} \tag{3.33}
$$

The current perturbations (3.29) can therefore be generated with the feedforward compensators:

$$
\begin{aligned}
\frac{i_d(s)}{p_l(s)} &= \frac{\sin\vartheta}{\kappa_1 \sin(2\vartheta)\, s - \kappa_2 \Omega_E \cos(2\vartheta)} \\
\frac{i_q(s)}{p_l(s)} &= -\frac{\cos\vartheta}{\kappa_1 \sin(2\vartheta)\, s - \kappa_2 \Omega_E \cos(2\vartheta)}
\end{aligned}
\tag{3.34}
$$

As for the voltage perturbations, these can be generated with the feedforward compensators:

$$
\begin{aligned}
\frac{u_d(s)}{p_l(s)} &= \frac{(L_d^{\text{diff}} s + r_s)\, \sin\vartheta}{\kappa_1 \sin(2\vartheta)\, s - \kappa_2 \Omega_E \cos(2\vartheta)} \\
\frac{u_q(s)}{p_l(s)} &= -\frac{(L_q^{\text{diff}} s + r_s)\, \cos\vartheta}{\kappa_1 \sin(2\vartheta)\, s - \kappa_2 \Omega_E \cos(2\vartheta)}
\end{aligned}
\tag{3.35}
$$

provided that the second terms on the right hand side of (3.30) are produced by the axes decoupler. Note that both (3.34) and (3.35) are BIBO stable compensators when the SynRM operates under MTPA conditions, for which the current vector angle is $\vartheta \in (0, \pi/2)$.

CHAPTER 4

Results and discussion

In this chapter the proposed control block diagram and experimental results are presented and discussed.

4.1 Control scheme

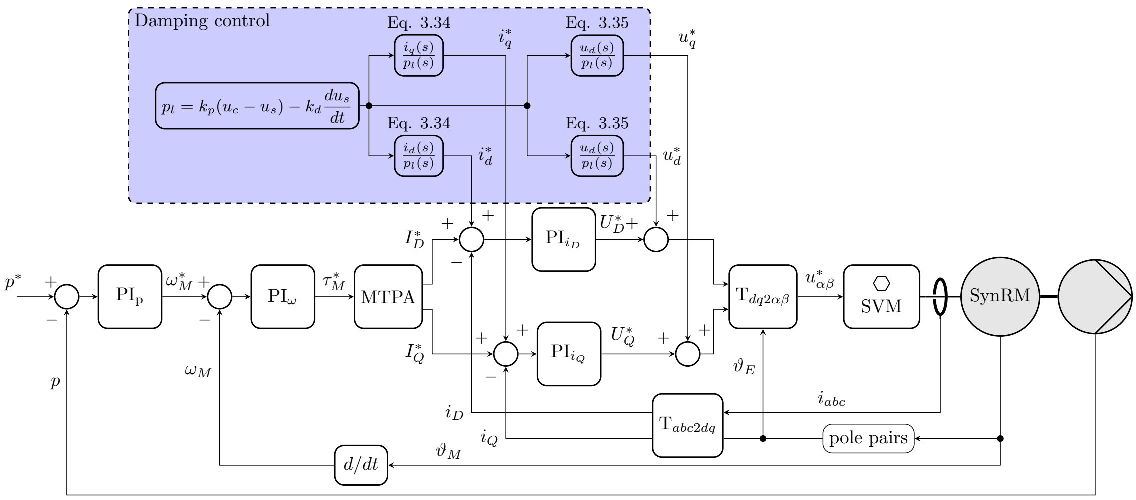


Figure 4.1 Block diagram of the proposed pressure control system for a centrifugal pump. Both the current control loop and the voltage feedforward actions are involved in the production of the commanded damping power p_l.

The block diagram of the proposed capacitorless ASD with pump load is shown in Fig. 4.1. The pressure reference signal p^* is compared with feedback from a sensor in the pump, and the resulting error e_p is processed by a proportional-integral controller (PI_p) to deliver the speed reference ω_M^*. The torque reference is translated into current references by an MTPA algorithm, and given as inputs to two PI-based control loops.

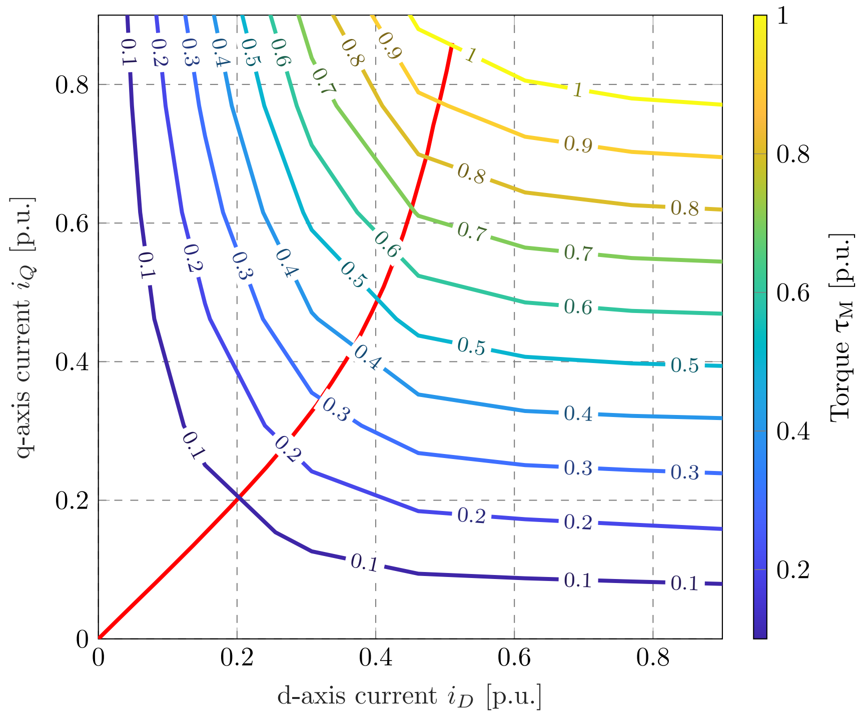


Figure 4.2 Maximum Torque Per Ampere trajectory (MTPA) adopted within the experiments, the plotted quantities are referred to machine rated values listed in Table 4.1.

The MTPA condition, depicted in Fig. 4.2, is computed by resorting to accurate nonlinear magnetic maps of the SynRM, which account also for the magnetic saturation effects.

The two loops are decoupled by adding the opposite of the cross-coupling terms that appear on the right-hand sides of (3.25) to each voltage reference. The current and voltage perturbations generated by the damping power algorithm described in Sect. 3.3 are added to the current and voltage references, namely (i_D^*, i_Q^*) and (u_D^*, u_Q^*) in Fig. 4.1, also taking the apparent and differential inductances depicted in Fig. 4.3 into account.

The voltages references are provided to a three phase voltage inverter, that feeds a SynRM whose nameplate data are reported in Table 4.1. The inverter duty cycles are computed after measuring the DC-Link voltage at each control cycle, so that the actual voltages applied to the SynRM are unaffected by the DC-Link voltage ripple.

4.1.1 Implementation hints

The synch generator, shown in Fig. 4.4, processes the measured DC-Link voltage $u_C(t)$ using an analog-to-digital converter (ADC), a band-pass filter at 300 Hz, and a zero detector. The synch signal is then used to synchronize the signal generator, which produces the estimated rectified voltage u_s and its derivative du_s/dt, both employed in the damping power algorithm described in (3.16).

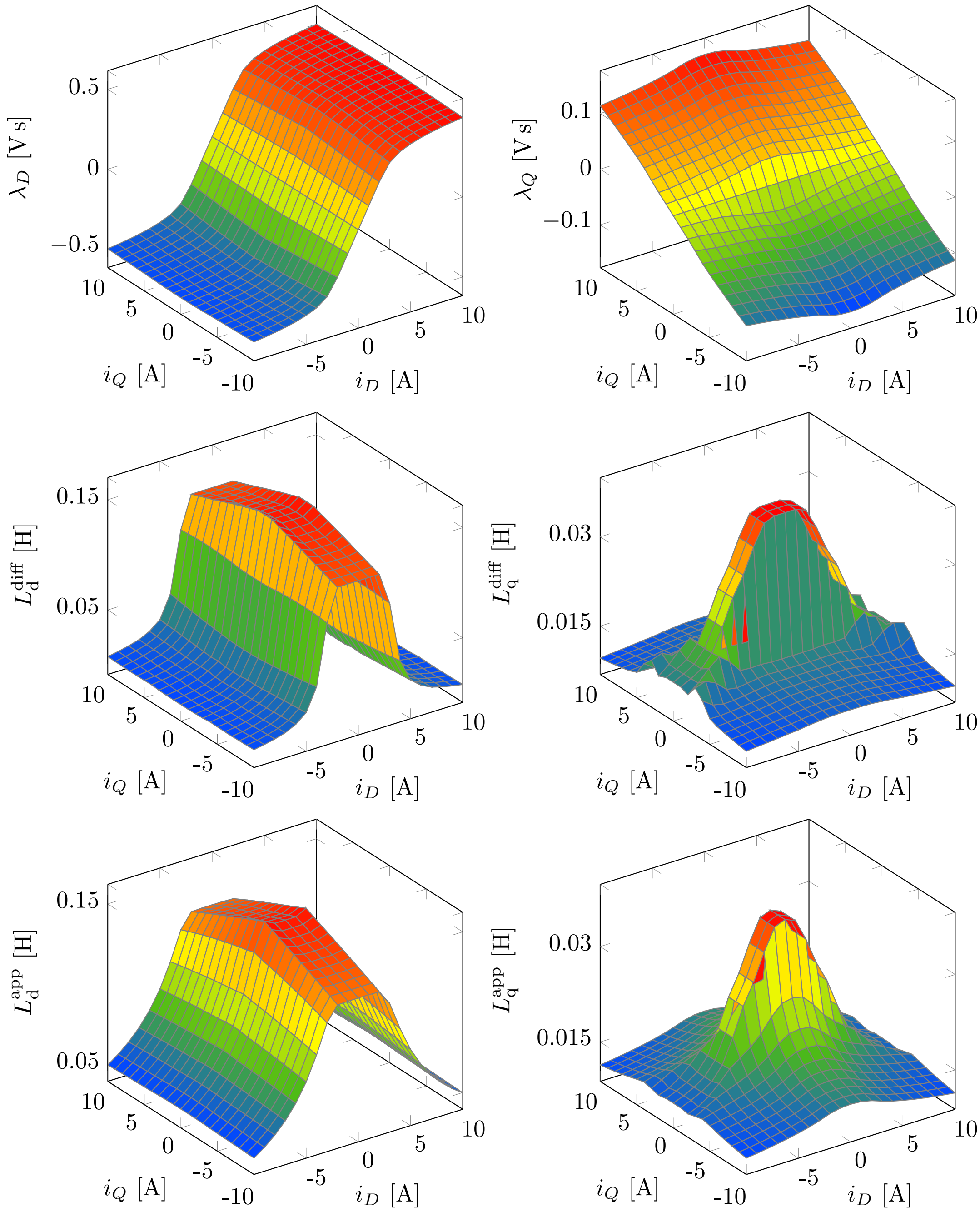


Figure 4.3 Flux linkages $\boldsymbol{\lambda}_{\mathrm{dq}}$, apparent inductances $\mathbf{L}_{\mathrm{dq}}^{\mathrm{app}}$, and differential inductances $\mathbf{L}_{\mathrm{dq}}^{\mathrm{diff}}$ of the SynRM used in the experiments (Table 4.1).

Table 4.1 SynRM nameplate parameters.

Symbol	Parameter	Value
P_n	Nominal mechanical power	917 W
T_n	Nominal torque	2.5 N m
I_n	Nominal current	4.6 $\mathrm{A_{rms}}$
p	Number of pole pairs	1
$\Omega_{m,n}$	Nominal speed	3500 rpm
f_B	Base frequency	58.33 Hz
r_s	Phase resistance	1.437 Ω
$L_{d,0}$	d-axis inductance (at no-load)	148 mH
$L_{q,0}$	q-axis inductance (at no-load)	37 mH

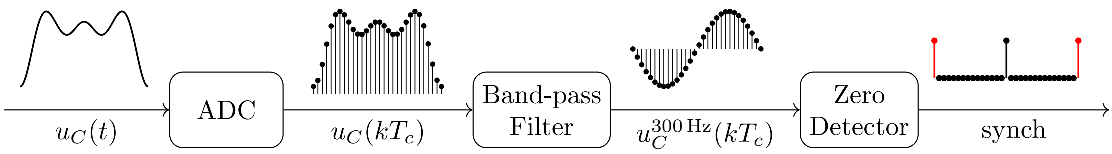


Figure 4.4 Block diagram of the synch generator. The input signal $u_C(t)$ is processed by an ADC, a band-pass filter, and a zero detector to produce the synch signal.

The `synch` signal starts a counter that is reset at the correct zero crossing and reaches its maximum value, computed as $f_{\mathrm{PWM}}/f_{1,u_C} = 16.66\,\mathrm{kHz}/300\,\mathrm{Hz} = 55.55 \approx 56$ steps. The two required signals, u_s and du_s/dt, are approximated using the following polynomials:

$$\begin{cases} u_s \simeq a\ \mathtt{cnt}^2 + b\ \mathtt{cnt} + c, \\ \dfrac{du_s}{dt} \simeq d\ \mathtt{cnt} + e, \end{cases} \tag{4.1}$$

where the coefficients $a = -0.0981$, $b = 5.4502$, $c = -50.4535$, $d = -0.1846$, and $e = 5.0358$, and `cnt` is the counter value, which ranges from 0 to 55. These coefficients are obtained by fitting the polynomials to the more accurate values from (3.2) using the Matlab function `polyfit`. Note that this fitting cannot be directly performed on a microcontroller.

A more refined implementation of the synch generator is shown in Fig. 4.5, where the counter is reset at the *negative-to-positive* zero crossing of the 300 Hz DC-Link voltage component.

The counter also determines the time instants when the estimated rectified voltage u_s and its derivative du_s/dt (red curves), used in the damping power algorithm described in (3.16), are computed.

The synch signal is essential to prevent incorrect feedforward actions. By observing the derivative of u_s in Fig. 4.5 and subsequently p_l in Fig. 4.7, it becomes evident that unsynchronized feedforward actions may result in generating positive damping power when a negative action is required, and vice versa.

The system performance could have been improved by reducing the sampling

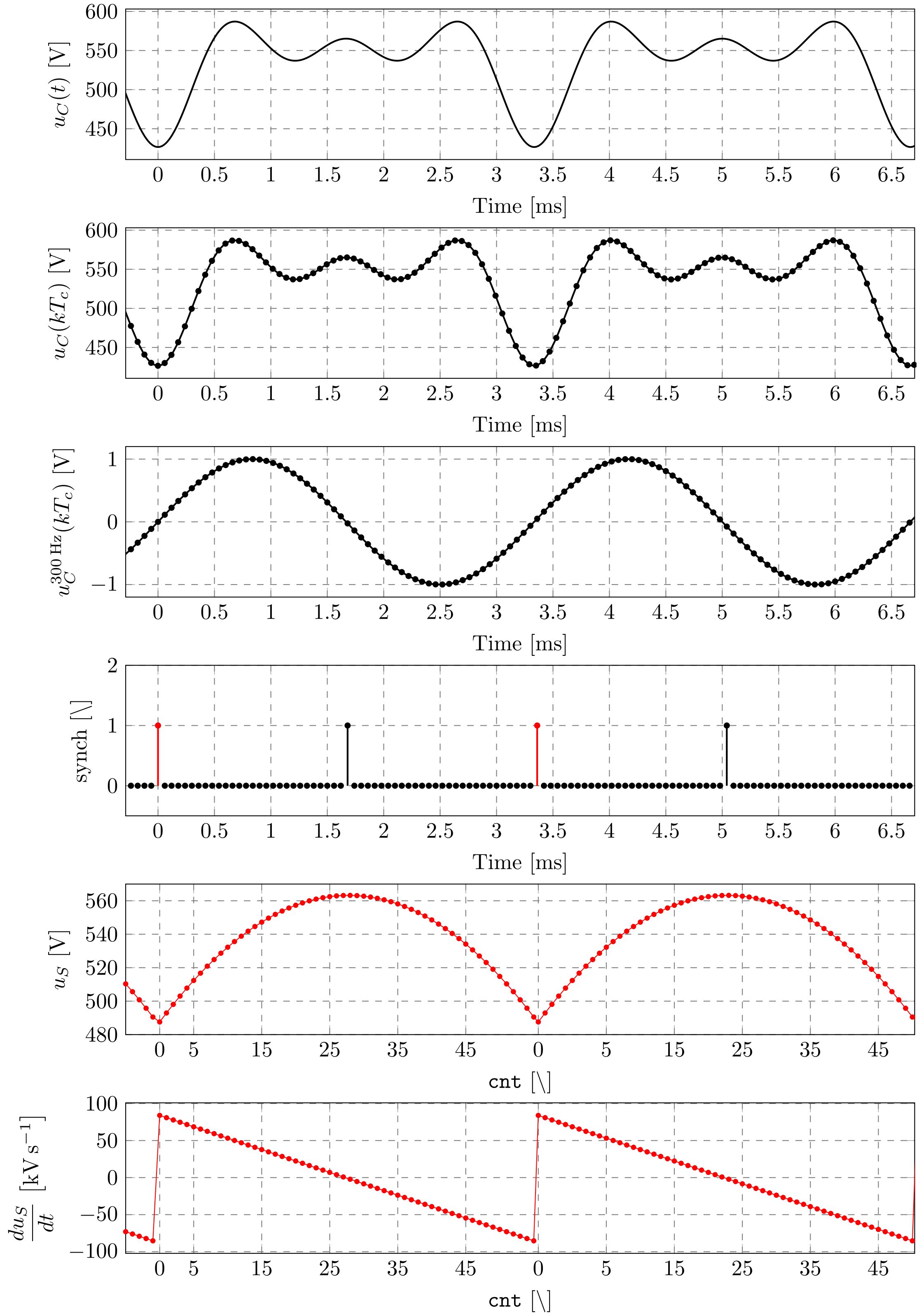
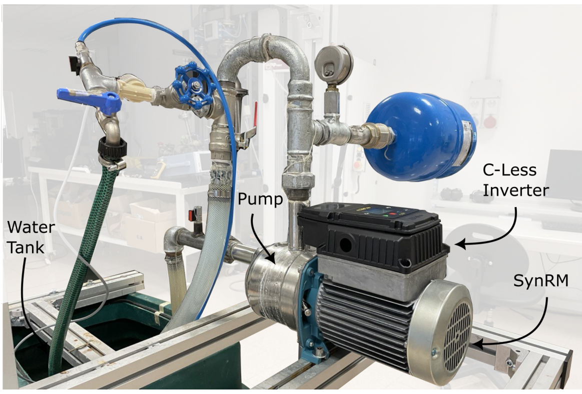


Figure 4.5 Processed and generated waveforms required for damping power generation p_l.

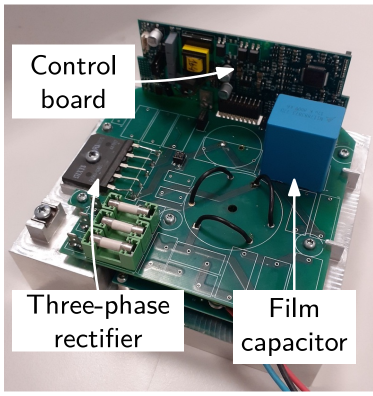
time adopted in the experiments (60 µs), as it coincides with the PWM period. A smaller sampling time would have enhanced zero-crossing detection accuracy and, consequently, feedforward actions. Additionally, it would have sharpened the voltage waveforms, as shown in Fig. 4.8d and Fig. 4.13e, generated by the inverter to meet the required damping power p_l.

However, reducing the sampling time was not feasible due to the limited computational power of the drive used for experimental tests, which were conducted in an industrial facility, specifically at TEKNEIDOS [35], an Italian company engaged in research and development of pumping systems.

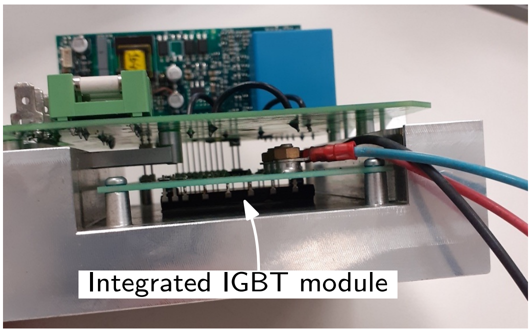
4.2 Test bench



(a) Test bench



(b) Capacitor-less inverter (top)



(c) Capacitor-less inverter (side)

Figure 4.6 Experimental setup available at the TEKNEIDOS facility [35].

The SynRM was coupled with a hydraulic centrifugal pump connected to a water tank via a pipe, as shown in Fig. 4.6a. The three-phase rectifier, the electrolytic capacitor-less inverter, and the control board were placed inside the black case. The control board features an `STM32F303` microcontroller, a common choice for embedded motor control applications.

The ASD was powered by a three-phase voltage generator with a line-to-line voltage of 400 V at 50 Hz. To emulate the grid inductance L_g of a medium-length connection, three inductors of 1.15 mH were connected in series between the power supply and the three-phase rectifier. Alternatively, the ASD could have been directly connected to the grid. However, at TEKNEIDOS, the voltage cabinet was located close to the industrial facility, resulting in a cable inductance too low to trigger resonance phenomena.

The equivalent inductance L was computed using the second expression in (3.4), while the resistance R was estimated by measuring the quantities I_S and U_C and evaluating the first expression in (3.6) at steady state.

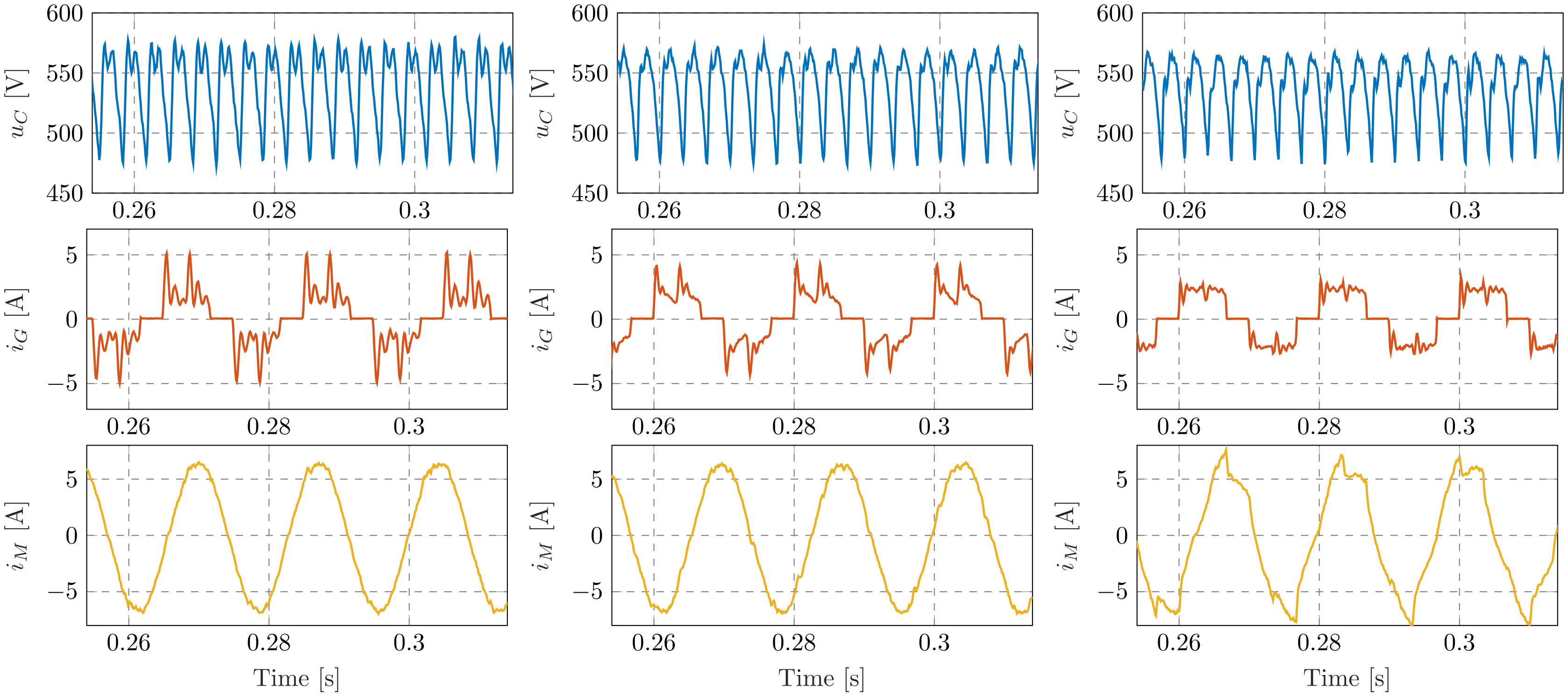
To meet the *capacitor-less* definition, the electrolytic capacitor bank was replaced with a smaller film capacitor of $C = 12\,\mu\mathrm{F}$. The resonance frequency was estimated to be approximately $f_n = 1/(2\pi\sqrt{LC}) = 958\,\mathrm{Hz}$. The equivalent circuit parameters of Fig. 3.3b are summarized in Table 3.1.

4.3 Results

From (3.15), a stability issue is most likely to occur when the capacitor C is small (as in the case of a capacitorless ASD) and the load power P_L is high. Therefore, experiments were conducted at two high load power levels, namely $P_{L,1}$ and $P_{L,2}$, as listed in Table 3.1. These values represent the electrical power absorbed by the motor, as measured by a power meter positioned between the inverter and the motor.

It is important to note that $P_{L,1}$ corresponds to the nominal operating point. However, it differs from the nominal mechanical power of the motor due to the non-unitary efficiency of the SynRM.

As illustrated in Fig. 3.6, the grid parameters in the experimental environment did not give rise to a clear instability issue.



(a) without damping strategy

(b) $k_p = 12$, $k_d = 0$

(c) $k_p = 12$, $k_d = 0.0064$

Figure 4.7 DC link voltage, line current and SynRM current at the working point $P_{L,1}$ (see Table 3.1): (a) without any damping strategy; (b) with proportional damping strategy ($k_p = 12$, $k_d = 0$); (c) with the proposed damping strategy ($k_p = 12$, $k_d = 0.0064$).

Figure 4.7a presents the waveforms of the bus voltage u_C, one of the line currents i_G, and one of the motor phase currents i_M at $P_{L,1}$ under three different power damping conditions: without damping (Fig. 4.7a), with proportional damping only (Fig. 4.7b), and with the full damping strategy described in (3.16) (Fig. 4.7c).

The proportional damping case can be regarded as representative of the conventional method discussed in the literature (e.g., [21]) and is included here for comparison purposes.

The derivative gain in (3.16) was set to $k_d = 0.0064$, following the design rule (3.18) with the average DC-Link voltage $U_C = 530\,\mathrm{V}$.

Regarding the proportional gain, it was selected to enhance the damping of the closed-loop poles, as suggested by the root locus in Fig. 3.6. A high value of k_p is generally recommended to shift the closed-loop poles closer to the negative real axis, leading to more damped locations.

However, an excessively high value would result in an unreasonably large inverter power absorption, as indicated by (3.16). Consequently, a trade-off was necessary. The value chosen for the experimental tests ($k_p = 12$) was determined according to this rationale after a series of trial-and-error experiments, which are documented in Table 4.2.

The improvement brought by the proposed technique is evident when comparing the line current i_G across the three cases. Oscillations in i_G decrease significantly from Fig. 4.7a to Fig. 4.7b and are further mitigated in Fig. 4.7c. Without damping, the current peak reaches 5 A. Employing proportional damping alone reduces this peak to 4 A, and with the addition of derivative action, the waveform becomes nearly flat. As anticipated, the line current also becomes smoother with the proposed approach.

The comparison between Fig. 4.7b and Fig. 4.7c highlights the advantage of incorporating the derivative term in the computation of the damping power (3.17).

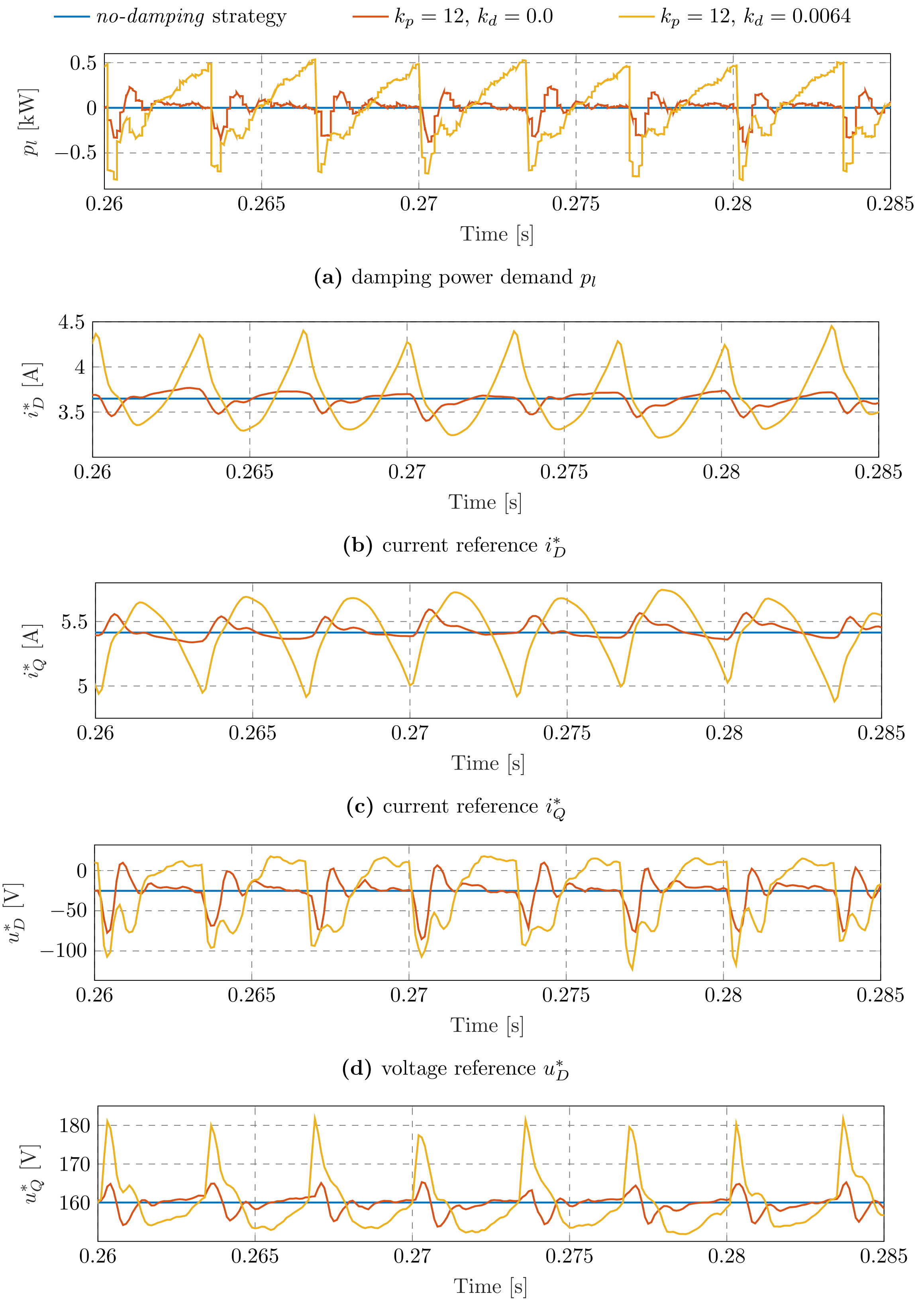
The waveforms of the damping power demand p_l under three conditions (without damping, with conventional proportional action only, and with the proposed technique) are shown in Fig. 4.8a.

It is important to highlight that the sharp waveforms in Fig. 4.8b and Fig. 4.8c would require a high current loop bandwidth if the power command were solely converted into an additional current reference. However, the proposed dual-way damping power generation method, based on (3.34) and (3.35), enables the current loops to operate within their bandwidth limits, while the voltage feedforward actions depicted in Fig. 4.8d and Fig. 4.8e automatically manage the remaining portion of the damping action.

The current control loops were tuned to achieve a bandwidth of approximately 550 Hz.

The torque response at the operating point $P_{L,1}$ under three different conditions (i.e., without damping, with proportional damping only, and with the proposed proportional-derivative damping) is shown in Fig. 4.9a.

The electromagnetic torque generated by a SynRM exhibits a non-linear dependency on the current vector amplitude, resulting in a corresponding increase in current ripple. Additionally, as observed for I_p, its spectrum is characterized by harmonics at multiples of six times the grid frequency ω_g.



(a) damping power demand p_l

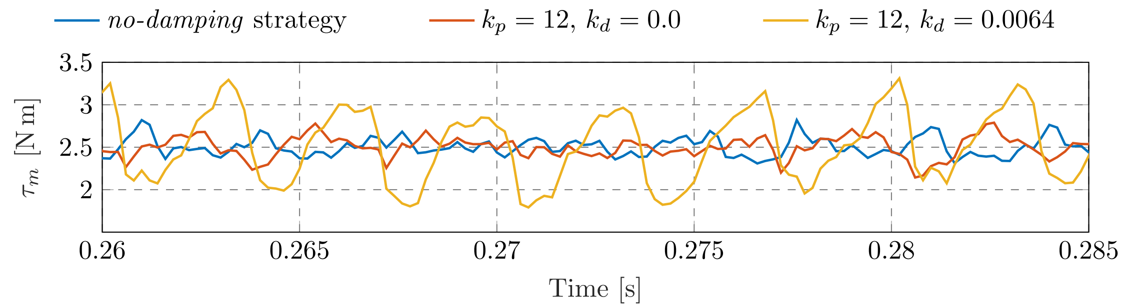
(b) current reference i_D^*

(c) current reference i_Q^*

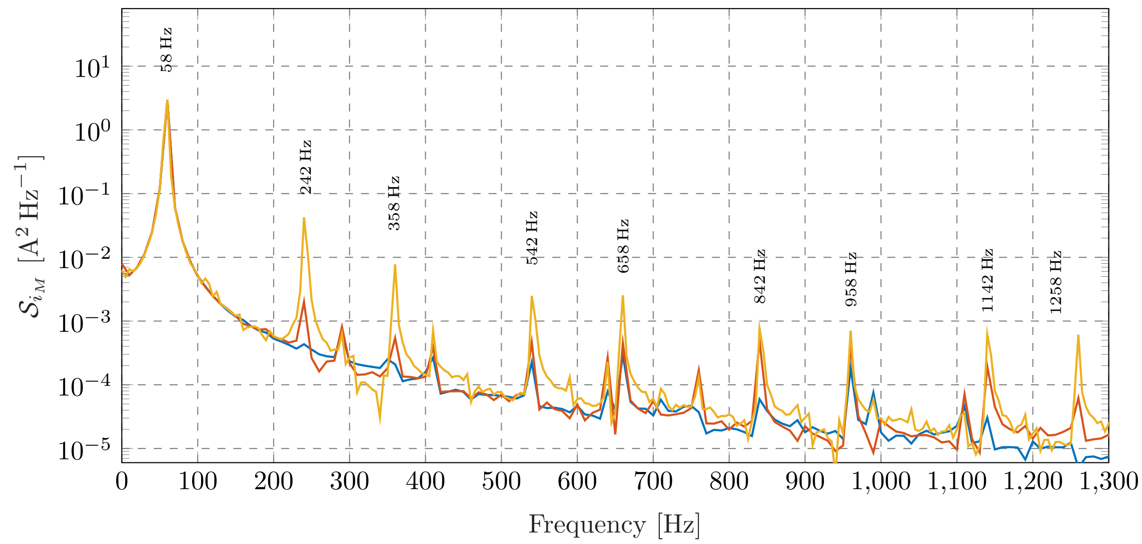
(d) voltage reference u_D^*

(e) voltage reference u_Q^*

Figure 4.8 Relevant SynRM quantities at the working point $P_{L,1}$: (a) damping power demand p_l; (b) d-axis current reference i_D^*; (c) q-axis current reference i_Q^*; (d) d-axis voltage reference u_D^*; (e) q-axis voltage reference u_Q^*;



(a) mechanical torque τ_m



(b) power spectral density of the motor current i_M

Figure 4.9 Relevant SynRM quantities at the working point $P_{L,1}$: (a) mechanical torque τ_m; (b) power spectral density of the motor current i_M.

For a grid frequency of 50 Hz, the fundamental harmonic of the torque ripple occurs at 300 Hz and its higher-order multiples. These frequencies are sufficiently high to render their impact on the speed of a conventional pump inertial load negligible. In fact, the mechanical pole of the machine was estimated to be around 1 Hz, confirming that the torque ripple does not affect the motor mechanical speed.

In the case of low-inertia loads, however, it becomes critical to balance the amplitude of the injected damping power against the associated benefits in harmonic content at the ASD input.

The residual speed ripple has a minimal impact on the drive acoustic noise. This is due to the dominance of noise generated by the pump mechanical components, coupled with the fact that the inverter switching frequency of 16.66 kHz is nearly outside the audible range for humans.

As with any injection-based technique, damping power injection negatively affects the motor current i_M, introducing distortion that becomes increasingly noticeable as the injected power increases. This distortion can be clearly observed in Fig. 4.7c. The power spectral density of the motor phase current under the three damping conditions is presented in Fig. 4.9b.

To further investigate the primary causes of the line current improvement, Fig. 4.10a illustrates the line current waveforms for the proposed full damping strategy, comparing single current/voltage feedforward and double feedforward implementations. For completeness, Fig. 4.10b presents the line current waveforms under the conventional computation of p_l (proportional damping only), while still comparing single current/voltage feedforward and double feedforward cases.

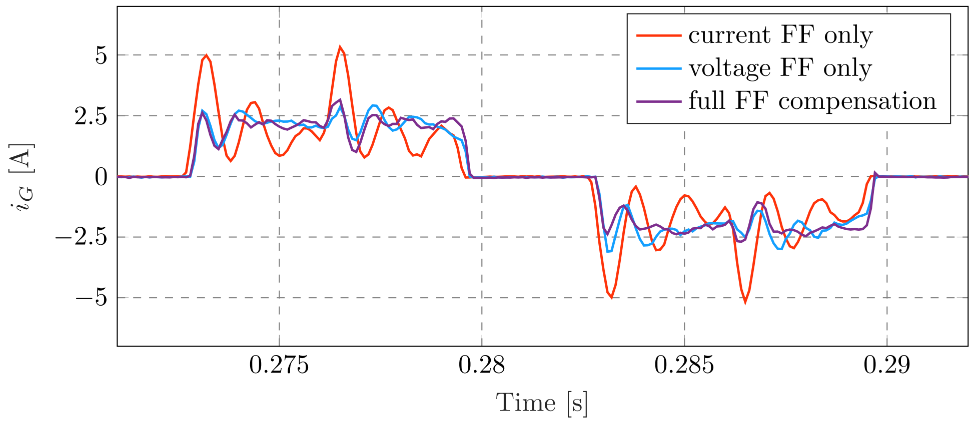
As anticipated, the results demonstrate superior performance of voltage injection compared to current injection, due to the limited bandwidth of the current control loop (refer to the discussion following (3.30)).

In all scenarios, the proposed double feedforward injection yields the best performance, offering significant improvements in line current behavior.

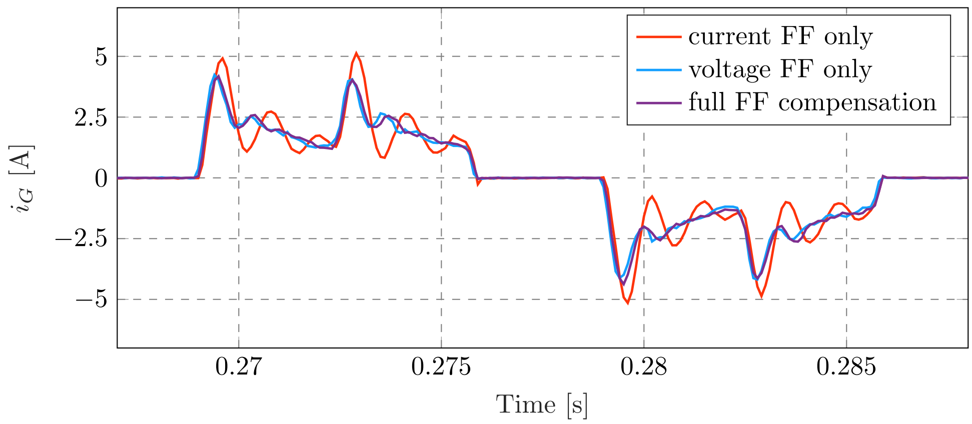
Table 4.2 Input power factor, max damping power, line current rms and THD at working point $P_{L,1}$ (see Table 3.1).

k_p	k_d	λ [/]	p_l^{max} [W]	i_G^{RMS} [A]	THD_{i_G} [%]
0	0	0.83	0	1.96	61.81
5	0	0.87	189	1.92	55.17
10	0	0.89	245	1.85	49.02
12	0	0.90	283	1.85	47.84
15	0	0.91	303	1.85	46.32
20	0	0.91	392	1.83	44.28
5	0.0064	0.95	518	1.75	32.42
10	0.0064	0.95	548	1.75	32.66
12	0.0064	0.95	560	1.77	32.35
15	0.0064	0.95	589	1.74	32.80
20	0.0064	0.95	621	1.78	32.30

The data presented in Table 4.2 show that the Total Harmonic Distortion (THD)



(a) proposed full damping strategy



(b) proportional (i.e. conventional) damping strategy

Figure 4.10 Line current waveform with: (a) proposed full damping strategy, under same experimental conditions of Fig. 4.7c; (b) proportional (i.e. conventional) damping strategy, under same experimental conditions of Fig. 4.7b.

of the line current decreases from 61.81% (without damping) to 47.84%, and finally to 32.35% when using the proposed technique. For completeness, the harmonic spectra of the single-line current i_G under the three different conditions are displayed in Fig. 4.11.

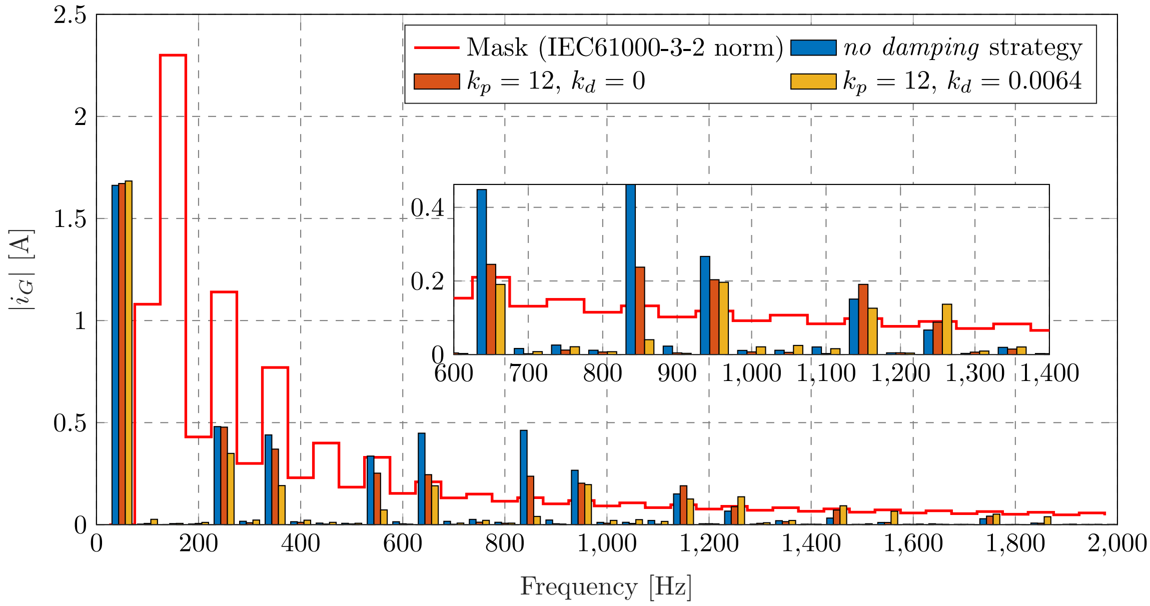


Figure 4.11 Harmonic content of the line current i_G with different values of k_p and k_d, at P_{L1}. Red mask refers to the limits imposed by IEC61000-3-2.

The figure also includes the limits (mask) specified by the IEC61000-3-2 standard. It can be observed that the proposed technique reduces most of the line current harmonics, bringing them very close to or even within the limits imposed by the standard. In particular, many low-frequency harmonics are significantly reduced, which justifies the previously noted improvement in THD.

This, in turn, eliminates the need for expensive Power Factor Correctors (PFCs) and reduces the requirement for additional filtering in domestic applications.

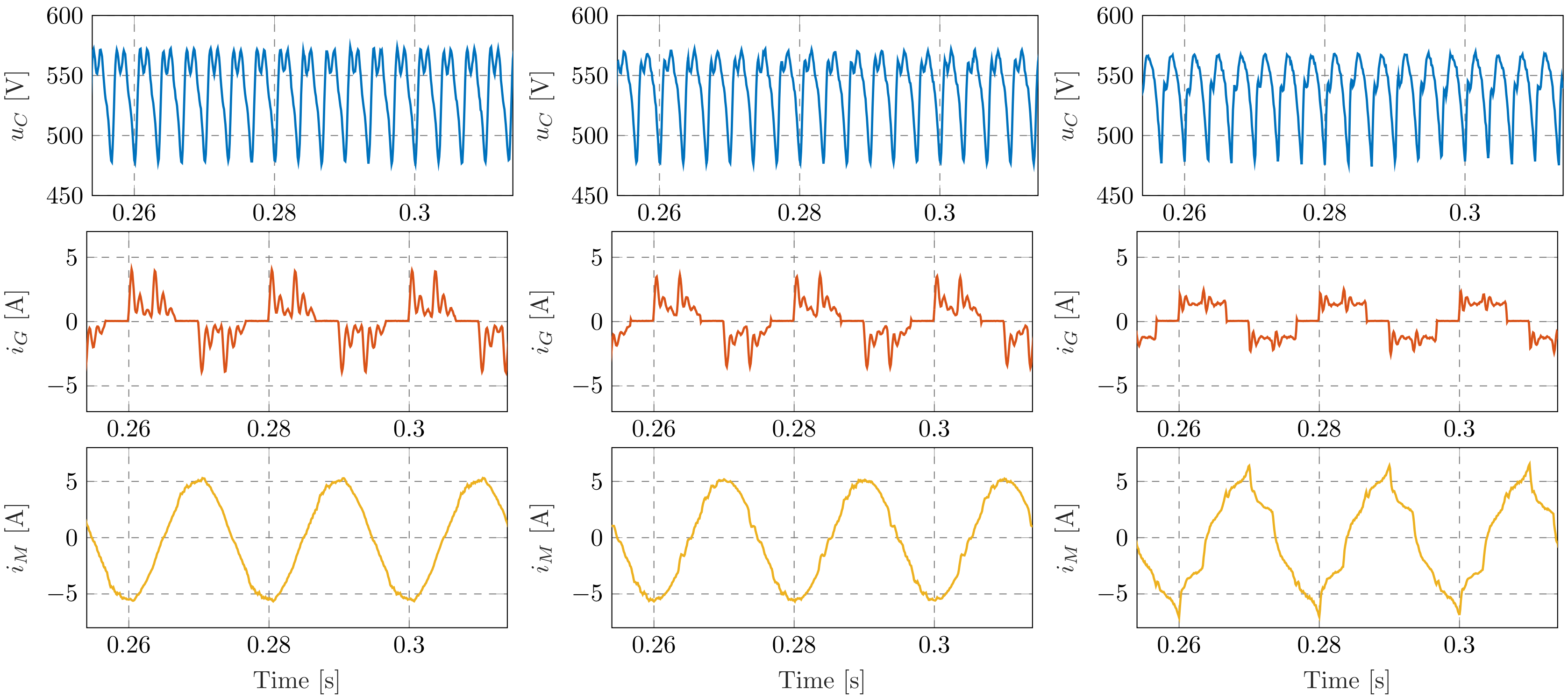
Table 4.2 also reports the input power factor λ [36, p. 885], defined as:

$$\lambda = \frac{P_{in}}{S_{in}} = \frac{\cos(\phi)}{\sqrt{1 + \mathrm{THD}^2}} \tag{4.2}$$

where P_{in} and S_{in} are the active and apparent power, respectively, and ϕ is the phase shift between the first harmonic of the line voltage and current. It is important to note that λ differs from the power factor of the SynRM, which remains unaffected by the proposed technique.

All the values necessary to compute (4.2) were obtained using a second power meter placed between the line inductances L_g and the three-phase rectifier (Fig. 3.1). As a positive side effect of the reduced THD, the proposed technique also significantly improves the power factor, leading to better efficiency in power transfer to the load.

The experiments conducted for the working point $P_{L,1}$ were repeated for the second working point $P_{L,2}$ (Table 3.1).



(a) without damping strategy

(b) $k_p = 12$, $k_d = 0$

(c) $k_p = 12$, $k_d = 0.0064$

Figure 4.12 DC link voltage, line current and SynRM current at the working point $P_{L,2}$ (see Table 3.1): (a) without any damping strategy; (b) with proportional damping strategy ($k_p = 12$, $k_d = 0$); (c) with the proposed damping strategy ($k_p = 12$, $k_d = 0.0064$).

Similarly to the approach taken for the working point $P_{L,1}$, Figures 4.12a, 4.12b, and 4.12c show the bus voltage u_C, line current i_G, and motor phase current i_M for the three cases of interest: without any damping technique, with conventional proportional action, and with the proposed technique. This time, the measurements were taken at the second working point $P_{L,2}$. For simplicity, the algorithm was maintained with the same value of the proportional gain, $k_p = 12$.

The waveform of the damping power demand p_l under three conditions (without any damping, with conventional proportional action only, and with the proposed technique) is shown in Fig. 4.13a, while the feedforward current and voltage references are displayed in Fig. 4.13b, Fig. 4.13c, and Fig. 4.13d, Fig. 4.13e, respectively.

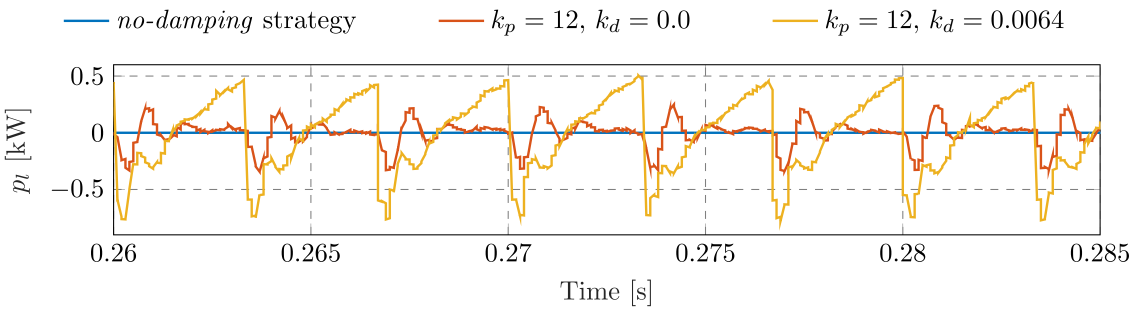
The torque response at the working point $P_{L,2}$ for three different operating conditions (i.e., without damping, with proportional damping only, and with the proportional-derivative damping proposed in this thesis) is presented in Fig. 4.14a, whereas the power spectral density of the motor phase current for the three different power damping conditions is shown in Fig. 4.14b.

For the sake of completeness, the harmonic spectra of a single-line current i_G at the working point $P_{L,2}$ under the three different conditions are shown in Fig. 4.15. As in the case of the working point $P_{L,1}$, the proposed technique brings most of the line current harmonics very close to, if not within, the limits imposed by the standard.

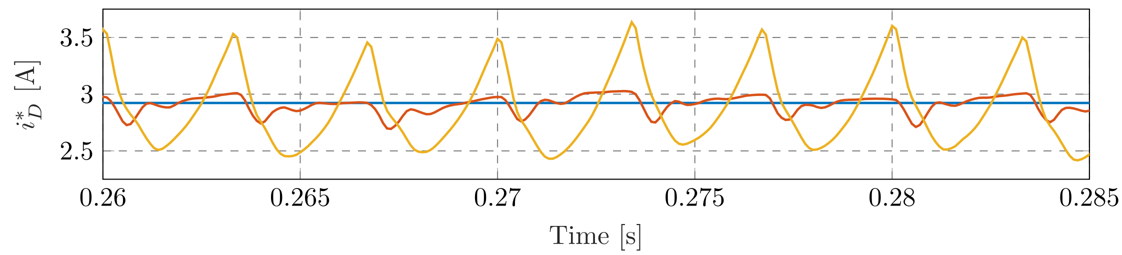
Table 4.3 Input power factor, max damping power, line current rms and THD at working point $P_{L,2}$ (see Table 3.1).

k_p	k_d	λ [/]	p_l^{max} [W]	i_G^{RMS} [A]	THD_{i_G} [%]
0	0	0.77	0	1.42	80.40
5	0	0.802	205	1.40	75.80
10	0	0.822	218	1.36	68.52
12	0	0.827	268	1.33	67.09
15	0	0.837	311	1.33	63.97
20	0	0.85	335	1.29	60.82
5	0.0064	0.935	489	1.19	36.12
10	0.0064	0.935	550	1.21	37.12
12	0.0064	0.935	545	1.19	36.65
15	0.0064	0.936	553	1.20	35.58
20	0.0064	0.934	560	1.21	37.42

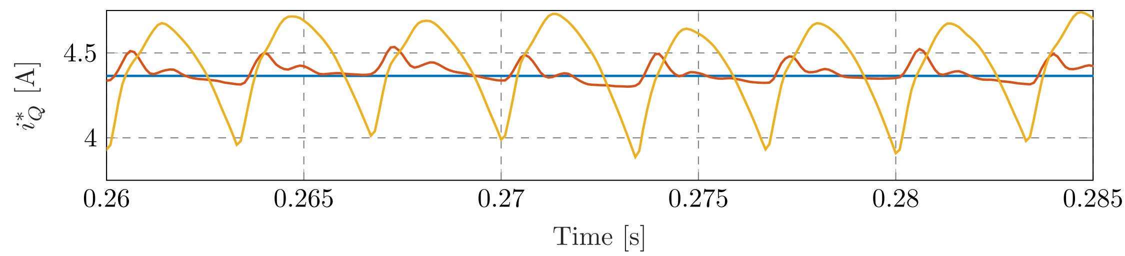
The data regarding the power factor and THD are reported in Table 4.3. Even with the use of $K_p = 12$ alone, there is a significant improvement in both the input power factor λ and the line current THD, which decreases from 80.40% in the case of no damping action to 67.09%. With the proposed technique, the THD decreases further to 36.65%, while the power factor improves from 0.832 to 0.947.



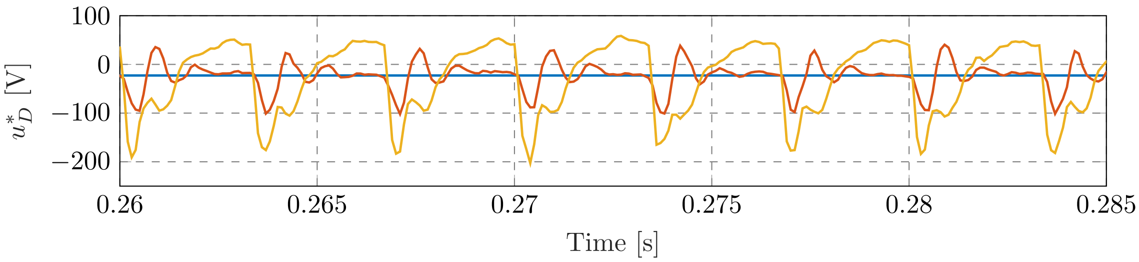
(a) damping power demand p_l



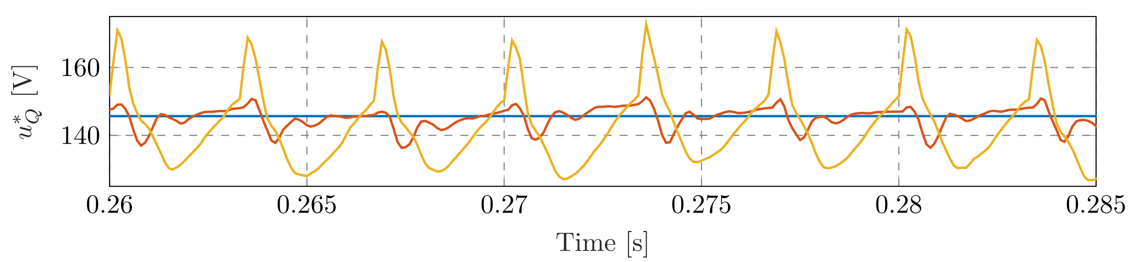
(b) current reference i_D^*



(c) current reference i_Q^*

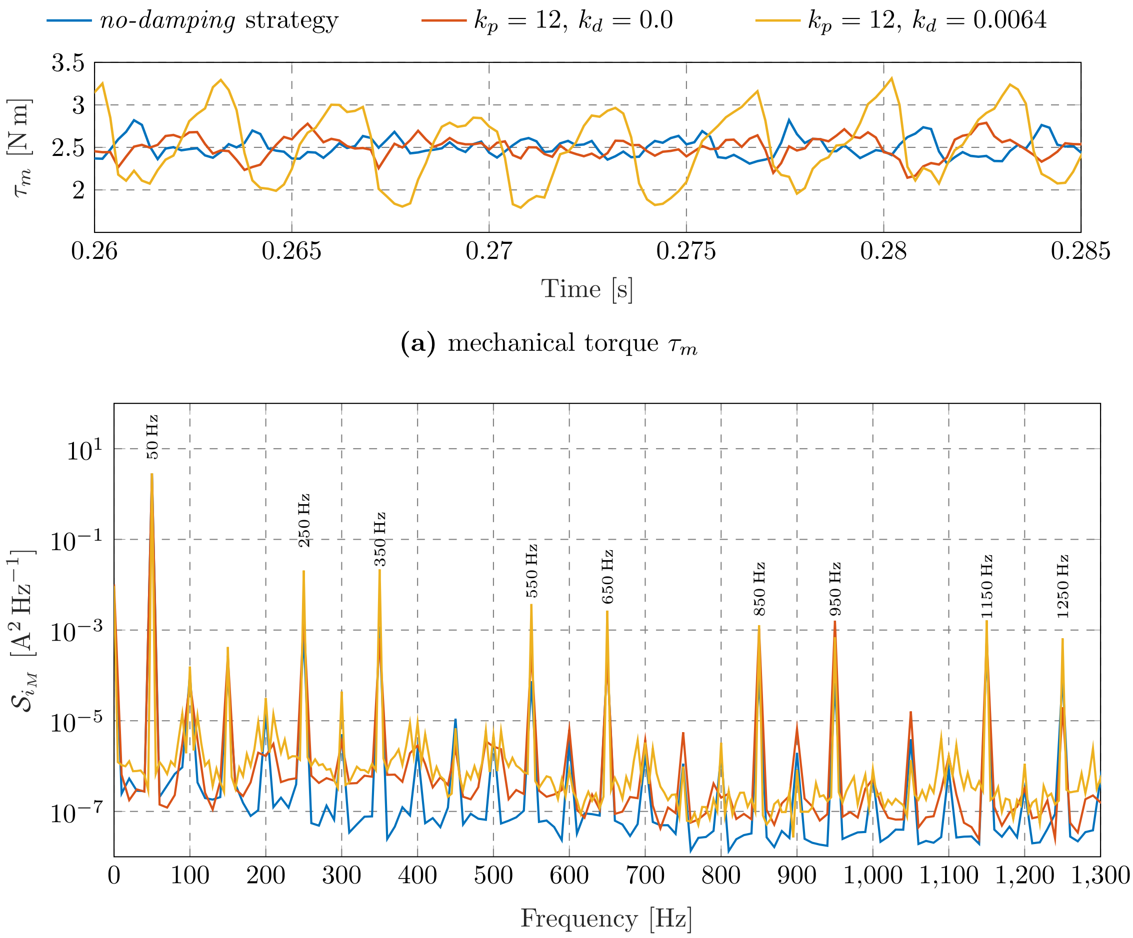


(d) voltage reference u_D^*



(e) voltage reference u_Q^*

Figure 4.13 Relevant SynRM quantities at the working point $P_{L,2}$: (a) damping power demand p_l; (b) d-axis current reference i_D^*; (c) q-axis current reference i_Q^*; (d) d-axis voltage reference u_D^*; (e) q-axis voltage reference u_Q^*;



(a) mechanical torque τ_m

(b) power spectral density of the motor current i_M

Figure 4.14 Relevant SynRM quantities at the working point $P_{L,2}$: (a) mechanical torque τ_m; (b) power spectral density of the motor current i_M.

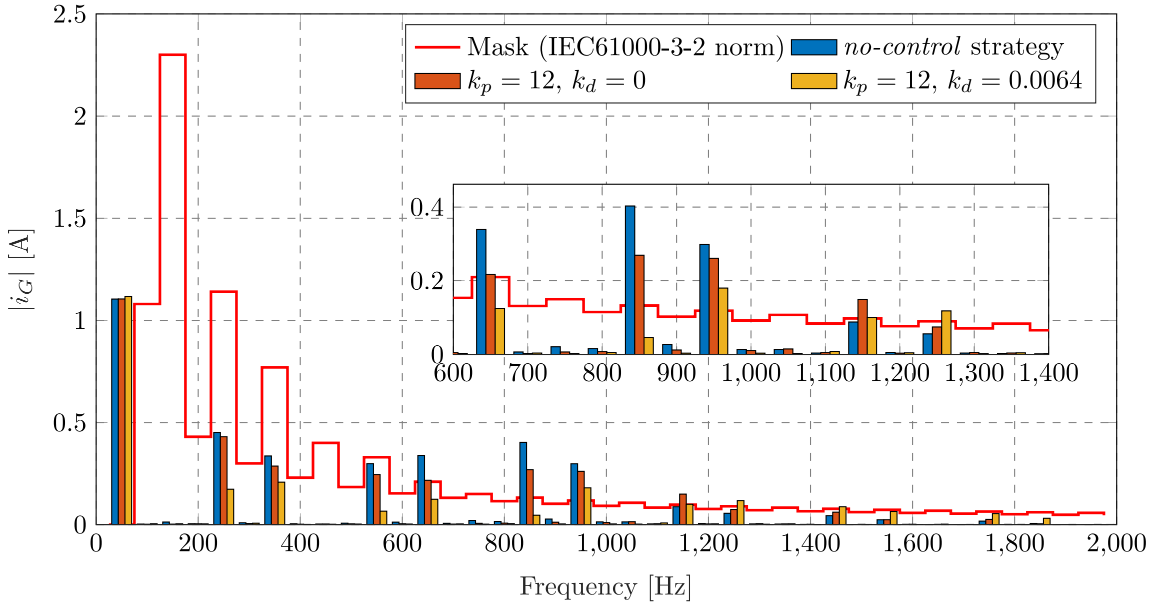


Figure 4.15 Harmonic content of the line current i_G with different values of k_p and k_d, at $P_{L,2}$. Red mask refers to the limits imposed by IEC61000-3-2.

CHAPTER 5

First part conclusions

The first part of my PhD research, conducted at the TEKNEIDOS premises, focused on the development of a capacitorless SynRM drive.

To summarize, this first part provides analytical insights into the complex interactions that arise when the conventional decoupling action of the DC-link capacitor bank is absent. Transfer functions were derived between DC-link disturbances and both line currents and voltages, enabling the computation of an optimal damping power reference. Furthermore, a novel dual-way PD-based method is proposed to enable the three-phase SynRM to absorb commanded power effectively. The proposed damping technique was experimentally validated, yielding results that met expectations: specifically, the elimination of oscillations, a significant reduction in line current harmonic content to 32%, and a noticeable improvement in the input power factor, reaching 0.95. Another important feature is that the proposed damping technique was implemented in an industrial environment by modifying the firmware of a commercial ASD. This is a key aspect, as it demonstrates the feasibility of the proposed technique in real-world applications.

The proposed capacitorless SynRM drive exhibits a harmonic content in the line current that is very close to, and sometimes within, the limits of the IEC 61000-3-2 standard. This eliminates the need for expensive power factor controllers and reduces reliance on line current filtering.

Future work should focus on two key aspects. First, the proposed damping technique should be further investigated using more advanced controllers than simple PIs, whose performance is limited by bandwidth constraints. For example, a Model Predictive Controller (MPC) could offer superior performance, potentially incorporating constraints on bus voltage or grid current into the control strategy.

Second, the proposed damping technique could be extended to other types of drives, such as single-phase drives, where reducing the DC-link capacitor bank is even more critical.

Part II

Adaptive Kalman Filters

CHAPTER 6

Introduction

In 1960, R.E. Kalman introduced his renowned stochastic estimator, demonstrating excellent capability in extracting desired hidden quantities from noisy measurements, such as dynamic states or model parameters. Sorenson describes the historical development of the Kalman Filter in [37], tracing its evolution from Gauss' ideas. Although the first applications to electrical drives emerged in the 1980s and 1990s, they were initially limited to simulations because of the heavy computational load, which presented an insurmountable obstacle at that time.

In recent years, KF-based algorithms have seen a resurgence, propelled by the availability of extremely fast and inexpensive processors. This resurgence is also driven by the introduction of new motors with complex, nonlinear magnetic structures that are ill-suited for linear state-space models.

Kalman filters are general observers that can be adapted to any system. One of the most common application in the field of electrical drives is the estimation of the rotor position of a synchronous machine, which is also called sensorless control. Unfortunately, the synchronous machine model, whether it is expressed in the synchronous reference or stationary reference is nonlinear when the machine speed is considered as a state variable, as in the case of a sensorless algorithm. Therefore the original version of KF is not directly applicable. The two main solutions to deal with non-linear systems are the Extended Kalman Filter (EKF) developed by NASA in 1966 [38] and the Unscented Kalman Filter (UKF) introduced by *Uhlmann et al.* in 1995 [39].

6.1 Nonlinear Kalman Filters

The most representative local filter, introduced by Schmidt [38], is named the Extended Kalman Filter (EKF). It finds widespread use in various electric drive applications due to its straightforward mathematical description [40], [41], [42]. The

EKF addresses nonlinear state update and output functions through a recursive linearization procedure around the previous state estimate [43], [44], [45]. These nonlinear functions are linearized through their derivatives with respect to the filter state. Once the so-called state and output Jacobian are obtained, the EKF equations coincide with those of the linear KF algorithm. The EKF is suitable for systems in which the model is known and not overly complicated; otherwise, the Jacobian calculation might be too difficult and prone to derivation errors.

In 1995, Uhlmann and Julier [39] argued that approximating a nonlinear function output through just one point, as the EKF does, is not always a reasonable approach. Alternatively, they proposed subjecting a few state points to the nonlinear function and taking the mean and covariance from them as estimation results. They named this procedure the *unscented*[1] transform, asserting that the Unscented Kalman Filter (UKF) is a more suitable observer for nonlinear control systems compared to the EKF [47]. These newer derivative-less filters eliminate the need to compute partial derivatives of system equations [43], [48]. Instead, they employ sigma-points, deterministically selected representative vector points for assumed Gaussian multivariate density distributions. These filters are collectively known as sigma-point Kalman filters (SPKF) [49], [50].

In the literature, most comparisons are limited to Permanent Magnet Synchronous Motors (PMSMs) and induction motors. Quite often they are based on simulations only and with some disagreement in the conclusions. Reference is often made to one filter over another, but no one clearly states which is the best choice.

In [51], a comparison between UKF and EKF is made by simulation on a PMSM with constant electrical parameters.

In [52], the author describes how the UKF behaves when matched with different mechanical models. The simulations focus on possible variations in the PM flux linkage, and the motor is a simple surface-mounted PMSM.

A comparison of EKF and UKF applied to an axial flux PMSM is presented in [53]. The conclusions are that EKF performs slightly better during motor start-up, while UKF is better at tracking the speed. However, the results were obtained only on simulation bases and the motor is represented by a simplified linear magnetic model.

Some experimental results are described in [54], but they are related only to a UKF implementation for the control of a non-saturated PMSM. The focus is not on a comparison with EKF, but on the advantage of adding an extra state variable that accounts for eventual inaccuracies in motor parameters.

With regard to existing comparisons for IMs, it is worth mentioning the work [55], in which the authors simulate the behavior of a UKF and EKF for sensorless direct field-oriented control of an IM.

In [56], an IM direct torque control algorithm is implemented, once again in a simulation environment only. The magnetic circuit IM is simplified to a linear

[1]The term *unscented* was arbitrarily chosen by Uhlmann, who explained that, initially, he referred to the new filter simply as "the new filter". When a more specific name was needed, it was briefly called the "Uhlmann filter" but he could not use that. One evening, while others were away, Uhlmann noticed the word *unscented* on a deodorant label and found it to be a fitting technical term. Although initially seen as absurd by his colleagues, he argued that people simply accept technical terms without questioning their origins, much like the term *tree*.[46]

one and the process and measurement noise matrices, $\boldsymbol{Q}$ and $\boldsymbol{R}$ respectively, are fixed by optimizing the mean square error of the estimated IM speed. In line with expectations, as the IM is described by a relatively simple model, the findings are that the EKF gives better state estimates than the UKF.

A notable work is reported in [57], which describes an experimental comparison between the UKF and EKF for an IM sensorless drive. A detailed IM observability analysis is carried out that considers several key aspects, such as model de-tuning, computational burden, and performance evaluation. Matrices $\boldsymbol{Q}$ and $\boldsymbol{R}$ are obtained by trial-and-error and taken to be equal for the two filters. It is a fair example of a comparison methodology, even though, once more it was applied to a simplified IM model with constant parameters, thus neglecting any magnetic saturation.

The aforementioned literature insight highlights that there is still a lack of indication about whether the UKF or EKF is best suited for SynRM sensorless applications. The UKF was specifically developed to address the limitations of the EKF in handling highly nonlinear systems. This raised the question of whether its application to nonlinear motors, such as the SynRM, would truly be advantageous. Although comparison studies exist in the literature for IM and PMSM drives, which feature relatively linear magnetic models, no studies have compared the two filters for the highly nonlinear SynRM machine.

Most of the cited works [51], [53], [54], [57] use fixed $\boldsymbol{Q}$ and $\boldsymbol{R}$ matrices, which are manually tuned by a trial-and-error procedure. This approach is definitely time-consuming and it requires a deep system knowledge and experience from the user. In addition, the fixed $\boldsymbol{Q}$ and $\boldsymbol{R}$ matrices are not able to adapt to the system changes, such as the motor parameters variations, the load torque changes, the temperature variations, etc. Therefore, it is mandatory to find solutions that tune or help to tune the filter. The collections of these algorithms are called *adaptive* Kalman Filters.

6.2 Adaptive filtering

Nowadays, while implementation complexity no longer represents a serious obstacle, achieving a proper filter tuning still remains a challenging and not entirely solved problem [58]. In particular, it is well known that the performance of the filter is closely related to how well the covariance matrices of the process and measurement noises in the estimation model match their actual values. Poor a priori knowledge of the noise statistics may seriously degrade the quality of the estimation or, even worse, cause the filter divergence [59].

In many practical applications, this information is difficult to obtain, especially regarding process noise, which is related to the quality of the model used for the estimation and the approximations introduced during its derivation, including errors arising from system linearization around the current estimated state.

Several methods have been proposed in literature to identify reasonable values to use for the noise covariance matrices. Manual tuning is the most straightforward approach, but it typically involves a time-consuming and somehow exhausting trial-and-error procedure [60], even when the effect of the design parameters on the filter performance has been partially clarified [58], [61].

Design guidelines can be provided to obtain an initial guess for the EKF covariance matrices, by starting from either the nominal [62] or normalized [63] values of the drive parameters. However, on-the-field refinements based on the well-known Bartlett's test are still required to achieve an optimal tuning for the specific application.

Alternative approaches based on offline optimization techniques have been proposed to further automate the tuning process. Meta-heuristic tools such as genetic [64], evolutionary [65] and swarm algorithms [66] are typically preferred over conventional optimization methods (e.g. gradient descent, Newton-Rhapson, etc.) because they do not require differentiation of the cost function [65]. Unfortunately, offline optimization remains as time-consuming as manual tuning, which is a significant drawback, especially considering that the filter should be re-tuned whenever the operating conditions change.

More effective tuning schemes involve some sort of online adaptation mechanism for the poorly known and potentially varying noise covariance matrices, which are regarded as extra time-variant parameters to be estimated alongside with the system state. The solutions proposed in literature for implementing such Adaptive Kalman Filter (AKF) algorithms can be broadly classified into three categories, namely multiple-model, covariance scaling, and innovation-based adaptive estimation methods [67], [68].

In multiple-model adaptive estimation (MMAE) methods, the state estimate is obtained as the weighted sum of the estimates produced by several individual filters running in parallel, each with slightly different model or noise assumptions that can be representative of a particular operating condition [69], [70]. An obvious drawback of such strategy is the high computational load required for the real-time concurrent execution of the filter bank [71], [72].

In contrast, covariance scaling methods are less demanding from a computational perspective. They consists of applying a forgetting (fading) factor to the state covariance matrix, in order to compensate for errors induced by an inaccurate model or poor filter tuning [73]. The approach is effective in preventing filtering divergence, but performances only improve during transients and in highly dynamic operating conditions [74]. Furthermore, the method is still dependent on the initial choice of the covariance matrices, and a poor selection may lead to suboptimal performance [75].

Innovation-based Adaptive Estimation (IAE) methods operate instead by adjusting the noise covariance matrices according to the observed innovation sequence. Correlation, covariance matching and Maximum-Likelihood Estimation (MLE) methods fall within this family [76]. In correlation methods the underlying idea is to relate the system parameters to the observed autocorrelation of either the system output or innovation sequence, and then solving for the unknown covariance matrices. A major drawback of such methods is that they are restricted to linear time-invariant systems.

Covariance matching techniques work instead by making the covariance of the innovation sequence consistent with its theoretical value. In case of any discrepancy, the noise covariance matrices are adjusted to close the gap. In [76] an output matrix with rank equal to the number of states is required to derive the estimate of the process noise covariance matrix. This limitation is overcome in [69], [77] by resorting to predicted or smoothing state residuals, but the stability of the resulting adaptive filter is not guaranteed.

The methods based on the MLE principle reformulate the adaptation problem in terms of maximization of the joint probability density function (likelihood function) of the system state and the unknown parameters [70]. They are those providing the best performance under dynamic state variations, and have also the advantage of providing full estimation of the unknown covariance matrices [78], [79]. The major challenge with adaptation methods is selecting the appropriate window size for the innovation sequence. This window length acts as a tuning parameter, as it depends on the specific application and the desired balance between tracking accuracy and adaptation speed. A shorter window may adapt faster but with reduced accuracy, while a longer window may improve accuracy but result in slower adaptation.

Another interesting approach based on innovation is the Master-Slave Kalman Filter (MSKF) algorithm, which is a combination of two simultaneously working Kalman filters, in which the former (master) estimates the quantities of interest, while the latter (slave) estimates the noise covariance matrices of the master. The primary advantage is that the problem of tuning is shifted from $\boldsymbol{Q}_m$ (subscript m refer to the master, subscript s to the slave) to the covariance matrix $\boldsymbol{Q}_s$ of the slave KF, which proves to be a considerably easier and more relaxed task.

The concept of the Master-Slave (MS) algorithm was initially introduced as an extension of the UKF [80]. Subsequently, the same research group introduced a simplified version of the algorithm, using a linear KF as a secondary (slave) filter [81], [82], [83].

6.3 Research objectives and structure

This part of the thesis addresses the comparison between the two most common nonlinear Kalman filters, EKF and UKF, in the context of SynRM sensorless applications. Additionally, it investigates two selected adaptive filtering algorithms, the Maximum-Likelihood Estimation Kalman Filter (MLEKF) and MSKF algorithms, for SynRM sensorless control.

Before introducing the Kalman filters, the state update and output functions common to all KF versions are presented in Ch. 7. The chapter includes the continuous-time and discrete-time mathematical models of the SynRM in Sect. 7.1 and provides the observability analysis in Sect. 7.2.

The following chapters provide a theoretical background for each filter, followed by experimental results. In particular, Ch. 8 addresses whether the UKF or EKF is better suited for SynRM sensorless applications. Ch. 9 introduces the MLE-based KF method for sensorless control of the SynRM drive, and Ch. 10 presents the MSKF method for sensorless control of the SynRM drive.

The conclusions of the second part of the thesis are discussed in Ch. 11.

CHAPTER 7

System model foundations

In this chapter, the theoretical foundation and derivation of both continuous-time and discrete-time system models are discussed. These models serve as the basis for all Kalman filter implementations in subsequent chapters.

The chapter comprises two main sections. Sect. 7.1 details the discrete-time model employed in the Kalman filter prediction step. Subsequently, Sect. 7.2 examines the observability of the adopted system.

7.1 Synchronous Machine Model

This section introduces the electrical and mechanical models of a general synchronous machine. The model is formulated in the synchronous reference frame and subsequently converted into the stationary reference frame. For a pure reluctance machine, such as the SynRM, the flux produced by the permanent magnet is absent. Consequently, the SynRM model is obtained by nullifying the terms in the equations related to the permanent magnet flux.

7.1.1 Electrical Model

The electrical model is crucial for deriving the current prediction equation used in the Kalman filter prediction step.

The synchronous machine's continuous-time linear dynamical system in the synchronous reference frame is considered. For simplicity, the explicit time dependency $x(t)$ is omitted in the subsequent equations.

$$\boldsymbol{u}_{DQ} = \boldsymbol{R}_s \boldsymbol{i}_{DQ} + \frac{\mathrm{d}\boldsymbol{\lambda}_{DQ}}{\mathrm{d}t} + \omega_E \boldsymbol{Z} \boldsymbol{\lambda}_{DQ} \tag{7.1}$$

where:

- $\boldsymbol{u}_{DQ} = [u_D,\ u_Q]^\top$ are the machine voltages in the synchronous reference.
- $\boldsymbol{i}_{DQ} = [i_D,\ i_Q]^\top$ are the machine currents in the synchronous reference.
- $\boldsymbol{R}_s = \begin{bmatrix} r_s & 0 \\ 0 & r_s \end{bmatrix}$, r_s is the stator resistance.
- $\boldsymbol{\lambda}_{DQ} = [\lambda_D,\ \lambda_Q]^\top$ are the machine flux linkages in the synchronous reference.
- ω_E is the electrical speed of the machine, $\omega_E = p\omega_M$, ω_M is the mechanical speed and p are the pole pairs.
- $\boldsymbol{Z} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.

It is possible to explicit the fluxes as function of inductances and currents, as a matter of facts (7.1) can be rewritten as

$$\boldsymbol{u}_{DQ} = \boldsymbol{R}_s \boldsymbol{i}_{DQ} + \boldsymbol{L}_{dq}^{\text{diff}} \frac{\mathrm{d}\boldsymbol{i}_{DQ}}{\mathrm{d}t} + \omega_E \boldsymbol{Z} \boldsymbol{L}_{dq}^{\text{app}} \boldsymbol{i}_{DQ} + \omega_E \boldsymbol{Z} \boldsymbol{\lambda}_{DQ}^{\text{mg}} \tag{7.2}$$

where:

- $\boldsymbol{L}_{dq}^{\text{diff}} = \begin{bmatrix} L_d^{\text{diff}} & L_{dq}^{\text{diff}} \\ L_{dq}^{\text{diff}} & L_q^{\text{diff}} \end{bmatrix} \triangleq \begin{bmatrix} \frac{\partial \lambda_D}{\partial i_D} & \frac{\partial \lambda_D}{\partial i_Q} \\ \frac{\partial \lambda_Q}{\partial i_D} & \frac{\partial \lambda_D}{\partial i_Q} \end{bmatrix}$, differential inductance matrix.
- $\boldsymbol{\lambda}_D^{\text{mg}} = [\Lambda_D^{\text{mg}}\ 0]^\top$, where Λ_D^{mg} is the permanent magnet d-axis flux constant component. In the case of a SynRM, $\Lambda_D^{\text{mg}} = 0$, but in the case of a Permanent Magnet Assisted SynRM (PMASynRM) $\Lambda_D^{\text{mg}} \neq 0$.
- $\boldsymbol{L}_{dq}^{\text{app}} = \begin{bmatrix} L_d^{\text{app}} & L_{dq}^{\text{app}} \\ L_{dq}^{\text{app}} & L_q^{\text{app}} \end{bmatrix}$, apparent inductance matrix, defined by $\boldsymbol{\lambda}_{DQ} = \boldsymbol{L}_{dq}^{\text{app}} \boldsymbol{i}_{DQ} + \boldsymbol{\lambda}_{DQ}^{\text{mg}}$.

To transition from the synchronous reference frame dq to the stationary reference frame $\alpha\beta$, the Park transformation matrices are defined as:

$$\boldsymbol{T}_{dq}^{\alpha\beta} = \begin{bmatrix} \cos(\vartheta_E) & -\sin(\vartheta_E) \\ \sin(\vartheta_E) & \cos(\vartheta_E) \end{bmatrix}, \qquad \boldsymbol{T}_{\alpha\beta}^{dq} = \begin{bmatrix} \cos(\vartheta_E) & \sin(\vartheta_E) \\ -\sin(\vartheta_E) & \cos(\vartheta_E) \end{bmatrix}. \tag{7.3}$$

By multiplying (7.2) by $\boldsymbol{T}_{dq}^{\alpha\beta}$ on both sides, the state-space model is transformed into the $\alpha\beta$ reference frame, where the matrices $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, and $\boldsymbol{D}$ are redefined accordingly.

$$\begin{aligned}
\boldsymbol{T}_{dq}^{\alpha\beta}\,\boldsymbol{u}_{DQ} \;&=\; \boldsymbol{R}_s\,\boldsymbol{T}_{dq}^{\alpha\beta}\,\boldsymbol{i}_{DQ} \,+\, \boldsymbol{T}_{dq}^{\alpha\beta}\,\boldsymbol{L}_{dq}^{\text{diff}}\frac{\mathrm{d}\,\boldsymbol{T}_{\alpha\beta}^{dq}\,\boldsymbol{i}_{\alpha\beta}}{\mathrm{d}t} \,+\, \omega_E\,\boldsymbol{T}_{dq}^{\alpha\beta}\,\boldsymbol{Z}\,\boldsymbol{L}_{dq}^{\text{app}}\,\boldsymbol{T}_{\alpha\beta}^{dq}\,\boldsymbol{i}_{\alpha\beta} \,+\, \cdots \\
&\cdots \,+\, \omega_E\,\boldsymbol{T}_{dq}^{\alpha\beta}\,\boldsymbol{Z}\,\boldsymbol{\lambda}_{DQ}^{\text{mg}}
\end{aligned}$$

$$\begin{aligned}
\boldsymbol{u}_{\alpha\beta} \;&=\; \boldsymbol{R}_s\,\boldsymbol{i}_{\alpha\beta} \,+\, \underbrace{\boldsymbol{T}_{dq}^{\alpha\beta}\,\boldsymbol{L}_{dq}^{\text{diff}}\,\boldsymbol{T}_{\alpha\beta}^{dq}}_{\boldsymbol{A}}\,\frac{\mathrm{d}\boldsymbol{i}_{\alpha\beta}}{\mathrm{d}t} \,+\, \underbrace{\boldsymbol{T}_{dq}^{\alpha\beta}\,\boldsymbol{L}_{dq}^{\text{diff}}\,\frac{\mathrm{d}\boldsymbol{T}_{\alpha\beta}^{dq}}{\mathrm{d}t}}_{\omega_E\boldsymbol{B}}\,\boldsymbol{i}_{\alpha\beta} \,+\, \cdots \\
&\cdots \,+\, \omega_E\,\underbrace{\boldsymbol{T}_{dq}^{\alpha\beta}\,\boldsymbol{Z}\,\boldsymbol{L}_{dq}^{\text{app}}\,\boldsymbol{T}_{\alpha\beta}^{dq}}_{\boldsymbol{C}}\,\boldsymbol{i}_{\alpha\beta} \,+\, \omega_E\,\underbrace{\boldsymbol{T}_{dq}^{\alpha\beta}\,\boldsymbol{Z}\,\boldsymbol{\lambda}_{DQ}^{\text{mg}}}_{\boldsymbol{D}}
\end{aligned}$$

$$\boldsymbol{u}_{\alpha\beta} \;=\; \boldsymbol{R_s}\,\boldsymbol{i}_{\alpha\beta} \,+\, \boldsymbol{A}\,\frac{\mathrm{d}\boldsymbol{i}_{\alpha\beta}}{\mathrm{d}t} + \omega_E\left((\boldsymbol{B} \,+\, \boldsymbol{C})\,\boldsymbol{i}_{\alpha\beta} + \boldsymbol{d}\right) \tag{7.4}$$

All the matrices and the vector considered in (7.4) are defined as follows. Detailed mathematical steps for deriving these elements are provided in Appendix A.1.

$$\begin{aligned}
\boldsymbol{A} &\triangleq \begin{bmatrix} L_{\Sigma}^{\text{diff}} + L_{\Delta}^{\text{diff}}\cos(2\vartheta_E) - L_{dq}^{\text{diff}}\sin(2\vartheta_E) & L_{\Delta}^{\text{diff}}\sin(2\vartheta_E) + L_{dq}^{\text{diff}}\cos(2\vartheta_E) \\ L_{\Delta}^{\text{diff}}\sin(2\vartheta_E) + L_{dq}^{\text{diff}}\cos(2\vartheta_E) & L_{\Sigma}^{\text{diff}} - L_{\Delta}^{\text{diff}}\cos(2\vartheta_E) + L_{dq}^{\text{diff}}\sin(2\vartheta_E) \end{bmatrix} \\
\boldsymbol{B} &\triangleq \begin{bmatrix} -L_{\Delta}^{\text{diff}}\sin(2\vartheta_E) - L_{dq}^{\text{diff}}\cos(2\vartheta_E) & L_{\Sigma}^{\text{diff}} + L_{\Delta}^{\text{diff}}\cos(2\vartheta_E) - L_{dq}^{\text{diff}}\sin(2\vartheta_E) \\ -L_{\Sigma}^{\text{diff}} + L_{\Delta}^{\text{diff}}\cos(2\vartheta_E) - L_{dq}^{\text{diff}}\sin(2\vartheta_E) & L_{\Delta}^{\text{diff}}\sin(2\vartheta_E) + L_{dq}^{\text{diff}}\cos(2\vartheta_E) \end{bmatrix} \\
\boldsymbol{C} &\triangleq \begin{bmatrix} -L_{\Delta}^{\text{app}}\sin(2\vartheta_E) - L_{dq}^{\text{app}}\cos(2\vartheta_E) & -L_{\Sigma}^{\text{app}} + L_{\Delta}^{\text{app}}\cos(2\vartheta_E) - L_{dq}^{\text{app}}\sin(2\vartheta_E) \\ L_{\Sigma}^{\text{app}} + L_{\Delta}^{\text{app}}\cos(2\vartheta_E) - L_{dq}^{\text{app}}\sin(2\vartheta_E) & L_{\Delta}^{\text{app}}\sin(2\vartheta_E) + L_{dq}^{\text{app}}\cos(2\vartheta_E) \end{bmatrix} \\
\boldsymbol{d} &\triangleq \begin{bmatrix} -\Lambda_D^{\text{mg}}\sin(2\vartheta_E) \\ \Lambda_D^{\text{mg}}\cos(2\vartheta_E) \end{bmatrix}
\end{aligned} \tag{7.5}$$

where

$$\begin{aligned}
L_{\Sigma}^{\text{diff}} &\triangleq \frac{L_d^{\text{diff}} + L_q^{\text{diff}}}{2} \qquad & L_{\Delta}^{\text{diff}} &\triangleq \frac{L_d^{\text{diff}} - L_q^{\text{diff}}}{2} \\
L_{\Sigma}^{\text{app}} &\triangleq \frac{L_d^{\text{app}} + L_q^{\text{app}}}{2} \qquad & L_{\Delta}^{\text{app}} &\triangleq \frac{L_d^{\text{app}} - L_q^{\text{app}}}{2}
\end{aligned} \tag{7.6}$$

The equation (7.4) describes the *electrical* subsystem, which yields the time derivative of the currents $\boldsymbol{i}_{\alpha\beta}$ used in the Kalman filter prediction step.

$$\boldsymbol{f}_{el}(\boldsymbol{x},\boldsymbol{u}) :\; \frac{\mathrm{d}\boldsymbol{i}_{\alpha\beta}}{\mathrm{d}t} \;=\; \boldsymbol{A}^{-1}\left[\boldsymbol{u}_{\alpha\beta} - \boldsymbol{R}_s\,\boldsymbol{i}_{\alpha\beta} \;-\; \omega_E\left((\boldsymbol{B} + \boldsymbol{C})\,\boldsymbol{i}_{\alpha\beta} + \boldsymbol{d}\right)\right] \tag{7.7}$$

7.1.2 Mechanical Model

The mechanical model of the machine is essential for deriving the speed and position prediction equations. The hypothesis is that the mechanical system is treated as a *single-mass* system. Thus, the interaction between the machine torque at the shaft τ_M and its speed ω_M is described by the following equation:

$$\tau_M \;=\; J\frac{\mathrm{d}\omega_M}{\mathrm{d}t} \,+\, B\omega_M \,+\, \tau_L \tag{7.8}$$

Here, J represents the combined inertia of the machine and the load, B is the viscous friction coefficient when the motor is coupled with the load, and τ_L is the load torque.

Two approaches can be used to model the mechanical system. The first approach assumes the *infinite* inertia hypothesis, where the mechanical system is modeled with J treated as infinite, resulting in constant speed:

$$\boldsymbol{f}_{mech,II}(\boldsymbol{x},\boldsymbol{u}) : \begin{cases} \dfrac{\mathrm{d}\omega_E}{\mathrm{d}t} &= 0 \\ \dfrac{\mathrm{d}\vartheta_E}{\mathrm{d}t} &= \omega_E \end{cases} \tag{7.9}$$

This model eliminates the need for mechanical parameters, as they are not required in (7.9). Although it may seem unsuitable for systems with varying speed, this approach is widely used in observer design for its simplicity. In such cases, speed updates are handled in the measurement update step, where the observer's gain is used to correct inaccurate speed predictions.

The *finite* inertia hypothesis requires knowledge of the total inertia J, the viscous friction coefficient B, the load torque τ_L, and the motor torque τ_M. In this case, the mechanical system in continuous time is described by the following equations:

$$f_{mech,FI}(\boldsymbol{x},\boldsymbol{u}) : \begin{cases} \dfrac{\mathrm{d}\omega_E}{\mathrm{d}t} &= \dfrac{p}{J}\left(\tau_M - \dfrac{B}{p}\omega_E - \tau_L\right) \\ \dfrac{\mathrm{d}\vartheta_E}{\mathrm{d}t} &= \omega_E \end{cases} \tag{7.10}$$

The electromagnetic torque generated by the motor can be written as

$$\begin{aligned}
\tau_M &= \frac{3}{2}p\left(\lambda_D i_Q - \lambda_Q i_D\right) \\
&= \frac{3}{2}\left(\left(L_d^{\mathrm{app}} - L_q^{\mathrm{app}}\right) i_D i_Q + \Lambda_D^{\mathrm{mg}} i_Q\right) \\
&= \frac{3}{2}p\,\left[\left(L_d^{\mathrm{app}} - L_q^{\mathrm{app}}\right)\left(i_\alpha\cos(\vartheta_E) + i_\beta\sin(\vartheta_E)\right)\left(-i_\alpha\sin(\vartheta_E) + i_\beta\cos(\vartheta_E)\right) + \cdots\right. \\
&\quad \left.\cdots + \Lambda_D^{\mathrm{mg}}\left(-i_\alpha\sin(\vartheta_E) + i_\beta\cos(\vartheta_E)\right)\right] \\
&= \frac{3}{2}p\,\left[\left(L_d^{\mathrm{app}} - L_q^{\mathrm{app}}\right)\left(-i_\alpha^2\cos(\vartheta_E)\sin(\vartheta_E) + i_\alpha i_\beta\cos^2(\vartheta_E) - i_\alpha i_\beta\sin^2(\vartheta_E) + \cdots\right.\right. \\
&\quad \left.\left.\cdots + i_\beta^2\sin(\vartheta_E)\cos(\vartheta_E)\right) + \Lambda_D^{\mathrm{mg}}\left(-i_\alpha\sin(\vartheta_E) + i_\beta\cos(\vartheta_E)\right)\right] \\
&= \frac{3}{2}p\,\left[\left(L_d^{\mathrm{app}} - L_q^{\mathrm{app}}\right)\left(\left(i_\beta^2 - i_\alpha^2\right)\cos(\vartheta_E)\sin(\vartheta_E) + i_\alpha i_\beta\left(\cos^2(\vartheta_E) + \cdots\right.\right.\right. \\
&\quad \left.\left.\left.\cdots - \sin^2(\vartheta_E)\right)\right) + \Lambda_D^{\mathrm{mg}}\left(-i_\alpha\sin(\vartheta_E) + i_\beta\cos(\vartheta_E)\right)\right] \\
&= \frac{3}{4}p\,\left[\left(L_d^{\mathrm{app}} - L_q^{\mathrm{app}}\right)\left(\left(i_\beta^2 - i_\alpha^2\right)\sin(2\vartheta_E) + 2i_\alpha i_\beta\cos(2\vartheta_E)\right) + \cdots\right. \\
&\quad \left.\cdots + 2\Lambda_D^{\mathrm{mg}}\left(-i_\alpha\sin(\vartheta_E) + i_\beta\cos(\vartheta_E)\right)\right] \\
&= \frac{3}{2}p\,\left[L_\Delta^{\mathrm{app}}\left(i_\beta^2 - i_\alpha^2\right)\sin(2\vartheta_E) + 2L_\Delta^{\mathrm{app}} i_\alpha i_\beta\cos(2\vartheta_E) + \cdots\right.
\end{aligned}$$

$$\cdots + \Lambda_D^{\text{mg}}\left(-i_\alpha \sin(\vartheta_E) + i_\beta \cos(\vartheta_E)\right)] \tag{7.11}$$

where the cross coupling apparent inductance L_{dq}^{app} has been disregarded in the aforementioned equations.

Including information about cross-coupling could, in principle, improve model accuracy. However, an overfitted model may degrade estimation performance by introducing noise, particularly if the cross-coupling data cannot be obtained with sufficient accuracy, as is the case here. To avoid this noise amplification, these mutual inductances are excluded from the model to achieve a more robust estimation.

This holds also for differential inductance $L_{dq}^{\text{diff}} \simeq 0$, which is not considered in the torque model but in the current model (7.7).

This is commonly done in the literature, where the cross-coupling effect in the observer model is often neglected. For instance, [84] approximates L_d, L_q inductances with empirical laws without considering any mutual inductance effect. [85] uses inductances as functions of both currents, $L_d(i_D, i_Q)$ and $L_q(i_D, i_Q)$, approximated using polynomials. No cross-coupling inductance is considered. [86] proposes an estimation method that can suppress position estimation errors caused by deviations in inductance parameters, by using dynamic and static inductances for each axis.

In [87], the authors provide a formula to link the cross-coupling inductance with rotor position estimation error:

$$\Delta\vartheta_E = \frac{1}{2}\arctan\left(\frac{2L_{dq}}{L_q - L_d}\right) \tag{7.12}$$

According to (7.12), neglecting the cross-coupling inductance maps of the SynRM used in the following chapters generates angle estimation errors of a few degrees. Unfortunately, this issue cannot be resolved by simply including the mutual inductance in the model because the cross-coupling inductance maps obtained experimentally [88] are highly noisy. Thus, their inclusion increases estimation noise without providing significant improvement.

However, it is worth noting that the magnetic models used in the experiments remain accurate in terms of self-inductances, as they employ four distinct inductance maps: L_d^{diff}, L_q^{diff}, L_d^{app}, and L_q^{app}, which are functions of both i_D and i_Q currents. Therefore, the cross-coupling effect is implicitly considered in the used model.

By substituting (7.11) into (7.10), the following augmented mechanical system is obtained:

$$\boldsymbol{f}_{mech,FI}(\boldsymbol{x},\boldsymbol{u}) : \begin{cases} \dfrac{\mathrm{d}\omega_E}{\mathrm{d}t} &= \dfrac{p}{J}\Bigg[\dfrac{3}{2}\left(L_\Delta^{\text{app}}\left(i_\beta^2 - i_\alpha^2\right)\sin(2\vartheta_E) + 2L_\Delta^{\text{app}} i_\alpha i_\beta \cos(2\vartheta_E) \;+\; \cdots\right. \\ & \qquad \left.\cdots + \Lambda_D^{\text{mg}}\left(-i_\alpha \sin(\vartheta_E) + i_\beta\cos(\vartheta_E)\right)\right) - \dfrac{B}{p}\omega_E - \tau_L\Bigg] \\ \dfrac{\mathrm{d}\vartheta_E}{\mathrm{d}t} &= \omega_E \\ \dfrac{\mathrm{d}\tau_L}{\mathrm{d}t} &= 0 \end{cases} \tag{7.13}$$

The same conclusion drawn for the infinite inertia system in (7.9) holds for the finite inertia system in (7.13). In the latter, the mechanical parameters are required in the prediction step, but the load torque is assumed to be constant. This assumption is reasonable for many applications, where the load torque is constant or slowly varying. Any discrepancies in the prediction will be corrected in the measurement update step by the observer gain and the difference between the predicted and actual filter output.

7.1.3 Discrete Time Model

Given the electrical continuous-time system in (7.7) and the mechanical systems in (7.10) and (7.13), the discrete time model is obtained through the *forward Euler* discretization method, in the form of the *infinite inertia* hypothesis (subscript $_{II}$) and the *finite inertia* hypothesis (subscript $_{FI}$).

Discrete time variables are denoted by the subscript k, and the sampling time is T_s.

Discrete time SynRM model with infinite inertia hypothesis

The discrete time model for the infinite inertia hypothesis is obtained by discretizing the electrical continuous time model in (7.7) and the mechanical continuous-time model in (7.9) using the *forward Euler* method.

The discrete-time model is expressed as $\boldsymbol{f}_{II,d}(\boldsymbol{x}_k, \boldsymbol{u}_k) : [\boldsymbol{f}_{el,d},\ \boldsymbol{f}_{mech,II,d}]^T$,

$$\boldsymbol{f}_{II,d}(\boldsymbol{x}_k, \boldsymbol{u}_k) : \begin{cases} \boldsymbol{i}_{\mathrm{AB},k+1} = \boldsymbol{i}_{\alpha\beta,k} + T_s\, \boldsymbol{A}_k^{-1} \left[\boldsymbol{u}_{\alpha\beta,k} - \boldsymbol{R}_s\, \boldsymbol{i}_{\alpha\beta,k} - \omega_{E,k} \left((\boldsymbol{B}_k + \cdots \right.\right. \\ \qquad\qquad \left.\left. \cdots + \boldsymbol{C}_k)\, \boldsymbol{i}_{\alpha\beta,k} + \boldsymbol{d}_k \right) \right] \\ \omega_{E,k+1} = \omega_{E,k} \\ \vartheta_{E,k+1} = \vartheta_{E,k} + T_s\, \omega_{E,k} \end{cases} \tag{7.14}$$

where the state vector is defined as $\boldsymbol{x}_k = [i_{\alpha,k},\ i_{\beta,k},\ \omega_{E,k},\ \vartheta_{E,k}]^T$ and the input vector is defined as $\boldsymbol{u}_k = [\boldsymbol{u}_{\alpha,k},\ \boldsymbol{u}_{\beta,k}]^T$.

Matrices $\boldsymbol{A}_k$, $\boldsymbol{B}_k$, $\boldsymbol{C}_k$, and the vector $\boldsymbol{d}_k$ are the matrices defined in (7.5), calculated at time instant k, whereas $\boldsymbol{A}_k^{-1}$ is defined as

$$\boldsymbol{A}_k^{-1} = \frac{1}{\det \boldsymbol{A}_k} \begin{bmatrix} a_{22,k} & -\,a_{21,k} \\ -\,a_{12,k} & a_{11,k} \end{bmatrix} \tag{7.15}$$

and the determinant of matrix $\boldsymbol{A}_k$ is calculated as

$$\det \boldsymbol{A}_k = a_{11,k}\, a_{22,k} - a_{12,k}\, a_{21,k} = L_{d,k}^{\mathrm{diff}}\, L_{q,k}^{\mathrm{diff}} - \left(L_{dq,k}^{\mathrm{diff}}\right)^2 \tag{7.16}$$

Details about the derivation of (7.16) are shown in Appendix A.

Discrete time SynRM model with finite inertia hypothesis

By considering the same electrical system in (7.7) and (7.13) for the mechanical part, the discrete time model $\boldsymbol{f}_{d,FI}(\boldsymbol{x}_k, \boldsymbol{u}_k) : [\boldsymbol{f}_{el,\,d},\, \boldsymbol{f}_{mech,\,FI,\,d}]^T$ becomes

$$\boldsymbol{f}_{FI,\,d}(\boldsymbol{x}_k, \boldsymbol{u}_k) : \begin{cases} \boldsymbol{i}_{\mathrm{AB},k+1} &= \boldsymbol{i}_{\alpha\beta,k} + T_s\ \boldsymbol{A}_k^{-1} [\boldsymbol{u}_{\alpha\beta,k} - \boldsymbol{R}_s\ \boldsymbol{i}_{\alpha\beta,k} - \omega_{E,k}\,((\boldsymbol{B}_k + \cdots \\ & \quad \cdots + \boldsymbol{C}_k)\,\boldsymbol{i}_{\alpha\beta,k} + \boldsymbol{d}_k]] \\ \omega_{E,k+1} &= \omega_{E,k} + T_s\,\dfrac{p}{J}\left[\dfrac{3}{2}\left(L_\Delta^{\mathrm{app}}\left(i_{\beta,k}^2 - i_{\alpha,k}^2\right)\sin(2\vartheta_{E,k}) + \cdots\right.\right. \\ & \quad \cdots + 2L_\Delta^{\mathrm{app}}\, i_{\alpha,k}\, i_{\beta,k}\,\cos(2\vartheta_{E,k}) + \Lambda_D^{\mathrm{mg}}\,(-i_{\alpha,k}\sin(\vartheta_{E,k}) + \cdots \\ & \quad \left.\cdots + i_{\beta,k}\cos(\vartheta_{E,k}))) - \dfrac{B}{p}\omega_{E,k} - \tau_{L,k}\right] \\ \vartheta_{E,k+1} &= \vartheta_{E,k} + T_s\ \omega_{E,k} \\ \tau_{L,k+1} &= \tau_{L,k} \end{cases} \tag{7.17}$$

where the state vector is defined as $\boldsymbol{x}_k = [i_{\alpha,k},\, i_{\beta,k},\, \omega_{E,k},\, \vartheta_{E,k},\, \tau_{L,k}]^T$ and the input vector is defined as $\boldsymbol{u}_k = [u_{\alpha,k},\, u_{\beta,k}]^T$.

When it comes to the output function $\boldsymbol{h}(\boldsymbol{x}_k)$, it is defined as the transformation of the state vector $\boldsymbol{x}_k$ into the output vector $\boldsymbol{y}_k$. In this case, the output function is defined as

$$\boldsymbol{y} = \boldsymbol{h}(\boldsymbol{x}) = [i_\alpha,\, i_\beta]^T \tag{7.18}$$

therefore the discrete version of h is simply given by

$$\boldsymbol{y}_k = \boldsymbol{h}_d(\boldsymbol{x}_k) = \boldsymbol{i}_{\alpha\beta,k} \tag{7.19}$$

From this point forward, the solely discrete model used in the following Kalman filter implementations is the one with the *infinite* inertia hypothesis in (7.14). This choice is motivated by the fact that the *finite* inertia model is more complex and requires the knowledge of the machine parameters, which are not always available. Moreover, (7.17) implies state augmentation with the load torque τ_L, whose estimation depends only on the filter correction step, as does the estimation of the speed ω_E in the infinite inertia model in (7.9).

7.2 Synchronous machine observability

The system observability is crucial to develop the nonlinear versions of Kalman filter [89]. The nonlinear state space system in the continuous time domain is

$$\begin{cases} \dot{\boldsymbol{x}}(t) = \boldsymbol{f}(\boldsymbol{x}(t),\, \boldsymbol{u}(t)) \\ \boldsymbol{y}(t) = \boldsymbol{h}(\boldsymbol{x}(t)) \end{cases} \tag{7.20}$$

and by considering the synchronous machine model (7.14) obtained in the previous section, the system state derivative is

$$\boldsymbol{f}(\boldsymbol{x},\boldsymbol{u}) : \begin{cases} \dfrac{\mathrm{d}\boldsymbol{i}_{\alpha\beta}}{\mathrm{d}t} = \boldsymbol{A}^{-1}\left[\boldsymbol{u}_{\alpha\beta} - \boldsymbol{R}_s\,\boldsymbol{i}_{\alpha\beta} - \omega_E\left((\boldsymbol{B}+\boldsymbol{C})\,\boldsymbol{i}_{\alpha\beta} + \boldsymbol{d}\right)\right] \\ \dfrac{\mathrm{d}\omega_E}{\mathrm{d}t} = 0 \\ \dfrac{\mathrm{d}\vartheta_E}{\mathrm{d}t} = \omega_E \end{cases} \tag{7.21}$$

whereas the output equation is

$$\boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{i}_{\alpha\beta} \tag{7.22}$$

From (7.21) and (7.22), the system state is $\boldsymbol{x} = [i_\alpha,\, i_\beta,\, \omega_E,\, \vartheta_E]^{\mathrm{T}}$ and system output is $\boldsymbol{y} = [i_\alpha,\, i_\beta]^{\mathrm{T}}$.

7.2.1 Observability for nonlinear systems

Given the nonlinear system (7.20), where $\boldsymbol{f}$ and $\boldsymbol{h}$ are the time update and the output $\mathbb{C}^\infty$ nonlinear functions, $\boldsymbol{x} \in \boldsymbol{X} \subset \mathbb{R}^n$ is the state vector, $\boldsymbol{u} \in \boldsymbol{U} \subset \mathbb{R}^m$ is the control vector (input), $\boldsymbol{y} \in \boldsymbol{Y} \subset \mathbb{R}^p$ is the output vector, the observability theory is stated with the following definitions and theorems clearly explained by *Hermann et al.* in [89] and recalled in [90], [91], [92].

Definition 1 (Input-Output map)**.** *Each admissible input* $(\boldsymbol{u}(t), [t_0, t_1])$ *gives rise to a solution* $(\boldsymbol{x}(t, [t_0, t_1]))$ *of* $\frac{\mathrm{d}x(t)}{\mathrm{d}t} = f(\boldsymbol{x}(t),\, \boldsymbol{u}(t))$ *satisfying the initial condition* $\boldsymbol{x}_0$*. This, in turn, defines an output* $(\boldsymbol{y}(t), [t_0, t_1])$ *by* $\boldsymbol{y}(t) = h(\boldsymbol{x}(t))$*. This map is denoted by*

$$\boldsymbol{\Sigma}_{\boldsymbol{x}_0} : (\boldsymbol{u}(t),\, [t_0,\, t_1]) \rightarrow (\boldsymbol{y}(t),\, [t_0,\, t_1]) \tag{7.23}$$

Definition 2 (Indistinguishability)**.** *A pair of points* $\boldsymbol{x}_0$ *and* $\boldsymbol{x}_1$ *are indistinguishable* I[1]*, denoted as* $\boldsymbol{x}_0 I \boldsymbol{x}_1$ *if* $(\boldsymbol{\Sigma}, \boldsymbol{x}_0)$ *and* $(\boldsymbol{\Sigma}, \boldsymbol{x}_1)$ *realize the input-output map, i. e. for every admissible input* $(\boldsymbol{u}(t), [t_0,\, t_1])$

$$\boldsymbol{\Sigma}_{\boldsymbol{x}_0}(\boldsymbol{x}(t),\ [t_0,\, t_1]) = \boldsymbol{\Sigma}_{\boldsymbol{x}_1}(\boldsymbol{u}(t),\, [t_0,\, t_1]) \tag{7.24}$$

Definition 3 (Observability)**.** $\boldsymbol{\Sigma}$ *is said to be observable at* $\boldsymbol{x}_0$ *if* $I(\boldsymbol{x}_0) = \boldsymbol{x}_0$*. Furthermore,* $\boldsymbol{\Sigma}$ *is observable if* $I(\boldsymbol{x}_0) = \boldsymbol{x}_0$ *for every* $\boldsymbol{x} \in \boldsymbol{X}$*.*

In other words, the first sentence of Def. 3 indicates that there are no indistinguishable points of $\boldsymbol{x}_0$, while the second one states that there are no indistinguishable points in the entire state space $\boldsymbol{X}$ [90]. Furthermore, since *observability* is a global concept [89] and may require a significant amount of time to distinguish between points in $\boldsymbol{X}$, the concept of *local weak observability* is introduced.

Definition 4 (Local weak observability)**.** $\boldsymbol{\Sigma}$ *is locally weak observable at* $\boldsymbol{x}_0$ *if there exists an open neighborhood* U *of* $\boldsymbol{x}_0$ *such that for every open neighborhood* V *of* $\boldsymbol{x}_0$ *contained in* U*,* $I_V(\boldsymbol{x}_0) = \boldsymbol{x}_0$ *and is locally weak observable if it is so at every* $\boldsymbol{x}$ *in* $\boldsymbol{X}$*.*

[1] The symbol I recalls the indistinguishability property, for instance: $x_0 I_U x_1$ means that x_0 is indistinguishable from x_1 in the neighborhood U, with $x_0,\, x_1 \subset U$

In conclusion, it is possible to instantly distinguish the initial state, but only from its neighboring points [90].

Before defining the local observability criterion, it is necessary to introduce the concepts of Lie derivatives and the observability matrix.

Definition 5 (Observability matrix and Lie derivative)**.** *From the system in* (7.20)*, the observability matrix* $\mathcal{O}(\boldsymbol{x})$ *in a point* $\boldsymbol{x}_0$ *of* $\boldsymbol{X}$ *is given by:*

$$\mathcal{O}(\boldsymbol{x})\big|_{\boldsymbol{x}_0} = \left.\begin{bmatrix} \frac{\partial}{\partial \boldsymbol{x}}\mathcal{L}_f^0 h(\boldsymbol{x}) \\ \frac{\partial}{\partial \boldsymbol{x}}\mathcal{L}_f^1 h(\boldsymbol{x}) \\ \frac{\partial}{\partial \boldsymbol{x}}\mathcal{L}_f^2 h(\boldsymbol{x}) \\ \cdot \\ \cdot \\ \cdot \end{bmatrix}\right|_{\boldsymbol{x}=\boldsymbol{x}_0} \tag{7.25}$$

Matrix $\mathcal{O}(x)$ *is made of Jacobian Lie derivatives, where the Lie derivative* $\mathcal{L}_f h(\boldsymbol{x})$ *is defined by*

$$\mathcal{L}_f h = \frac{\partial \boldsymbol{h}(\boldsymbol{x})}{\partial \boldsymbol{x}} f(\boldsymbol{x}, \boldsymbol{u}) \tag{7.26}$$

In other words, the Lie derivative calculates the change in the output equation $\boldsymbol{y} = h(\boldsymbol{x})$, where the dynamics of $\boldsymbol{x}$ are determined by the vector $\dot{\boldsymbol{x}} = f(\boldsymbol{x}, \boldsymbol{u})$ [90]. One important property of Lie derivatives is that the derivative of order zero is defined by:

$$\mathcal{L}_f^0 h = \boldsymbol{h}(x) \tag{7.27}$$

and higher order derivatives are defined recursively by

$$\mathcal{L}_f^n h = \frac{\partial \mathcal{L}_f^{n-1} h}{\partial \boldsymbol{x}} \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}) \tag{7.28}$$

where n is the system order.

Theorem 6. *If* $\boldsymbol{\Sigma}$ *satisfies the observability rank condition at* $\boldsymbol{x}_0$*,*

$$rank(\mathcal{O}(\boldsymbol{x})\big|_{\boldsymbol{x}_0}) = n \tag{7.29}$$

then Σ *is locally weak observable*

Checking the rank of the observability matrix, as provided by (7.25), is a practical approach to proving local observability for nonlinear systems. This is because applying the Cayley-Hamilton theorem, as shown below for clarity, is impossible in this context, as the theorem is suitable only for linear systems [90].

Theorem 7 (Cayley-Hamilton)**.** *Consider a linear time-invariant system described by the state-space equations:*

$$\frac{\mathrm{d}\boldsymbol{x}(t)}{\mathrm{d}t} = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{B}\boldsymbol{u}(t), \quad \boldsymbol{y}(t) = \boldsymbol{C}\boldsymbol{x}(t),$$

where $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, $\boldsymbol{B} \in \mathbb{R}^{n\times m}$, *and* $\boldsymbol{C} \in \mathbb{R}^{p\times n}$. *The system is observable if and only if the observability matrix*

$$\mathcal{O} = \begin{bmatrix} \boldsymbol{C} \\ \boldsymbol{CA} \\ \boldsymbol{CA}^2 \\ \vdots \\ \boldsymbol{CA}^{n-1} \end{bmatrix}$$

has full rank, i.e., $\mathrm{rank}(\mathcal{O}) = n$.

7.2.2 Observability matrix for the SynRM

In order to apply (7.25) to the SynRM mathematical model in (7.21), the Lie derivatives appearing in (7.25) can be rewritten according to definition (7.28) [90] as:

$$\begin{aligned}
\mathcal{L}_f^0\ h &= \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{y} \\
\mathcal{L}_f^1\ h &= \frac{\partial h(\boldsymbol{x})}{\partial \boldsymbol{x}}\boldsymbol{f}(\boldsymbol{x},u) = \frac{\partial h(\boldsymbol{x})}{\partial \boldsymbol{x}}\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \frac{\mathrm{d}\boldsymbol{y}}{\mathrm{d}t} \\
\mathcal{L}_f^2\ h &= \frac{\partial \mathcal{L}_f^1}{\partial \boldsymbol{x}}\boldsymbol{f}(\boldsymbol{x},u) = \frac{\partial}{\partial \boldsymbol{x}}\left(\frac{\partial h(\boldsymbol{x})}{\partial t}\right)\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \frac{\mathrm{d}^2\boldsymbol{y}}{\mathrm{d}t^2} \\
\mathcal{L}_f^3\ h &= \frac{\partial \mathcal{L}_f^2}{\partial \boldsymbol{x}}\boldsymbol{f}(\boldsymbol{x},u) = \frac{\partial}{\partial \boldsymbol{x}}\left(\frac{\partial}{\partial \boldsymbol{x}}\left(\frac{\partial h(\boldsymbol{x})}{\partial t}\right)\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}\right)\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \frac{\mathrm{d}^3\boldsymbol{y}}{\mathrm{d}t^3} \\
&\cdots
\end{aligned}$$

It is important to recall that the output equation is $\boldsymbol{y} = \boldsymbol{i}_{\alpha\beta}$, while the state vector is $\boldsymbol{x} = [i_\alpha,\, i_\beta,\, \omega_E,\, \vartheta_E]^{\mathrm{T}}$. By considering the system in (7.21), in order to satisfy the local weak observability theorem, the rank of $\mathcal{O}(\boldsymbol{x})$ must be:

$$\mathrm{rank}(\mathcal{O}(\boldsymbol{x})) = 4 \tag{7.30}$$

The condition (7.30) can be proved by searching for a regular matrix assembled from any four rows of matrix (7.25). For the sake of simplicity, the matrix formed by the first four rows is considered, and it is denoted by $\mathcal{O}(\boldsymbol{x})_{1234}$, as shown in (7.31).

$$\mathcal{O}(\boldsymbol{x})_{1234} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \dfrac{\partial \mathcal{L}_f^1 h_1}{\partial i_\alpha} & \dfrac{\partial \mathcal{L}_f^1 h_1}{\partial i_\beta} & \dfrac{\partial \mathcal{L}_f^1 h_1}{\partial \omega_E} & \dfrac{\partial \mathcal{L}_f^1 h_1}{\partial \vartheta_E} \\ \dfrac{\partial \mathcal{L}_f^1 h_2}{\partial i_\alpha} & \dfrac{\partial \mathcal{L}_f^1 h_2}{\partial i_\beta} & \dfrac{\partial \mathcal{L}_f^1 h_2}{\partial \omega_E} & \dfrac{\partial \mathcal{L}_f^1 h_2}{\partial \vartheta_E} \end{bmatrix} \tag{7.31}$$

since matrix (7.31) is a block lower triangular matrix, its determinant can be easily calculated as:

$$\det \mathcal{O}_{1234} = \frac{\partial \mathcal{L}_f^1 h_1}{\partial \omega_E}\frac{\partial \mathcal{L}_f^1 h_2}{\partial \vartheta_E} - \frac{\partial \mathcal{L}_f^1 h_1}{\partial \vartheta_E}\frac{\partial \mathcal{L}_f^1 h_2}{\partial \omega_E} \tag{7.32}$$

The determinant (7.32) is obtained using a symbolic computation tool in MATLAB to handle its complexity. The result is presented after substituting the $\alpha\beta$ quantities with dq ones, according to:

$$i_D = i_\alpha \cos(\vartheta_E) + i_\beta \sin(\vartheta_E) \tag{7.33}$$
$$i_Q = -i_\alpha \sin(\vartheta_E) + i_\beta \cos(\vartheta_E) \tag{7.34}$$
$$u_D = u_\alpha \cos(\vartheta_E) + u_\beta \sin(\vartheta_E) \tag{7.35}$$
$$u_Q = -u_\alpha \sin(\vartheta_E) + u_\beta \cos(\vartheta_E) \tag{7.36}$$

and by assuming the nullity of the cross inductances L_{dq}^{app} and L_{dq}^{diff}, as explained in Subsect. 7.1.2, the determinant (7.32) becomes

$$\begin{aligned}
\det \mathcal{O}_{1234} = \big(&\omega_E (\Lambda_D^{\text{mg}})^2 (L_d^{\text{diff}})^2 + 2\omega_E \Lambda_D^{\text{mg}} L_d^{\text{app}} (L_d^{\text{diff}})^2 i_D - \omega_E \Lambda_D^{\text{mg}} (L_d^{\text{diff}})^2 L_q^{\text{diff}} i_D + \cdots \\
&\cdots + r_S \Lambda_D^{\text{mg}} (L_d^{\text{diff}})^2 i_Q - u_Q \Lambda_D^{\text{mg}} (L_d^{\text{diff}})^2 - \omega_E \Lambda_D^{\text{mg}} L_d^{\text{diff}} L_q^{\text{app}} L_q^{\text{diff}} i_D + \cdots \\
&\cdots - r_S \Lambda_D^{\text{mg}} L_d^{\text{diff}} L_q^{\text{diff}} i_Q + u_Q \Lambda_D^{\text{mg}} L_d^{\text{diff}} L_q^{\text{diff}} + \omega_E (L_d^{\text{app}})^2 (L_d^{\text{diff}})^2 i_D^2 + \cdots \\
&\cdots - \omega_E L_d^{\text{app}} (L_d^{\text{diff}})^2 L_q^{\text{diff}} i_D^2 + \omega_E L_d^{\text{app}} (L_d^{\text{diff}})^2 L_q^{\text{diff}} i_Q^2 + \cdots \\
&\cdots + r_S L_d^{\text{app}} (L_d^{\text{diff}})^2 i_D i_Q - u_Q L_d^{\text{app}} (L_d^{\text{diff}})^2 i_D - \omega_E L_d^{\text{app}} L_d^{\text{diff}} L_q^{\text{app}} L_q^{\text{diff}} i_D^2 + \cdots \\
&\cdots - \omega_E L_d^{\text{app}} L_d^{\text{diff}} L_q^{\text{app}} L_q^{\text{diff}} i_Q^2 - r_S L_d^{\text{app}} L_d^{\text{diff}} L_q^{\text{diff}} i_D i_Q + u_Q L_d^{\text{app}} L_d^{\text{diff}} L_q^{\text{diff}} i_D + \cdots \\
&\cdots - 2 r_S (L_d^{\text{diff}})^2 L_q^{\text{diff}} i_D i_Q + u_Q (L_d^{\text{diff}})^2 L_q^{\text{diff}} i_D + u_D (L_d^{\text{diff}})^2 L_q^{\text{diff}} i_Q + \cdots \\
&\cdots - \omega_E L_d^{\text{diff}} L_q^{\text{app}} (L_q^{\text{diff}})^2 i_D^2 - \omega_E L_d^{\text{diff}} L_q^{\text{app}} (L_q^{\text{diff}})^2 i_Q^2 + r_S L_d^{\text{diff}} L_q^{\text{app}} L_q^{\text{diff}} i_D i_Q + \cdots \\
&\cdots - u_D L_d^{\text{diff}} L_q^{\text{app}} L_q^{\text{diff}} i_Q + 2 r_S L_d^{\text{diff}} (L_q^{\text{diff}})^2 i_D i_Q - u_Q L_d^{\text{diff}} (L_q^{\text{diff}})^2 i_D + \cdots \\
&\cdots - u_D L_d^{\text{diff}} (L_q^{\text{diff}})^2 i_Q + \omega_E (L_q^{\text{app}})^2 (L_q^{\text{diff}})^2 i_Q^2 - r_S L_q^{\text{app}} (L_q^{\text{diff}})^2 i_D i_Q + \cdots \\
&\cdots + u_D L_q^{\text{app}} (L_q^{\text{diff}})^2 i_Q \big) / ((L_d^{\text{diff}})^2 (L_q^{\text{diff}})^2)
\end{aligned} \tag{7.37}$$

which, of course, is not immediately understandable. From this point on, since within this thesis, and also in the experimental phase, only a pure reluctance machine is considered, the permanent magnet flux is absent, yielding $\Lambda_D^{\text{mg}} = 0$. Moreover, in order to simplify the determinant (7.37), the following substitutions are made:

$$\begin{aligned}
u_D &= r_S i_D + L_d^{\text{diff}} \frac{di_D}{dt} - \omega_E L_q^{\text{app}} i_Q \\
u_Q &= r_S i_Q + L_q^{\text{diff}} \frac{di_Q}{dt} + \omega_E L_d^{\text{app}} i_D
\end{aligned} \tag{7.38}$$

The determinant (7.37) becomes:

$$\det \mathcal{O}_{1234} = \Big(\omega_E (L_d^{\text{app}})^2 i_D^2 - \frac{di_Q}{dt} L_d^{\text{app}} L_d^{\text{diff}} i_D + \omega_E L_d^{\text{app}} L_d^{\text{diff}} i_Q^2 - \omega_E L_d^{\text{app}} L_q^{\text{app}} i_D^2 + \cdots$$
$$\cdots - \omega_E L_d^{\text{app}} L_q^{\text{app}} i_Q^2 - \omega_E L_d^{\text{app}} L_q^{\text{diff}} i_D^2 + \frac{di_Q}{dt} L_d^{\text{app}} L_q^{\text{diff}} i_D + \cdots$$
$$\cdots + \frac{di_D}{dt} (L_d^{\text{diff}})^2 i_Q - \omega_E L_d^{\text{diff}} L_q^{\text{app}} i_Q^2 - \frac{di_D}{dt} L_d^{\text{diff}} L_q^{\text{app}} i_Q + \cdots$$
$$\cdots + \frac{di_Q}{dt} L_d^{\text{diff}} L_q^{\text{diff}} i_Q - \frac{di_D}{dt} L_d^{\text{diff}} L_q^{\text{diff}} i_Q + \omega_E (L_q^{\text{app}})^2 i_Q^2 + \cdots$$
$$\cdots + \omega_E L_q^{\text{app}} L_q^{\text{diff}} i_D^2 + \frac{di_D}{dt} L_q^{\text{app}} L_q^{\text{diff}} i_Q - \frac{di_Q}{dt} (L_q^{\text{diff}})^2 i_D \Big) / (L_d^{\text{diff}} L_q^{\text{diff}}) \tag{7.39}$$

which can be resorted to the following form:

$$\det \mathcal{O}_{1234} = \big[\omega_E \, \big(i_D^2 \, ((L_d^{\text{app}})^2 - L_d^{\text{app}} L_q^{\text{app}} - L_d^{\text{app}} L_q^{\text{diff}} + L_q^{\text{app}} L_q^{\text{diff}}) + \cdots$$
$$\cdots + i_Q^2 \, ((L_q^{\text{app}})^2 - L_d^{\text{app}} L_q^{\text{app}} - L_d^{\text{diff}} L_q^{\text{app}} + L_d^{\text{app}} L_d^{\text{diff}})\big) + \cdots$$
$$\cdots + \frac{di_D}{dt} i_Q \, ((L_d^{\text{app}})^2 - L_d^{\text{app}} L_q^{\text{app}} - L_d^{\text{app}} L_q^{\text{diff}} + L_q^{\text{app}} L_q^{\text{diff}}) + \cdots$$
$$\cdots - \frac{di_Q}{dt} i_D \, ((L_q^{\text{app}})^2 - L_d^{\text{app}} L_q^{\text{app}} - L_d^{\text{diff}} L_q^{\text{app}} + L_d^{\text{app}} L_d^{\text{diff}})\Big] / (L_d^{\text{diff}} L_q^{\text{diff}})$$
$$= \omega_E \, (c_1 i_D^2 \, + \, c_2 i_Q^2) \, + \, c_1 \frac{di_D}{dt} i_Q \, - \, c_2 \frac{di_Q}{dt} i_D \tag{7.40}$$

where c_1 and c_2 are coefficients depending on the apparent and differential inductances.

It is worth noting that in the case of linear magnetic model, i. e. when:

$$L_d^{\text{app}} = L_d^{\text{diff}} = L_d$$
$$L_q^{\text{app}} = L_q^{\text{diff}} = L_q$$

where L_d and L_q are constant, the observability matrix determinant (7.40) reduces to:

$$\det \mathcal{O}_{1234}^{\text{lin}} = \frac{(L_d - L_q)^2}{L_d \, L_q} \left(\omega_E \, (i_D^2 \, + \, i_Q^2) \, + \, \frac{di_D}{dt} i_Q \, - \, \frac{di_Q}{dt} i_D\right) \tag{7.41}$$

which is the same as the one obtained in [90], [91], [92].

By further analyzing (7.40), it is possible to distinguish two situations in which the system is observable. The first is when the mechanical speed is zero or very close to zero, and the second is when the mechanical speed is different from zero.

- $\omega_E = 0$ or $\omega_E \simeq 0$: in this case (7.40) becomes:

$$\det \mathcal{O}_{1234} = c_1 \frac{di_D}{dt} i_Q \, - \, c_2 \frac{di_Q}{dt} i_D \tag{7.42}$$

and to ensure the full rank condition (7.30), it needs to be $\det \mathcal{O}_{1234} \neq 0$. Under these conditions, (7.42) affirms that the stator currents (i_D, i_Q) must be time-varying, i.e., the magnitude of the current vector must vary even if the mechanical speed is zero. Therefore, it means that the system is not observable if the SynRM is starting from zero speed with constant dq current references; the only way to make the system observable is to inject high-frequency currents [91].

- $\omega_E \neq 0$: by imposing to $\det \mathcal{O}_{1234} \neq 0$, (7.40) can be rewritten as:

$$\omega_E \neq \frac{c_1 \dfrac{di_D}{dt} i_Q \, - \, c_2 \dfrac{di_Q}{dt} i_D}{c_1 i_D^2 \, + \, c_2 i_Q^2} \tag{7.43}$$

This requirement is met in steady state when the derivative components are zero. During transients, the condition is not satisfied when the electric mechanical speed is equal to the second term of (7.43). In any case, this does not represent a cause for concern since the system's observability is restored as soon as the system passes through these points during the transient [92].

CHAPTER 8

EKF and UKF-based Sensorless SynRM drives

This chapter presents a critical comparison of sensorless algorithms based on two main nonlinear Kalman Filters: the EKF and UKF. The objective is to highlight the theoretical and practical advantages and disadvantages of each method when applied to the speed control of a SynRM, culminating in a definitive decision regarding the best choice. The magnetic model of the SynRM is significantly more nonlinear than that of induction motors, which are predominantly addressed in existing literature. This study aims to fill the gap by answering the question: "Is the extended Kalman filter still the best choice for nonlinear electric motors?" The answer is provided through an extensive set of experiments, including speed and load torque tests, zero-speed standstill starts, parameter sensitivity analyses, and evaluations of the computational burdens for both Kalman filter algorithms.

In Sect. 8.1, the structures of the EKF and UKF are presented, followed by the Monte Carlo Simulation (MCS) analysis of the nonlinear prediction function in Sect. 8.2. The experimental results are detailed in Sect. 8.3.

8.1 EKF and UKF equations

Consider the following nonlinear dynamical system:

$$\begin{cases} \boldsymbol{x}_{k+1} = \boldsymbol{f}\left(\boldsymbol{x}_k,\, \boldsymbol{u}_k,\, \boldsymbol{w}_k\right) & \boldsymbol{x}_k \in \mathbb{R}^{n\times 1}\,, \quad \boldsymbol{u}_k \in \mathbb{R}^{m\times 1} \\ \boldsymbol{y}_k = \boldsymbol{h}\left(\boldsymbol{x}_k,\, \boldsymbol{v}_k\right) & \boldsymbol{y}_k \in \mathbb{R}^{p\times 1} \end{cases} \tag{8.1}$$

where:

- the nonlinear functions $\boldsymbol{f}(\cdot)$, $\boldsymbol{h}(\cdot)$ and the input $\boldsymbol{u}_k$ are known a priori.

- $\boldsymbol{w}_k$ and $\boldsymbol{v}_k$ are zero-mean Gaussian random processes with known covariances:

$$\boldsymbol{w}_k \sim \mathcal{N}(\mathbf{0}, \boldsymbol{Q}_k), \qquad \boldsymbol{v}_k \sim \mathcal{N}(\mathbf{0}, \boldsymbol{R}_k) \tag{8.2}$$

- $\boldsymbol{w}_k$ and $\boldsymbol{v}_k$ are white noises:

$$\mathbb{E}\left[\boldsymbol{w}_k \boldsymbol{w}_h^T\right] = \boldsymbol{Q}_k\, \delta[k-h], \qquad \mathbb{E}\left[\boldsymbol{v}_k \boldsymbol{v}_h^T\right] = \boldsymbol{R}_k\, \delta[k-h] \tag{8.3}$$

  where k and h are the discrete time indices.

- $\boldsymbol{w}_k$ and $\boldsymbol{v}_k$ are mutually uncorrelated:

$$\mathbb{E}\left[\boldsymbol{w}_k \boldsymbol{v}_h^T\right] = \mathbf{0} \tag{8.4}$$

- the initial state $\boldsymbol{x}_0$ is a normally distributed random variable with mean $\bar{\boldsymbol{x}}_0$ and covariance $\boldsymbol{P}_{0|0}$:

$$\boldsymbol{x}_0 \sim \mathcal{N}(\bar{\boldsymbol{x}}_0, \boldsymbol{P}_{0|0}) \tag{8.5}$$

- $\boldsymbol{w}_k$ and $\boldsymbol{v}_k$ are uncorrelated with the initial state $\boldsymbol{x}_0$:

$$\mathbb{E}\left[\boldsymbol{x}_0 \boldsymbol{w}_k^T\right] = \mathbf{0}, \qquad \mathbb{E}\left[\boldsymbol{x}_0 \boldsymbol{v}_k^T\right] = \mathbf{0} \tag{8.6}$$

In the following KF structures, the functions $\boldsymbol{f}(\cdot)$ and $\boldsymbol{h}(\cdot)$ of (8.1) are those presented in (7.14) and (7.19), respectively, and are briefly restated here for clarity:

$$\boldsymbol{f}(\boldsymbol{x}_k, \boldsymbol{u}_k): \begin{cases} \boldsymbol{i}_{\alpha\beta,k+1} = \boldsymbol{i}_{\alpha\beta,k} + T_s\, \boldsymbol{A}_k^{-1} [\boldsymbol{u}_{\alpha\beta,k} - \boldsymbol{R}_s\, \boldsymbol{i}_{\alpha\beta,k} - \omega_{E,k} ((\boldsymbol{B}_k + \cdots \\ \qquad \cdots + \boldsymbol{C}_k)\, \boldsymbol{i}_{\alpha\beta,k} + \boldsymbol{d}_k]] \\ \omega_{E,k+1} = \omega_{E,k} \\ \vartheta_{E,k+1} = \vartheta_{E,k} + T_s\, \omega_{E,k} \end{cases} \tag{8.7}$$

$$\boldsymbol{y}_k = \boldsymbol{h}(\boldsymbol{x}^k) = \boldsymbol{i}_{\alpha\beta,k} \tag{8.8}$$

The state vector is defined as $\boldsymbol{x}_k = [i_{\alpha,k},\ i_{\beta,k},\ \omega_{E,k},\ \vartheta_{E,k}]^{\mathrm{T}}$ $(n = 4)$, the input voltages as $\boldsymbol{u}_k = [u_{\alpha,k},\ u_{\beta,k}]^{\mathrm{T}}$ $(m = 2)$, and the measured output vector as $\boldsymbol{y}_k = [i_{\alpha,k},\ i_{\beta,k}]^{\mathrm{T}}$ $(p = 4)$.

In terms of notation, the subscript $_{k|k}$ denotes the *corrected* quantity, whereas the subscript $_{k|k-1}$ indicates the *predicted* variable at the considered step k. For example, $\hat{\boldsymbol{x}}_{k|k-1}$ represents the predicted estimate of the state at step k, given all the measurements up to step $k-1$.

8.1.1 The Extended Kalman Filter

The EKF was introduced by *Schmidt* in 1966 [93], and its well-known structure consists of a linearization of the nonlinear functions in (8.1) with respect to the state vector $\boldsymbol{x}_k$. Considering all the aforementioned assumptions, the EKF algorithm is based on the following equations:

Initialization

$$\hat{\boldsymbol{x}}_{0|0} = \mathbb{E}\left[\boldsymbol{x}_0\right] \tag{8.9}$$

$$\boldsymbol{P}_{0|0} = \mathbb{E}\left[\left(\boldsymbol{x}_0 - \mathbb{E}\left[\boldsymbol{x}_0\right]\right)\ \left(\boldsymbol{x}_0 - \mathbb{E}\left[\boldsymbol{x}_0\right]\right)^T\right] \tag{8.10}$$

Time update

$$\hat{\boldsymbol{x}}_{k|k-1} = \boldsymbol{f}\left(\hat{\boldsymbol{x}}_{k-1|k-1}, \boldsymbol{u}_{k-1}\right) \tag{8.11}$$

$$\boldsymbol{\Phi}_k = \left.\frac{\partial \boldsymbol{f}}{\partial \hat{\boldsymbol{x}}}\right|_{\boldsymbol{x}_{k-1|k-1}, \boldsymbol{u}_{k-1}} \tag{8.12}$$

$$\boldsymbol{P}_{k|k-1} = \boldsymbol{\Phi}_k \boldsymbol{P}_{k-1|k-1} \boldsymbol{\Phi}_k^T + \boldsymbol{Q}_k \tag{8.13}$$

Measurement update

$$\boldsymbol{H}_k = \left.\frac{\partial \boldsymbol{h}}{\partial \hat{\boldsymbol{x}}}\right|_{\boldsymbol{x}_{k|k-1}} \tag{8.14}$$

$$\boldsymbol{K}_k = \boldsymbol{P}_{k|k-1} \boldsymbol{H}_k^{\mathrm{T}} \left(\boldsymbol{H}_k \boldsymbol{P}_{k|k-1} \boldsymbol{H}_k^{\mathrm{T}} + \boldsymbol{R}_k\right)^{-1} \tag{8.15}$$

$$\hat{\boldsymbol{x}}_{k|k} = \hat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}_k \left(\boldsymbol{y}_k - \boldsymbol{h}(\hat{\boldsymbol{x}}_{k|k-1})\right) \tag{8.16}$$

$$\boldsymbol{P}_{k|k} = \left(\boldsymbol{I}_{n\times n} - \boldsymbol{K}_k \boldsymbol{H}_k\right) \boldsymbol{P}_{k|k-1} \tag{8.17}$$

Where the matrices $\boldsymbol{\Phi}_k$ and $\boldsymbol{H}_k$ represent the state and output Jacobians, respectively, they pose significant challenges in the implementation of the EKF, as their computation requires the evaluation of the partial derivatives of the nonlinear functions $\boldsymbol{f}(\cdot)$ and $\boldsymbol{h}(\cdot)$ with respect to the state vector $\boldsymbol{x}_k$. When the model is highly complex, calculating these matrices can become cumbersome and prone to errors.

State Jacobian $\boldsymbol{\Phi}_k$

By considering the time update function $\boldsymbol{f}(\boldsymbol{x}_k, \boldsymbol{u}_k)$ in (7.14) and deriving it with respect to the state vector $\boldsymbol{x}_k$, we obtain the $n \times n$ state Jacobian matrix

$$\boldsymbol{\Phi}_k = \left.\frac{\partial \boldsymbol{f}(\boldsymbol{x}_k, \boldsymbol{u}_k)}{\partial \boldsymbol{x}_k}\right|_{\boldsymbol{x}_{k-1|k-1},\, \boldsymbol{u}_{k-1}} = \begin{bmatrix} \phi_{11,k} & \phi_{12,k} & \phi_{13,k} & \phi_{14,k} \\ \phi_{21,k} & \phi_{22,k} & \phi_{23,k} & \phi_{24,k} \\ \phi_{31,k} & \phi_{32,k} & \phi_{33,k} & \phi_{34,k} \\ \phi_{41,k} & \phi_{42,k} & \phi_{43,k} & \phi_{44,k} \end{bmatrix} \tag{8.18}$$

where the elements $\phi_{ij,k}$ (for $i, j = 1, \ldots, n$) are grouped into matrices for the sake of brevity. For instance, the matrix $\boldsymbol{\Phi}_{m-n,r-s,k}$ is composed as follows:

$$\boldsymbol{\Phi}_{m-n,r-s,k} = \begin{bmatrix} \phi_{m1,k} & \cdots & \phi_{mr,k} \\ \vdots & \cdots & \vdots \\ \phi_{n1,k} & \cdots & \phi_{ns,k} \end{bmatrix} \tag{8.19}$$

The sub-matrices of $\boldsymbol{\Phi}_k$ are:

$$
\begin{aligned}
\boldsymbol{\Phi}_{1-2,1-2} &= \boldsymbol{I}_{2\times 2} + T_s\, \boldsymbol{A}_k^{-1} \left[-\boldsymbol{R}_s - \omega_{E,k} \left(\boldsymbol{B}_k + \boldsymbol{C}_k \right) \right] \\
\boldsymbol{\Phi}_{1-2,3} &= -T_s\, \boldsymbol{A}_k^{-1} \left[\left(\boldsymbol{B}_k + \boldsymbol{C}_k \right) \boldsymbol{i}_{\alpha\beta,k} + \boldsymbol{d}_k \right] \\
\boldsymbol{\Phi}_{1-2,4} &= T_s\, \frac{\partial \boldsymbol{A}_k^{-1}}{\partial \vartheta_{E,k}} \left[\boldsymbol{u}_{\alpha\beta,k} - \boldsymbol{R}_s\, \boldsymbol{i}_{\alpha\beta,k} - \omega_{E,k} \left(\left(\boldsymbol{B}_k + \boldsymbol{C}_k \right) \boldsymbol{i}_{\alpha\beta,k} + \boldsymbol{d}_k \right) \right] + \cdots \\
&\quad \cdots + T_s\, \boldsymbol{A}_k^{-1} \left[-\omega_{E,k} \left(\left(\frac{\partial \boldsymbol{B}_k}{\partial \vartheta_{E,k}} + \frac{\partial \boldsymbol{C}_k}{\partial \vartheta_{E,k}} \right) \boldsymbol{i}_{\alpha\beta,k} + \frac{\partial \boldsymbol{d}_k}{\partial \vartheta_{E,k}} \right) \right] \\
\boldsymbol{\Phi}_{3-4,1-4} &= \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & T_s & 1 \end{bmatrix}
\end{aligned}
$$

where matrices $\frac{\partial \boldsymbol{A}_k^{-1}}{\partial \vartheta_{E,k}}$, $\frac{\partial \boldsymbol{B}_k}{\partial \vartheta_{E,k}}$, $\frac{\partial \boldsymbol{C}_k}{\partial \vartheta_{E,k}}$ and $\frac{\partial \boldsymbol{d}_k}{\partial \vartheta_{E,k}}$ are calculated in Appendix A.2.

Output Jacobian $\boldsymbol{H}_k$

By considering the state update function $h(\boldsymbol{x}_k)$ in (7.19) and deriving it with respect to the state vector $\boldsymbol{x}_k$, we obtain the $p \times n$ output Jacobian matrix

$$
\boldsymbol{H}_k = \left. \frac{\partial h(\boldsymbol{x}_k)}{\partial \boldsymbol{x}_k} \right|_{\boldsymbol{x}_{k|k-1}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \tag{8.20}
$$

which in this case is a constant and simple matrix that does not depend on the state vector $\boldsymbol{x}_k$.

8.1.2 Unscented Kalman Filter

The UKF was introduced by *Uhlmann et al.* in 1995 [39]. They observed that approximating a nonlinear function using only a single point (as in the EKF case) was insufficient for capturing the true behavior of the system. Instead, they proposed the use of multiple points, known as *sigma points*, which undergo a nonlinear transformation. The mean and covariance are then computed from these transformed points. This approach, known as the *unscented transform*, accurately captures the true mean and covariance up to the 2nd order for any nonlinearity, providing superior accuracy compared to the first-order approximation of the EKF. Importantly, the UKF achieves this improved accuracy while maintaining a computational complexity comparable to that of the EKF [94].

Although this method bears a superficial resemblance to particle filters, there are several fundamental differences [43]. The primary distinction is that the sigma points are not drawn at random; they are deterministically chosen to exhibit certain specific properties (e.g., a given mean and covariance). As a result, higher-order information about the distribution can be captured using a fixed, small number of points [49].

The UKF structure, outlined in the following equations, employs three tuning parameters associated with the distribution of sigma points, as discussed in [49]:

- α: a scaling factor that determines the spread of sigma points around $\boldsymbol{x}_{k-1|k-1}$. It is typically set as $1 \times 10^{-4} \leq \alpha < 1$.
- n: the state dimension, which, for the systems considered in (7.14) and (7.19), is $n = 4$.
- γ: another scaling factor. [49] suggests setting $\gamma = 0$, while [57] sets $\gamma = 3 - n$.
- β: related to prior knowledge of the probability density function of the state $\boldsymbol{x}$. For a Gaussian distribution, $\beta = 2$ is optimal.

All these tuning parameters are combined into the composite pair (λ, w_p), which is essential for defining the weights of the sigma points:

$$\lambda = \alpha^2(n + \gamma) - n, \quad w_p = \sqrt{n + \lambda}. \tag{8.21}$$

The parameters α, β, γ, and λ are all required to calculate the first weight w_p, which is used to create the matrix $\boldsymbol{\chi}$ containing the sigma points.

The covariance matrix $\boldsymbol{P}$ is guaranteed to be positive definite and can be decomposed into $\sqrt{\boldsymbol{P}}(\sqrt{\boldsymbol{P}})^T$ using the Cholesky algorithm [39], [49]. The columns of $\sqrt{\boldsymbol{P}}$ are used to construct a matrix $\boldsymbol{\chi}$ containing $2n + 1$ sigma points (or vectors):

$$\hat{\boldsymbol{\chi}}_{k-1|k-1} = \hat{\boldsymbol{X}}_{k-1|k-1} + w_p \left[0 \quad \sqrt{\boldsymbol{P}_{k-1|k-1}} \quad -\sqrt{\boldsymbol{P}_{k-1|k-1}} \right]. \tag{8.22}$$

Here, $\hat{\boldsymbol{X}}_{k-1|k-1}$ is an $(n \times (2n + 1))$ matrix formed by replicating the $\boldsymbol{x}_{k-1|k-1}$ state vector in $2n + 1$ columns. It can be observed that the first column of the sigma points matrix $\hat{\boldsymbol{\chi}}_{k-1|k-1}$ corresponds to the previous state estimate $\hat{\boldsymbol{x}}_{k-1|k-1}$, while the remaining $2n$ columns are the state estimate perturbed by the square root of the covariance matrix $\boldsymbol{P}_{k-1|k-1}$. Furthermore, the weight w_p is used to scale the perturbation of the covariance matrix.

There is another type of weight, w_i^m and w_i^c, used throughout the UKF algorithm. These weights are calculated during the initialization phase using the following formulas:

$$\begin{aligned} w_0^m &= \frac{\lambda}{n + \lambda}, \quad w_0^c = w_0^m + (1 - \alpha^2 + \beta), \\ w_i^m &= w_i^c = \frac{1}{2(n + \lambda)}, \quad i = 1, \ldots, 2n. \end{aligned} \tag{8.23}$$

These weights are responsible for weighting the average of the nonlinear function outputs and the covariance matrices. The sigma points matrix $\hat{\boldsymbol{\chi}}_{k-1|k-1}$ is then propagated through the nonlinear function $\boldsymbol{f}(\cdot)$ in the time update step, resulting in the predicted sigma points matrix $\hat{\boldsymbol{\chi}}_{k|k-1}$. The mean and covariance are then derived from this matrix to obtain the predicted state $\hat{\boldsymbol{x}}_{k|k-1}$ and the covariance matrix $\boldsymbol{P}_{k|k-1}$, respectively. Not all predicted sigma points are weighted equally in the empirical mean and covariance calculations, as the weights w_i^m and w_i^c depend on the sigma point index i (8.23).

In the following equations, the UKF algorithm is presented in detail:

Initialization

$$\hat{\boldsymbol{x}}_{0|0} = \mathbb{E}\left[\boldsymbol{x}_0\right] \tag{8.24}$$

$$\boldsymbol{P}_{0|0} = \mathbb{E}\left[\left(\boldsymbol{x}_0 - \mathbb{E}\left[\boldsymbol{x}_0\right]\right)\ \left(\boldsymbol{x}_0 - \mathbb{E}\left[\boldsymbol{x}_0\right]\right)^T\right] \tag{8.25}$$

$$\hat{\boldsymbol{\chi}}_{0|0} = \hat{\boldsymbol{X}}_{0|0} + w_p\left[0,\ \sqrt{\boldsymbol{P}_{0|0}},\ -\sqrt{\boldsymbol{P}_{0|0}}\right] \tag{8.26}$$

Time update

$$\hat{\boldsymbol{\chi}}_{k|k-1} = \boldsymbol{f}\left(\hat{\boldsymbol{\chi}}_{k-1|k-1}, \boldsymbol{u}_{k-1}\right) \tag{8.27}$$

$$\hat{\boldsymbol{x}}_{k|k-1} = \sum_{i=0}^{2n} w_i^m\ \hat{\boldsymbol{\chi}}_{k|k-1}(:, i) \tag{8.28}$$

$$\boldsymbol{P}_{k|k-1} = \sum_{i=0}^{2n} w_i^{\mathrm{c}}\left(\hat{\boldsymbol{\chi}}^-_{k|k-1}(:, i) - \hat{\boldsymbol{x}}_{k|k-1}\right)\left(\hat{\boldsymbol{\chi}}^-_{k|k-1}(:, i) - \hat{\boldsymbol{x}}_{k|k-1}\right)^{\mathrm{T}} + \boldsymbol{Q}_k \tag{8.29}$$

$$\hat{\boldsymbol{\Upsilon}}_{k|k-1} = \boldsymbol{f}\left(\hat{\boldsymbol{\chi}}^+_{k-1}\right) \tag{8.30}$$

$$\hat{\boldsymbol{y}}_{k|k-1} = \sum_{i=0}^{2n} w_i^m\ \hat{\boldsymbol{\Upsilon}}_{k|k-1}(:, i) \tag{8.31}$$

where $\hat{\boldsymbol{\chi}}^-_{k|k-1}$ and $\hat{\boldsymbol{\Upsilon}}_{k|k-1}$ are the predicted state and output sigma points matrices, respectively.

Measurement update

$$\boldsymbol{P}_k^{yy} = \sum_{i=0}^{2n} w_i^{\mathrm{c}}\left(\hat{\boldsymbol{\Upsilon}}_{k|k-1}(:, i) - \hat{\boldsymbol{y}}_{k|k-1}\right)\left(\hat{\boldsymbol{\Upsilon}}_{k|k-1}(:, i) - \hat{\boldsymbol{y}}_{k|k-1}\right)^{\mathrm{T}} + \boldsymbol{R}_k \tag{8.32}$$

$$\boldsymbol{P}_k^{xy} = \sum_{i=0}^{2n} w_i^{\mathrm{c}}\left(\hat{\boldsymbol{\chi}}^-_{k|k-1}(:, i) - \hat{\boldsymbol{x}}_{k|k-1}\right)\left(\hat{\boldsymbol{\Upsilon}}_{k|k-1}(:, i) - \hat{\boldsymbol{y}}_{k|k-1}\right)^{\mathrm{T}} \tag{8.33}$$

$$\boldsymbol{K}_k = \boldsymbol{P}_k^{xy}\left(\boldsymbol{P}_k^{yy}\right)^{-1} \tag{8.34}$$

$$\hat{\boldsymbol{x}}_{k|k} = \hat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}_k\left(\boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1}\right) \tag{8.35}$$

$$\boldsymbol{P}_{k|k} = \boldsymbol{P}_{k|k-1} - \boldsymbol{K}_k \boldsymbol{P}_{k|k-1} \boldsymbol{K}_k^{\mathrm{T}} \tag{8.36}$$

8.2 Monte Carlo Simulation

Before implementing the algorithm in practice, a Monte Carlo simulation is conducted to assess the estimation quality of both the EKF and UKF. The approach followed in this experiment aligns with the methodology presented in [43] for the polar-to-Cartesian transformation. In this context, the nonlinear function $\boldsymbol{f}(\cdot)$ in (7.14) exhibits nonlinearity solely in the first two rows, corresponding to the $(i_\alpha,\ i_\beta)$ time update. Conversely, the output function $\boldsymbol{h}(\cdot)$ in (7.19) is linear. Therefore, this analysis focuses exclusively on the first two rows of the dynamics function $\boldsymbol{f}(\cdot)$.

The magnetic model depicted in Fig. 8.1, derived through an elaboration of a series of measurements [88], serves as one of the inputs for the EKF/UKF prediction function, as evident in (7.14).

A SynRM working point with rated speed and rated torque is selected. Following the MTPA trajectory in Fig. 8.5, the corresponding state vector around which the Monte Carlo analysis is performed is determined as follows:

$$\boldsymbol{x}_k = \begin{bmatrix} I_n \cos(\vartheta_i^{\mathrm{MTPA}}) & I_n \sin(\vartheta_i^{\mathrm{MTPA}}) & p\,\omega_{M,n} & 0 \end{bmatrix}^T \tag{8.37}$$

where I_n and $\omega_{E,n}$ are the rated current and the rated speed of the machine, shown in Table 8.3.

The simulation starts by generating 1×10^4 state vectors normally distributed around (8.37), assuming the following standard deviations:

$$\boldsymbol{\sigma}_{\boldsymbol{x}_k} = \begin{bmatrix} 0.01 & 0.01 & 2.09 & 1 \end{bmatrix}^T \tag{8.38}$$

The standard deviation of the currents is set to 0.01 mA, approximately 20 rpm = 2 rad s^{-1} for the electrical speed, and 1 deg for the electrical position. These values were chosen to ensure that the generated state vectors are close to the working point, as the standard deviations are relatively small compared to the rated values, but also based on field experience.

The generated state vectors are processed using the function $\boldsymbol{f}(\cdot)$ in (7.14). The mean and covariance of these output vectors are then computed and compared with the results obtained from the EKF and UKF algorithms, as described in Subsect. 8.1.1 and Subsect. 8.1.2, respectively. The EKF and UKF algorithms are tuned using the same parameters listed in Subsect. 8.3.2.

The comparison is repeated by arbitrarily varying the standard deviation vector. However, it is observed that only the standard deviation of the electrical position, σ_{ϑ_E}, significantly affects the estimation quality.

Several simulations are performed to emphasize the discrepancies between the true values derived from the samples (MCS) and the results from the EKF or UKF, achieved by varying the standard deviations $\boldsymbol{\sigma}_{\boldsymbol{x}k}$ of the generated vector samples. The comparison does not reveal substantial differences between the mean and variance values derived from MCS and those obtained from the EKF or UKF, regardless of variations in all the standard deviations except for the standard deviation associated exclusively with the electrical position, σ_{ϑ_E}.

The tables presented in Table 8.1 and Table 8.2 summarize the results obtained with increasing values of σ_{ϑ_E} under different operating conditions. Each row begins with the UKF tuning parameter settings, followed by additional specifications.

In the first row of Table 8.1, a comparison is made between EKF and UKF results when a precise nonlinear magnetic model (NMM) is used for the Monte Carlo simulations (MCS). In contrast, a simple linear model (LMM) with constant

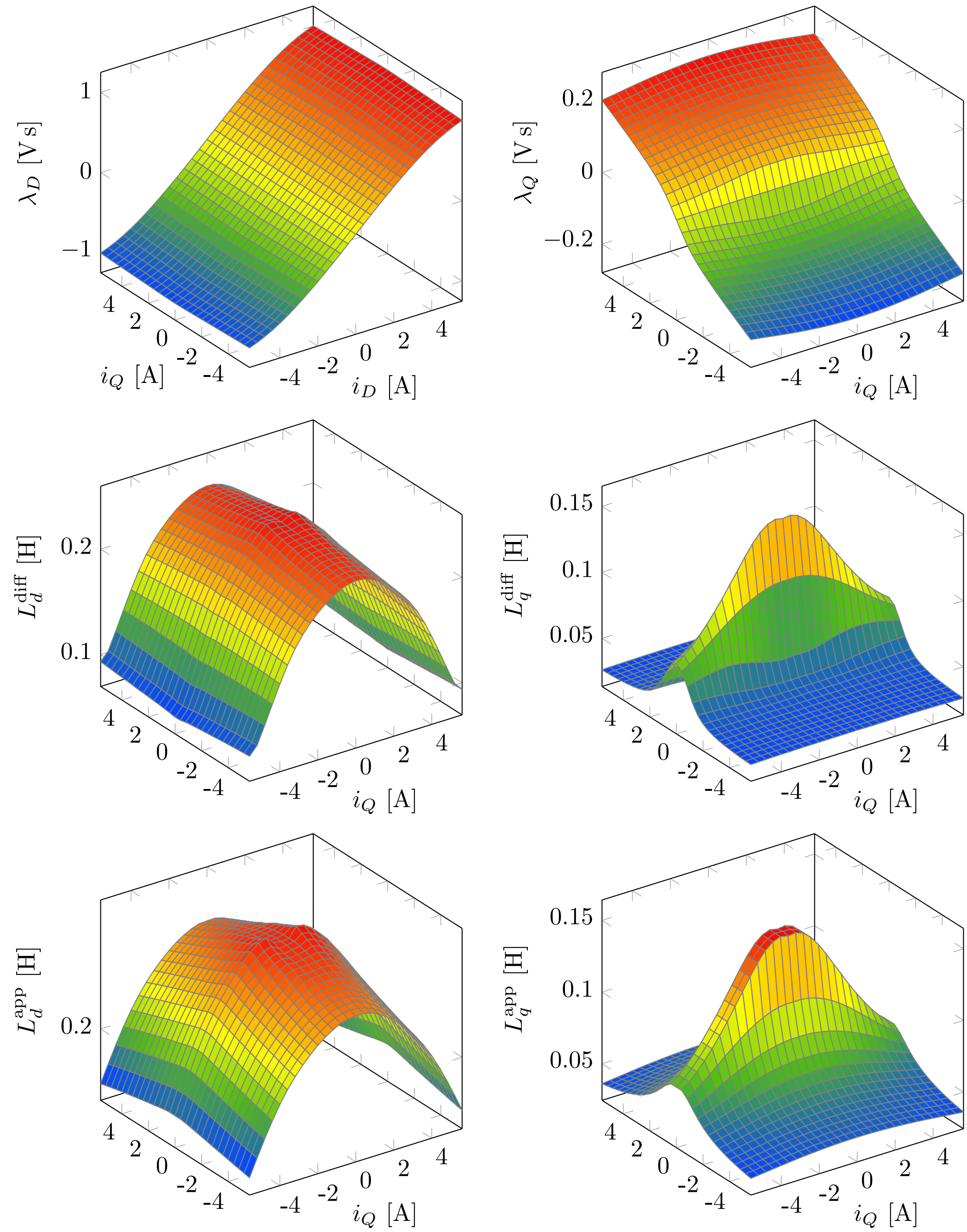


Figure 8.1 Flux linkages $\boldsymbol{\lambda}_{\mathrm{dq}}$, apparent inductances $\mathbf{L}_{\mathrm{dq}}^{\mathrm{app}}$, and differential inductances $\mathbf{L}_{\mathrm{dq}}^{\mathrm{diff}}$ of the SynRM used in the experiments (Table 8.3).

parameters is employed in the second row for both the Extended Kalman Filter (EKF) and Unscented Kalman Filter (UKF). The LMM implies that the prediction function $\boldsymbol{f}(\cdot)$ in (7.14) does not involve the four apparent and differential inductances $\boldsymbol{L}_{dq}^{\text{app}}$ and $\boldsymbol{L}_{dq}^{\text{diff}}$ maps shown in Fig. 8.1. Instead, only two constant values are used for the inductances:

$$L_d^{\text{app}} = L_d^{\text{diff}} = L_d = 230\,\text{mH} \tag{8.39}$$

$$L_q^{\text{app}} = L_q^{\text{diff}} = L_q = 60\,\text{mH} \tag{8.40}$$

as specified in the machine parameters in Table 8.3.

The first row of Table 8.2 introduces a resistance mismatch between the actual resistance value used in the MCS model and the prediction model employed in both the EKF and UKF. Specifically, a 20% smaller resistance value is applied in the KF prediction function. The second row of Table 8.2 documents the case where the dynamic state function $\boldsymbol{f}(\cdot)$ is integrated using the Matlab solver function `ode45` to propagate the states in MCS, while the EKF and UKF utilize the forward Euler discretization method with a sampling interval of $T_s = 125\,\mu\text{s}$.

The results merit some discussion. Across all cases, the UKF mean and variance remain closer to the true values than those of the EKF. This is particularly evident when the position standard deviation is high, which typically occurs at high speeds where the speed estimate is uncertain due to the absence of a precise mechanical model. Nevertheless, the discrepancies in the EKF remain negligible for practical purposes.

In conclusion, the comparison of the two nonlinear estimators using MCS was performed as part of due diligence; however, the results are not definitive. Unlike the case of the polar-to-Cartesian transformation cited in [43], where system nonlinearity highlighted the UKF's advantages over the EKF, the differences in the context of electric drives are less pronounced. This transition leads to the next section, where the comparison is extended to a real SynRM drive, addressing different yet equally important aspects.

Table 8.1 The mean and standard deviation ellipses for the true statistics (MCS, red), those calculated through linearization (EKF, blue) and those calculated by the Unscented Kalman Filter (UKF, black).

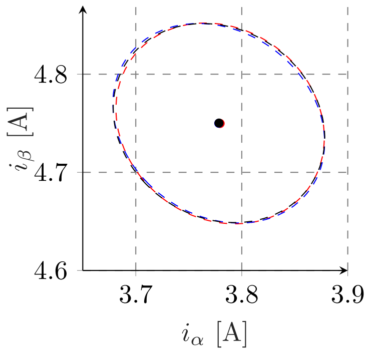
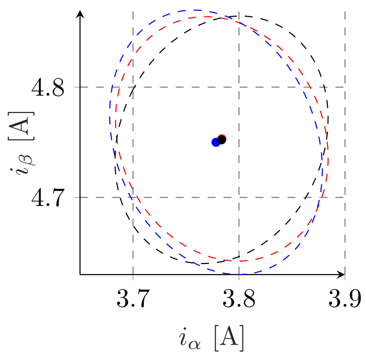
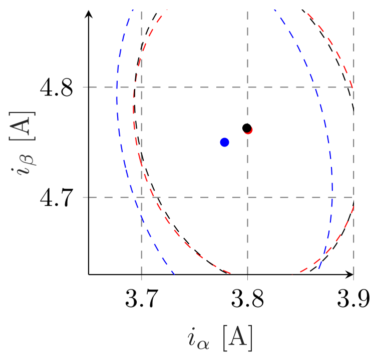
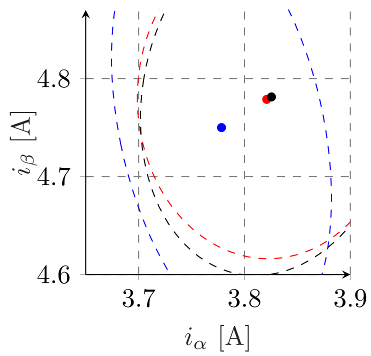
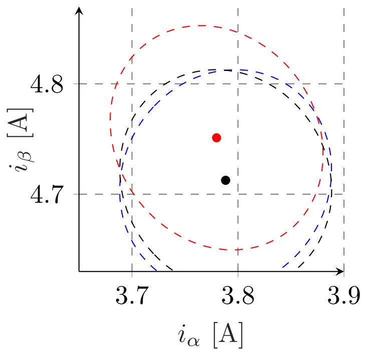
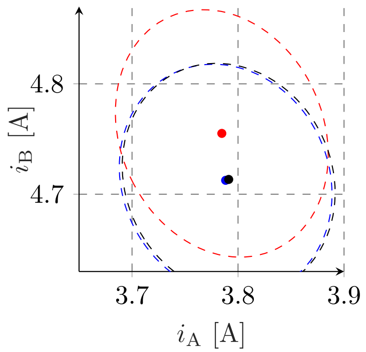
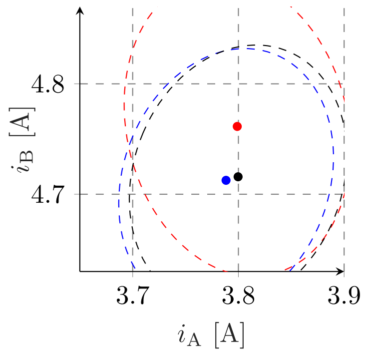
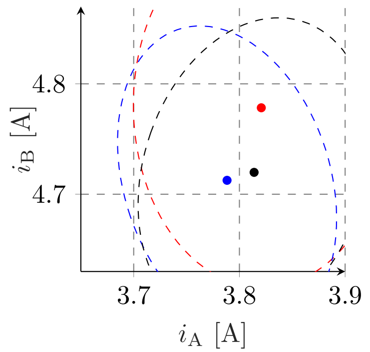
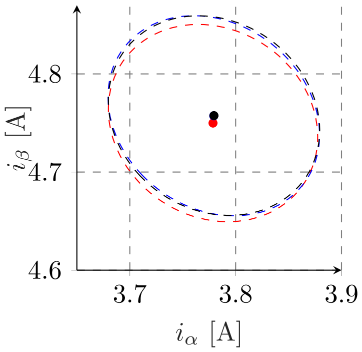
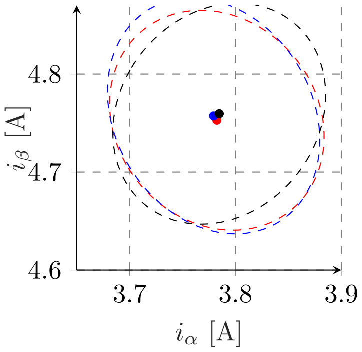
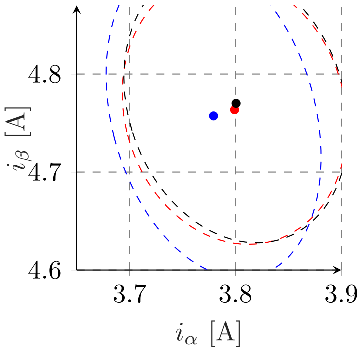
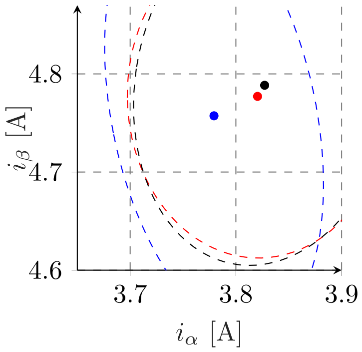
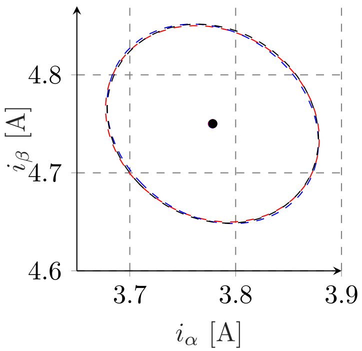
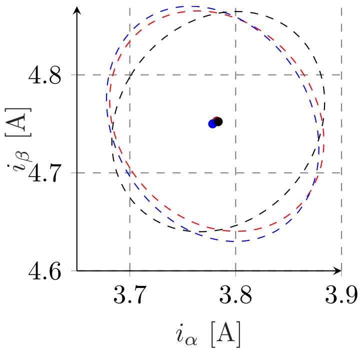
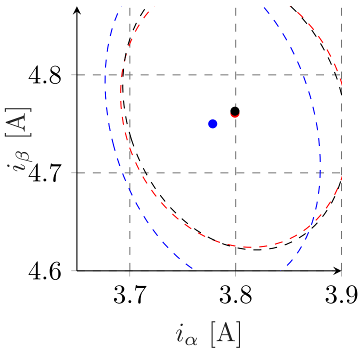
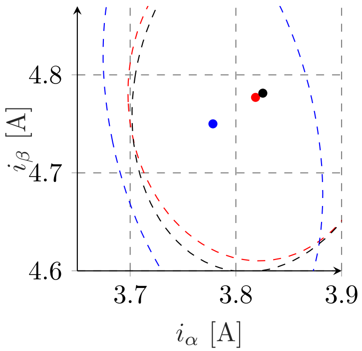
Settings	$\sigma_{\vartheta_E} = 1°$	$\sigma_{\vartheta_E} = 5°$	$\sigma_{\vartheta_E} = 10°$	$\sigma_{\vartheta_E} = 15°$
• $\alpha = 0.1$ • $\beta = 2$, • $k = -1$. • NMM				
• $\alpha = 0.1$ • $\beta = 2$, • $k = -1$. • LMM				

Table 8.2 The mean and standard deviation ellipses for the true statistics (MCS, red), those calculated through linearization (EKF, blue) and those calculated by the Unscented Kalman Filter (UKF, black).

Settings	$\sigma_{\vartheta_E} = 1°$	$\sigma_{\vartheta_E} = 5°$	$\sigma_{\vartheta_E} = 10°$	$\sigma_{\vartheta_E} = 15°$
• $\alpha = 0.1$ • $\beta = 2,$ • $k = -1.$ • $r_s^{\mathrm{KF}} = 0.8\ r_s$				
• $\alpha = 0.1$ • $\beta = 2,$ • $k = -1.$ • `ode45`				

8.3 Experimental results

In this section, the performance of the EKF and UKF in a sensorless SynRM drive is evaluated. The tests were conducted under typical operating conditions, including speed inversion, load handling, and four-quadrant operations. The results are compared in terms of the Mean Squared Error (MSE) of the estimated state variables. The robustness of the two filters is also assessed by evaluating their sensitivity to stator resistance mismatches and their ability to handle the rated torque at minimum speed. The execution time of the control algorithm elements is also analyzed.

8.3.1 Test bench

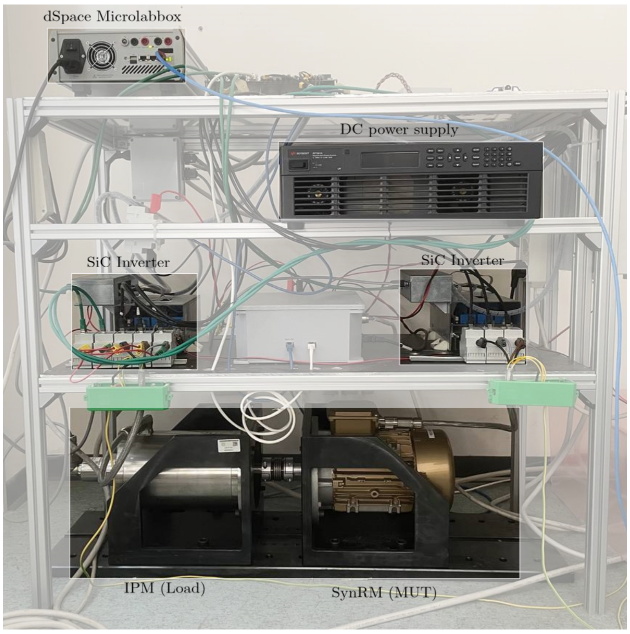


Figure 8.2 The experimental setup.

The experiments were conducted using the test bench depicted in Fig. 8.2, which includes a SynRM drive. The nameplate data of the SynRM are provided in Table 8.3. The SynRM was rigidly coupled to an IPMSM motor, which acted as a torque-controlled virtual load. Both motors were powered by three-phase voltage inverters, sharing a 450 V DC link, and operated with two-level space vector modulation at a switching frequency of 8 kHz. The control for both the SynRM and the IPMSM was implemented using a MicroLabBox dSpace Fast Control Prototyping (FCP) system.

8.3.2 Control Design

The drive block schematic is shown in Fig. 8.3. The speed control loop features a PI controller with a bandwidth of 5 Hz, while the current control loops are managed by two gain-scheduling PI controllers [95]. Their gains were tuned to achieve a constant

Table 8.3 System and SynRM parameters

Parameters	Symbol	Unit	Value
Nominal torque	τ_n	N m	9
Nominal current	I_n	A	6.5
Nominal speed	$\omega_{M,n}$	$\min^{-1}$	1000
Stator resistance	r_s	Ω	1.72
unsat. d-axis inductance	L_d	mH	230
unsat. q-axis inductance	L_q	mH	60
pole pairs	p	\	2
3-phase SiC Inverter (module: SKiiP 26ACM12V17)			
DC voltage	u_C	V	450
Inverted dead-time	t_d	ns	300
Switching frequency	f_s	kHz	8
Control Platform	dSPACE MicroLabBox		
Control frequency	f_c	kHz	8
Sampling time	T_s	µs	125

bandwidth of 200 Hz and a phase margin of 70 deg. All experiments were performed using estimated quantities as feedback.

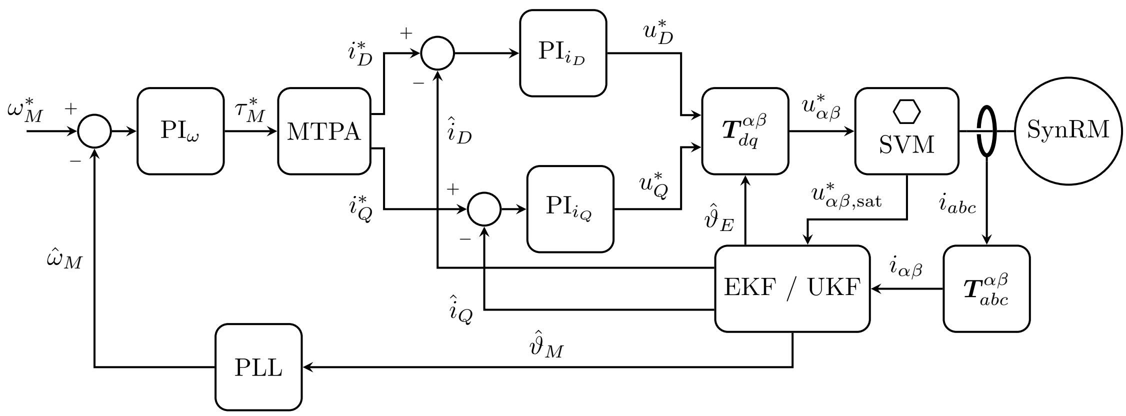


Figure 8.3 The proposed sensorless control of the SynRM.

The speed estimate is worth particular attention. The mathematical model of the Kalman Filter is based on the infinite inertia assumption, simplifying the speed differential equation in (7.10) to a null value. Consequently, the speed estimate relies solely on the EKF gain correction, resulting in both a noisy speed estimate and high sensitivity to the $\boldsymbol{Q}$-matrix weight.

As inferred from (7.10), the mechanical model of the EKF assumes infinite inertia, which simplifies the speed differential equation in (7.10) to a null value. Consequently, the speed update relies entirely on EKF gain correction, leading to a noisy speed estimate and high sensitivity to the $\boldsymbol{Q}$-matrix weight. Although such a dynamic speed estimate does not affect position estimation, it is not smooth enough for direct feedback control. To address this, a Phase-Locked Loop (PLL) filter is

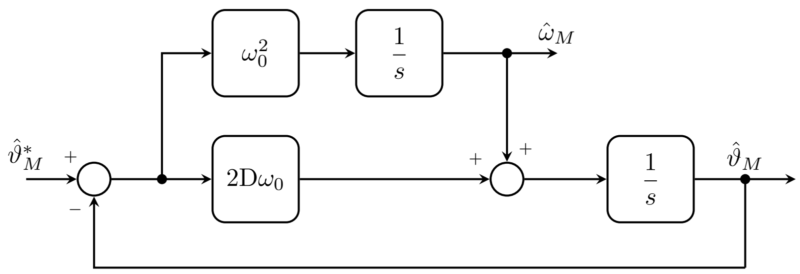


Figure 8.4 The PLL block diagram used within the experiments [96].

used to recover the speed $\hat{\omega}_M$ from the estimated position $\hat{\vartheta}_M$. The PLL was tuned with a characteristic frequency of $\omega_0 = 2\pi \cdot 50\,\mathrm{Hz}$ and a damping factor of $\mathrm{D} = 1$ [96]. For clarity, the PLL speed is used only for feedback and not in the EKF or UKF model. The block diagram of the PLL is shown in Fig. 8.4.

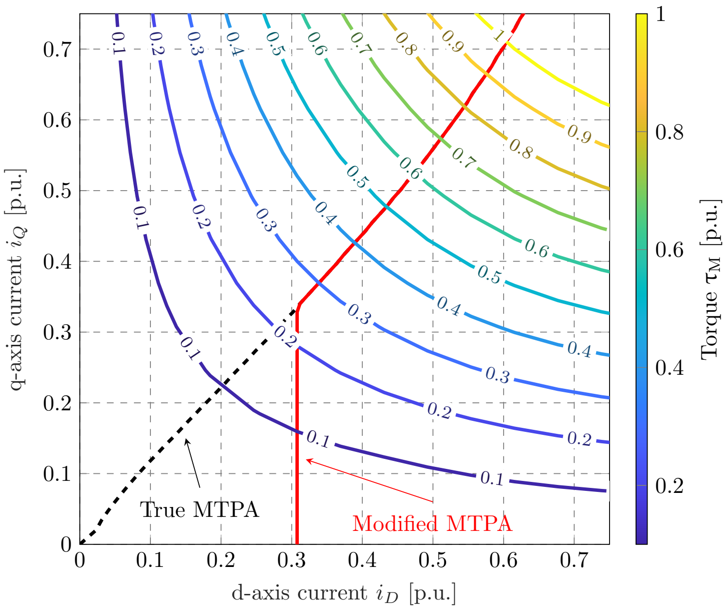


Figure 8.5 Real Maximum Torque Per Ampere trajectory (black curve), and modified MTPA trajectory (red curve) adopted within the experiments.

The current references are derived from the output of the speed controller using the standard MTPA curve, which can be easily calculated from an accurate motor model. The availability of this model is a prerequisite for this work. The MTPA curve is obtained by analyzing the current-torque map to identify the point where maximum torque is achieved with the minimum current magnitude.

As stated in [42], [97], the correct functioning of the sensorless algorithm requires supplying the SynRM with a minimum d-axis current value. In this experiment, this value was set to $I_D^{\min} = 2\,\mathrm{A} \simeq I_n/3$.

This value is necessary around zero speed and no load conditions. Under these conditions, without a minimum current vector, the KF innovation, whether it belongs to the EKF or the MSKF, is either zero or close to zero. As a result, the KF cannot recover any meaningful information from it, since the innovation becomes indistinguishable from the measurement noise, leading to unreliable estimation. To overcome this issue, setting a minimum current vector, below which the innovation can no longer be zero, provides a valid solution to enhance the observer reliability.

The chosen limit $I_D^{\min} = 2\,\mathrm{A} \simeq I_n/3$ was initially validated through simulations and subsequently confirmed during the experiments. Fig. 8.5 illustrates the modified MTPA trajectory used throughout the experiments.

In [42], an interesting investigation is presented on the relationship between the minimum current vector and the minimum bearable speed. It is shown that as the current limit increases, the minimum bearable speed decreases.

The EKF and UKF were fine-tuned using a trial-and-error approach, resulting in the following constant process noise covariance matrices:

$$\mathbf{Q}_{\mathrm{EKF}} = \mathrm{diag}(1 \times 10^{-5}, 1 \times 10^{-5}, 1, 0), \quad \mathbf{Q}_{\mathrm{UKF}} = \mathrm{diag}(1 \times 10^{-4}, 1 \times 10^{-4}, 1, 0).$$

The first two weights correspond to the $\alpha\beta$ currents, the third to the electrical speed, and the fourth to the electrical position. The trial-and-error procedure involves initially setting the weights to a small value and gradually increasing them until the estimation is fast enough to track the reference without introducing instability, yet not so noisy as to affect control.

The measurement noise covariance matrices are identical for both filters since the current measurements are shared between the two observers:

$$\mathbf{R}_{\mathrm{EKF}} = \mathbf{R}_{\mathrm{UKF}} = \mathrm{diag}(1 \times 10^{-3}, 1 \times 10^{-3}).$$

These values were determined by acquiring the output of the current sensors in the absence of current flow, when the system was off but the measurement was active. Subsequently, the empirical covariance was calculated.

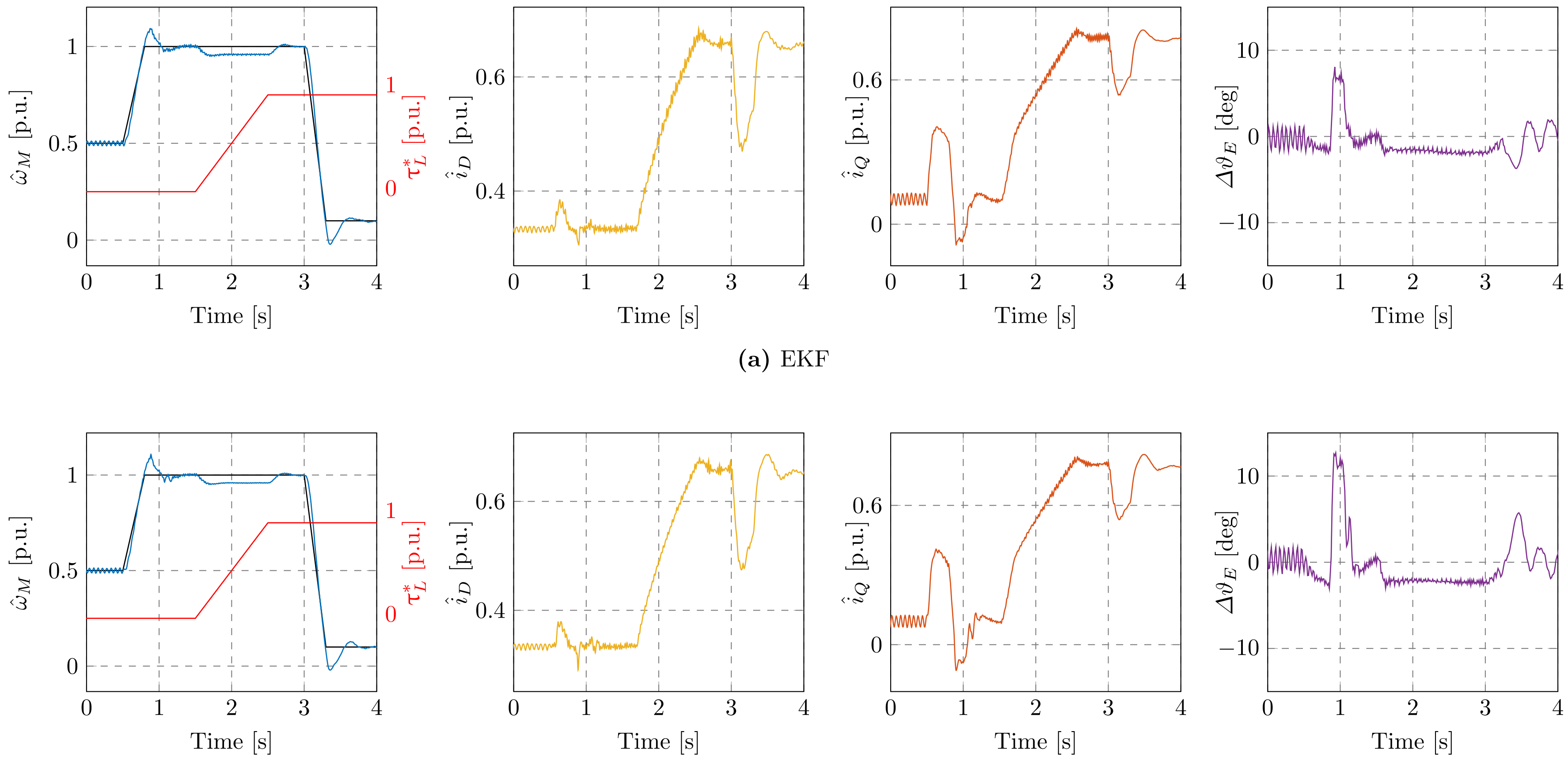
The UKF tuning parameters were determined after extensive testing and are set as follows: $\alpha = 0.1$, $k = -1$, and $\beta = 2$.

8.3.3 Performance Comparison

EKF and UKF are tested as position estimators under typical operating conditions of a modern electric drive. The tests include speed inversion, load handling, and four-quadrant operations. All estimated and measured state variables are collected at the control frequency rate (see Table 8.3) for post-processing. For each test and for each variable $[\omega_M,\ i_D,\ i_Q,\ \vartheta_E]$, the MSE is calculated as follows:

$$\mathrm{MSE}_{x_i} = \frac{1}{\mathrm{N}} \sum_{\mathrm{i}=1}^{N} (\hat{x}_{\mathrm{i}} - x_{\mathrm{i}}^{\mathrm{meas}})^2 \tag{8.41}$$

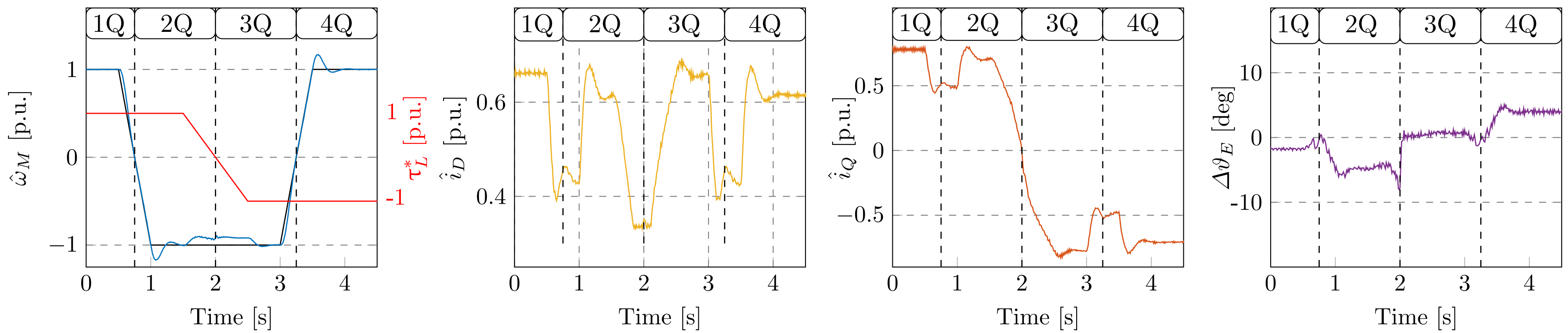
where N is the total number of samples during the test. The results of the experiments are grouped in Table 8.4. Comments on the MSE are included in the description of each test.



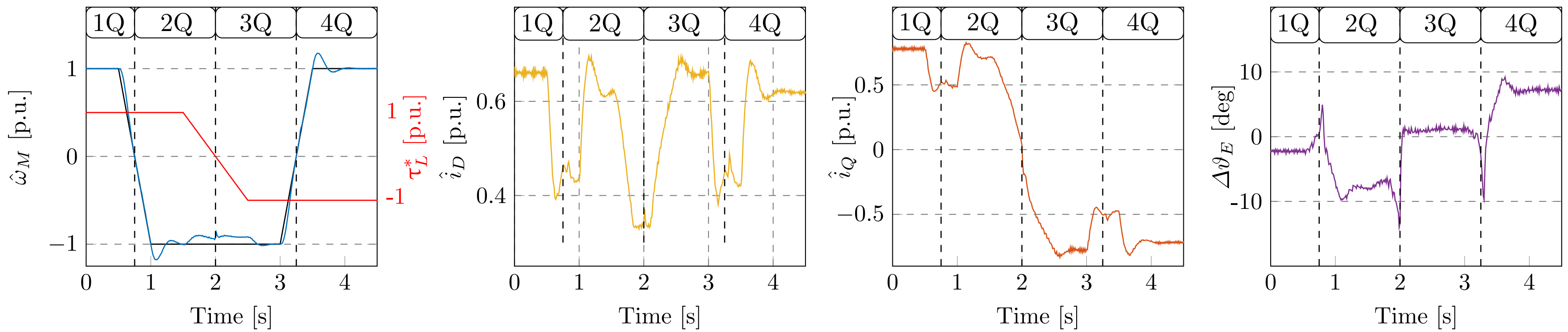
(a) EKF

(b) UKF

Figure 8.6 Minimum Speed at Rated Torque (MSRT) test for conventional EKF (upper row) and UKF (lower row). From left to right: mechanical speed $\hat{\omega}_M$, d-axis $\hat{i}_D$ and q-axis $\hat{i}_Q$ currents, electrical position error $\Delta\vartheta_E$. All values refer to their respective nominal values as listed in Table 8.3.



(a) EKF



(b) UKF

Figure 8.7 Four-Quadrant Operation (FQO) test for conventional EKF (upper row) and UKF (lower row). From left to right: mechanical speed ω_M, d-axis i_D and q-axis i_Q currents, electrical position error $\Delta\vartheta_E$. All values refer to their respective nominal values as listed in Table 8.3.

The Minimum Speed at Rated Torque (MSRT) test

The limitations of EKF and UKF techniques are most evident at low speeds. The lowest speed at rated torque is a common benchmark and is discussed in the sequel. The test results are presented in Fig. 8.6. In the leftmost figure, the black line represents the speed reference provided to the SynRM, while the red line depicts the torque reference applied to the IPM, which serves as a virtual load. The blue trace corresponds to the estimated speed, denoted as $\hat{\omega}_M$ and obtained from the PLL block of Fig. 8.3.

The test starts by setting a positive speed reference without load. Before initiating the transient load torque at $t = 0.5\,\mathrm{s}$, the SynRM accelerates from 0.5 p.u. to 1 p.u., still under no load conditions. Subsequently, the speed reference is reduced to 10% of the rated speed. This minimum threshold was determined through repeated tests until uncontrollable oscillations occurred at low speed.

Continuing from left to right, the other figures show the estimated d-q currents and the electrical position error $\Delta\vartheta_E = \hat{\vartheta}_E - \vartheta_E^{\mathrm{meas}}$, during the above speed and torque transients.

At full load and the lowest stable speed, both KFs exhibit position estimate oscillations (see Fig. 8.6a). The position error oscillates within $\pm 3\,\mathrm{deg}$ for both the EKF and the UKF. In both cases, the control effectively manages the speed downward ramp. Any further reduction in speed leads to instability with both KFs.

Referring to Table 8.4, it is evident that for the EKF, the MSE is consistently lower than for the UKF. This difference is particularly notable for the electrical position estimate, where the EKF MSE is less than half that of the UKF.

Table 8.4 Mean Squared Error (MSE)

Variable	Symbol	Unit	Test	EKF	UKF	Fig.
Mechanical speed	ω_M	rpm^2	MSRT	50.33	63.56	8.6
			FQO	149.37	164.03	8.7
			SSS	405.06	255	8.8
d-axis current	i_D	mA^2	MSRT	9642	15807	8.6
			FQO	47933	139912	8.7
			SSS	6938	1549	\
q-axis current	i_Q	mA^2	MSRT	9613	17553	8.6
			FQO	43084	138590	8.7
			SSS	9621	6255	\
Electrical position	ϑ_E	deg^2	MSRT	3.54	7.79	8.6
			FQO	10.95	32.56	8.7
			SSS	6.03	4.69	8.8

The Four-Quadrant Operation (FQO) test

The test results are presented in Fig. 8.7. The test begins with a positive nominal speed reference and a rated load torque, τ_L^*. While maintaining the same torque, a speed reversal is initiated at $t = 0.5\,\mathrm{s}$, transitioning from the first quadrant (labeled as 1Q) to the second quadrant (2Q). At $t = 1.5\,\mathrm{s}$, a load torque reversal is triggered, moving from $1,$ p.u. to $-1,$ p.u.. Finally, at $t = 3\,\mathrm{s}$, a second speed reversal is imposed to exit the third quadrant (3Q) and enter the fourth quadrant (4Q).

It should be noted that the electrical position estimation error tends to increase during speed zero crossings, a well-known limitation in sensorless techniques that do not utilize injected signals for estimation. For more details about observability, refer to Sect. 7.2.

However, when comparing Fig.8.7a and Fig.8.7b, it becomes evident that the EKF exhibits a lower $\Delta\vartheta_E$ than the UKF.

In the FQO test, the EKF also provides a better MSE compared to the UKF, with a difference of only $15\,\mathrm{rpm}^2$ for speed. In relation to the other variables, the MSE is one-third for i_D, i_Q, and for ϑ_E.

It is worth highlighting that the speed errors in Fig. 8.6 and Fig. 8.7 are present only during transients, i.e. when the speed reference signal changes or when the requested load torque varies. This is due to the PI controller and is not caused by the Kalman filters.

The Step Starting from Standstill (SSS) test

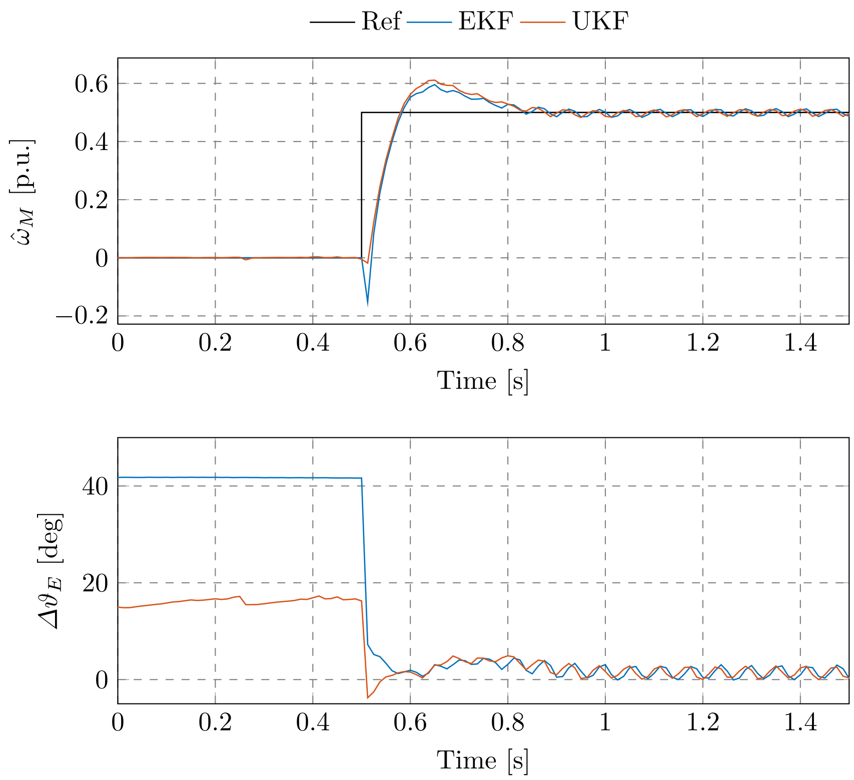


Figure 8.8 Step Starting from Standstill (SSS) test: mechanical speed (top) and electrical position error (bottom)

The test results are presented in Fig.8.8. The experiment evaluates the drive performance when the motor starts from a complete standstill and a step change in the speed reference is applied. In particular, at $t = 0.5\,\mathrm{s}$ a speed step is executed up to half of the nominal speed.

Both EKF and UKF are capable of initiating motion from a standstill condition and exhibit nearly identical responses. Prior to the step command, the idle EKF position ϑ_e appears stable, with an error of approximately 40 deg. In contrast, the position error derived from the UKF oscillates around 17 deg. However, following the speed step, the position errors for both the EKF and the UKF stabilize at around 2 deg.

At standstill, with zero speed and at no load, the system is not observable, as detailed in Sect. 7.2, therefore the position error $\varDelta\vartheta_E$ before the speed step at $t = 0.5\,\mathrm{s}$ is random. Anyway, as soon as the system is solicited by a speed reference change, the estimation errors rapidly converge to zero.

It should be noted that the EKF estimated speed at $t = 0.5\,\mathrm{s}$ becomes negative. If this were the actual speed of the machine, it could pose a problem for certain applications where the load is coupled with the motor and only one direction of rotation is allowed. However, the actual speed of the machine remains at zero, and the negative speed is merely a transient phenomenon caused by poor estimation at the beginning.

When it come to the MSE Table 8.4, the MSE UKF values are lower than in the EKF case, because the EKF MSE value is significantly influenced by the incorrect estimated speed at the beginning of the speed step. It should be noted that the MSE values are not computed throughout the duration of the test, but only for the time interval where the speed step is applied, since the system is not observable at standstill.

Sensitivity to resistance mismatch

Although an accurate magnetic model is assumed to be available, it is also assumed that the resistance may vary with temperature.

Table 8.5 Resistance mismatch test

R_s^{KF}/R_s	Mean ($\Delta\vartheta_E$)		Unit	MSE (ϑ_e)		Unit
	EKF	UKF		EKF	UKF	
1	−0.01	−0.63	deg	1.58	1.92	deg^2
0.9	3.88	3.72		16.6	15.46	
0.8	7.23	6.96		53.9	49.99	

The test results, including the mean electrical position error and MSE, are presented in Table 8.5. The experiment aims at evaluating the robustness of both the EKF and UKF against stator resistance mismatches. The resistance mismatch arises from temperature variations, which causes an increase in the actual stator resistance compared to the value used in the model (7.14). Three experiments are carried out under steady-state conditions, with a speed reference of 0.1 p.u. and rated load torque, each with a progressively increasing resistance mismatch. The resistance mismatch is expressed as the ratio r_s^{KF}/r_s, with obvious symbol meaning. r_s corresponds to the nominal value documented in Table 8.3.

When the resistance mismatch is null, that is, $r_s^{\mathrm{KF}}/r_s = 1$, the EKF provides more accurate results. However, as the mismatch increases, the UKF offers slightly more precise estimates in terms of position error. The mean value shown in Table 8.5 is calculated by averaging the electrical position error over 1600 samples (0.2 s). It should be noted that both the mean and the MSE of $\Delta\vartheta_E$ are significantly affected by a resistance mismatch, and this effect is the same for both KFs. Additional tests were performed by considering a resistance variation up to 40%. In case of copper, this corresponds to an ambient temperature variation of 100°C, which is not uncommon in certain SynRM applications, such as pumps located inside car hoods. Both the EKF and UKF fail to maintain convergence, and the tests is subsequently aborted. For such large resistance variations, either an online parameter tracking technique or an adaptive approach to the Kalman filter parameters is necessary. These techniques, which are currently receiving significant research attention, should be considered for a complete implementation in cases involving extreme temperature ranges.

Linear Magnetic model

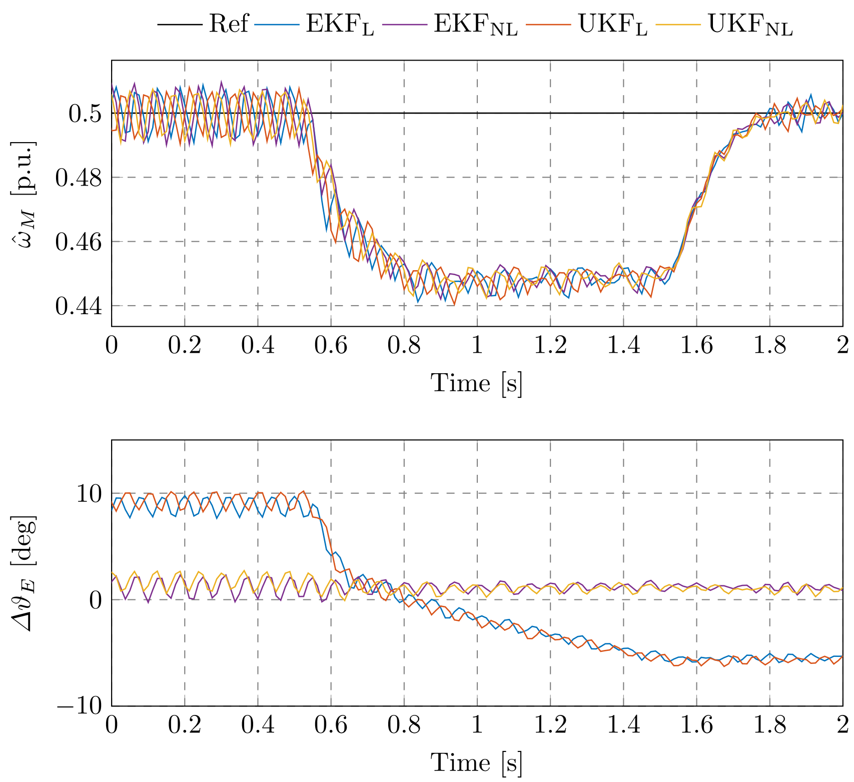


Figure 8.9 Constant Magnetic model test: mechanical speed (top) and electrical position error (bottom). This test is performed in open loop. Here the subscripts L and NL refer to the linear and nonlinear magnetic models, respectively.

In this test, the motor inductances L_d^x, L_q^x with (x = diff, app) are kept constant in the prediction function $\boldsymbol{f}(\cdot)$ in (7.14). Their values are equal to those shown in Table 8.3.

Fig. 8.9 depicts the experiment conducted using the measured speed and position as feedback in the control loops. This was necessary because, when using the linear magnetic model, both filters fail the test as the load torque increases. The position estimation error grows larger causing the control to be lost.

From $t = 0.5\,\mathrm{s}$ to $t = 1.5\,\mathrm{s}$ the load torque is increased from $0\,\mathrm{N\,m}$ to $9\,\mathrm{N\,m}$ (rated value). The results show that the estimation accuracy is higher when the inductances are allowed to vary, as expected. In detail, with constant L_d and L_q, during the torque transient the estimation error increases in both filters because the motor inductances undergo saturation, while the model does not account for it. With the aid of the nonlinear magnetic model, both KFs exhibit smaller (and constant) the position error, in the range of 1 deg.

Therefore, the inclusion of an accurate magnetic model in a SynRM observer is not a choice but a necessity.

Pull-Out Curve (POC) test

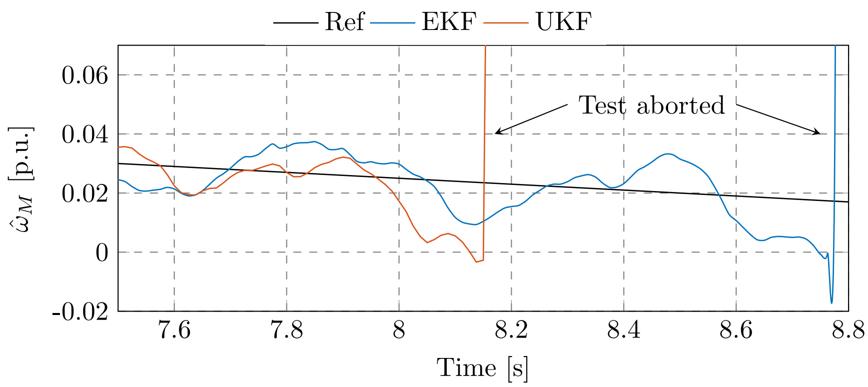


Figure 8.10 Pull-Out Curve (POC) test: mechanical speed. All values are referenced to their respective nominal values as listed in Table 8.3.

The test results are presented in Fig. 8.10. This experiment aims to evaluate the lowest stable steady-state speed at rated torque. An downhill slow speed ramp is taken as reference, starting from $\omega_\mathrm{m} = 0.1$ p.u. ($100\,\mathrm{rpm}$) and decreasing to $\omega_\mathrm{m} = 0\,\mathrm{rpm}$ over a duration of $10\,\mathrm{s}$, while the virtual load maintains the rated torque of $\tau_\mathrm{M} = 9\,\mathrm{N\,m}$. Analysis of the results shown in Fig. 8.10 reveals that the UKF works correctly until $\omega_\mathrm{m} = 0.033$ p.u. ($33\,\mathrm{rpm}$). In contrast, the EKF maintains the rated torque even further, up to $\omega_\mathrm{m} = 0.019$ p.u. ($19\,\mathrm{rpm}$).

Execution Time

The execution times are listed in Table 8.6. Although these times depend on the hardware platform, in our case a dSpace Microlabbox, the comparison reveals that the UKF execution time is 4.5 times greater than that of the EKF. This difference stems from the computational overhead of the UKF, which involves evaluating the nonlinear prediction function (7.14) for each sigma point. In contrast, the EKF computation entails calculating the 4×4 Jacobian matrix in (8.12). Considering

Table 8.6 Execution times of the control algorithm elements

Routine	Execution Time
Control cycle	125 µs
Speed & Current Control	22 µs
EKF	14 µs
UKF	63 µs

the KF calculation times alongside the time required for control tasks, it becomes evident that the control period of an electric drive using the EKF can be significantly shorter, offering various well-known associated benefits.

CHAPTER 9

Maximum Likelihood Estimation Kalman Filter

One of the main difficulties faced in commissioning a sensorless AC motor drive based on an EKF still remains the fine tuning of the noise covariance matrices of the model used for the state estimation. An inadequate choice penalizes the overall system performance, and it is the most common cause of the filter divergence.

Aimed to overcome these limitations, in this chapter the advantages of using an Adaptive Extended Kalman Filter (AEKF) algorithm based on Maximum Likelihood Estimation (MLE) for the simultaneous estimation of the system state and the unknown noise covariance matrices, are presented. By continuously adjusting the filter parameters to the actual operating conditions, the proposed system performs as conventional solutions relying on manually tuned EKFs.

In Sect. 9.1, the Maximum Likelihood Estimation method is introduced, and the formula for the online estimation of the process noise covariance matrix is presented. Then, in Sect. 9.2, the complete filter equations are provided. Finally, the experimental results are discussed in Sect. 9.3.

9.1 Maximum Likelihood Estimation

To ensure clarity and readability, only the essential formulas are presented, while the detailed derivation is provided in Appendix B.

Consider the following nonlinear dynamical system:

$$\begin{cases} \boldsymbol{x}_{k+1} = \boldsymbol{f}\left(\boldsymbol{x}_k,\, \boldsymbol{u}_k,\, \boldsymbol{w}_k\right) & \boldsymbol{x}_k \in \mathbb{R}^{n\times 1}, \quad \boldsymbol{u}_k \in \mathbb{R}^{m\times 1} \\ \boldsymbol{y}_k = \boldsymbol{h}\left(\boldsymbol{x}_k,\, \boldsymbol{v}_k\right) & \boldsymbol{y}_k \in \mathbb{R}^{p\times 1} \end{cases} \tag{9.1}$$

where:

- the nonlinear functions $\boldsymbol{f}(\cdot)$, $\boldsymbol{h}(\cdot)$ and the input $\boldsymbol{u}_k$ are known a priori.
- $\boldsymbol{w}_k$ and $\boldsymbol{v}_k$ are zero-mean Gaussian random processes with known covariances:

$$\boldsymbol{w}_k \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q}_k)\,, \qquad \boldsymbol{v}_k \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}_k) \tag{9.2}$$

- $\boldsymbol{w}_k$ and $\boldsymbol{v}_k$ are white noises:

$$\mathbb{E}\left[\boldsymbol{w}_k \boldsymbol{w}_h^T\right] = \boldsymbol{Q}_k\, \delta[k-h]\,, \qquad \mathbb{E}\left[\boldsymbol{v}_k \boldsymbol{v}_h^T\right] = \boldsymbol{R}_k\, \delta[k-h] \tag{9.3}$$

  where $\delta[k-h]$ is the Kronecker delta function, which is equal to 1 if $k = h$ and 0 otherwise.
- $\boldsymbol{w}_k$ and $\boldsymbol{v}_k$ are mutually uncorrelated:

$$\mathbb{E}\left[\boldsymbol{w}_k \boldsymbol{v}_h^T\right] = \boldsymbol{0} \tag{9.4}$$

- the initial state $\boldsymbol{x}_0$ is a normally distributed random variable with mean $\bar{\boldsymbol{x}}_0$ and covariance $\boldsymbol{P}_{0|0}$:

$$\boldsymbol{x}_0 \sim \mathcal{N}(\bar{\boldsymbol{x}}_0, \boldsymbol{P}_{0|0}) \tag{9.5}$$

- $\boldsymbol{w}_k$ and $\boldsymbol{v}_k$ are uncorrelated with the initial state $\boldsymbol{x}_0$:

$$\mathbb{E}\left[\boldsymbol{x}_0 \boldsymbol{w}_k^T\right] = \boldsymbol{0}\,, \qquad \mathbb{E}\left[\boldsymbol{x}_0 \boldsymbol{v}_k^T\right] = \boldsymbol{0} \tag{9.6}$$

9.1.1 Maximum Likelihood function

The problem considered in this section is to obtain a simultaneous and optimal estimation of the system state and the covariance matrices of the process noise. In the following, the $\boldsymbol{Q}_k$ matrix is assumed *diagonal*:

$$\boldsymbol{Q}_k = \begin{bmatrix} q_{11,k} & & 0 \\ & \ddots & \\ 0 & & q_{nn,k} \end{bmatrix} \tag{9.7}$$

and the parameters of the process noise covariance matrix is grouped into the following vector of *process noise covariance parameters*:

$$\boldsymbol{\xi}^T \triangleq [q_{11,k}\,,\ \ldots\,,\ q_{nn,k}]^T \tag{9.8}$$

The objective is to estimate the system state $\boldsymbol{x}_k$ and the noise covariance parameters $\boldsymbol{\xi}$ by maximizing the following Probability Density Function (PDF), collectively referred to as the likelihood function (see equation (B.60) in Appendix B):

$$\begin{aligned} \ell_k(\boldsymbol{x}_k,\, \boldsymbol{\xi},\, \mathcal{U}_k,\, \mathcal{Y}_k) &\triangleq p(\boldsymbol{x}_k,\, \mathcal{Y}_k \,|\, \boldsymbol{\xi},\, \mathcal{U}_k) \\ &= p(\boldsymbol{x}_k \,|\, \boldsymbol{\xi},\, \mathcal{U}_k, \mathcal{Y}_k)\; p(\mathcal{Y}_k \,|\, \boldsymbol{\xi},\, \mathcal{U}_k) \end{aligned} \tag{9.9}$$

where $\mathcal{U}_k$ and $\mathcal{Y}_k$ are the sets of inputs and outputs up to time k, respectively, defined as:

$$\mathcal{U}_k \triangleq \{u_i\,:\, i = 0, \ldots, k\}\,, \qquad \mathcal{Y}_k \triangleq \{y_i\,:\, i = 0, \ldots, k\} \tag{9.10}$$

In (9.9), the PDFs of the state $\boldsymbol{x}_k$ given $\boldsymbol{\xi}$, $\mathcal{U}_k$, $\mathcal{Y}_k$ is

$$p(\boldsymbol{x}_k \,|\, \boldsymbol{\xi},\, \mathcal{U}_k,\, \mathcal{Y}_k) \;=\; \frac{1}{\sqrt{(2\pi)^n \det \boldsymbol{P}_{k|k}}} \exp\left[-\tfrac{1}{2}(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k})^T \boldsymbol{P}_{k|k}^{-1}(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k})\right] \quad (9.11)$$

and the PDF of the output $\boldsymbol{y}_k$ given $\boldsymbol{\xi}$, $\mathcal{U}_{k-1}$, $\mathcal{Y}_{k-1}$ is

$$p(\boldsymbol{y}_k \,|\, \boldsymbol{\xi},\, \mathcal{U}_{k-1},\, \mathcal{Y}_{k-1}) \;=\; \frac{1}{\sqrt{(2\pi)^p \det \boldsymbol{\Lambda}_k}} \exp\left[-\tfrac{1}{2}(\boldsymbol{y}_k - \boldsymbol{H}_k\hat{\boldsymbol{x}}_{k|k-1})^T \boldsymbol{\Lambda}_k^{-1}(\boldsymbol{y}_k - \boldsymbol{H}_k\hat{\boldsymbol{x}}_{k|k-1})\right] \quad (9.12)$$

where $\boldsymbol{\Lambda}_k$ represents the covariance matrix of the filter innovation $\tilde{\boldsymbol{y}}_k$, defined in (B.46) as

$$\boldsymbol{\Lambda}_k \;=\; \boldsymbol{H}_k \boldsymbol{P}_{k|k} \boldsymbol{H}_k^T, +, \boldsymbol{R}_k \quad (9.13)$$

The complete derivation of (9.11) and (9.12) is provided in Sect. B.2.

Similarly to (9.9), all measurements starting from the initial time must be considered in the maximization process. This approach, however, results in a filter with *growing-length memory*, which is impractical for online estimation [68]. To facilitate the implementation of estimators with *fixed-length memory*, the conditional PDF in (9.9) is modified to account for only a finite set of measurements. This approach is detailed in equation (10-19) and discussed on pp. 76-77 of [70].

$$\ell_k(\boldsymbol{x}_k,\, \boldsymbol{\xi},\, \mathcal{U}_k^N,\, \mathcal{Y}_k^N) \;\triangleq\; p(\boldsymbol{x}_k \,|\, \boldsymbol{\xi},\, \mathcal{U}_k, \mathcal{Y}_k) \prod_{i=k-N+1}^{k} p(\boldsymbol{y}_i \,|\, \boldsymbol{\xi},\, \mathcal{U}_{i-1},\, \mathcal{Y}_{i-1}) \quad (9.14)$$

where, the sets of inputs and outputs are limited to the most recent N samples

$$\mathcal{U}_k^N \triangleq \{u_i \,:\, i = k-N+1,\, \ldots,\, k\}\ , \qquad \mathcal{Y}_k^N \triangleq \{y_i \,:\, i = k-N+1,\, \ldots,\, k\} \quad (9.15)$$

In order to find the solution which maximizes (9.14), its natual logarithm is considered

$$\begin{aligned} L_k(\boldsymbol{x}_k,\, \boldsymbol{\xi},\, \mathcal{U}_k,\, \mathcal{Y}_k) \;&\triangleq\; \ln \ell_k(\boldsymbol{x}_k,\, \boldsymbol{\xi},\, \mathcal{U}_k,\, \mathcal{Y}_k) \\ &\triangleq\; -\tfrac{1}{2}\ln\det \boldsymbol{P}_{k|k} \;-\; \tfrac{1}{2}\tilde{\boldsymbol{x}}_{k|k}^T\, \boldsymbol{P}_{k|k}^{-1}\, \tilde{\boldsymbol{x}}_{k|k} \;+\; \cdots \\ &\quad\; \cdots + \sum_{i=k-N+1}^{k} \left[-\tfrac{1}{2}\ln\det \boldsymbol{\Lambda}_i - \tfrac{1}{2}\, \tilde{\boldsymbol{y}}_{i|i-1}^T\, \boldsymbol{\Lambda}_i^{-1}\, \tilde{\boldsymbol{y}}_{i|i-1}\right] \;+\; \cdots \\ &\quad\; \cdots + \text{const} \end{aligned} \quad (9.16)$$

9.1.2 Determination of the maximum likelihood solution

In the following we consider the maximization of the likelihood function (9.16). The value $(\hat{\boldsymbol{x}}_{k|k},\, \hat{\boldsymbol{\xi}}_k)$ that maximizes the function is determined by equating its derivatives with respect to $\boldsymbol{x}_k$ and $\boldsymbol{\xi}$ to zero.

The derivative with respect to $\boldsymbol{x}_k$ is equal to[1]:

$$\frac{\partial L_k(\boldsymbol{x}_k, \boldsymbol{\xi})}{\partial \boldsymbol{x}_k} \;=\; -\boldsymbol{P}_{k|k}^{-1}\,(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k}) \quad (9.17)$$

[1]In the following, the notation $L_k(\boldsymbol{x}_k, \boldsymbol{\xi}, \mathcal{U}_k,\, \mathcal{Y}_k)$ is shortened as $L_k(\boldsymbol{x}_k, \boldsymbol{\xi})$.

ad equating it to zero yields:

$$\frac{\partial L_k(\boldsymbol{x}_k, \boldsymbol{\xi})}{\partial \boldsymbol{x}_k} = 0 \qquad \Rightarrow \qquad \boldsymbol{x}_k = \hat{\boldsymbol{x}}_{k|k} \tag{9.18}$$

The derivative of (9.16) with respect to the j^{th} component of $\boldsymbol{\xi}$, denoted as ξ_j is equal to

$$\begin{aligned} \frac{\partial L_k}{\partial \xi_j} = -\frac{1}{2}\Bigg\{ & \mathrm{Tr}\left[\left(\boldsymbol{P}_{k|k}^{-1} - \boldsymbol{P}_{k|k}^{-1}\tilde{\boldsymbol{x}}_{k|k}\tilde{\boldsymbol{x}}_{k|k}^T\boldsymbol{P}_{k|k}^{-1}\right)\frac{\partial \boldsymbol{P}_{k|k}}{\partial \xi_j}\right] - 2\,\tilde{\boldsymbol{x}}_{k|k}^T\,\boldsymbol{P}_{k|k}^{-1}\,\frac{\partial \hat{\boldsymbol{x}}_{k|k}}{\partial \xi_j} + \cdots \\ & \cdots + \sum_{i=k-N+1}^{k} \mathrm{Tr}\left[\left(\boldsymbol{\Lambda}_i^{-1} - \boldsymbol{\Lambda}_i^{-1}\tilde{\boldsymbol{y}}_{i|i-1}\tilde{\boldsymbol{y}}_{i|i-1}^T\boldsymbol{\Lambda}_i^{-1}\right)\frac{\partial \boldsymbol{\Lambda}_i}{\partial \xi_j}\right] + \cdots \\ & \cdots - 2\,\tilde{\boldsymbol{y}}_{i|i-1}^T\,\boldsymbol{\Lambda}_i^{-1}\,\boldsymbol{H}_i\,\frac{\partial \hat{\boldsymbol{x}}_{i|i-1}}{\partial \xi_j}\Bigg\} \end{aligned} \tag{9.19}$$

by equating it to zero it follows that:

$$\begin{aligned} \frac{\partial L_k(\boldsymbol{x}_k, \boldsymbol{\xi})}{\partial \xi_j} = -\frac{1}{2}\Bigg\{ & \mathrm{Tr}\left[\boldsymbol{P}_{k|k}^{-1}\frac{\partial \boldsymbol{P}_{k|k}}{\partial \xi_j}\right] + \cdots \\ & \cdots + \sum_{i=k-N+1}^{k} \mathrm{Tr}\left[\left(\boldsymbol{\Lambda}_i^{-1} - \boldsymbol{\Lambda}_i^{-1}\tilde{\boldsymbol{y}}_{i|i-1}\tilde{\boldsymbol{y}}_{i|i-1}^T\boldsymbol{\Lambda}_i^{-1}\right)\frac{\partial \boldsymbol{\Lambda}_i}{\partial \xi_j}\right] - \cdots \\ & \cdots - 2\,\tilde{\boldsymbol{y}}_{i|i-1}^T\,\boldsymbol{\Lambda}_i^{-1}\,\boldsymbol{H}_i\,\frac{\partial \hat{\boldsymbol{x}}_{i|i-1}}{\partial \xi_j}\Bigg\} = 0 \end{aligned} \tag{9.20}$$

Unfortunately, (9.20) can only be solved using numerical methods, which require computational times that are incompatible with online estimation. Alternatively, simplifications can be introduced to address this issue [98], [99]. In the next section, an explicit suboptimal solution is proposed for the online estimation of process and measurement noise covariance matrices.

9.1.3 Explicit suboptimal solution

In this approach, the equation (9.20) is simplified by assuming that:

$$\mathrm{Tr}\left[\boldsymbol{P}_{k|k}^{-1}\frac{\partial \boldsymbol{P}_{k|k}}{\partial \xi_j}\right] = 0\,, \qquad \frac{\partial \hat{\boldsymbol{x}}_{i|i-1}}{\partial \xi_j} = 0 \tag{9.21}$$

namely the 1$^{\text{st}}$ term in (9.20) is dominated by the 2$^{\text{nd}}$, and the state prediction $\hat{\boldsymbol{x}}_{i|i-1}$ is almost unaffected by the parameter ξ_j [98], [99]. The approximated equation is:

$$\begin{aligned} \frac{\partial \tilde{L}_k(\boldsymbol{x}_k, \boldsymbol{\xi})}{\partial \xi_j} &= -\frac{1}{2}\sum_{i=k-N+1}^{k} \mathrm{Tr}\left[\left(\boldsymbol{\Lambda}_i^{-1} - \boldsymbol{\Lambda}_i^{-1}\tilde{\boldsymbol{y}}_{i|i-1}\tilde{\boldsymbol{y}}_{i|i-1}^T\boldsymbol{\Lambda}_i^{-1}\right)\frac{\partial \boldsymbol{\Lambda}_i}{\partial \xi_j}\right] \\ &= -\frac{1}{2}\sum_{i=k-N+1}^{k} \mathrm{Tr}\left[\Delta\boldsymbol{\Lambda}_i^{-1}\frac{\partial \boldsymbol{\Lambda}_i}{\partial \xi_j}\right] = 0 \end{aligned} \tag{9.22}$$

where $\Delta\boldsymbol{\Lambda}_i$ is defined as:

$$\Delta\boldsymbol{\Lambda}_i^{-1} \triangleq \boldsymbol{\Lambda}_i^{-1} - \boldsymbol{\Lambda}_i^{-1}\tilde{\boldsymbol{y}}_{i|i-1}\tilde{\boldsymbol{y}}_{i|i-1}^T\boldsymbol{\Lambda}_i^{-1} \tag{9.23}$$

and by explicitly setting $\xi_j = q_{j,k}$, where $q_{j,k}$ represents the generic element of $\boldsymbol{Q}_k$ (which is diagonal by definition (9.7)), it follows that:

$$\frac{\partial\boldsymbol{\Lambda}_i}{\partial q_{j,k}} = \frac{\partial}{\partial q_{j,k}}\left[\boldsymbol{H}_i\boldsymbol{P}_{i|i-1}\boldsymbol{H}_i^T + \boldsymbol{R}_i\right] \tag{9.24}$$

$$= \boldsymbol{H}_i\frac{\partial\boldsymbol{P}_{i|i-1}}{\partial q_{j,k}}\boldsymbol{H}_i^T \tag{9.25}$$

With the aid of (9.23) and (9.25), equation (9.22) becomes to

$$\frac{\partial\tilde{L}_k}{\partial q_{j,k}} = -\frac{1}{2}\sum_{i=k-N+1}^{k}\mathrm{Tr}\left[\Delta\boldsymbol{\Lambda}_i^{-1}\boldsymbol{H}_i\frac{\partial\boldsymbol{P}_{i|i-1}}{\partial q_{j,k}}\boldsymbol{H}_i^T\right] = 0 \tag{9.26}$$

After some algebraic manipulations, detailed in Appendix B.3 and equation (B.96), it follows that (9.26) simplifies to

$$\frac{\partial\tilde{L}_k}{\partial q_{j,k}} = -\frac{1}{2}\sum_{i=k-N+1}^{k}\left[\boldsymbol{P}_{i|i-1}^{-1}\left(\boldsymbol{U}_i + \boldsymbol{Q}_i - \boldsymbol{P}_{i|i} - \Delta\boldsymbol{x}_i\Delta\boldsymbol{x}_i^T\right)\boldsymbol{P}_{i|i-1}^{-1}\right]_{jj} = 0 \tag{9.27}$$

where it has been set:

$$\boldsymbol{U}_i \triangleq \boldsymbol{\Phi}_{i-1}\boldsymbol{P}_{i-1|i-1}\boldsymbol{\Phi}_{i-1}^T \tag{9.28}$$

$$\Delta\boldsymbol{x}_i \triangleq \boldsymbol{K}_i\tilde{\boldsymbol{y}}_{i|i-1} \tag{9.29}$$

Assuming again that the estimation process is *time invariant over the most recent N steps* (see p. 123 of [70]), namely:

$$\boldsymbol{P}_{i|i-1} = \boldsymbol{P}_{k|k-1}\,, \qquad \boldsymbol{Q}_i = \boldsymbol{Q}_k \qquad k-N+1 \le i \le k \tag{9.30}$$

then (9.27) reduces to:

$$\frac{\partial\tilde{L}_k}{\partial q_{j,k}} = -\frac{1}{2}\left[\boldsymbol{P}_{k|k-1}\left(\sum_{i=k-N+1}^{k}\boldsymbol{U}_i + \boldsymbol{Q}_k - \boldsymbol{P}_{i|i} - \Delta\boldsymbol{x}_i\Delta\boldsymbol{x}_i^T\right)\boldsymbol{P}_{k|k-1}\right]_{jj} = 0 \tag{9.31}$$

which is satisfied if and only if:

$$\sum_{i=k-N+1}^{k}\boldsymbol{U}_i + \boldsymbol{Q}_k - \boldsymbol{P}_{i|i} - \Delta\boldsymbol{x}_i\Delta\boldsymbol{x}_i^T = \boldsymbol{0} \tag{9.32}$$

or, alternatively:

$$\boldsymbol{Q}_k = \frac{1}{N}\sum_{i=k-N+1}^{k}\boldsymbol{P}_{i|i} - \boldsymbol{U}_i + \Delta\boldsymbol{x}_i\Delta\boldsymbol{x}_i^T \tag{9.33}$$

The estimate of the j^{th} element on the diagonal of $\boldsymbol{Q}$ is therefore:

$$\hat{q}_{j,k} = \frac{1}{N} \sum_{i=k-N+1}^{k} \left(\boldsymbol{P}_{i|i} - \boldsymbol{U}_i + \Delta\boldsymbol{x}_i \Delta\boldsymbol{x}_i^T\right)_{jj} \tag{9.34}$$

The equation (9.33) is the main result of this section, and it provides an explicit solution for the online estimation of the process noise covariance matrix $\boldsymbol{Q}_k$.

9.2 Maximum Likelihood Estimation Kalman Filter equations

The structure employed is essentially an EKF with an update of the process noise covariance matrix $\boldsymbol{Q}$ using the MLE method (9.33), while the measurement noise covariance matrix $\boldsymbol{R}$ remains fixed. Henceforth, the proposed algorithm will be referred to as the MLEKF.

Initialization

$$\hat{\boldsymbol{x}}_{0|0} = \mathbb{E}\left[\boldsymbol{x}_0\right] \tag{9.35}$$

$$\boldsymbol{P}_{0|0} = \mathbb{E}\left[\left(\boldsymbol{x}_0 - \mathbb{E}\left[\boldsymbol{x}_0\right]\right)\ \left(\boldsymbol{x}_0 - \mathbb{E}\left[\boldsymbol{x}_0\right]\right)^T\right] \tag{9.36}$$

Time update

$$\hat{\boldsymbol{x}}_{k|k-1} = \boldsymbol{f}(\hat{\boldsymbol{x}}_{k|k-1}, \boldsymbol{u}_{k-1}) \tag{9.37}$$

$$\hat{\boldsymbol{Q}}_{k-1} = \frac{1}{N} \sum_{i=k-N}^{k-1} \boldsymbol{P}_{i|i} - \boldsymbol{U}_i + \Delta\boldsymbol{x}_i \Delta\boldsymbol{x}_i^T \tag{9.38}$$

$$\boldsymbol{P}_{k|k-1} = \boldsymbol{\Phi}_{k-1} \boldsymbol{P}_{k-1|k-1} \boldsymbol{\Phi}_{k-1}^{\mathrm{T}} + \hat{\boldsymbol{Q}}_{k-1} \tag{9.39}$$

Measurement update

$$\boldsymbol{K}_k = \boldsymbol{P}_{k|k-1} \boldsymbol{H}_k^{\mathrm{T}} \left(\boldsymbol{H}_k \boldsymbol{P}_{k|k-1} \boldsymbol{H}_k^{\mathrm{T}} + \boldsymbol{R}_{k-1}\right)^{-1} \tag{9.40}$$

$$\hat{\boldsymbol{x}}_{k|k} = \hat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}_k \left[\boldsymbol{y}_k - \boldsymbol{h}(\hat{\boldsymbol{x}}_{k|k-1}, \boldsymbol{u}_{k-1})\right] \tag{9.41}$$

$$\boldsymbol{P}_{k|k} = \left(\boldsymbol{I} - \boldsymbol{K}_k \boldsymbol{H}_k\right) \boldsymbol{P}_{k|k-1} \tag{9.42}$$

The state vector is defined as $\boldsymbol{x}_k = \left[i_{\alpha,k},\ i_{\beta,k},\ \omega_{E,k},\ \vartheta_{E,k}\right]^{\mathrm{T}}$ $(n = 4)$, the input voltages as $\boldsymbol{u}_k = \left[u_{\alpha,k},\ u_{\beta,k}\right]^{\mathrm{T}}$ $(m = 2)$, and the measured output vector as $\boldsymbol{y}_k = \left[i_{\alpha,k},\ i_{\beta,k}\right]^{\mathrm{T}}$ $(p = 4)$.

The nonlinear functions $\boldsymbol{f}(\cdot)$ (9.37) and $\boldsymbol{h}(\cdot)$ (9.41) correspond to those presented in (7.14) and (7.19), respectively, and are briefly restated here for clarity:

$$\boldsymbol{f}(\boldsymbol{x}_k, \boldsymbol{u}_k) : \begin{cases} \boldsymbol{i}_{\alpha\beta,k+1} = \boldsymbol{i}_{\alpha\beta,k} + T_s\, \boldsymbol{A}_k^{-1} [\boldsymbol{u}_{\alpha\beta,k} - \boldsymbol{R}_s\, \boldsymbol{i}_{\alpha\beta,k} - \omega_{E,k}\, ((\boldsymbol{B}_k + \cdots \\ \qquad \cdots + \boldsymbol{C}_k)\, \boldsymbol{i}_{\alpha\beta,k} + \boldsymbol{d}_k]] \\ \omega_{E,k+1} = \omega_{E,k} \\ \vartheta_{E,k+1} = \vartheta_{E,k} + T_s\, \omega_{E,k} \end{cases} \tag{9.43}$$

$$\boldsymbol{y}_k = \boldsymbol{h}(\boldsymbol{x}^k) = \boldsymbol{i}_{\alpha\beta,k} \tag{9.44}$$

The state transition $\boldsymbol{\Phi}_k$ matrix in (9.39) and observation matrix $\boldsymbol{H}_k$ in (9.42) are known as the following Jacobians, already defined in (8.12) and (8.14):

$$\boldsymbol{\Phi}_k = \frac{\partial \boldsymbol{f}(\hat{\boldsymbol{x}}_{k|k}, \boldsymbol{u}_k)}{\partial x} \qquad \boldsymbol{H}_k = \frac{\partial \boldsymbol{h}(\hat{\boldsymbol{x}}_{k|k}, \boldsymbol{u}_k)}{\partial x} \tag{9.45}$$

9.2.1 Filter convergence conditions

With regards to the convergence of the proposed method, the following considerations are the same as the usually made for the standard EKF, whose convergency is guaranteed if the following assumptions hold [100], [101]:

- **Assumption** 1
  1. The covariance matrices $\hat{\boldsymbol{Q}}_k$ and $\boldsymbol{R}$ are bounded, ensuring that $\hat{\boldsymbol{Q}}_k$ is always positive semi-definite and $\boldsymbol{R}$ is positive definite. Specifically, there exist real numbers $\underline{q}$, $\overline{q}$, $\underline{r}$, and $\overline{r}$ such that
  $$\underline{q}\boldsymbol{I}_{4\times 4} \leq \hat{\boldsymbol{Q}}_k \leq \overline{q}\boldsymbol{I}_{4\times 4} \tag{9.46}$$
  $$\underline{r}\boldsymbol{I}_{2\times 2} \leq \boldsymbol{R} \leq \overline{r}\boldsymbol{I}_{2\times 2} \tag{9.47}$$
  for each time step $k \geq 0$. This condition is satisfied because $\boldsymbol{R}$ is fixed, with its diagonal composed of positive real numbers. Meanwhile, the diagonal elements of $\hat{\boldsymbol{Q}}_k$, estimated through (9.38), are bounded in implementation to lie between zero and positive real numbers. These boundaries are discussed in detail in Subsect. 9.3.3.
  2. The matrices $\boldsymbol{\Phi}_k$ and $\boldsymbol{H}_k$, equations (8.12) and (8.14) respectively, satisfy the uniform observability condition [90], [92], please refer to Sect. 7.2 for more details about the system observability.
  3. The initial error covariance matrix $\boldsymbol{P}_{0|0}$ is positive definite.

- **Assumption** 2
  1. There are positive real numbers $\overline{\gamma}_1$, $\overline{\gamma}_2$, $\underline{p}$ and $\overline{q}$ such that the following inequalities hold for each time step $k \geq 0$.
  $$||\boldsymbol{\Phi}_k|| \leq \overline{\gamma}_1 \tag{9.48}$$
  $$||\boldsymbol{H}_k|| \leq \overline{\gamma}_1 \tag{9.49}$$
  $$\underline{p}\boldsymbol{I}_{4\times 4} \leq \boldsymbol{P}_{k|k} \leq \overline{p}\boldsymbol{I}_{4\times 4} \tag{9.50}$$

2. $\mathbf{\Phi}_k$, eq. (8.12) is non-singular for each time $k \geq 0$.

The conditions listed in the aforementioned assumptions are checked online within the algorithm code to prevent prohibited operations, such as matrix inversion, when the matrices are singular or near singularity. Regarding the bounding conditions, as previously mentioned, the values in (9.46) are detailed in Subsect. 9.3.3, while (9.48) is satisfied by preventing the state $\boldsymbol{x}_{k|k}$, through which $\mathbf{\Phi}_k$ is calculated, from diverging or becoming equal to `NaN` (Not a Number). Equation (9.49) is satisfied by definition since $\boldsymbol{H}_k$ is a constant matrix, as shown in (8.14). Finally, (9.50) is satisfied by ensuring that the master error covariance matrix $\boldsymbol{P}_{k|k}$ remains positive definite. This can be checked online using a specific MATLAB function that calculates its determinant and verifies if it is greater than zero. However, this function increases the computational cost of the algorithm, so it was used only for debugging purposes. In the final implementation, the positive definiteness of $\boldsymbol{P}_{k|k}$ was ensured by observing the other boundaries.

9.3 Experimental results

In this section, the performance of the EKF and MLEKF in a sensorless SynRM drive is evaluated. The tests are conducted under typical operating conditions, including speed inversion, load handling, and four-quadrant operations. The results are compared in terms of the MSE of the estimated state variables. The robustness of the two filters is also assessed by evaluating their sensitivity to stator resistance mismatches and magnetic model absence. The execution time of the control algorithm elements is also analyzed.

9.3.1 Test bench

The experiments were conducted using a test bench depicted in Fig. 9.1 comprising a SynRM drive, whose nameplate data are provided in Table 9.1. The SynRM was rigidly coupled to an IPMSM, which acted as a torque-controlled virtual load. Both motors were powered by three-phase voltage inverters, sharing a 450 V DC-Link, with a switching frequency of 8 kHz. The control for both SynRM and IPMSM was implemented through a MicroLabBox dSpace Fast Control Prototype (FCP) system.

The magnetic model of the SynRM is depicted in Fig. 9.2, these maps were obtained through an elaboration of a series of measurements [88] and they are used in the time update step in (9.37).

9.3.2 Control Design

The drive block schematic is shown in Fig. 9.3. The speed control loop features a PI controller with a bandwidth of 5 Hz, while the current control loops are managed by two gain-scheduling PI controllers [95]. Their gains were tuned to achieve a constant bandwidth of 200 Hz and a phase margin of 70 deg. All experiments were performed using estimated quantities as feedback.

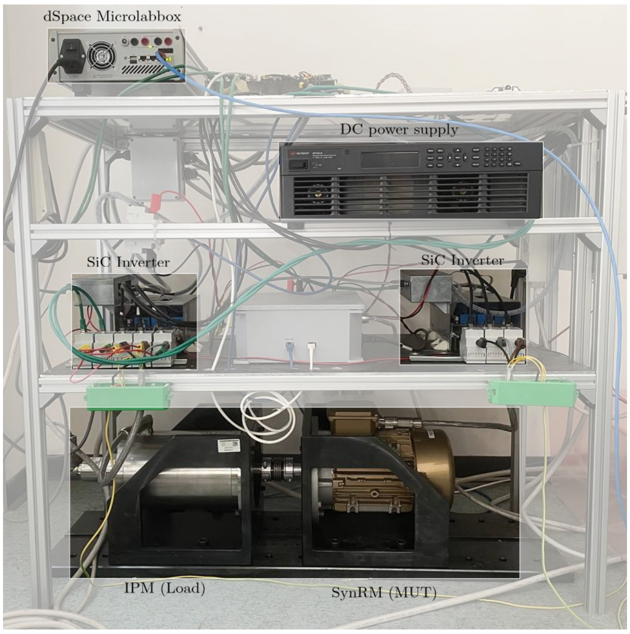


Figure 9.1 The experimental setup.

Table 9.1 System and SynRM parameters

Parameters	Symbol	Unit	Value
Nominal torque	τ_n	N m	9
Nominal current	I_n	A	6.5
Nominal speed	$\omega_{M,n}$	$\min^{-1}$	1000
Stator resistance	r_s	Ω	1.72
unsat. d-axis inductance	L_d	mH	230
unsat. q-axis inductance	L_q	mH	60
pole pairs	p	\	2
3-phase SiC Inverter (module: SKiiP 26ACM12V17)			
DC voltage	u_C	V	450
Inverted dead-time	t_d	ns	300
Switching frequency	f_s	kHz	8
Control Platform	dSPACE MicroLabBox		
Control frequency	f_c	kHz	8
Sampling time	T_s	µs	125

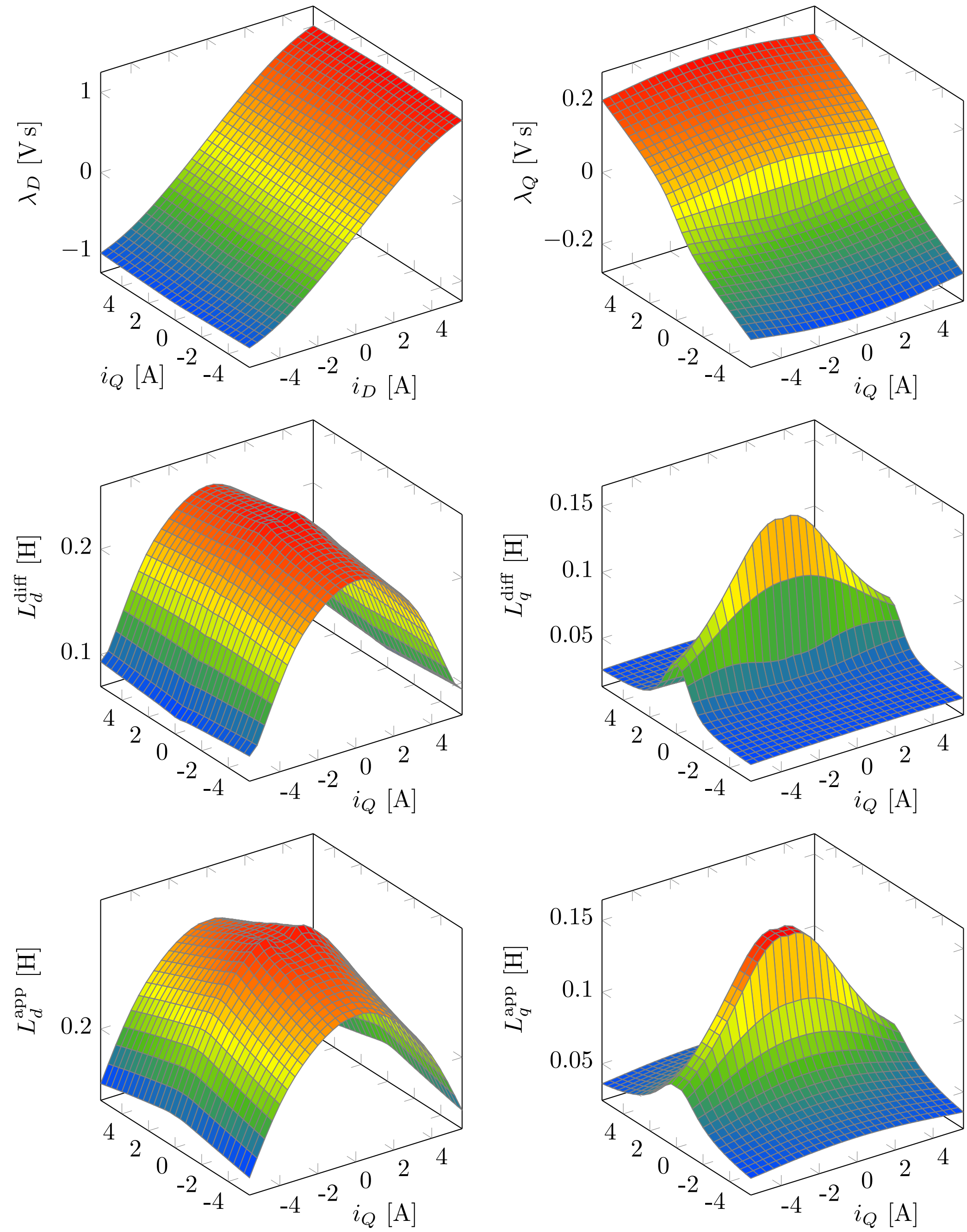


Figure 9.2 Flux linkages $\boldsymbol{\lambda}_{\mathrm{dq}}$, apparent inductances $\mathbf{L}_{\mathrm{dq}}^{\mathrm{app}}$, and differential inductances $\mathbf{L}_{\mathrm{dq}}^{\mathrm{diff}}$ of the SynRM used in the experiments (Table 9.1).

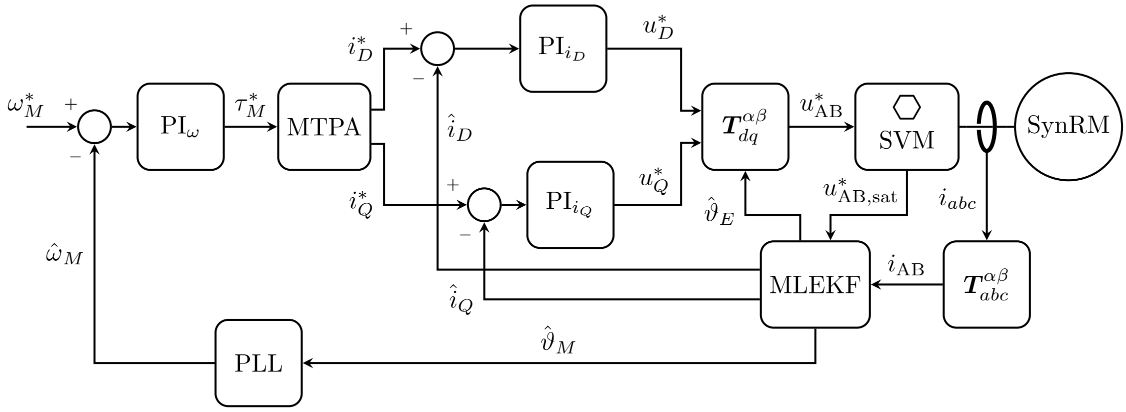


Figure 9.3 The proposed sensorless control of the SynRM.

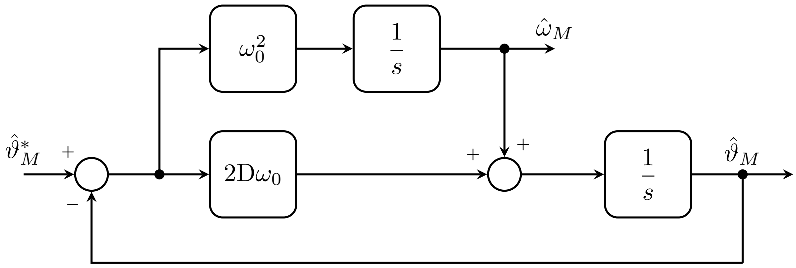


Figure 9.4 The PLL block diagram used within the experiments [96].

The speed estimate is worth particular attention. The mechanical model of the EKF is based on the assumption of infinite inertia, which simplifies the differential equation of speed in (7.10) to a null value. Consequently, the speed update relies solely on the EKF gain correction, resulting in both a noisy speed estimate and high sensitivity to the weight of the $\boldsymbol{Q}$ matrix. Although such a highly dynamic speed estimate does not affect the position estimation, it is not smooth enough to be used directly as feedback in the speed control loop. To address these issues, a PLL filter is employed to recover the speed $\hat{\omega}_M$ from the estimated position $\hat{\vartheta}_M$. The PLL has been tuned with a characteristic frequency of $\omega_0 = 2\pi \cdot 50\,\mathrm{Hz}$ and a damping factor of D = 1 [96]. For clarity, the PLL speed is used only for feedback and not in the EKF or MLEKF model. The block diagram of the PLL is shown in Fig. 9.4.

The current references are obtained from the output of the speed controller using the usual MTPA curve, which can be easily derived from the accurate motor model. The availability of this model is a prerequisite. According to [42], [97], ensuring the correct functioning of the sensorless algorithm requires supplying the SynRM with a minimum d-axis current value, which in the experiment was set to $I_D^{\min} = 2\,\mathrm{A} \simeq I_n/3$. This value was first tested by simulations and then validated during the experiments. Fig. 9.5 shows the modified MTPA trajectory used during all the following experiments.

The $I_D^{\min}$ value is essential under zero-speed and no-load conditions. In these scenarios, without a minimum current vector, the KF innovation, whether in the EKF or the MLEKF, becomes either zero or negligibly small. Consequently, the KF fails to extract meaningful information, as the innovation becomes indistinguishable

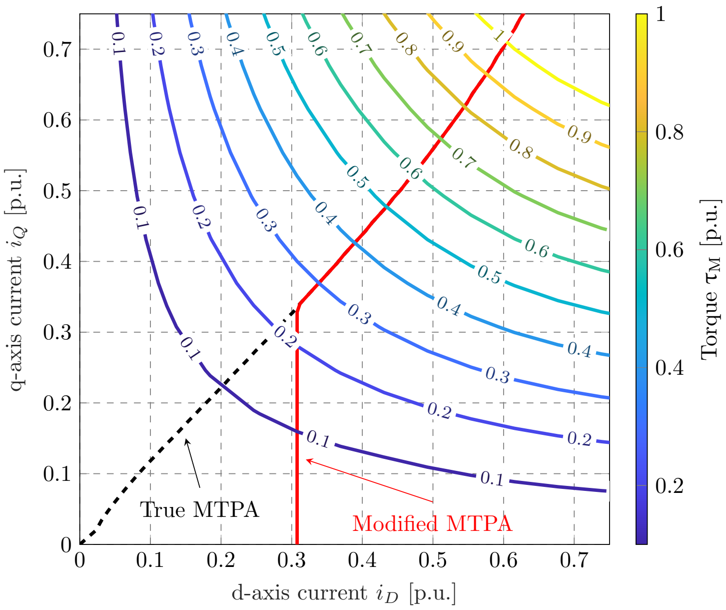


Figure 9.5 Real Maximum Torque Per Ampere trajectory (black curve), and modified MTPA trajectory (red curve) adopted within the experiments.

from measurement noise, resulting in unreliable estimations. To address this issue, imposing a minimum current vector, below which the innovation cannot be zero, offers an effective solution to improve observer reliability.

In [42], an interesting investigation is presented on the relationship between the minimum current vector and the minimum bearable speed. It is shown that as the current limit increases, the minimum bearable speed decreases.

The EKF is tuned to their best using the trial-and-error method, resulting in the following constant noise covariance matrices of the process:

$$\boldsymbol{Q}_{\text{EKF}} = \text{diag}\left\{1 \times 10^{-5}, 1 \times 10^{-5}, 1, 0\right\} \tag{9.51}$$

The first two weights refer to the $\alpha\beta$ currents, the third one refers to the electrical speed, and the latter to the electrical position. The measurement noise matrices are equal for both the EKF and MLEKF, since the current measurements are the same for the two observers and the MLEKF updates online the solely $\boldsymbol{Q}$ matrix, i. e.

$$\boldsymbol{R}_{\text{EKF}} = \boldsymbol{R}_{\text{MLEKF}} = \text{diag}\left\{1 \times 10^{-3}, 1 \times 10^{-3}\right\} \tag{9.52}$$

9.3.3 Window size tuning and MLE bounds

Regarding the window size N in (9.33), there is no general rule to determine its value, and must be chosen according to the specific application. Typically, a small value of N is preferred to ensure a fast adaptation and less computational burden, but this may lead to a noisy estimate of the process noise covariance matrix, or the filter divergence [59]. Moreover, to further reduce computational complexity, the covariance matrix in (9.33) is forced to be diagonal, and only the elements q_{ii} on the main diagonal are adapted. An analysis of the influence of this parameter on the speed estimation error in a sensorless IM drive can be found in [78].

In the experiments, the window size N was set to $N = 200$. Window size values were tried with lower values, but the $\boldsymbol{Q}$ estimation was too noisy, resulting in an unreliable estimation. The N in (9.33) is considered a tuning parameter, since it is not possible to determine its value a priori and must be chosen according to the test dynamics the application is subjected to.

Regardless of the choice of window size, it has been pointed out in [67] that an AEKF implementation based on (9.33) performs poorly under non-dynamic operating conditions. In fact, the elements of the process noise covariance matrix $\boldsymbol{Q}_k$ tend to shrink when the motor operates under quasi-static conditions for a prolonged time, preventing the filter from waking up at any subsequent variation. To avoid this issue and keep the filter always "alive", the elements q_{ii} of $\boldsymbol{Q}_k$ are lower bounded by a small value $\underline{q}_{ii}$ selected mostly by trial and error. An upper bound $\overline{q}_{ii}$ is also considered to prevent a unbounded growth of the elements of $\boldsymbol{Q}_k$ in the event of sudden changes in the dynamics of the system. The lower and upper bounds were set to

$$\underline{\boldsymbol{q}} = \left[1 \times 10^{-7},\, 1 \times 10^{-7},\, 0.1,\, 0\right] \tag{9.53}$$

$$\overline{\boldsymbol{q}} = \left[1 \times 10^{-4},\, 1 \times 10^{-4},\, 10,\, 1 \times 10^{-6}\right] \tag{9.54}$$

The currents variance lower bounds $\underline{q}_{11}$ and $\underline{q}_{22}$ were set to keep the filter alive under steady state condition, while $\overline{q}_{11}$ and $\overline{q}_{22}$ were set to prevent the filter from diverging in the event of a sudden change in the system dynamics.

The position lower weight $\underline{q}_{44}$ was set equal to the weight used by the EKF. Experiments have evidenced that the position weight should be kept close to zero in order to obtain the best estimation result. In particular, a too high position weights leads to steady-state estimation error in the speed. For this reason, the position variance upper bound $\overline{q}_{44}$ was set to $1 \times 10^{-6}\,\mathrm{rad}^2 = 3.3 \times 10^{-3}\,\mathrm{deg}^2$.

When it comes to the speed, the upper bound $\overline{q}_{33}$ was set to $10\,\mathrm{rad}^2\,\mathrm{s}^{-2}$, while the lower bound $\underline{q}_{33}$ was set to $0.1\,\mathrm{rad}^2\,\mathrm{s}^{-2}$. The lower bound is essential to keep the filter alive under steady-state conditions, in particular when the speed reference crosses the zero value, the most struggling condition for a Kalman filter sensorless application. The upper bound is set to prevent the filter from estimating a too noisy speed value. It should be noted that this issues does not directly affect the speed control loop, since the speed is estimated by the PLL block, which recovers the speed from the estimated position $\hat{\vartheta}_E$. However the problem involves the filter itself, since the estimated speed is used in the prediction function of the filter.

9.3.4 Performance Comparison

The EKF and the MLEKF are tested as position estimators under typical operating conditions of a modern electric drive. The tests include speed inversion, load handling, and four-quadrant operations. All estimated and measured state variables are collected at the control frequency rate (see Table 9.1) for post-processing. The MSE was computed as in (8.41), and all these values are gathered in Table 9.2. The MSE related comments are included within the description of each test.

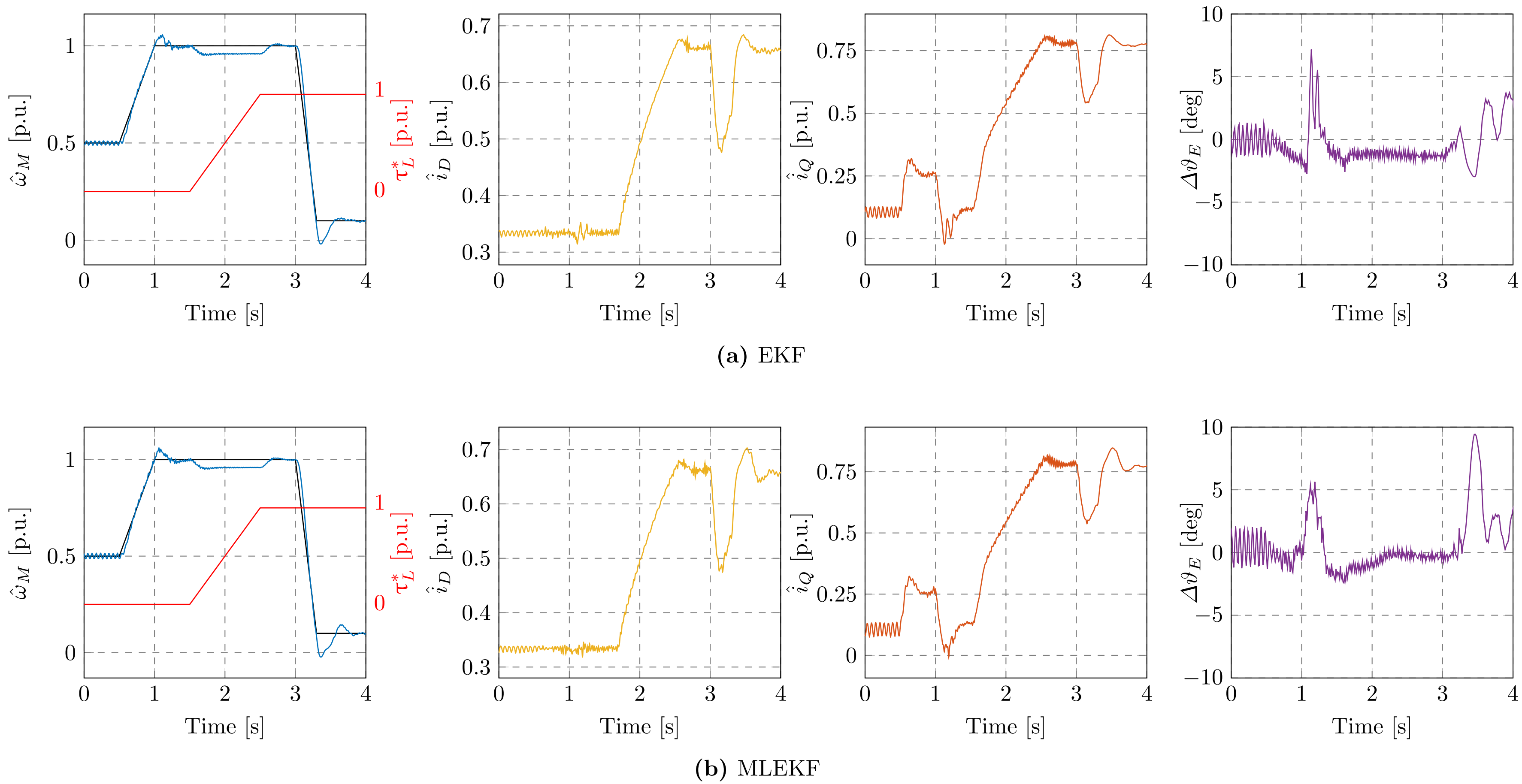
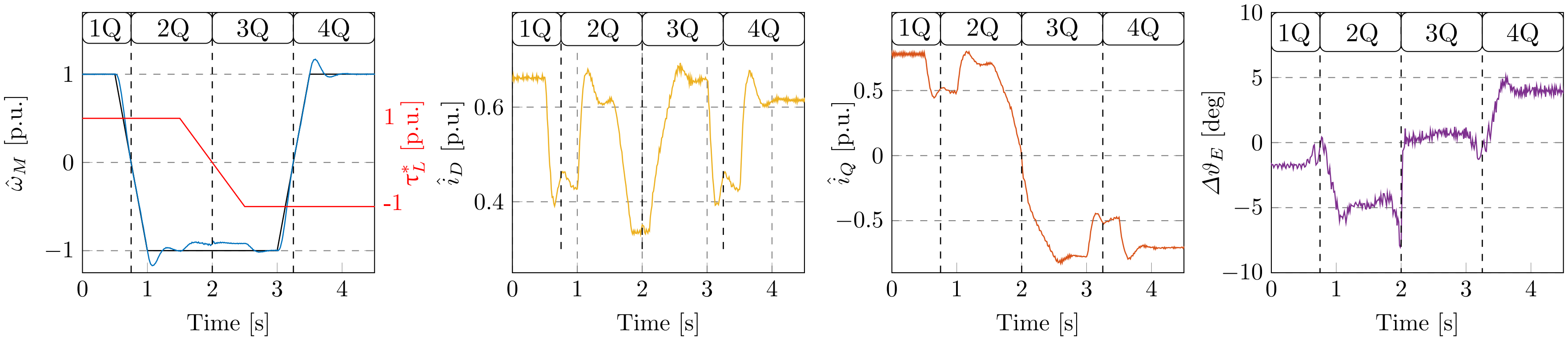
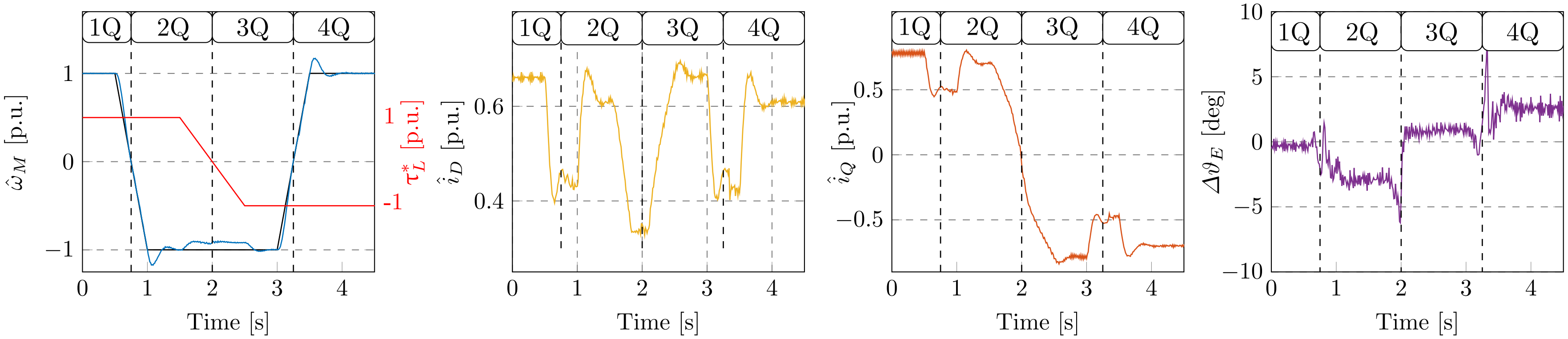


Figure 9.6 Minimum Speed at Rated Torque (MSRT) test for conventional EKF (upper row) and MLEKF (lower row). From left to right: mechanical speed $\hat{\omega}_M$, d-axis $\hat{i}_D$ and q-axis $\hat{i}_Q$ currents, electrical position error $\Delta\vartheta_E$. All values refer to their respective nominal values as listed in Table 9.1.



(a) EKF



(b) MLEKF

Figure 9.7 Four-Quadrant Operation (FQO) test for conventional EKF (upper row) and MLEKF (lower row). From left to right: mechanical speed ω_M, d-axis i_D and q-axis i_Q currents, electrical position error $\Delta\vartheta_E$. All values refer to their respective nominal values as listed in Table 9.1.

The Minimum Speed at Rated Torque (MSRT) test

The limitations of EKF and MLEKF techniques are most evident at low speeds. The lowest speed at rated torque is a common benchmark and is discussed in the sequel. The test results are presented in Fig. 9.6. In the leftmost figure, the black line represents the speed reference provided to the SynRM, while the red line depicts the torque reference applied to the IPM, which serves as a virtual load. The blue trace corresponds to the estimated speed, denoted as $\hat{\omega}_M$ and obtained from the PLL block of Fig. 9.3.

The test starts by setting a positive speed reference without load. Before initiating the transient load torque at $t = 0.5\,\mathrm{s}$, the SynRM accelerates from 0.5 p.u. to 1 p.u., still under no load conditions. Subsequently, the speed reference is reduced to 10% of the rated speed at full load torque. This minimum threshold was determined through repeated tests until uncontrollable oscillations occurred at low speed.

Continuing from left to right, the other figures show the estimated d-q currents and the electrical position error $\Delta\vartheta_E = \hat{\vartheta}_E - \vartheta_E^{\mathrm{meas}}$, during the above speed and torque transients.

At full load and the lowest stable speed, both KFs exhibit position estimate oscillations (see Fig. 9.6a). The position error averagely lies on $\sim 2\,\mathrm{deg}$ for both the EKF and the MLEKF. In both cases, the control effectively manages the speed downward ramp, and the estimated speed remains close to the actual value.

Referring to Table 9.2, it is evident that for the EKF case, the MSE collected values are lower than for the MLEKF. This difference is not significant for all the quantities, therefore it emerges that the MLEKF is a valid alternative to the EKF in this specific test case.

It is worth noting the estimated variances of the process noise covariance matrix, $\hat{\boldsymbol{Q}}$, shown in Fig. 9.8. In particular, at $t = 0.5\,\mathrm{s}$, all the variances appear to be at their lower bound, set to $\underline{\boldsymbol{q}}$ in (9.54). However, as the speed reference changes, the variances do not increase immediately; instead, it takes some time for them to adapt to the new operating point. This behavior is attributed to the large window size N used in the (9.33). During this phase of the test, setting the lower bound was crucial to ensure the filter remained operational.

The Four-Quadrant Operation (FQO) test

The test results are presented in Fig. 9.7. The test begins with a positive nominal speed reference and a rated load torque, τ_L^*. While maintaining the same torque, a speed reversal is initiated at $t = 0.5\,\mathrm{s}$, transitioning from the first quadrant (labeled as 1Q) to the second quadrant (2Q). At $t = 1.5\,\mathrm{s}$, a load torque reversal is triggered, moving from $1, \mathrm{p.u.}$ to $-1, \mathrm{p.u.}$. Finally, at $t = 3\,\mathrm{s}$, a second speed reversal is imposed to exit the third quadrant (3Q) and enter the fourth quadrant (4Q).

It should be noted that the electrical position estimation error tends to increase in the second and fourth quadrants—specifically when the SynRM is accelerated by the imposed load torque. In this condition, the SynRM provides less torque to maintain the speed operating point, resulting in lower currents, which in turn degrade the observer's performance. Conversely, as soon as the speed reference crosses the zero value (first and third quadrants), the load torque brakes the SynRM. In this scenario,

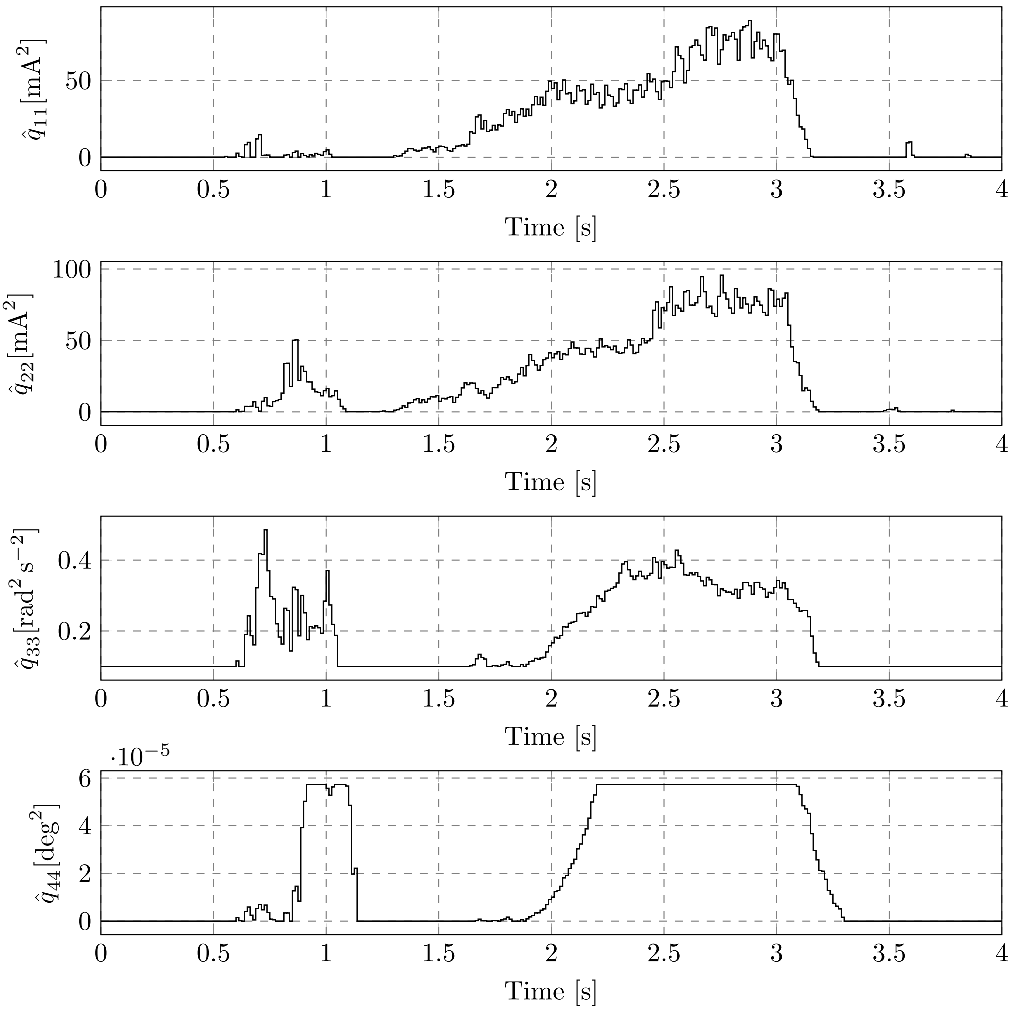


Figure 9.8 Estimated diagonal elements of $\hat{\boldsymbol{Q}}^m$ in the MSRT test, depicted in Fig. 9.6.

Table 9.2 Mean Squared Error (MSE)

Variable	Symbol	Unit	Test	EKF	MLEKF	Fig.
Mechanical speed	ω_M	rpm^2	MSRT	48.60	56.7	9.6
			FQO	149.37	164.28	9.7
			SSS	583	193.76	9.10
d-axis current	i_D	mA^2	MSRT	15519	33961	9.6
			FQO	47933	20839	9.7
			SSS	11089	651	\
q-axis current	i_Q	mA^2	MSRT	11607	20675	9.6
			FQO	43084	18189	9.7
			SSS	16016	2899	\
Electrical position	ϑ_E	deg^2	MSRT	3.32	5.64	9.6
			FQO	10.95	5.13	9.7
			SSS	9.37	2.29	9.10

the currents are higher, and the position estimation error decreases to zero.

A comparison between Fig.9.7a and Fig.9.7b reveals that the MLEKF demonstrates a lower $\Delta\vartheta_E$ compared to the EKF, as also corroborated by Table 9.2. However, the MLEKF exhibits higher variance in the estimated speed. Despite these differences, both filters effectively handle speed reversal and load torque reversal without any issues.

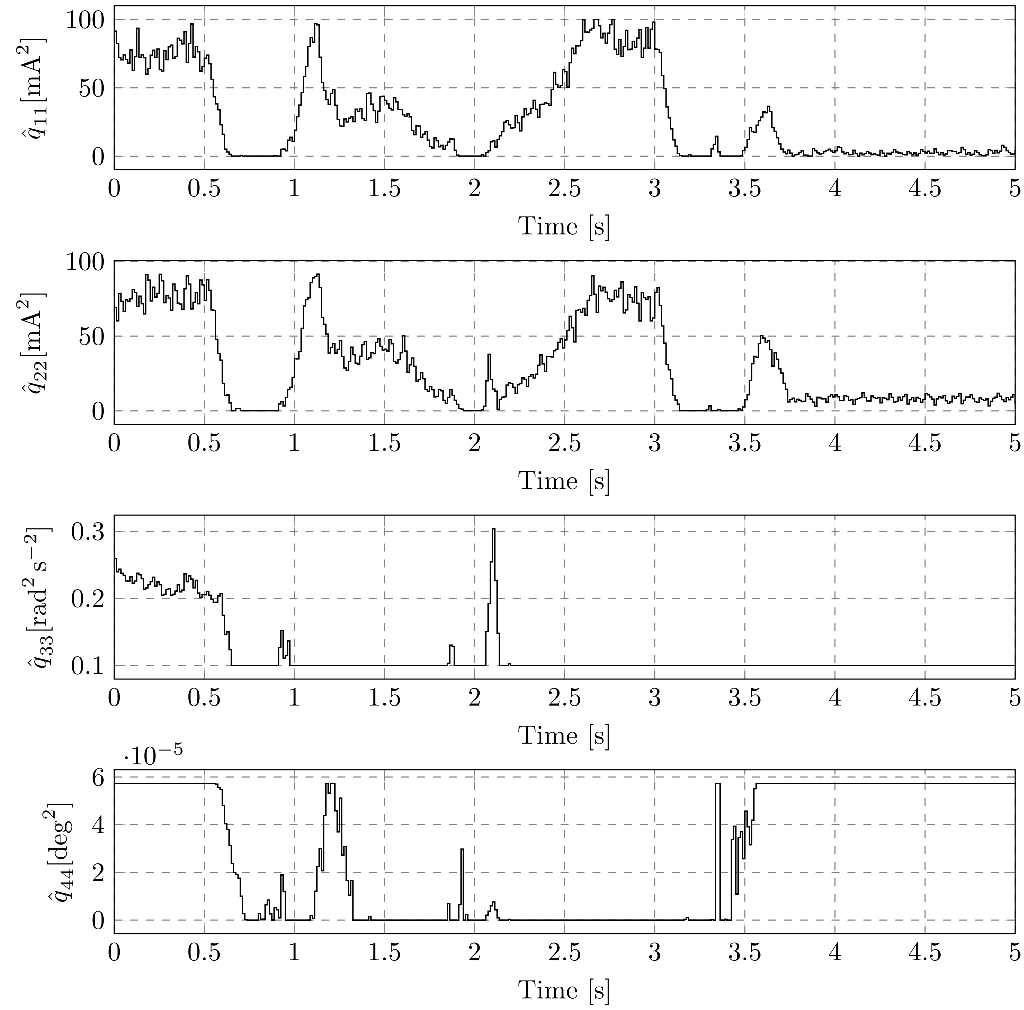


Figure 9.9 Estimated diagonal elements of $\hat{\boldsymbol{Q}}^m$ in the FQO test, depicted in Fig. 9.7.

The estimated variances of the process noise covariance matrix, $\hat{\boldsymbol{Q}}$, are presented in Fig. 9.9. The same *delay* effect observed in the MSRT test (Fig. 9.8) is also evident in this test, particularly at $t = 0.5\,\text{s}$ and $t = 3.0\,\text{s}$, when the speed reference changes.

An additional noteworthy observation is that the estimated variance associated with the electrical position, $\hat{q}_{44}$, reaches its maximum allowable value, $\overline{q}_{44}$, set to $1 \times 10^{-6}\,\text{rad}^2 = 3.3 \times 10^{-3}\,\text{deg}^2$, in both Fig. 9.8 and Fig. 9.9. This constraint reduces the *adaptive* capabilities of the proposed technique. If $\hat{q}_{44}$ were allowed to evolve to higher values, the filter output would deviate, resulting in an incorrect estimated speed.

The Step Starting from Standstill (SSS) test

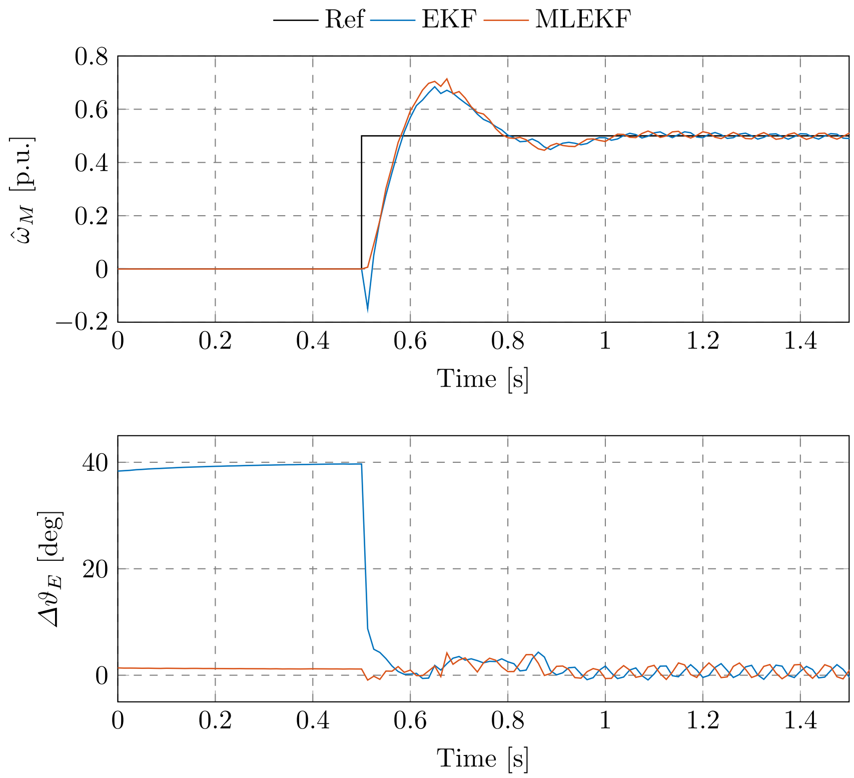


Figure 9.10 Step Starting from Standstill (SSS) test: mechanical speed (top) and electrical position error (bottom).

The test results are presented in Fig. 9.10. The experiment evaluates the drive performance when the motor starts from a complete standstill and a step change in the speed reference is applied. In particular, at $t = 0.5\,\mathrm{s}$ a speed step is executed up to half of the nominal speed.

Both EKF and MLEKF are capable of initiating motion from a standstill condition and exhibit nearly identical responses. Prior to the step command, the idle EKF position ϑ_e appears stable, with an error of approximately 40 deg. In contrast, the position error derived from the MLEKF oscillates around 3 deg. However, following the speed step, the position errors for both the EKF and the MLEKF stabilize at around 0.5 deg.

At standstill, with zero speed and at no load, the system is not observable, as detailed in Sect. 7.2, therefore the position error $\Delta\vartheta_E$ before the speed step at $t = 0.5\,\mathrm{s}$ is random. Anyway, as soon as the system is solicited by a speed reference change, the estimation errors rapidly converge to zero.

It should be noted that the EKF estimated speed at $t = 0.5\,\mathrm{s}$ becomes negative. If this were the actual speed of the machine, it could pose a problem for certain applications where the load is coupled with the motor and only one direction of rotation is allowed. However, the actual speed of the machine remains at zero, and the negative speed is merely a transient phenomenon caused by poor estimation at the beginning.

When it come to the MSE Table 9.2, the MSE MLEKF values are lower than in the EKF case, because the EKF MSE value is significantly influenced by the incorrect estimated speed at the beginning of the speed step. It should be noted that the MSE values are not computed throughout the duration of the test, but only for the time interval where the speed step is applied, since the system is not observable at standstill.

Sensitivity to resistance mismatch

Table 9.3 Resistance mismatch test

r_s^{KF}/r_s	Mean ($\Delta\vartheta_E$)		Unit	MSE (ϑ_E)		Unit
	EKF	MLEKF		EKF	MLEKF	
1	−0.01	0.24	deg	1.58	1.51	deg^2
0.9	3.88	4.11		16.6	18.25	
0.8	7.23	8.14		53.9	67.22	

The test results, including the mean electrical position error and MSE, are presented in Table 9.3. The experiment aimed at evaluating the robustness of both the EKF and MLEKF against stator resistance mismatches. The resistance mismatch arises from temperature variations, which causes an increase in the actual stator resistance compared to the value used in the model (7.14). Three experiments are carried out under steady-state conditions, with a speed reference of 0.1 p.u. and rated load torque, each with a progressively increasing resistance mismatch. The resistance mismatch is expressed as the ratio r_s^{KF}/r_s. r_s corresponds to the nominal value documented in Table 9.1.

When the resistance mismatch is null or a reduced resistance is used in the filter equations, the EKF provides slightly more accurate results, as shown in Table 9.3. The MSE values are calculated by averaging the electrical position error over 1600 samples (0.2 s). It should be noted that both the mean and the MSE of $\Delta\vartheta_E$ are significantly affected by a resistance mismatch, and this effect is the same for both EKF and MLEKF. Additional tests were performed by considering a resistance variation up to 30%, but both the EKF and MLEKF fail to maintain convergence, and the tests is subsequently aborted. All things considered, both the filters exhibit similar robustness to resistance mismatches.

Constant Magnetic model

In this test, the motor inductances, L_d^x and L_q^x (with x = diff, app), were kept constant in the prediction function $\boldsymbol{f}(\cdot)$ in (7.14). Their values are equal to those listed in Table 9.1.

From $t = 0.5$ s to $t = 2.5$ s, the load torque is increased from 0 N m to its rated value of 9 N m. With constant L_d and L_q, the estimation error increases during the torque transient in both filters because the motor inductances experience saturation, which is not accounted for in the model.

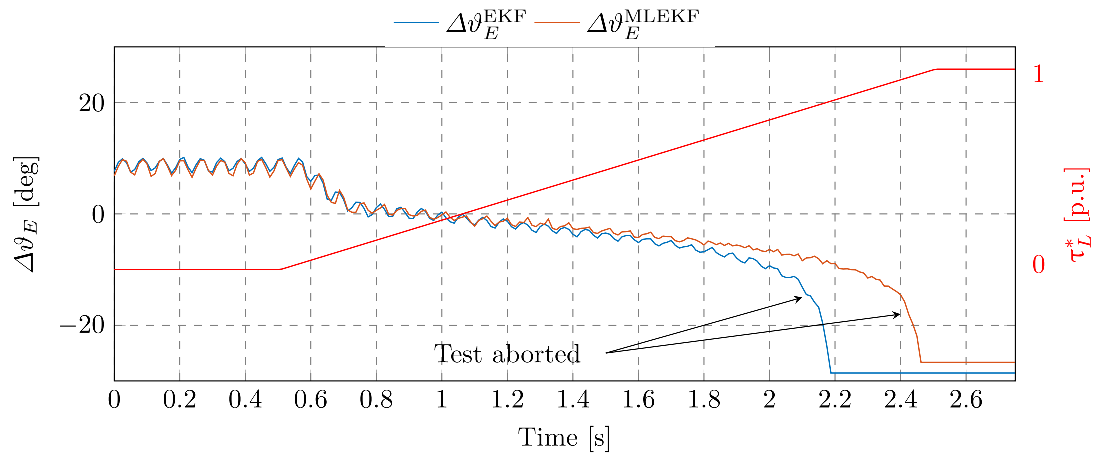


Figure 9.11 Electrical position error, performance in case of constant inductance model (conventional EKF vs. proposed MLEKF).

As anticipated, both filters fail to complete the test. As the load torque increases, the saturation effect becomes more significant, causing both filters to diverge. The MLEKF demonstrates slightly better performance compared to the EKF, managing a load torque of up to approximately 0.9, τ_N before diverging, while the EKF diverges at around 0.82, τ_N.

This test underscores the importance of incorporating a magnetic model for sensorless control of a SynRM. Without accounting for inductance variation due to saturation, both filters are unable to handle high load torques effectively.

Execution Time

Table 9.4 Execution times of the control algorithm elements

Routine	Execution Time
Control cycle	125 µs
Speed & Current Control	22 µs
EKF	14 µs
MLEKF	23 µs

The execution times of the proposed sensorless drives are summarized in Table 9.4. The control cycle, which corresponds also to the sample time, is repeated every 80 µs, of which 22 µs are used for the speed and current control routines. The execution of the EKF and AEKF algorithms require 14 µs and 23 µs, respectively. The AEKF requires more time than the EKF because of the additional computation involved with the updating of the $\boldsymbol{Q}$ matrix according to (9.33). The execution time depends on the window length N. It can be potentially reduced by resorting to a recursive formula for (9.33), as also proposed in [78].

CHAPTER 10

Master Slave Kalman Filter

As already mentioned in Ch. 9, the EKF is a widely used algorithm for estimating the state of nonlinear dynamical systems. However, it is highly sensitive to random noise, and inaccurate characterization of this noise can result in reduced filtering accuracy or even divergence. Consequently, variance estimation remains a critical challenge. Despite significant research over the past few decades, this issue has not been conclusively resolved and continues to be an active area of investigation.

This chapter proposes the automatic tuning of the EKF through a second Kalman Filter (KF) in a master-slave (MS) configuration. The two KFs work concurrently: the first one, called master, estimates the required quantities for machine control, while the second, called slave, updates the process noise statistics of the first KF. The slave KF is significantly easier to tune, requiring only one non-critical parameter. Experimental results confirm the validity of this approach.

In Sect. 10.1, the Master-Slave KF structure is presented, followed by the experimental results detailed in Sect. 10.2.

10.1 Master Slave KF Equations

The MSKF structure was originally introduced by *Song et al.* [102] for an UKF application. The estimation principle was described by the same author in [103], but the adaptive rule was developed using the MIT rule, which is computationally cumbersome as an online estimation technique. Therefore, the structure considered within this thesis is the one first presented in [102] and later refined by the same author in [104]. A more detailed description of the MSKF structure is given by *Ritter* in [81], [105], particularly in the Appendix of [81], where *Ritter* explains how the MSKF equations are adapted for an EKF application instead of a UKF one.

In the study of literature, [104] examines the dynamics of helicopter robots, [82] focuses on actuator failure estimation for unmanned aerial vehicles, and [81] investigates adaptive filters for wind turbine applications. While each of these works employs Unscented or Cubature Kalman Filters, they present only simulation results and address state estimation for systems with dynamics that differ significantly from those of SynRM drives.

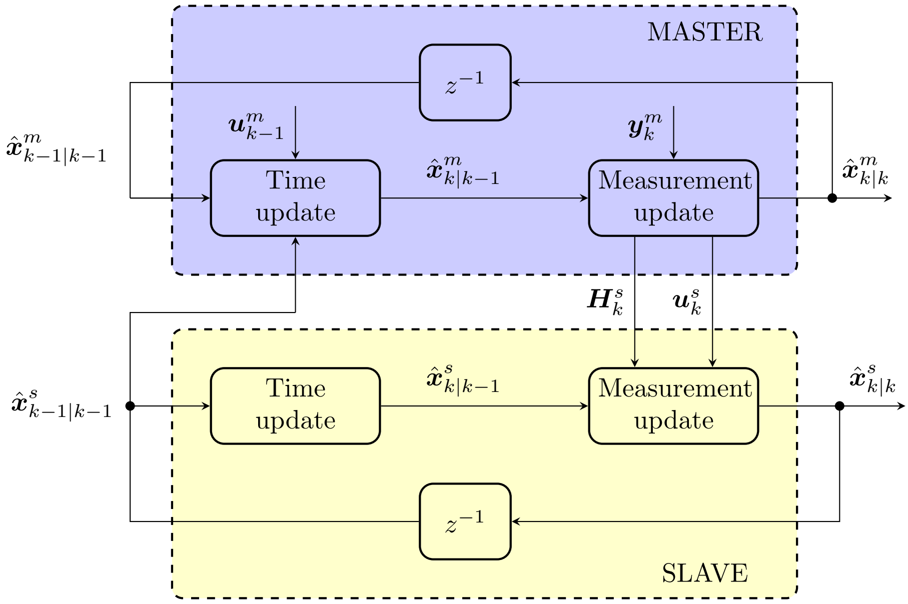


Figure 10.1 Master-Slave Kalman Filter structure.

The MSKF structure is depicted in Fig. 10.1.

In terms of notation, superscripts m and s refer to the master EKF and the slave KF, respectively.

The system premises from which the MSKF is derived are the same as those described in 8.1. However, they will be restated here for clarity. Consider the following linear dynamical system:

$$\begin{cases} \boldsymbol{x}^m_{k+1} = \boldsymbol{f}^m\left(\boldsymbol{x}^m_k,\, \boldsymbol{u}^m_k,\, \boldsymbol{w}^m_k\right) & \boldsymbol{x}^m_k \in \mathbb{R}^{n\times 1}, \quad \boldsymbol{u}^m_k \in \mathbb{R}^{m\times 1} \\ \boldsymbol{y}^m_k = \boldsymbol{h}^m\left(\boldsymbol{x}^m_k,\, \boldsymbol{v}^m_k\right) & \boldsymbol{y}^m_k \in \mathbb{R}^{p\times 1} \end{cases} \tag{10.1}$$

where:

- the functions $\boldsymbol{f}(\cdot)$, $\boldsymbol{h}(\cdot)$ and the input $\boldsymbol{u}_k$ are known a priori.
- $\boldsymbol{w}^m_k$ and $\boldsymbol{v}^m_k$ are zero-mean Gaussian random processes with known covariances:

$$\boldsymbol{w}^m_k \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q}^m_k)\,, \qquad \boldsymbol{v}^m_k \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}^m_k) \tag{10.2}$$

- $\boldsymbol{w}^m_k$ and $\boldsymbol{v}^m_k$ are white noises:

$$\mathbb{E}\left[\boldsymbol{w}^m_k (\boldsymbol{w}^m_h)^T\right] = \boldsymbol{Q}^m_k\, \delta[k-h]\,, \qquad \mathbb{E}\left[\boldsymbol{v}^m_k (\boldsymbol{v}^m_h)^T\right] = \boldsymbol{R}^m_k\, \delta[k-h] \tag{10.3}$$

  where $\delta[k-h]$ is the Kronecker delta function, which is equal to 1 when $k = h$ and 0 otherwise.

- $\boldsymbol{w}_k^m$ and $\boldsymbol{v}_k^m$ are mutually uncorrelated:

$$\mathbb{E}\left[\boldsymbol{w}_k^m (\boldsymbol{v}_h^m)^T\right] \;=\; \mathbf{0} \tag{10.4}$$

- the initial state $\boldsymbol{x}_0^m$ is a normally distributed random variable with mean $\bar{\boldsymbol{x}}_0^m$ and covariance $\boldsymbol{P}_{0|0}^m$:

$$\boldsymbol{x}_0^m \sim \mathcal{N}(\bar{\boldsymbol{x}}_0^m, \boldsymbol{P}_{0|0}^m) \tag{10.5}$$

- $\boldsymbol{w}_k^m$ and $\boldsymbol{v}_k^m$ are uncorrelated with the initial state $\boldsymbol{x}_0^m$:

$$\mathbb{E}\left[\boldsymbol{x}_0^m (\boldsymbol{w}_k^m)^T\right] \;=\; \mathbf{0}\,, \qquad \mathbb{E}\left[\boldsymbol{x}_0^m (\boldsymbol{v}_k^m)^T\right] \;=\; \mathbf{0} \tag{10.6}$$

In the following KF structures, the functions $\boldsymbol{f}^m(\cdot)$ and $\boldsymbol{h}^m(\cdot)$ correspond to those presented in (7.14) and (7.19), respectively, and are briefly restated here for clarity:

$$\boldsymbol{f}^m(\boldsymbol{x}_k^m, \boldsymbol{u}_k^m) : \begin{cases} \boldsymbol{i}_{\mathrm{AB},k+1} \;=\; \boldsymbol{i}_{\alpha\beta,k} \;+\; T_s\; \boldsymbol{A}_k^{-1} \left[\boldsymbol{u}_{\alpha\beta,k} \;-\; \boldsymbol{R}_s\; \boldsymbol{i}_{\alpha\beta,k} \;\; - \omega_{E,k}\left((\boldsymbol{B}_k \;\; + \cdots \right.\right. \\ \qquad\qquad\quad \left.\left. \cdots + \boldsymbol{C}_k)\, \boldsymbol{i}_{\alpha\beta,k} \;+\; \boldsymbol{d}_k\right]\right] \\ \omega_{E,k+1} \;=\; \omega_{E,k} \\ \vartheta_{E,k+1} \;=\; \vartheta_{E,k} \;+\; T_s\; \omega_{E,k} \end{cases} \tag{10.7}$$

$$\boldsymbol{y}_k^m = \boldsymbol{h}^m(\boldsymbol{x}_k^m) = \boldsymbol{i}_{\alpha\beta,k} \tag{10.8}$$

The state vector is defined as $\boldsymbol{x}_k^m = [i_{\alpha,k},\; i_{\beta,k},\; \omega_{E,k},\; \vartheta_{E,k}]^{\mathrm{T}}$ $(n = 4)$, the input voltages as $\boldsymbol{u}_k = \left[u_{\alpha,k}^m,\; u_{\beta,k}\right]^{\mathrm{T}}$ $(m = 2)$, and the measured output vector as $\boldsymbol{y}_k^m = [i_{\alpha,k},\; i_{\beta,k}]^{\mathrm{T}}$ $(p = 4)$.

10.1.1 Master EKF

The master EKF is responsible for estimating the electrical position ϑ_E of the SynRM. Its equations are the same as those presented in 8.1.1, with the addition of the update of the process noise covariance matrix $\boldsymbol{Q}_m$.

Initialization

$$\hat{\boldsymbol{x}}_{0|0}^m = \mathbb{E}\left[\boldsymbol{x}_0^m\right] \tag{10.9}$$

$$\boldsymbol{P}_{0|0}^m = \mathbb{E}\left[\left(\boldsymbol{x}_0^m - \mathbb{E}\left[\boldsymbol{x}_0^m\right]\right)\;\left(\boldsymbol{x}_0^m - \mathbb{E}\left[\boldsymbol{x}_0^m\right]\right)^T\right] \tag{10.10}$$

Time update

$$\hat{\boldsymbol{x}}^m_{k|k-1} = \boldsymbol{f}^m\left(\hat{\boldsymbol{x}}^m_{k-1|k-1}, \boldsymbol{u}^m_{k-1}\right) \tag{10.11}$$

$$\hat{\boldsymbol{Q}}^m_k = \hat{\boldsymbol{x}}^s_{k-1|k-1}\boldsymbol{I}_{n\times n} \tag{10.12}$$

$$\boldsymbol{\Phi}^m_k = \left.\frac{\partial \boldsymbol{f}^m}{\partial \boldsymbol{x}^m}\right|_{\boldsymbol{x}^m_{k-1|k-1}, \boldsymbol{u}^m_{k-1}} \tag{10.13}$$

$$\boldsymbol{P}^m_{k|k-1} = \boldsymbol{\Phi}^m_k \boldsymbol{P}^m_{k-1|k-1}(\boldsymbol{\Phi}^m_k)^T + \hat{\boldsymbol{Q}}^m_k \tag{10.14}$$

Measurement update

$$\boldsymbol{H}^m_k = \left.\frac{\partial \boldsymbol{h}^m}{\partial \hat{\boldsymbol{x}}^m}\right|_{\boldsymbol{x}^m_{k|k-1}} \tag{10.15}$$

$$\boldsymbol{\Lambda}^m_k = \boldsymbol{H}^m_k \boldsymbol{P}^m_{k|k-1}(\boldsymbol{H}^m_k)^T + \boldsymbol{R}^m_k \tag{10.16}$$

$$\boldsymbol{K}^m_k = \boldsymbol{P}^m_{k|k-1}(\boldsymbol{H}^m_k)^T\,(\boldsymbol{\Lambda}^m_k)^{-1} \tag{10.17}$$

$$\tilde{\boldsymbol{y}}^m_k = \boldsymbol{y}^m_k - \boldsymbol{h}^m(\hat{\boldsymbol{x}}^m_{k|k-1}) \tag{10.18}$$

$$\hat{\boldsymbol{x}}^m_{k|k} = \hat{\boldsymbol{x}}^m_{k|k-1} + \boldsymbol{K}^m_k \tilde{\boldsymbol{y}}^m_k \tag{10.19}$$

$$\boldsymbol{P}^m_{k|k} = \left(\boldsymbol{I}_{n\times n} - \boldsymbol{K}^m_k \boldsymbol{H}^m_k\right)\boldsymbol{P}^m_{k|k-1} \tag{10.20}$$

where matrices $\boldsymbol{\Phi}^m_k$ and $\boldsymbol{H}^m_k$ are the state (8.12) and output (8.14) Jacobians, respectively, previously calculated in (8.12) and (8.14). It is worth mentioning that the estimated process noise covariance matrix $\hat{\boldsymbol{Q}}^m_k$ is updated at each time step k by taking the slave KF estimates $\hat{\boldsymbol{x}}^s_{k-1}$ as its diagonal elements, as shown in (10.12).

10.1.2 Slave KF

Under the assumption (10.2), the matrices $\boldsymbol{Q}_m$ and $\boldsymbol{R}_m$ are diagonal [102], [104], meaning that estimating the noise covariances corresponds to estimating the diagonal elements of these matrices. Given that measurement noise knowledge can be somewhat obtained from the accuracy of the current sensors, the focus is placed on process noise estimation. Therefore, the state vector of the slave filter corresponds to the diagonal elements of the $\boldsymbol{Q}_m$ matrix of the master EKF, expressed as

$$\hat{\boldsymbol{x}}^s = \mathrm{diag}\left\{\hat{\boldsymbol{Q}}^m\right\} = [\hat{q}^m_{11},\ \hat{q}^m_{22},\ \hat{q}^m_{33},\ \hat{q}^m_{44}]^T \tag{10.21}$$

Due to the uncertainty in the process noise dynamics, the discrete *time update* of the slave KF is given by

$$\hat{q}^m_{ii,\,k|k-1} = \hat{q}^m_{ii,\,k-1|k-1} \qquad i = 1, \ldots, 4 \tag{10.22}$$

For the slave measurement update, the algorithm processes the master innovation $\tilde{\boldsymbol{y}}_k^m$ and its covariance $\boldsymbol{\Lambda}_k^m$, as defined in (10.18) and (10.16), respectively. In particular, the slave filter aims to match the master error covariance matrix $\boldsymbol{\Lambda}_k^m$ with the empirical one $\hat{\boldsymbol{\Lambda}}_k^m$, defined as

$$\hat{\boldsymbol{\Lambda}}_k^m = \frac{1}{N} \sum_{j=k-W+1}^{k} \left(\tilde{\boldsymbol{y}}_j^m - \tilde{\boldsymbol{Y}}_k^m \right) \left(\tilde{\boldsymbol{y}}_j^m - \tilde{\boldsymbol{Y}}_k^m \right)^T \tag{10.23}$$

Here, N represents the window length, $\tilde{\boldsymbol{y}}_j^m$ denotes the innovation at time j, and $\tilde{\boldsymbol{Y}}_k^m$ is the average innovation over the window. The empirical innovation covariance, $\hat{\boldsymbol{\Lambda}}_k^m$, is calculated using (10.23) by storing the innovations $\tilde{\boldsymbol{y}}_j^m$ in a moving buffer of length N.

Based on these premises, the slave innovation $\tilde{\boldsymbol{y}}_k^s$ is defined as

$$\tilde{\boldsymbol{y}}_k^s = \text{diag}\left\{ \hat{\boldsymbol{\Lambda}}_k^m \right\} - \text{diag}\left\{ \boldsymbol{\Lambda}_k^m \right\} \tag{10.24}$$

Thus, the *measurement* provided to the slave filter consists of the diagonal elements of the empirical $\hat{\boldsymbol{\Lambda}}_k^m$, while the estimated output is the diagonal of the master KF internally-computed $\boldsymbol{\Lambda}_k^m$ [81], [104], [105].

The slave KF output model is obtained by considering the standard KF error covariance matrix update equation (B.52) or (8.17):

$$\boldsymbol{P}_{k|k}^m = \boldsymbol{P}_{k|k-1}^m - \boldsymbol{K}_k^m \boldsymbol{P}_{k|k-1}^m (\boldsymbol{K}_k^m)^T \tag{10.25}$$

From (10.25), we extract $\boldsymbol{\Lambda}_k^m$ and consider only its diagonal elements:

$$\text{diag}\left\{ \boldsymbol{\Lambda}_k^m \right\} = \text{diag}\left\{ \boldsymbol{K}_k^{m\#} (\boldsymbol{P}_{k|k-1}^m - \boldsymbol{P}_{k|k}^m)(\boldsymbol{K}_k^{m\#})^T \right\} \tag{10.26}$$

where $\boldsymbol{K}_k^{m\#}$ represents the Moore-Penrose inverse of the master Kalman gain $\boldsymbol{K}_k^m$.

The predicted outputs of the slave system are thus expressed as

$$\hat{\boldsymbol{y}}_k^s = \text{diag}\left\{ \boldsymbol{\Lambda}_k^m \right\} = \text{diag}\left\{ \boldsymbol{K}_k^{m\#} (\boldsymbol{P}_{k|k-1}^m - \boldsymbol{P}_{k|k}^m)(\boldsymbol{K}_k^{m\#})^T \right\} \tag{10.27}$$

The matrix $\boldsymbol{P}_{k|k}^m$ can be expressed through (8.13) and substituted into (10.27) to obtain:

$$\begin{aligned}
\hat{\boldsymbol{y}}_k^s &= \text{diag}\left\{ \boldsymbol{K}_k^{m\#} (\boldsymbol{P}_{k|k-1}^m - \boldsymbol{P}_{k|k}^m)(\boldsymbol{K}_k^{m\#})^T \right\} \\
&= \text{diag}\left\{ \boldsymbol{K}_k^{m\#} (\boldsymbol{\Phi}_k^m \boldsymbol{P}_{k|k-1}^m (\boldsymbol{\Phi}_k^m)^T + \hat{\boldsymbol{Q}}_k^m - \boldsymbol{P}_{k|k}^m)(\boldsymbol{K}_k^{m\#})^T \right\} \\
&= \text{diag}\left\{ \boldsymbol{K}_k^{m\#} \hat{\boldsymbol{Q}}_k^m (\boldsymbol{K}_k^{m\#})^T \right\} + \text{diag}\Big\{ \boldsymbol{K}_k^{m\#} \Big(\boldsymbol{\Phi}_k^m \boldsymbol{P}_{k|k-1}^m (\boldsymbol{\Phi}_k^m)^T + \cdots \\
&\qquad \cdots - \boldsymbol{P}_{k|k}^m \Big) (\boldsymbol{K}_k^{m\#})^T \Big\} \\
&= \left(\boldsymbol{K}_k^{m\#} \circ \boldsymbol{K}_k^{m\#} \right) \text{diag}\left\{ \hat{\boldsymbol{Q}}_k^m \right\} + \text{diag}\Big\{ \boldsymbol{K}_k^{m\#} \Big(\boldsymbol{\Phi}_k^m \boldsymbol{P}_{k|k-1}^m (\boldsymbol{\Phi}_k^m)^T + \cdots \\
&\qquad \cdots - \boldsymbol{P}_{k|k}^m \Big) (\boldsymbol{K}_k^{m\#})^T \Big\} \\
&= \boldsymbol{H}_k^s \hat{\boldsymbol{x}}_{k|k-1}^s + \boldsymbol{u}_k^s
\end{aligned} \tag{10.28}$$

where the states $\hat{\boldsymbol{x}}_k^s$ and outputs $\hat{\boldsymbol{y}}_k^s$ together form the state space model for estimating the $\boldsymbol{Q}_k^m$-matrix and $\boldsymbol{H}_k^s$ is the slave output matrix. They are defined as follows:

$$\hat{\boldsymbol{x}}_{k|k-1}^s = \operatorname{diag}\left\{\hat{\boldsymbol{Q}}_k^m\right\} \tag{10.29}$$

$$\mathbf{H}_k^s = \boldsymbol{K}_k^{m\#} \circ \boldsymbol{K}_k^{m\#} \tag{10.30}$$

$$\boldsymbol{u}_k^s = \operatorname{diag}\left\{\boldsymbol{K}_k^{m\#}\left(\boldsymbol{\Phi}_k^m \boldsymbol{P}_{k|k-1}^m (\boldsymbol{\Phi}_k^m)^T - \boldsymbol{P}_{k|k}^m\right)(\boldsymbol{K}_k^{m\#})^T\right\} \tag{10.31}$$

in which the symbol $\circ$ denotes the element-wise multiplication.

After considering the state and output models, the slave KF structure is summarized as follows:

Initialization

$$\hat{\boldsymbol{x}}_{0|0}^s = \mathbb{E}\left[\boldsymbol{x}_0^s\right] \tag{10.32}$$

$$\boldsymbol{P}_{0|0}^s = \mathbb{E}\left[\left(\boldsymbol{x}_0^s - \mathbb{E}\left[\boldsymbol{x}_0^s\right]\right)\ \left(\boldsymbol{x}_0^s - \mathbb{E}\left[\boldsymbol{x}_0^s\right]\right)^T\right] \tag{10.33}$$

Time Update

$$\hat{\boldsymbol{x}}_{k|k-1}^s = \hat{\boldsymbol{x}}_{k-1|k-1}^s \tag{10.34}$$

$$\boldsymbol{P}_{k|k-1}^s = \boldsymbol{P}_{k-1|k-1}^s + \boldsymbol{Q}_k^s \tag{10.35}$$

$$\hat{\mathbf{y}}_{k|k-1}^s = \boldsymbol{H}_k^s \hat{\boldsymbol{x}}_{k|k-1}^s + \boldsymbol{u}_k^s \tag{10.36}$$

where $\boldsymbol{H}_k^s$ and $\boldsymbol{u}_k^s$ are defined in (10.30) and (10.31), respectively.

Measurement Update

$$\boldsymbol{K}_k^s = \boldsymbol{P}_{k|k-1}^s (\boldsymbol{H}_k^s)^T \left(\boldsymbol{H}_k^s \boldsymbol{P}_{k|k-1}^s (\boldsymbol{H}_k^s)^T + \boldsymbol{R}_k^s\right)^{-1} \tag{10.37}$$

$$\hat{\boldsymbol{x}}_{k|k}^s = \hat{\boldsymbol{x}}_{k|k-1}^s + \boldsymbol{K}_k^s \left(\boldsymbol{y}_k^s - \hat{\boldsymbol{y}}_{k|k-1}^s\right) \tag{10.38}$$

$$\boldsymbol{P}_{k|k}^s = \left(\boldsymbol{I}_{n\times n} - \boldsymbol{K}_k^s \boldsymbol{H}_k^s\right) \boldsymbol{P}_{k|k-1}^s \tag{10.39}$$

where $\boldsymbol{I}_{n\times n}$ is the identity matrix and $\boldsymbol{y}_k^s$ corresponds to the diagonal of the *empirical* innovation covariance defined in (10.23), whereas $\hat{\boldsymbol{y}}_{k|k-1}^s$ is defined in (10.28).

Fig. 10.1 illustrates the interconnection between the master EKF and the slave KF, as defined by equations (10.12), (10.30), (10.31) and (10.23).

Fig. 10.2 shows the flowchart of the MSKF algorithm, it might help to understand the sequence of operations performed by the master EKF and the slave KF.

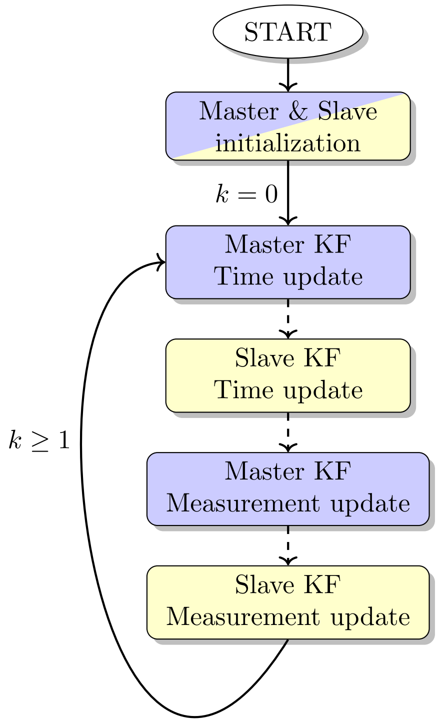


Figure 10.2 Master-Slave Kalman Filter flow-chart.

10.1.3 Filter convergence conditions

Regarding the convergence of the Master-Slave KF algorithm, the following considerations are nearly identical to those typically made for the standard EKF [67]. This similarity arises because the master KF functions as a standard EKF with a variable $\hat{\boldsymbol{Q}}_k^m$ and a constant $\boldsymbol{R}_k^m$, while the slave KF operates as a conventional linear KF with fixed noise covariances.

Therefore, only the convergence of the master EKF will be analyzed. Its convergence is ensured if the following assumptions hold [100], [101]:

- **Assumption** 1

  1. The covariance matrices $\hat{\boldsymbol{Q}}_k^m$ and $\boldsymbol{R}^m$ are bounded, ensuring that $\hat{\boldsymbol{Q}}_k^m$ is always positive semi-definite and $\boldsymbol{R}^m$ is positive definite. Specifically, there exist real numbers $\underline{q}^m$, $\overline{q}^m$, $\underline{r}^m$, and $\overline{r}^m$ such that

$$\underline{q}^m \boldsymbol{I}_{4\times4} \leq \hat{\boldsymbol{Q}}_k^m \leq \overline{q}^m \boldsymbol{I}_{4\times4} \tag{10.40}$$

$$\underline{r}^m \boldsymbol{I}_{2\times2} \leq \boldsymbol{R}^m \leq \overline{r}^m \boldsymbol{I}_{2\times2} \tag{10.41}$$

  for each time step $k \geq 0$. This condition is satisfied because $\boldsymbol{R}^m$ is fixed, with its diagonal composed of positive real numbers. Meanwhile, the diagonal elements of $\hat{\boldsymbol{Q}}_k^m$, estimated by the slave KF, are bounded in implementation to lie between zero and positive real numbers. These boundaries are discussed in detail in Subsect. 10.2.3.

2. The matrices $\boldsymbol{\Phi}_k^m$ and $\boldsymbol{H}_k^m$, equations (8.12) and (8.14) respectively, satisfy the uniform observability condition [90], [92], please refer to Sect. 7.2 for more details about the system observability.
3. The initial error covariance matrix $\boldsymbol{P}_{0|0}^m$ is positive definite.

- **Assumption** 2

1. There are positive real numbers $\overline{\gamma}_1^m$, $\overline{\gamma}_2^m$, $\underline{p}^m$ and $\overline{q}^m$ such that the following inequalities hold for each time step $k \geq 0$.

$$||\boldsymbol{\Phi}_k^m|| \leq \overline{\gamma}_1^m \tag{10.42}$$

$$||\boldsymbol{H}_k^m|| \leq \overline{\gamma}_1^m \tag{10.43}$$

$$\underline{p}^m \boldsymbol{I}_{4\times 4} \leq \boldsymbol{P}_{k|k}^m \leq \overline{p}^m \boldsymbol{I}_{4\times 4} \tag{10.44}$$

2. $\boldsymbol{\Phi}_k^m$, eq. (8.12) is non-singular for each time $k \geq 0$.

The conditions listed in the aforementioned assumptions are checked online within the algorithm code to prevent prohibited operations, such as matrix inversion, when the matrices are singular or near singularity. Regarding the bounding conditions, as previously mentioned, the values in (10.40) are detailed in Subsect. 10.2.3, while (10.42) is satisfied by preventing the state $\boldsymbol{x}_{k|k}^m$, through which $\boldsymbol{\Phi}_k^m$ is calculated, from diverging or becoming equal to `NaN` (Not a Number). Equation (10.43) is satisfied by definition since $\boldsymbol{H}_k^m$ is a constant matrix, as shown in (8.14). Finally, (10.44) is satisfied by ensuring that the master error covariance matrix $\boldsymbol{P}_{k|k}^m$ remains positive definite. This can be checked online using a specific MATLAB function that calculates its determinant and verifies if it is greater than zero. However, this function increases the computational cost of the algorithm, so it was used only for debugging purposes. In the final implementation, the positive definiteness of $\boldsymbol{P}_{k|k}^m$ was ensured by observing the other boundaries.

10.2 Experimental Results

In this section, the performance of the MSKF and a manually tuned EKF in a sensorless SynRM drive is evaluated. The tests are conducted under typical operating conditions, including speed inversion, load handling, and four-quadrant operations. The results are compared in terms of the MSE of the estimated state variables. The robustness of both filters is also assessed by evaluating their sensitivity to stator resistance mismatches and their ability to handle the pull-out curve. Additionally, the execution time of the control algorithm components is analyzed.

10.2.1 Test bench

The experiments were conducted using a test bench comprising a SynRM, the nameplate data of which are detailed in Table 10.1, and a PMSM mechanically connected to the SynRM, serving as a virtual load. Both motors were powered by Space Vector Modulation (SVM) inverters, sharing a 400 V DC-Link, with a switching frequency

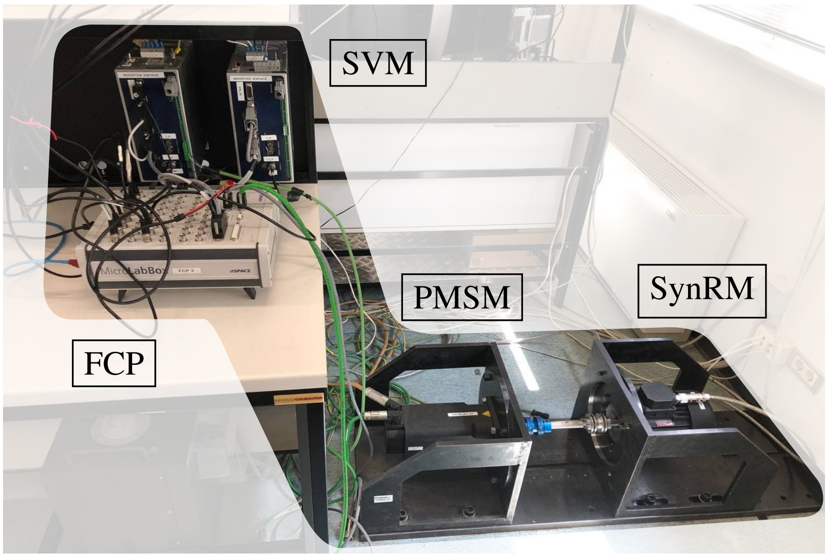


Figure 10.3 The experimental test bench.

Table 10.1 System and SynRM parameters

Parameters	Symbol	Unit	Value
Nominal torque	τ_n	N m	3.5
Nominal current	I_n	A	3
Nominal speed	ω_n	min^{-1}	1200
Stator resistance	r_s	Ω	4.72
unsat. d-axis inductance	L_d	mH	380
unsat. q-axis inductance	L_q	mH	85
pole pairs	p	\	2
Voltage Source Inverter, 2-levels IGBT			
DC voltage	u_C	V	400
Dead-time	t_d	µs	2
Switching frequency	f_s	kHz	8
Control Platform	dSPACE MicroLabBox		
Control frequency	f_c	kHz	8

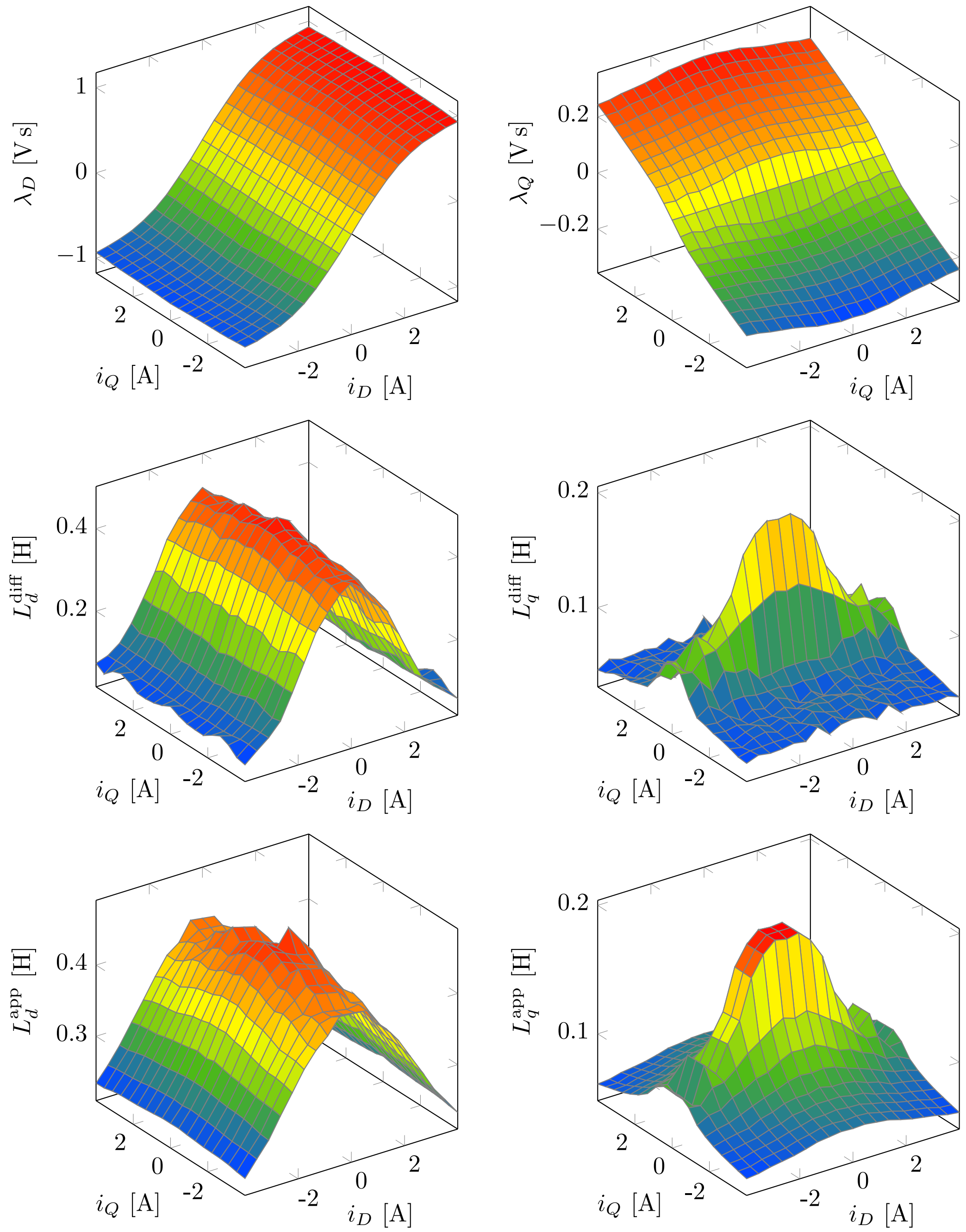


Figure 10.4 Flux linkages $\boldsymbol{\lambda}_{DQ}$, apparent inductances $\boldsymbol{L}_{dq}^{\text{app}}$, and differential inductances $\boldsymbol{L}_{dq}^{\text{diff}}$ of the SynRM used in the experiments (Table 10.1).

of 8 kHz. Drive control was implemented using a MicroLabBox dSpace Fast Control Prototype (FCP) system. The test bench is illustrated in Fig. 10.3.

The magnetic model of the SynRM is depicted in Fig. 10.4, these maps were obtained through an elaboration of a series of measurements [88] and they are used in the time update step in (10.11).

10.2.2 Control Design

The current control consists of two gain-scheduling PI controllers, whose implementation details are described in [95]. The PI gains were tuned to achieve a constant 150 Hz bandwidth with 70 deg phase margin. This eases the design of the outer speed control loop, which consists of a conventional PI controller, with antiwind-up and fixed bandwidth of 5 Hz. The current and speed control block diagram is reported in Fig. 10.5. All experiments were performed using the quantities estimated by the MSKF as feedback.

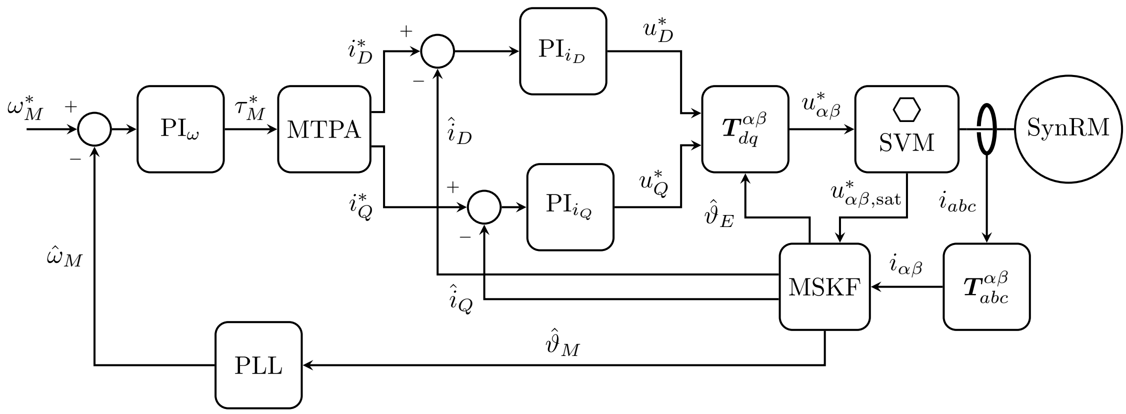


Figure 10.5 The proposed MSKF-based sensorless control of the SynRM.

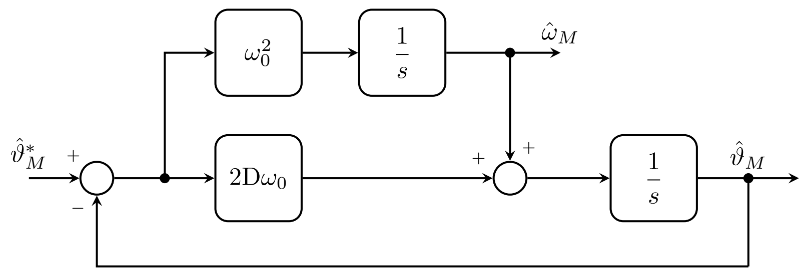


Figure 10.6 The PLL block diagram used within the experiments [96].

As inferable from (7.10), the mechanical model of the EKF is based on the assumption of infinite inertia, which simplifies the differential equation of speed in (7.10) to a null value. Consequently, the speed update relies solely on the EKF gain correction, resulting in both a noisy speed estimate and high sensitivity to the weight of the $\boldsymbol{Q}^m$ matrix. Although such a highly dynamic speed estimate does not affect the position estimation, it is not smooth enough to be used directly as feedback in the speed control loop. To address these issues, a PLL filter is employed to recover the speed $\hat{\omega}_M$ from the estimated position $\hat{\vartheta}_M$. The PLL has been tuned with a characteristic frequency of $\omega_0 = 2\pi \cdot 50\,\mathrm{Hz}$ and a damping factor of $\mathrm{D} = 1$ [96]. For clarity, the PLL speed is used only for feedback and not in the EKF or MSKF model. The block diagram of the PLL is shown in Fig. 10.6.

An approximate MTPA strategy was implemented by dividing the output of the speed Proportional-Integral (PI_{ω_m}) controller into equal current references, denoted i_D^* and i_Q^*, which corresponds to setting the current vector phase θ_i^{MTPA} to 45 deg. As in the comparison between EKF and UKF in Sect. 8.3, the sensorless algorithm is improved by supplying the SynRM with a minimum d-axis current even under no-load conditions [97]. In this case, a minimum d-axis current of $I_D^{\mathrm{min}} = I_n/3$ was established.

This value is essential under zero-speed and no-load conditions. In these scenarios, without a minimum current vector, the KF innovation, whether in the EKF or the MSKF, becomes either zero or negligibly small. Consequently, the KF fails to extract meaningful information, as the innovation becomes indistinguishable from measurement noise, resulting in unreliable estimations. To address this issue, imposing a minimum current vector, below which the innovation cannot be zero, offers an effective solution to improve observer reliability.

In [42], an interesting investigation is presented on the relationship between the minimum current vector and the minimum bearable speed. It is shown that as the current limit increases, the minimum bearable speed decreases.

10.2.3 Tuning of the proposed MSKF

The experiments involved a performance comparison between a manually tuned EKF and the proposed MSKF. In the case of the EKF, the trial-and-error approach yielded the values

$$\boldsymbol{Q}^{\mathrm{EKF}} = \mathrm{diag}\,\{0.01, , 0.01, , 20, , 0.001\} \tag{10.45}$$

$$\boldsymbol{R}^{\mathrm{EKF}} = \mathrm{diag}\,\{0.001, , 0.001\} \tag{10.46}$$

which remained constant throughout the experiments. It is important to note that the measurement noise matrix $\boldsymbol{R}$ represents the noise introduced by the current sensors and was set based on the measured variance of the current sensor output noise.

For the MSKF, the master filter matrix $\boldsymbol{Q}_k^m$ was tuned and updated online by the slave KF, while the matrix $\boldsymbol{R}_k^m$ was set equal to $\boldsymbol{R}^{\mathrm{EKF}}$, since the output model and measurement noise are identical for both filters.

In the slave KF, the initial state $\hat{\boldsymbol{x}}_{0|0}^s$ was arbitrarily set to $\hat{\boldsymbol{x}}_{0|0}^s = [1, 1, 1, 1]^T$, and consequently, $\hat{\boldsymbol{Q}}_{0|0}^m$ was initialized to the identity matrix. The covariance matrices were configured as $\boldsymbol{R}_k^s = r^s \cdot \mathrm{diag}\,\{1, 1\}$ and $\boldsymbol{Q}_k^s = q^s \cdot \mathrm{diag}\,\{1, 1, 1, 1\}$. The ratio q^s/r^s significantly influences the slave KF estimation and, consequently, the master EKF matrix $\hat{\boldsymbol{Q}}_k^m$. It was found that a higher ratio leads to quicker (but noisier) convergence of the slave KF. In simulations, a ratio of $q^s/r^s = 100$ provided a satisfactory trade-off. Experimental results confirmed this observation, but also highlighted that the value is not critical and can vary without significantly affecting the system. For instance, a ratio of $q^s/r^s = 10$ was also tested, yielding similar results.

As for the initial covariance matrices of the master and slave filters, $\boldsymbol{P}^m_{0|0}$ and $\boldsymbol{P}^s_{0|0}$, they were set to the identity matrix. These initial values had little impact on the system, as the filter quickly converged to the correct solution, even when different initial conditions were used, such as $10 \cdot \boldsymbol{I}n \times n$ or $0.1 \cdot \boldsymbol{I}n \times n$.

When dealing with unmodeled disturbances, such as dead times, highly nonlinear systems, or significant nonzero mean measurement noise, the predicted covariance change, $\boldsymbol{P}^m_{k|k-1} - \boldsymbol{P}^m_{k|k}$, may lose its positive definiteness. This issue compromises the validity of the covariance matrix, as it would imply no reduction in uncertainty during the correction step, potentially leading to estimation divergence. To maintain consistency, the state estimate $\hat{\boldsymbol{x}}^s_{k|k-1}$ of the slave KF was constrained to stay positive, ensuring a semi-positive definite process noise covariance matrix $\boldsymbol{Q}^m_k$. This constraint was implemented by setting the minimum value of the state estimate to $\hat{\boldsymbol{x}}^s_{k|k-1} \geq \boldsymbol{0}$.

Preliminary simulation tests, later validated experimentally, showed that the q weight of the speed state variable, $\hat{q}^m_{33,k}$, must have a minimum value greater than 0. Without this, under steady-state conditions, the speed estimate ω_E becomes unresponsive to variations in speed reference or load torque. This leads to growing estimation errors until the slave KF adjusts $\hat{q}^m_{33,k}$. To prevent this, a minimum weight constraint was enforced during experiments:

$$\hat{\boldsymbol{Q}}^m_k \succeq \operatorname{diag}\{0,\ 0,\ 5,\ 0\} \tag{10.47}$$

The value $\underline{q}^m_{33} = 5$ was determined during the experimental phase, specifically to enable the MSKF to handle the four quadrant operation test, which will be discussed later. This test involves speed inversions. The constraint was necessary to prevent the speed estimate from becoming stuck at zero, which would result in the loss of control. For this reason, $\underline{q}^m_{33}$ can be considered both a critical value for the proposed MSKF and a tuning parameter.

The upper bounds were set to $\hat{\boldsymbol{Q}}^m_k \preceq \operatorname{diag}\{1,\ 1,\ 100, ,\ 1\}$, which were found to be sufficiently large to avoid constraint violations. It is important to note that these upper limits were never exceeded during the experiments, meaning they were not necessary for the test cases considered.

Lastly, the length of the window N in (10.23), which determines the number of innovation samples used by the slave filter, plays a crucial role. A shorter window increases the filter's responsiveness to load or speed changes, as reflected in the innovations, while a longer window makes the system more robust to disturbances but slower in following rapid transients. Through extensive simulations, a window size of $N = 10$ was found to offer a good trade-off between responsiveness and robustness.

10.2.4 Performance Comparison

A comprehensive measurement campaign was conducted to compare the manually tuned EKF with the proposed MSKF topology, which incorporates self-adaptive online tuning. The objective was to demonstrate that the MSKF can match the performance of the EKF perfectly, while partially eliminating the need for manual tuning.

The experimental results, summarized in Table 10.2, are presented in the following sections. In each test, the quantities estimated by both the EKF and the MSKF (after coordinate transformation), namely $\hat{\boldsymbol{x}} = [\omega_M, i_D, i_Q, \vartheta_E]$, were collected at the control frequency listed in Table 10.1, alongside the measured values. It is important to note that the measurements were available in the prototype solely for comparison purposes. The MSE was calculated as described in Subsect. 8.3.3.

The Minimum Speed at Rated Torque (MSRT) test

The test begins by setting a positive speed reference under no-load conditions. Before initiating the transient load torque at $t = 0.5\,\mathrm{s}$, the SynRM accelerates from 0.5 p.u. to 1 p.u. speed, still under no load. Subsequently, the reference speed is reduced to 5% of the rated speed. This minimum threshold was determined through repeated tests until uncontrollable oscillations occurred at low speeds.

The results are presented in Fig. 10.9. In the leftmost figure, the black line represents the speed reference provided to the SynRM, while the red line depicts the torque reference applied to the PMSM, which serves as a virtual load. The blue trace corresponds to the estimated speed, denoted as $\hat{\omega}_M$, and obtained from the PLL block shown in Fig. 10.5.

From left to right, the other figures display the estimated current on the d-axis $\hat{i}_D$, the estimated current on the q-axis $\hat{i}_Q$, and the electrical position error $\varDelta\vartheta_E = \hat{\vartheta}_E - \vartheta_E^{\mathrm{meas}}$ during the speed and torque transients.

At full load and stable minimum speed, both the EKF and the MSKF exhibit oscillations in the position estimate (see Fig. 10.9a). The position error remains limited within $[0-3°]$ for the proposed MSKF and within $[0-5°]$ for the conventional EKF. In both cases, the control effectively manages the speed's downward ramp, but any further reduction in speed leads to instability.

Referring to Table 10.2, it is evident that the MSE of the speed is lower in the EKF than in the MSKF. However, the MSKF provides more accurate results for both currents and the electrical position.

The estimated diagonal elements of $\hat{\boldsymbol{Q}}^m$ are shown in Fig. 10.7. These diagonal elements of $\hat{\boldsymbol{Q}}^m$ are updated online by the slave KF, which adapts the master filter to the varying conditions of the SynRM.

The Four-Quadrant Operation (FQO) test

The test begins with a positive nominal speed reference and a rated load torque, τ_L^*. While maintaining the same torque, a speed reversal is initiated at $t = 0.5\,\mathrm{s}$, transitioning from the first quadrant (labeled as 1Q) to the second quadrant (2Q), as shown in Fig. 10.10. At $t = 1.5\,\mathrm{s}$, a load torque reversal is triggered, transitioning from $1,$p.u. to $-1,$p.u.. Finally, at $t = 3\,\mathrm{s}$, a second speed reversal is imposed to exit the third quadrant (3Q) and enter the fourth quadrant (4Q).

It is worth noting that the electrical position estimation error tends to increase during speed zero-crossings, a well-known limitation in sensorless techniques that do not utilize injected signals for estimation [90], [92]. Nevertheless, when comparing

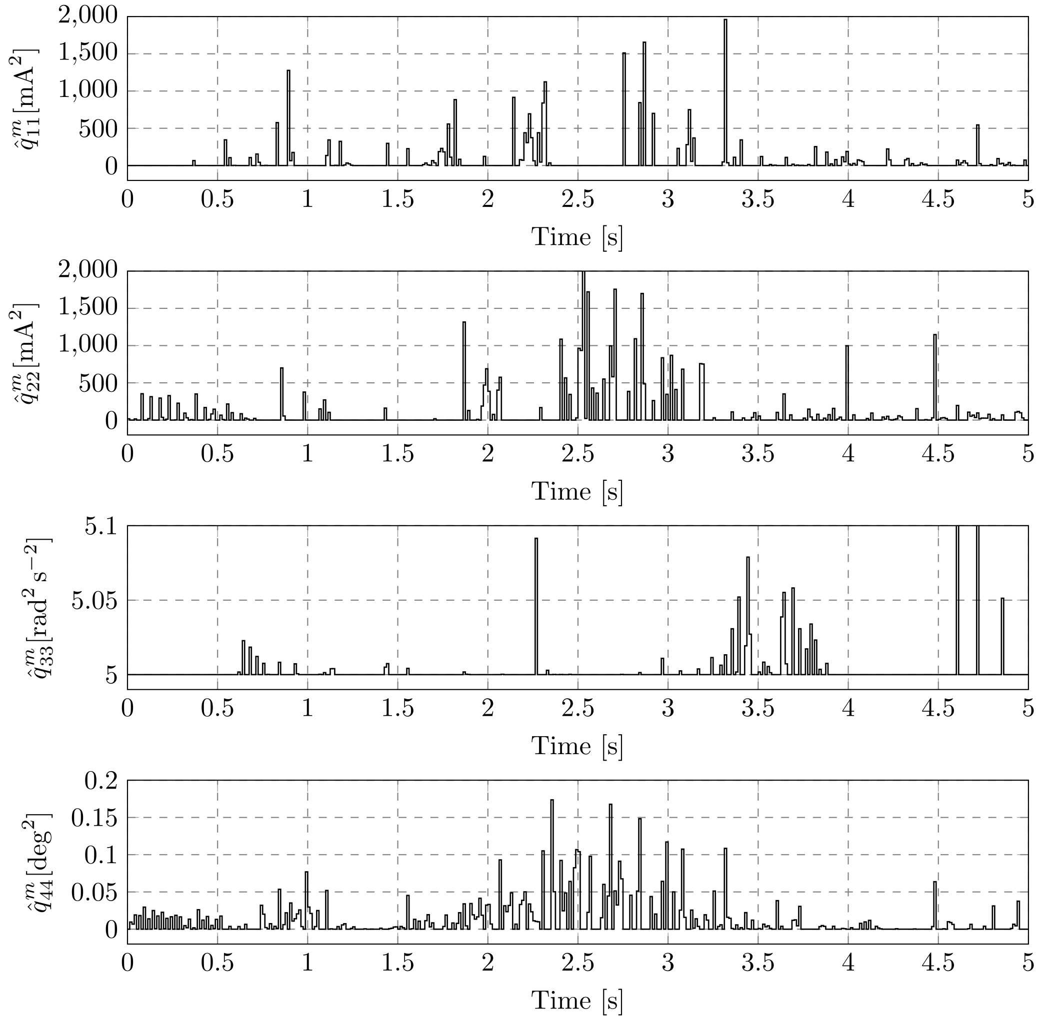


Figure 10.7 Estimated diagonal elements of $\hat{\boldsymbol{Q}}^m$ in the MSRT test, depicted in Fig. 10.9.

Fig. 10.10a and Fig.10.10b, it becomes evident that the MSKF exhibits a lower $\Delta\vartheta_E$ than the EKF.

The manually tuned EKF provides a slightly better speed MSE compared to the MSKF, with a difference of only 12 rpm^2. Again, the MSKF outperforms the EKF in the other state variables (the MSE is half for i_D), one third for i_Q and just one fifth for ϑ_E.

The slight improvements shown in Fig. 10.9 and in Fig. 10.10 can be attributed to the adaptive nature of the proposed tuning procedure, which accommodates varying environmental conditions.

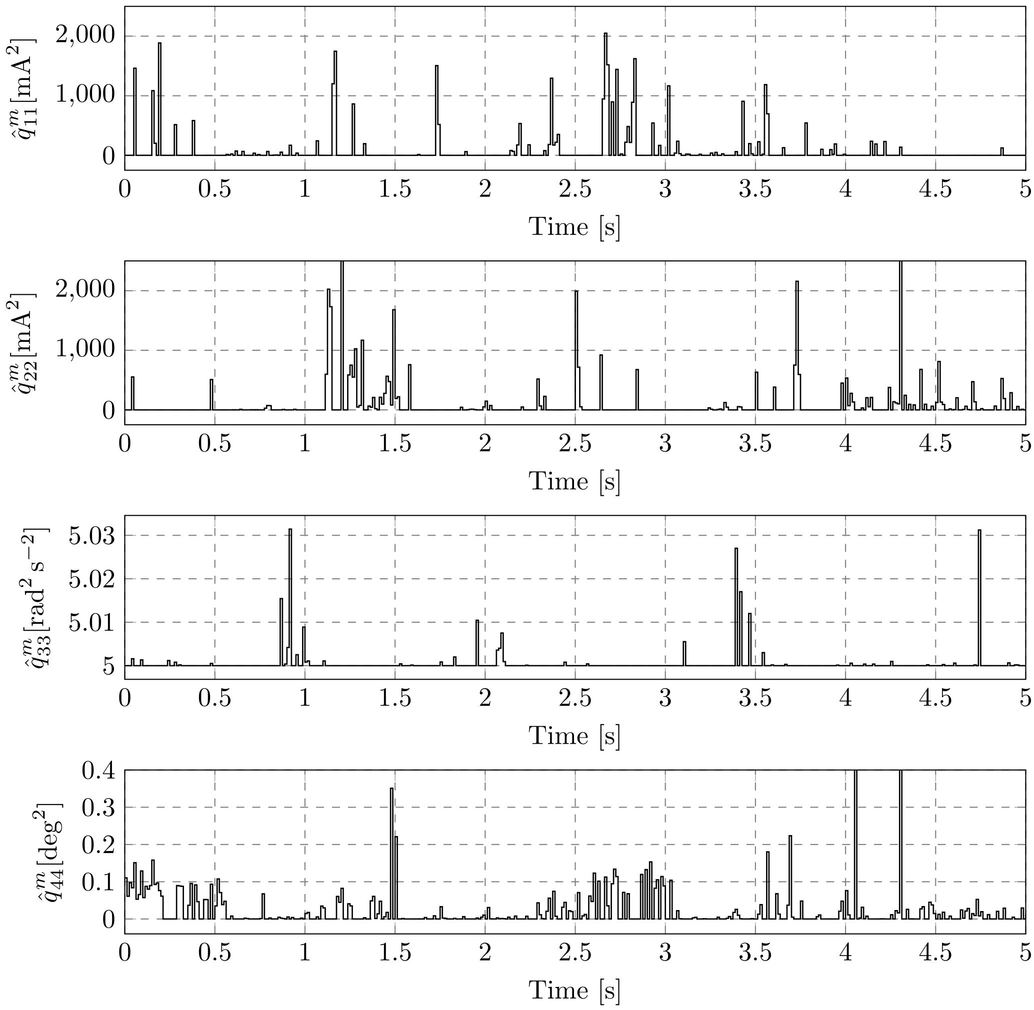
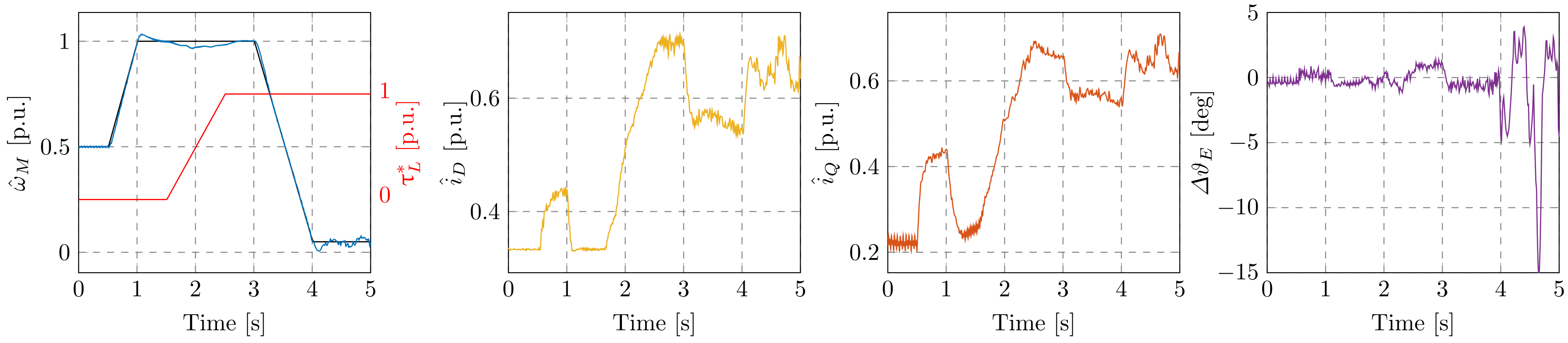
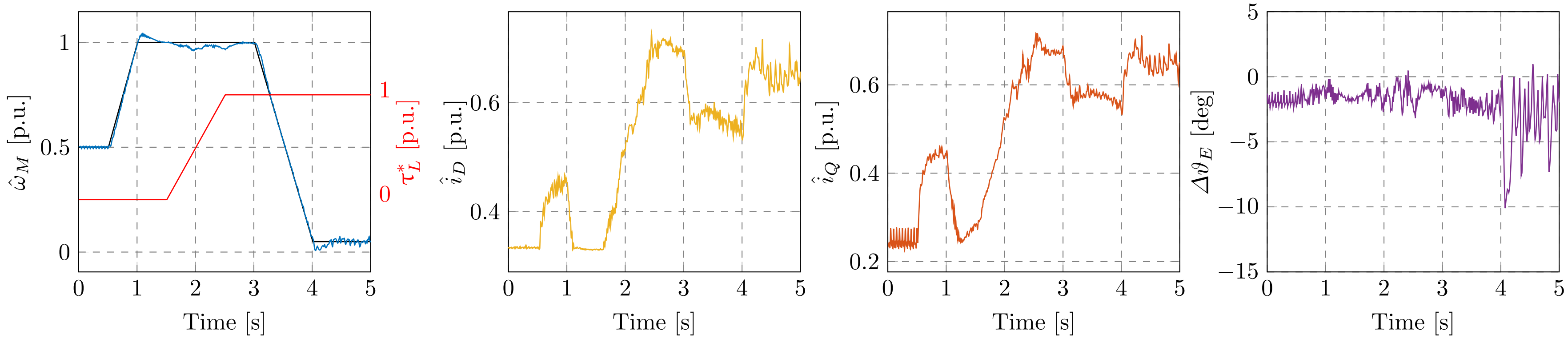


Figure 10.8 Estimated diagonal elements of $\hat{\boldsymbol{Q}}^m$ in the FQO test, depicted in Fig. 10.10.

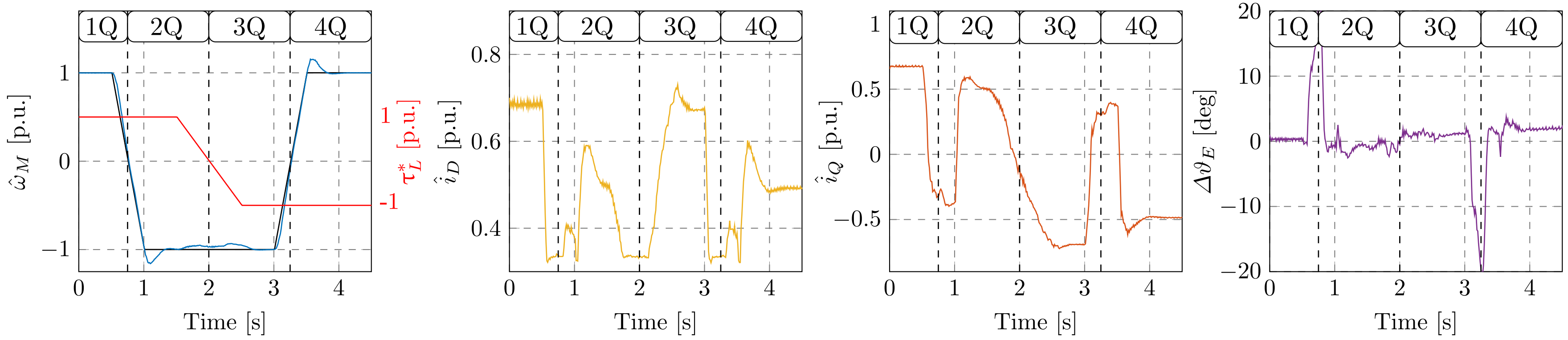


(a) Conventional EKF (manually tuned)

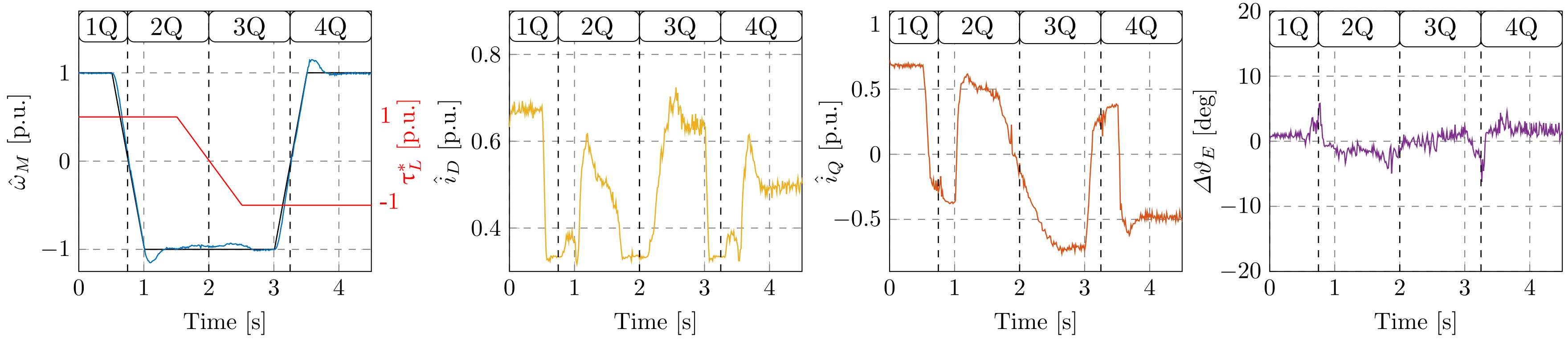


(b) Proposed MSKF (self-tuned)

Figure 10.9 Minimum Speed at Rated Torque (MSRT) test for conventional EKF (upper row) and MSKF (lower row). From left to right: mechanical speed ω_M, currents on the d-axis i_D and the q-axis i_Q, electrical position error $\Delta\vartheta_E$. All values refer to their respective nominal values as listed in Table 10.1.



(a) Conventional EKF (manually tuned)



(b) Proposed MSKF (self-tuned)

Figure 10.10 Four Quadrant Operation (FQO) test for conventional EKF (upper row) and MSKF (lower row). From left to right: mechanical speed ω_M, currents on the d-axis i_D and the q-axis i_Q, electrical position error $\Delta\vartheta_E$. All values refer to their respective nominal values as listed in Table 10.1.

The Step Starting from Standstill (SSS) test

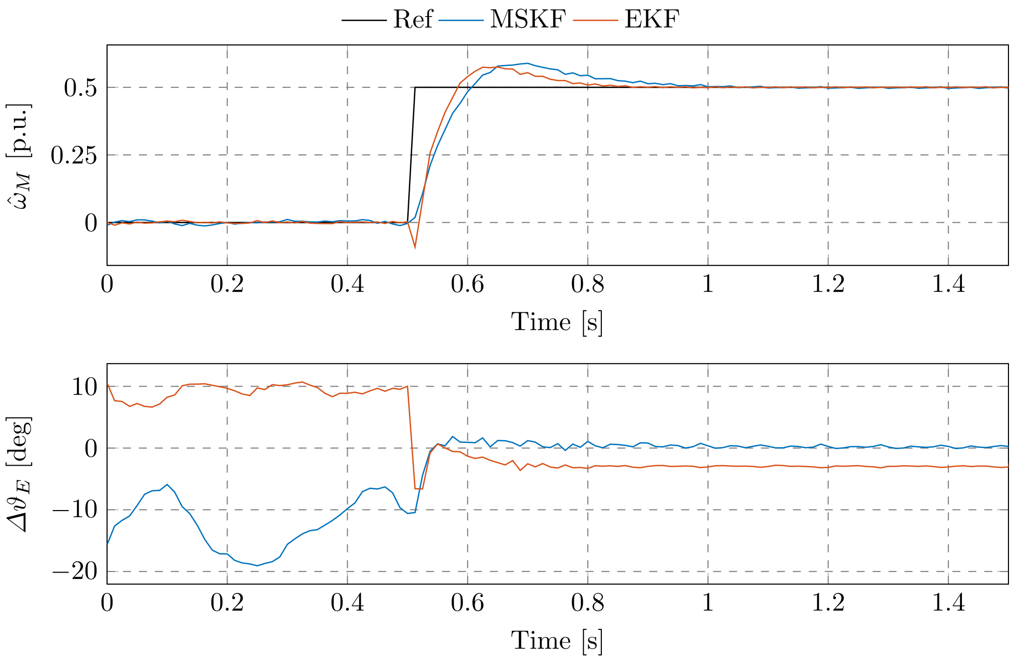


Figure 10.11 Step Starting from Standstill (SSS) test: mechanical speed (top) and electrical position error (bottom). The speed is referenced to its respective nominal value as listed in Table 10.1.

The test evaluates the drive performance when the motor starts from a complete standstill and a sudden change in speed reference is applied. In particular, at $t = 0.5$, a speed step is executed up to half the nominal speed. The state of the proposed MSKF is initialized randomly.

The test shows that both EKF and MSKF are capable of initiating motion from a standstill condition. In terms of speed, as shown in Fig. 10.11, they exhibit nearly identical responses. Note that the EKF is already tuned, while the MSKF relies just on the self-tuning at standstill and no load.

Regarding the electrical position error $\Delta\vartheta_E$ displayed in Fig. 10.11, prior to the step command, the EKF position ϑ_E appears stable, with an error of approximately 10 deg. In contrast, the position error derived from the MSKF appears less stable, oscillating between -7 deg and -20 deg. However, following the speed step, the position error for the EKF stabilizes around -3 deg, while the MSKF solution settles approximately at 0.4 deg.

In contrast to previous tests, the EKF exhibits an MSE higher than that of the MSKF for the speed variable, while it performs better than that of the MSKF when considering $\boldsymbol{i}_{DQ}$ and ϑ_E. To understand this behavior, it is important to note that the MSE calculations are performed over the entire test time. As shown in Fig. 10.11, before the initiation of the speed step, $\Delta\vartheta_E^{\mathrm{MSKF}}$ is greater than in the case of the EKF, and this behavior significantly influences the MSE calculation. Between $t = 1.2$ and $t = 1.5$, the MSE related to electrical position drops to $8.92\,\mathrm{deg}^2$ for the EKF and to $0.27\,\mathrm{deg}^2$ for the MSKF.

Table 10.2 Mean Squared Error (MSE)

Variable	Symbol	Unit	Test	EKF	MSKF	Fig.
Mechanical speed	ω_M	rpm^2	MSRT	38.6	100.7	10.9
			FQO	160.5	172.2	10.10
			SSS	268.7	65.9	10.11
d-axis current	i_D	mA^2	MSRT	4783	2583	10.9
			FQO	4471	2754	10.10
			SSS	3799	3058	\
q-axis current	i_Q	mA^2	MSRT	6652	2447	10.9
			FQO	9208	3108	10.10
			SSS	14680	21955	\
Electrical position	ϑ_E	deg^2	MSRT	4.63	2.37	10.9
			FQO	22.38	4.62	10.10
			SSS	35.24	57.56	10.11

Sensitivity to parameter mismatch

Table 10.3 Resistance mismatch test

r_s^{KF}/r_s	Mean ($\Delta\vartheta_E$)		Unit	MSE (ϑ_E)		Unit
	EKF	MSKF		EKF	MSKF	
1	-1.24	-2.45		3.68	7.72	
0.9	1.2	-2.69		4.8	9.17	
0.8	4.41	-2.98	deg	20.59	11.73	deg^2
0.7	6.22	-2.68		39.65	13.06	
0.6	8.21	-3.64		68.46	17.69	
0.5	9.85	-3.72		98.97	20.16	

The measurement noise is associated with the current sensors, which are assumed to be known. The adaptation proposed in this thesis focuses solely on the $\boldsymbol{Q}^m$ matrix, i.e., on the stochastic properties of process noise. The effect of process noise variation was studied by introducing parameter mismatches, specifically in resistance and flux linkage. The resistance mismatch is attributed to temperature variations, which cause the actual stator resistance to increase compared to the value used in the model (7.14).

Six experiments were carried out under steady-state conditions with a speed reference of $0.1,$p.u. and rated load torque. Each experiment involved progressively increasing resistance mismatch. The results, including the mean electrical position error and MSE, are summarized in Table 10.3. The resistance mismatch is represented as the ratio r_s^{KF}/r_s, where r_s^{KF} is the value used in the filter model, and r_s is the nominal value listed in Table 10.1.

For resistance mismatches within the range of $r_s^{\mathrm{KF}} = 0.9, r_s$, the EKF yields more accurate results. However, as the mismatch increases, the MSKF provides more precise estimations of the position error. Notably, the mean value of the electrical position error, $\Delta\vartheta_E$, is significantly affected by the resistance value used in the filter. However, in the MSKF case, the mean value remains almost constant across the range of resistance mismatches, demonstrating reduced sensitivity to resistance estimation. The mean value shown in Table 10.3 was calculated by averaging the electrical position error over 1600 samples.

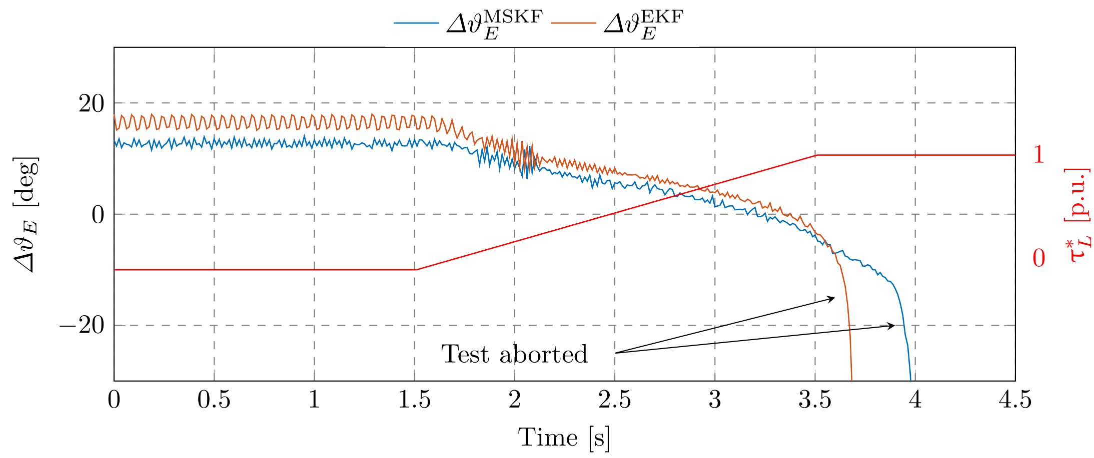


Figure 10.12 Electrical position error, performance in case of linear magnetic model (conventional EKF vs. proposed MSKF).

A second set of experiments was conducted to evaluate the robustness against potential discrepancies in the magnetic model used in the Kalman filter equations (7.14). The test involved replacing the magnetic maps shown in Fig. 8.1 with constant unsaturated inductance values listed in Table 10.1 (referred to as *linear model*). The speed was maintained at a constant value of 0.5 p.u., while a ramp up of the load torque to the full load torque was applied, as illustrated in Fig. 10.12. Before the onset of the load torque, the electrical position error was approximately 17 deg for the EKF and 13 deg for the MSKF, with less oscillation. Both the EKF and the MSKF proved unable to cope with the substantial magnetic mismatch at full load due to motor saturation. The MSKF exhibited a delayed decline, which offered limited improvement. However, setting the load torque ramp to 95% of the full load, both filters passed the test.

The inclusion of an accurate magnetic model in a SynRM observer is not optional but essential, whether using a conventional EKF or the proposed MSKF. While the MSKF can adapt to varying process noise conditions and, in theory, even cope with a mismatched magnetic model, for a loaded SynRM, the model mismatch is too significant to be effectively compensated by the proposed self-tuning mechanism.

Execution Time

As one would expect, the computational load of the MSKF is heavier than that of the conventional EKF, as it involves two simultaneously operating Kalman Filters. In the experiments, we used an integrated test bench for fast prototyping, which

Table 10.4 Execution times of the control algorithm elements

Routine	Magnetic Model	Execution Time
Control cycle	\	125 µs
Speed & Current Control	\	22 µs
Conventional EKF	Linear	14.7 µs
	Nonlinear	15.6 µs
Proposed MSKF	Linear	16.6 µs
	Nonlinear	17.5 µs

primarily consists of a dSpace control board and a PC-based Matlab-Simulink environment. Once the system is programmed in this environment, a real-time interface compiles the code and uploads it to the dSpace board. The compilation utilizes linear algebra libraries designed for real-time embedded systems, exploiting properties like sparsity to accelerate matrix computations. Additionally, floating-point acceleration units help offload tasks from the microcontroller.

Considering that there are already examples of working industrial products including EKF implementations, an additional 12% of calculations will not pose a difficult challenge in actual microcontroller-based implementations of possible final products powered by MSKF. Available low-cost microcontrollers will be able to handle the optimized code in any case, since it is possible to offload parallelizable code like the slave KF structure or control algorithm calculations to secondary processor cores.

Table 10.4 reports the execution time. It is worth to note that the slave KF adds only 1.9 µs to the execution time with respect to the EKF, which in percentage corresponds to just a 12% increment. This amount of time seems reasonable considering the low complexity of the slave KF. It consists of a linear KF, and the time update equation (10.34) is absent. The remaining time is due to the computation of the Kalman gain and the empirical innovation covariance matrix (10.23).

It is worth noting that the use of magnetic maps in Fig. 10.4 through Look-Up Tables (LUTs) requires an additional 0.9 µs due to memory access and subsequent value interpolation.

CHAPTER 11

Second part conclusions

The second part of this PhD thesis was dedicated to addressing and proposing solutions for bridging the gap between Kalman filtering techniques for sensorless control of synchronous reluctance machines (SynRMs) and their real-world applications. A first set of experiments demonstrated the superiority of the Extended Kalman Filter (EKF) in synchronous motor drive applications. A second set of experiments thus focused on the EKF and its adaptive extensions.

Although the Unscented Kalman Filter (UKF) demonstrated slightly better results in a Monte Carlo analysis of the nonlinear prediction function, the EKF outperformed it in terms of estimation error across almost all tests conducted. Additionally, the EKF exhibited significantly lower execution times compared to the UKF. In conclusion, for the specific case of controlling electrical drives, even in the presence of pronounced magnetic nonlinearities, the systematic approach of this research underscores that the EKF is the optimal choice for position estimation.

Building on this foundation, the Master-Slave Kalman Filter (MSKF) was introduced to address the cumbersome trial-and-error tuning process of the EKF's process noise covariance matrix. The proposed MSKF algorithm eliminates the need for manual tuning of the covariance matrix $\boldsymbol{Q}^m$ by delegating the task to a slave Kalman Filter (KF), which is efficiently tuned using the ratio q^s/r^s.

Comparisons between the traditional EKF and MSKF show that simplifying the tuning process does not compromise estimation quality. The slight increase in processing time is considered reasonable in exchange for eliminating the labor-intensive manual tuning of the master EKF.

Finally, another adaptive Kalman filter technique, based on the Maximum Likelihood Estimation (MLE) principle, was proposed, namely the Maximum Likelihood Estimation Kalman Filter (MLEKF). Compared to the conventional EKF, MLEKF can simultaneously estimate the system state and the unknown process noise covariance matrix by maximizing the joint probability density function.

Experimental results showed that the MLEKF can achieve performance on par with the conventional EKF, which, as mentioned, relies on time-consuming trial-and-error tuning.

A comparison of the MSKF and MLEKF reveals similar performance in terms of estimation error and computational time. Thus, they are both viable alternatives to the traditional EKF.

However, when applied in real-world scenarios, where the KF estimates are actively used to close control loops, rather than being merely evaluated in simulations or alongside a sensor-based feedback implementation, some delicate issues arise. The selection of the window size and the bounding policy within the adaptation mechanism, which applies to both the MSKF and MLEKF, is still determined through trial-and-error tuning. While this approach proves effective in maintaining filter responsiveness during prolonged stationary conditions and preventing divergence of the estimated noise covariances, it undermines the intrinsic adaptive nature of the proposed techniques.

The experience gained during these years of PhD research suggests that the MSKF is more robust and less sensitive to the choice of window size and bounding policy. In contrast, the MLEKF does not require the tuning of an additional KF (i.e., the slave KF in the MSKF), although this task is relatively straightforward. On the software side, specifically the time required for implementing the algorithms, the MSKF proved to be more plug-and-play compared to the MLEKF.

Ultimately, although both the MSKF and MLEKF algorithms are valid alternatives to the traditional EKF, the former is likely the best choice for sensorless SynRM drive applications. As is often the case, however, several factors can influence the decision. For instance, the MLEKF might be easier to implement in real-time applications using languages like C, where matrix operations are less straightforward than in MATLAB (used in the experiments), as it requires only one additional formula alongside the standard EKF code. Available low-cost microcontrollers will be able to handle the optimized code for both algorithms in any case, as parallelizable code, such as the Maximum Likelihood algorithm or the slave KF structure, can be offloaded to secondary processor cores.

In addition, another technique was considered, belonging to the class of covariance scaling methods: applying a forgetting (fading) factor to the state covariance matrix. While these techniques show potential, some attempts in this regard have revealed certain implementation difficulties. Nevertheless, they represent a possible avenue for further research.

Altogether, the results achieved over these years of valuable experience in various laboratories at a pump manufacturer and at universities in Italy and Germany represent a significant step forward in the development of more user-friendly Kalman filter-based sensorless methodologies, paving the way for future advancements in autonomous electrical drive systems. In the context of HVAC and pump systems, the proposed techniques have the potential to enhance the performance of sensorless control systems, improving reliability while reducing overall system costs.

APPENDIX A

Synchronous model matrices

All the matrices and the vector considered in (7.4) are detailed in the following pages, specifically in Sect. A.1. The derivatives of the matrices in (7.4) with respect to the state variables, which constitute the state Jacobian (8.12), are presented in Sect. A.2.

A.1 State update matrices

Matrix $\boldsymbol{A}$

$$
\begin{aligned}
\boldsymbol{A} &= \boldsymbol{T}_{dq}^{\alpha\beta}\, \boldsymbol{L}_{dq}^{\mathrm{diff}}\, \boldsymbol{T}_{\alpha\beta}^{dq} \\
&= \begin{bmatrix} \cos(\vartheta_E) & -\sin(\vartheta_E) \\ \sin(\vartheta_E) & \cos(\vartheta_E) \end{bmatrix} \begin{bmatrix} L_d^{\mathrm{diff}} & L_{dq}^{\mathrm{diff}} \\ L_{dq}^{\mathrm{diff}} & L_q^{\mathrm{diff}} \end{bmatrix} \begin{bmatrix} \cos(\vartheta_E) & \sin(\vartheta_E) \\ -\sin(\vartheta_E) & \cos(\vartheta_E) \end{bmatrix} \\
&= \begin{bmatrix} L_d^{\mathrm{diff}}\cos(\vartheta_E) - L_{dq}^{\mathrm{diff}}\sin(\vartheta_E) & L_{dq}^{\mathrm{diff}}\cos(\vartheta_E) - L_q^{\mathrm{diff}}\sin(\vartheta_E) \\ L_d^{\mathrm{diff}}\sin(\vartheta_E) + L_{dq}^{\mathrm{diff}}\cos(\vartheta_E) & L_{dq}^{\mathrm{diff}}\sin(\vartheta_E) + L_q^{\mathrm{diff}}\cos(\vartheta_E) \end{bmatrix} \begin{bmatrix} \cos(\vartheta_E) & \sin(\vartheta_E) \\ -\sin(\vartheta_E) & \cos(\vartheta_E) \end{bmatrix} \\
&= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}
\end{aligned}
$$

where each element of the matrix $\boldsymbol{A}$ is given by:

$$
\begin{aligned}
a_{11} &= L_d^{\mathrm{diff}}\cos^2(\vartheta_E) - L_{dq}^{\mathrm{diff}}\cos(\vartheta_E)\sin(\vartheta_E) - L_{dq}^{\mathrm{diff}}\cos(\vartheta_E)\sin(\vartheta_E) + \cdots \\
&\quad \cdots + L_q^{\mathrm{diff}}\sin^2(\vartheta_E) \\
&= \frac{L_d^{\mathrm{diff}} + L_q^{\mathrm{diff}}}{2} + \frac{L_d^{\mathrm{diff}} - L_q^{\mathrm{diff}}}{2}\cos(2\vartheta_E) + L_{dq}^{\mathrm{diff}}\sin(2\vartheta_E) \\
&= L_\Sigma^{\mathrm{diff}} + L_\Delta^{\mathrm{diff}}\cos(2\vartheta_E) - L_{dq}^{\mathrm{diff}}\sin(2\vartheta_E)
\end{aligned}
$$

$$
\begin{aligned}
a_{12} &= L_d^{\text{diff}} \cos(\vartheta_E)\sin(\vartheta_E) \,-\, L_{dq}^{\text{diff}} \sin^2(\vartheta_E) \,+\, L_{dq}^{\text{diff}} \cos^2(\vartheta_E) \,+\, \cdots \\
&\quad \cdots \,-\, L_q^{\text{diff}} \sin^2(\vartheta_E) \\
&= \frac{L_d^{\text{diff}} \,-\, L_q^{\text{diff}}}{2} \sin(2\vartheta_E) \,+\, L_{dq}^{\text{diff}} \cos(2\vartheta_E) \\
&= L_\Delta^{\text{diff}} \sin(2\vartheta_E) \,+\, L_{dq}^{\text{diff}} \cos(2\vartheta_E)
\end{aligned}
$$

$$
\begin{aligned}
a_{21} &= L_d^{\text{diff}} \cos(\vartheta_E)\sin(\vartheta_E) \,-\, L_{dq}^{\text{diff}} \sin^2(\vartheta_E) \,+\, L_{dq}^{\text{diff}} \cos^2(\vartheta_E) \,+\, \cdots \\
&\quad \cdots \,-\, L_q^{\text{diff}} \cos(\vartheta_E)\sin(\vartheta_E) \\
&= \frac{L_d^{\text{diff}} \,-\, L_q^{\text{diff}}}{2} \sin(2\vartheta_E) \,+\, L_{dq}^{\text{diff}} \cos(2\vartheta_E) \\
&= L_\Delta^{\text{diff}} \sin(2\vartheta_E) \,+\, L_{dq}^{\text{diff}} \cos(2\vartheta_E)
\end{aligned}
$$

$$
\begin{aligned}
a_{22} &= L_d^{\text{diff}} \sin^2(\vartheta_E) \,+\, L_{dq}^{\text{diff}} \cos(\vartheta_E)\sin(\vartheta_E) \,+\, L_{dq}^{\text{diff}} \cos(\vartheta_E)\sin(\vartheta_E) \,+\, \cdots \\
&\quad \cdots \,+\, L_q^{\text{diff}} \cos^2(\vartheta_E) \\
&= \frac{L_d^{\text{diff}} \,+\, L_q^{\text{diff}}}{2} \,-\, \frac{L_d^{\text{diff}} - L_q^{\text{diff}}}{2} \cos(2\vartheta_E) \,+\, L_{dq}^{\text{diff}} \sin(2\vartheta_E) \\
&= L_\Sigma^{\text{diff}} \,-\, L_\Delta^{\text{diff}} \cos(2\vartheta_E) \,+\, L_{dq}^{\text{diff}} \sin(2\vartheta_E)
\end{aligned}
$$

Matrix $\boldsymbol{B}$

$$
\omega_E \; \boldsymbol{B} \;=\; \boldsymbol{T}_{dq}^{\alpha\beta} \; \boldsymbol{L}_{dq}^{\text{diff}} \; \frac{\mathrm{d}\boldsymbol{T}_{\alpha\beta}^{dq}}{\mathrm{d}t}
$$

$$
\cancel{\omega_E} \; \boldsymbol{B} \;=\; \begin{bmatrix} \cos(\vartheta_E) & -\sin(\vartheta_E) \\ \sin(\vartheta_E) & \cos(\vartheta_E) \end{bmatrix} \begin{bmatrix} L_d^{\text{diff}} & L_{dq}^{\text{diff}} \\ L_{dq}^{\text{diff}} & L_q^{\text{diff}} \end{bmatrix} \begin{bmatrix} \cos(\vartheta_E) & \sin(\vartheta_E) \\ -\sin(\vartheta_E) & \cos(\vartheta_E) \end{bmatrix} \cancel{\omega_E}
$$

$$
\begin{aligned}
\boldsymbol{B} &= \begin{bmatrix} L_d^{\text{diff}} \cos(\vartheta_E) - L_{dq}^{\text{diff}} \sin(\vartheta_E) & L_{dq}^{\text{diff}} \cos(\vartheta_E) - L_q^{\text{diff}} \sin(\vartheta_E) \\ L_d^{\text{diff}} \sin(\vartheta_E) + L_{dq}^{\text{diff}} \cos(\vartheta_E) & L_{dq}^{\text{diff}} \sin(\vartheta_E) + L_q^{\text{diff}} \cos(\vartheta_E) \end{bmatrix} \begin{bmatrix} -\sin(\vartheta_E) & \cos(\vartheta_E) \\ -\cos(\vartheta_E) & -\sin(\vartheta_E) \end{bmatrix} \\
&= \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}
\end{aligned}
$$

where each element of the matrix $\boldsymbol{B}$ is given by:

$$
\begin{aligned}
b_{11} &= -L_d^{\text{diff}} \cos(\vartheta_E)\sin(\vartheta_E) \,+\, L_{dq}^{\text{diff}} \sin^2(\vartheta_E) \,-\, L_{dq}^{\text{diff}} \cos^2(\vartheta_E) \,+\, \cdots \\
&\quad \cdots \,+\, L_q^{\text{diff}} \cos(\vartheta_E)\sin(\vartheta_E) \\
&= -\frac{L_d^{\text{diff}} - L_q^{\text{diff}}}{2} \sin(2\vartheta_E) \,-\, L_{dq}^{\text{diff}} \cos(2\vartheta_E) \\
&= -L_\Delta^{\text{diff}} \sin(2\vartheta_E) \,-\, L_{dq}^{\text{diff}} \cos(2\vartheta_E)
\end{aligned}
$$

$$
\begin{aligned}
b_{12} &= L_d^{\text{diff}} \cos^2(\vartheta_E) \,-\, L_{dq}^{\text{diff}} \cos(\vartheta_E)\sin(\vartheta_E) \,-\, L_{dq}^{\text{diff}} \cos(\vartheta_E)\sin(\vartheta_E) \,+\, \cdots \\
&\quad \cdots \,+\, L_q^{\text{diff}} \sin^2(\vartheta_E) \\
&= \frac{L_d^{\text{diff}} \,+\, L_q^{\text{diff}}}{2} \,+\, \frac{L_d^{\text{diff}} \,-\, L_q^{\text{diff}}}{2} \cos(2\vartheta_E) \,-\, L_{dq}^{\text{diff}} \sin(2\vartheta_E)
\end{aligned}
$$

$$= L_\Sigma^{\text{diff}} + L_\Delta^{\text{diff}} \cos(2\vartheta_E) - L_{dq}^{\text{diff}} \sin(2\vartheta_E)$$

$$\begin{aligned}
b_{21} &= -L_d^{\text{diff}} \cos^2(\vartheta_E) - L_{dq}^{\text{diff}} \cos(\vartheta_E)\sin(\vartheta_E) - L_{dq}^{\text{diff}} \cos(\vartheta_E)\sin(\vartheta_E) + \cdots \\
&\quad \cdots - L_q^{\text{diff}} \sin^2(\vartheta_E) \\
&= -\frac{L_d^{\text{diff}} + L_q^{\text{diff}}}{2} + \frac{L_d^{\text{diff}} - L_q^{\text{diff}}}{2}\cos(2\vartheta_E) - L_{dq}^{\text{diff}} \sin(2\vartheta_E) \\
&= -L_\Sigma^{\text{diff}} + L_\Delta^{\text{diff}} \cos(2\vartheta_E) - L_{dq}^{\text{diff}} \sin(2\vartheta_E)
\end{aligned}$$

$$\begin{aligned}
b_{22} &= L_d^{\text{diff}} \cos(\vartheta_E)\sin(\vartheta_E) + L_{dq}^{\text{diff}} \sin^2(\vartheta_E) - L_{dq}^{\text{diff}} \cos^2(\vartheta_E) + \cdots \\
&\quad \cdots - L_q^{\text{diff}} \cos(\vartheta_E)\sin(\vartheta_E) \\
&= \frac{L_d^{\text{diff}} - L_q^{\text{diff}}}{2}\sin(2\vartheta_E) + L_{dq}^{\text{diff}} \cos(2\vartheta_E) \\
&= L_\Delta^{\text{diff}} \sin(2\vartheta_E) + L_{dq}^{\text{diff}} \cos(2\vartheta_E)
\end{aligned}$$

Matrix $\boldsymbol{C}$

$$\begin{aligned}
\boldsymbol{C} &= \boldsymbol{T}_{dq}^{\alpha\beta}\, \boldsymbol{J}\, \boldsymbol{L}_{dq}^{\text{app}}\, \boldsymbol{T}_{\alpha\beta} \\
&= \begin{bmatrix} \cos(\vartheta_E) & -\sin(\vartheta_E) \\ \sin(\vartheta_E) & \cos(\vartheta_E) \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} L_d^{\text{app}} & L_{dq}^{\text{app}} \\ L_{dq}^{\text{app}} & L_q^{\text{app}} \end{bmatrix} \begin{bmatrix} \cos(\vartheta_E) & \sin(\vartheta_E) \\ -\sin(\vartheta_E) & \cos(\vartheta_E) \end{bmatrix} \\
&= \begin{bmatrix} -\sin(\vartheta_E) & -\cos(\vartheta_E) \\ \cos(\vartheta_E) & -\sin(\vartheta_E) \end{bmatrix} \begin{bmatrix} L_d^{\text{app}} & L_{dq}^{\text{app}} \\ L_{dq}^{\text{app}} & L_q^{\text{app}} \end{bmatrix} \begin{bmatrix} \cos(\vartheta_E) & \sin(\vartheta_E) \\ -\sin(\vartheta_E) & \cos(\vartheta_E) \end{bmatrix} \\
&= \begin{bmatrix} L_d^{\text{app}}\cos(\vartheta_E) - L_{dq}^{\text{app}}\sin(\vartheta_E) & L_{dq}^{\text{app}}\cos(\vartheta_E) - L_q^{\text{app}}\sin(\vartheta_E) \\ L_d^{\text{app}}\sin(\vartheta_E) + L_{dq}^{\text{app}}\cos(\vartheta_E) & L_{dq}^{\text{app}}\sin(\vartheta_E) + L_q^{\text{app}}\cos(\vartheta_E) \end{bmatrix} \begin{bmatrix} -\sin(\vartheta_E) & \cos(\vartheta_E) \\ -\cos(\vartheta_E) & -\sin(\vartheta_E) \end{bmatrix} \\
&= \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}
\end{aligned}$$

where each element of the matrix $\boldsymbol{C}$ is given by:

$$\begin{aligned}
c_{11} &= -L_d^{\text{app}} \cos(\vartheta_E)\sin(\vartheta_E) - L_{dq}^{\text{app}} \cos^2(\vartheta_E) + L_{dq}^{\text{app}} \sin^2(\vartheta_E) + \cdots \\
&\quad \cdots + L_q^{\text{app}} \cos(\vartheta_E)\sin(\vartheta_E) \\
&= -\frac{L_d^{\text{app}} - L_q^{\text{app}}}{2}\sin(2\vartheta_E) - L_{dq}^{\text{app}} \cos(2\vartheta_E) \\
&= -L_\Delta^{\text{app}} \sin(2\vartheta_E) - L_{dq}^{\text{app}} \cos(2\vartheta_E)
\end{aligned}$$

$$\begin{aligned}
c_{12} &= -L_d^{\text{app}} \sin^2(\vartheta_E) - L_{dq}^{\text{app}} \cos(\vartheta_E)\sin(\vartheta_E) - L_{dq}^{\text{app}} \cos(\vartheta_E)\sin(\vartheta_E) + \cdots \\
&\quad \cdots - L_q^{\text{app}} \cos^2(\vartheta_E) \\
&= -\frac{L_d^{\text{app}} + L_q^{\text{app}}}{2} + \frac{L_d^{\text{app}} - L_q^{\text{app}}}{2}\cos(2\vartheta_E) - L_{dq}^{\text{app}} \sin(2\vartheta_E) \\
&= -L_\Sigma^{\text{app}} + L_\Delta^{\text{app}} \cos(2\vartheta_E) - L_{dq}^{\text{app}} \sin(2\vartheta_E)
\end{aligned}$$

$$c_{21} = L_d^{\text{app}} \cos^2(\vartheta_E) - L_{dq}^{\text{app}} \cos(\vartheta_E)\sin(\vartheta_E) - L_{dq}^{\text{app}} \cos(\vartheta_E)\sin(\vartheta_E) + \cdots$$

$$
\begin{aligned}
& \cdots + L_q^{\mathrm{app}} \sin^2(\vartheta_E) \\
&= \frac{L_d^{\mathrm{app}} + L_q^{\mathrm{app}}}{2} + \frac{L_d^{\mathrm{app}} - L_q^{\mathrm{app}}}{2} \cos(2\vartheta_E) - L_{dq}^{\mathrm{app}} \sin(2\vartheta_E) \\
&= L_\Sigma^{\mathrm{app}} + L_\Delta^{\mathrm{app}} \cos(2\vartheta_E) - L_{dq}^{\mathrm{app}} \sin(2\vartheta_E)
\end{aligned}
$$

$$
\begin{aligned}
c_{22} &= L_d^{\mathrm{app}} \cos(\vartheta_E) \sin(\vartheta_E) - L_{dq}^{\mathrm{app}} \sin^2(\vartheta_E) + L_{dq}^{\mathrm{app}} \cos^2(\vartheta_E) + \cdots \\
& \cdots - L_q^{\mathrm{app}} \cos(\vartheta_E) \sin(\vartheta_E) \\
&= \frac{L_d^{\mathrm{app}} - L_q^{\mathrm{app}}}{2} \sin(2\vartheta_E) + L_{dq}^{\mathrm{app}} \cos(2\vartheta_E) \\
&= L_\Delta^{\mathrm{app}} \sin(2\vartheta_E) + L_{dq}^{\mathrm{app}} \cos(2\vartheta_E)
\end{aligned}
$$

Vector $\boldsymbol{d}$

$$
\begin{aligned}
\boldsymbol{d} &= \boldsymbol{T}_{dq}^{\alpha\beta} \, \boldsymbol{J} \, \boldsymbol{\lambda}_d^{\mathrm{mg}} \\
&= \begin{bmatrix} \cos(\vartheta_E) & -\sin(\vartheta_E) \\ \sin(\vartheta_E) & \cos(\vartheta_E) \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \Lambda_d^{\mathrm{mg}} \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} -\sin(\vartheta_E) & -\cos(\vartheta_E) \\ \cos(\vartheta_E) & -\sin(\vartheta_E) \end{bmatrix} \begin{bmatrix} \Lambda_d^{\mathrm{mg}} \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} \\
&= \begin{bmatrix} -\Lambda_d^{\mathrm{mg}} \sin(\vartheta_E) \\ = \Lambda_d^{\mathrm{mg}} \cos(\vartheta_E) \end{bmatrix}
\end{aligned}
$$

Determinant of matrix $\boldsymbol{A}$

$$
\begin{aligned}
\det \boldsymbol{A} &= a_{11} \, a_{22} - a_{12} \, a_{21} \\
&= L_\Sigma^{\mathrm{diff},2} - L_\Delta^{\mathrm{diff},2} \cos^2(2\vartheta_E) - L_{dq}^{\mathrm{diff},2} \sin^2(2\vartheta_E) - \cancel{L_\Sigma^{\mathrm{diff}} L_\Delta^{\mathrm{diff}} \cos(2\vartheta_E)} + \cdots \\
& \cdots + \cancel{L_\Sigma^{\mathrm{diff}} L_{dq}^{\mathrm{diff}} \sin(2\vartheta_E)} + \cancel{L_\Sigma^{\mathrm{diff}} L_\Delta^{\mathrm{diff}} \cos(2\vartheta_E)} + \cdots \\
& \cdots + \cancel{L_\Delta^{\mathrm{diff}} L_{dq}^{\mathrm{diff}} \cos(2\vartheta_E) \sin(2\vartheta_E)} - \cancel{L_\Sigma^{\mathrm{diff}} L_{dq}^{\mathrm{diff}} \sin(2\vartheta_E)} + \cdots \\
& \cdots + \cancel{L_\Delta^{\mathrm{diff}} L_{dq}^{\mathrm{diff}} \cos(2\vartheta_E) \sin(2\vartheta_E)} - L_\Delta^{\mathrm{diff},2} \sin^2(2\vartheta_E) + \cdots \\
& \cdots - L_{dq}^{\mathrm{diff},2} \cos^2(2\vartheta_E) - \cancel{2 \, L_\Delta^{\mathrm{diff}} L_{dq}^{\mathrm{diff}} \cos(2\vartheta_E) \sin(2\vartheta_E)} \\
&= L_\Sigma^{\mathrm{diff},2} - L_\Delta^{\mathrm{diff},2} \big(\cos^2(2\vartheta_E) + \sin^2(2\vartheta_E) \big) \cos^2(2\vartheta_E) + \cdots \\
& \cdots + L_{dq}^{\mathrm{diff},2} \big(\cos^2(2\vartheta_E) + \sin^2(2\vartheta_E) \big) \\
&= L_\Sigma^{\mathrm{diff},2} - L_\Delta^{\mathrm{diff},2} - L_{dq}^{\mathrm{diff},2} \\
&= L_d^{\mathrm{diff}} L_q^{\mathrm{diff}} - L_{dq}^{\mathrm{diff},2}
\end{aligned}
$$

A.2 State Jacobian

The derivatives of matrices $\boldsymbol{A}^{-1}$, $\boldsymbol{B}$, $\boldsymbol{C}$, and vector $\boldsymbol{d}$ with respect to electrical position $\vartheta_{E,k}$ are derived as

Matrix $\frac{\partial \boldsymbol{A}_k^{-1}}{\partial \vartheta_{E,k}}$

$$\begin{aligned}
\left(\frac{\partial \boldsymbol{A}_k^{-1}}{\partial \vartheta_{E,k}}\right)_{11} &= \frac{1}{\det \boldsymbol{A}_k}\left[2L_\Delta^{\text{diff}}\sin(2\vartheta_{E,k}) + 2L_{dq}^{\text{diff}}\cos(2\vartheta_{E,k})\right]\\
\left(\frac{\partial \boldsymbol{A}_k^{-1}}{\partial \vartheta_{E,k}}\right)_{12} &= \frac{1}{\det \boldsymbol{A}_k}\left[-2L_\Delta^{\text{diff}}\cos(2\vartheta_{E,k}) + 2L_{dq}^{\text{diff}}\sin(2\vartheta_{E,k})\right]\\
\left(\frac{\partial \boldsymbol{A}_k^{-1}}{\partial \vartheta_{E,k}}\right)_{21} &= \frac{1}{\det \boldsymbol{A}_k}\left[-2L_\Delta^{\text{diff}}\cos(2\vartheta_{E,k}) + 2L_{dq}^{\text{diff}}\sin(2\vartheta_{E,k})\right]\\
\left(\frac{\partial \boldsymbol{A}_k^{-1}}{\partial \vartheta_{E,k}}\right)_{22} &= \frac{1}{\det \boldsymbol{A}_k}\left[-2L_\Delta^{\text{diff}}\sin(2\vartheta_{E,k}) + 2L_{dq}^{\text{diff}}\cos(2\vartheta_{E,k})\right]
\end{aligned}$$

Matrix $\frac{\partial \boldsymbol{B}_k}{\partial \vartheta_{E,k}}$

$$\begin{aligned}
\left(\frac{\partial \boldsymbol{B}_k}{\partial \vartheta_{E,k}}\right)_{11} &= -2L_\Delta^{\text{diff}}\cos(2\vartheta_{E,k}) + 2L_{dq}^{\text{diff}}\sin(2\vartheta_{E,k})\\
\left(\frac{\partial \boldsymbol{B}_k}{\partial \vartheta_{E,k}}\right)_{12} &= -2L_\Delta^{\text{diff}}\sin(2\vartheta_{E,k}) - 2L_{dq}^{\text{diff}}\cos(2\vartheta_{E,k})\\
\left(\frac{\partial \boldsymbol{B}_k}{\partial \vartheta_{E,k}}\right)_{21} &= -2L_\Delta^{\text{diff}}\sin(2\vartheta_{E,k}) - 2L_{dq}^{\text{diff}}\cos(2\vartheta_{E,k})\\
\left(\frac{\partial \boldsymbol{B}_k}{\partial \vartheta_{E,k}}\right)_{22} &= 2L_\Delta^{\text{diff}}\cos(2\vartheta_{E,k}) - 2L_{dq}^{\text{diff}}\sin(2\vartheta_{E,k})
\end{aligned}$$

Matrix $\frac{\partial \boldsymbol{C}_k}{\partial \vartheta_{E,k}}$

$$\begin{aligned}
\left(\frac{\partial \boldsymbol{C}_k}{\partial \vartheta_{E,k}}\right)_{11} &= -2L_\Delta^{\text{app}}\cos(2\vartheta_{E,k}) + 2L_{dq}^{\text{app}}\sin(2\vartheta_{E,k})\\
\left(\frac{\partial \boldsymbol{C}_k}{\partial \vartheta_{E,k}}\right)_{12} &= -2L_\Delta^{\text{app}}\sin(2\vartheta_{E,k}) - 2L_{dq}^{\text{app}}\cos(2\vartheta_{E,k})\\
\left(\frac{\partial \boldsymbol{C}_k}{\partial \vartheta_{E,k}}\right)_{21} &= -2L_\Delta^{\text{app}}\sin(2\vartheta_{E,k}) - 2L_{dq}^{\text{app}}\cos(2\vartheta_{E,k})\\
\left(\frac{\partial \boldsymbol{C}_k}{\partial \vartheta_{E,k}}\right)_{22} &= 2L_\Delta^{\text{app}}\cos(2\vartheta_{E,k}) - 2L_{dq}^{\text{app}}\sin(2\vartheta_{E,k})
\end{aligned}$$

Vector $\frac{\partial \boldsymbol{d}_k}{\partial \vartheta_{E,k}}$

$$\begin{aligned}
\left(\frac{\partial \boldsymbol{d}_k}{\partial \vartheta_{E,k}}\right)_1 &= -\Lambda_d^{mg}\cos(\vartheta_{E,k}) \\
\left(\frac{\partial \boldsymbol{d}_k}{\partial \vartheta_{E,k}}\right)_2 &= -\Lambda_d^{mg}\sin(\vartheta_{E,k})
\end{aligned}$$

APPENDIX B

Maximum Likelihood Estimation notes

In the following, a brief overview of the Maximum Likelihood Estimation (MLE) theory is provided, which is used in the context of the Maximum Likelihood Estimation Kalman Filter (MLEKF) presented in Chapter 9. This appendix begins with the concept of Bayesian estimation and ends with the derivation of the formulas for the online estimation of the process and noise covariance matrices of the MLEKF algorithm.

B.1 Bayesian State Estimation

Consider the following linear dynamical system:

$$\begin{cases} \boldsymbol{x}_{k+1} \,=\, \boldsymbol{\Phi}_k \boldsymbol{x}_k \,+\, \boldsymbol{\Gamma}_k \boldsymbol{u}_k + \boldsymbol{w}_k & \boldsymbol{x}_k \,\in\, \mathbb{R}^{n\times 1}\,, \quad \boldsymbol{u}_k \,\in\, \mathbb{R}^{m\times 1} \\ \quad \boldsymbol{y}_k \,=\, \boldsymbol{H}_k \boldsymbol{x}_k \,+\, \boldsymbol{v}_k & \boldsymbol{y}_k \,\in\, \mathbb{R}^{p\times 1} \end{cases} \tag{B.1}$$

where:

- the matrices $\boldsymbol{\Phi}_k$, $\boldsymbol{\Gamma}_k$, $\boldsymbol{H}_k$ and the input $\boldsymbol{u}_k$ are known a priori.
- $\boldsymbol{w}_k$ and $\boldsymbol{v}_k$ are zero-mean normal (gaussian) random processes with known covariances:
$$\boldsymbol{w}_k \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q}_k)\,, \qquad \boldsymbol{v}_k \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}_k) \tag{B.2}$$
- $\boldsymbol{w}_k$ and $\boldsymbol{v}_k$ are white noises:

$$\mathbb{E}\,\boldsymbol{w}_k \boldsymbol{w}_h^T \,=\, \boldsymbol{Q}_k\, \delta[k-h]\,, \qquad \mathbb{E}\,\boldsymbol{v}_k \boldsymbol{v}_h^T \,=\, \boldsymbol{R}_k\, \delta[k-h] \tag{B.3}$$

- $\boldsymbol{w}_k$ and $\boldsymbol{v}_k$ are mutually uncorrelated:

$$\mathbb{E}\,\boldsymbol{w}_k\boldsymbol{v}_h^T = \boldsymbol{0} \tag{B.4}$$

- the initial state $\boldsymbol{x}_0$ is a normally distributed random variable with mean $\bar{\boldsymbol{x}}_0$ and covariance $\boldsymbol{P}_{0|0}$:

$$\boldsymbol{x}_0 \sim \mathcal{N}(\bar{\boldsymbol{x}}_0, \boldsymbol{P}_{0|0}) \tag{B.5}$$

- $\boldsymbol{w}_k$ and $\boldsymbol{v}_k$ are uncorrelated with the initial state $\boldsymbol{x}_0$:

$$\mathbb{E}\,\boldsymbol{x}_0\boldsymbol{w}_k^T = \boldsymbol{0}\,, \qquad \mathbb{E}\,\boldsymbol{x}_0\boldsymbol{v}_k^T = \boldsymbol{0} \tag{B.6}$$

The state-space representation is based on the assumption that the probabilistic model for state evolution is a *first-order Markov process*, which means that value of the state vector at the instant $k+1$ depends only on the value at time instant k, but not on its values in previous time instants. Therefore, the state transition equation in (B.1) can be rewritten as a state-transition PDF of the type:

$$p(\boldsymbol{x}_{k+1}\,|\,\boldsymbol{x}_k,\,\boldsymbol{u}_k) \tag{B.7}$$

Similarly, the measurement equation in (B.1) can be represented as a PDF of the type:

$$p(\boldsymbol{y}_k\,|\,\boldsymbol{x}_k,\,\boldsymbol{u}_k) \tag{B.8}$$

A *filter* uses the input and output measurements available up to time instant k, i.e.

$$\mathcal{U}_k \triangleq \{u_i\,:\,i=0,\ldots,k\}\,, \qquad \mathcal{Y}_k \triangleq \{y_i\,:\,i=0,\ldots,k\} \tag{B.9}$$

to infer an estimate $\hat{\boldsymbol{x}}_{k|k}$ of the state at time k.

The *Bayesian framework* provides a general unifying framework for sequential state estimation. The only assumption that is made is that the evolution of the state is Markovian. The main idea is to start from an initial PDF for the state vector, $p(\boldsymbol{x}_0)$, and recursively calculate the *posterior* PDF, $p(\boldsymbol{x}_k\,|\,\mathcal{U}_k, \mathcal{Y}_k)$, based on the measurements. This can be done by a filtering algorithm that includes two-stages, a *state prediction* and a *state update*:

- *state prediction*: according to the *Chapman-Kolmogorov equation* it holds that:

$$p(\boldsymbol{x}_{k+1}\,|\,\mathcal{U}_k,\,\mathcal{Y}_k) = \int p(\boldsymbol{x}_{k+1}\,|\,\boldsymbol{x}_k,\,\boldsymbol{u}_k)\,p(\boldsymbol{x}_k\,|\,\mathcal{U}_k,\,\mathcal{Y}_k)\,d\boldsymbol{x}_k \tag{B.10}$$

  where:

  - $p(\boldsymbol{x}_{k+1}\,|\,\mathcal{U}_k,\,\mathcal{Y}_k)$: *predictive* PDF of the state $\boldsymbol{x}_{k+1}$, given the measurements up to the time instant k.
  - $p(\boldsymbol{x}_{k+1}\,|\,\boldsymbol{x}_k,\,\boldsymbol{u}_k)$: state-transition PDF; also referred as *prior* PDF of state $\boldsymbol{x}_{k+1}$, given the immediate past state $\boldsymbol{x}_k$ and input $\boldsymbol{u}_k$.
  - $p(\boldsymbol{x}_k\,|\,\mathcal{U}_k,\,\mathcal{Y}_k)$: old *posterior* PDF of the state $\boldsymbol{x}_k$ – see next.

- *state update*: according to the *Bayes' rule*:

$$
\begin{aligned}
p(\boldsymbol{x}_{k+1}\,|\,\mathcal{U}_{k+1},\,\mathcal{Y}_{k+1}) &= \frac{p(\boldsymbol{x}_{k+1},\mathcal{U}_{k+1},\,\mathcal{Y}_{k+1})}{p(\mathcal{U}_{k+1},\,\mathcal{Y}_{k+1})} \\
&= \frac{p(\boldsymbol{x}_{k+1},\mathcal{U}_{k+1},\,\boldsymbol{y}_{k+1},\,\mathcal{Y}_{k})}{p(\mathcal{U}_{k+1},\,\boldsymbol{y}_{k+1},\,\mathcal{Y}_{k})} \\
&= \frac{p(\boldsymbol{x}_{k+1},\,\boldsymbol{y}_{k+1}\,|\,\mathcal{U}_{k+1},\,\mathcal{Y}_{k})\,\cancel{p(\mathcal{U}_{k+1},\,\mathcal{Y}_{k})}}{p(\boldsymbol{y}_{k+1}\,|\,\mathcal{U}_{k+1},\,\mathcal{Y}_{k})\,\cancel{p(\mathcal{U}_{k+1},\,\mathcal{Y}_{k})}} \\
&= p(\boldsymbol{y}_{k+1}\,|\,\boldsymbol{x}_{k+1},\,\mathcal{U}_{k+1},\,\mathcal{Y}_{k})\,\frac{p(\boldsymbol{x}_{k+1}\,|\,\mathcal{U}_{k+1},\,\mathcal{Y}_{k})}{p(\boldsymbol{y}_{k+1}\,|\,\mathcal{U}_{k+1},\,\mathcal{Y}_{k})}
\end{aligned} \tag{B.11}
$$

  where:

  - $p(\boldsymbol{x}_{k+1}\,|\,\mathcal{U}_{k+1},\,\mathcal{Y}_{k+1})$: *posterior* PDF of the state $\boldsymbol{x}_{k+1}$, given the measurements up to the time instant $k+1$.
  - $p(\boldsymbol{y}_{k+1}\,|\,\boldsymbol{x}_{k+1},\,\mathcal{U}_{k+1},\,\mathcal{Y}_{k})$: output measurement PDF; also referred as *likelihood* function.
  - $p(\boldsymbol{y}_{k+1}\,|\,\mathcal{U}_{k+1},\,\mathcal{Y}_{k})$: *normalizing constant*:

$$
p(\boldsymbol{y}_{k+1}\,|\,\mathcal{U}_{k+1},\,\mathcal{Y}_{k}) = \int p(\boldsymbol{y}_{k+1}\,|\,\boldsymbol{x}_{k+1},\,\boldsymbol{u}_{k+1})p(\boldsymbol{x}_{k+1}\,|\,\mathcal{U}_{k+1},\,\mathcal{Y}_{k})\,d\boldsymbol{x}_{k+1} \tag{B.12}
$$

    It ensures that the total volume under the multidimensional curve of the posterior PDF $p(\boldsymbol{x}_{k+1}\,|\,\mathcal{U}_{k+1},\,\mathcal{Y}_{k+1})$ is unity.

  The previous expression can be further simplified by noting that:

$$
p(\boldsymbol{y}_{k+1}\,|\,\boldsymbol{x}_{k+1},\,\mathcal{U}_{k+1},\,\mathcal{Y}_{k}) = p(\boldsymbol{y}_{k+1}\,|\,\boldsymbol{x}_{k+1},\,\boldsymbol{u}_{k+1}) \tag{B.13}
$$

$$
p(\boldsymbol{x}_{k+1}\,|\,\mathcal{U}_{k+1},\,\mathcal{Y}_{k}) = p(\boldsymbol{x}_{k+1}\,|\,\mathcal{U}_{k},\,\mathcal{Y}_{k}) \tag{B.14}
$$

$$
p(\boldsymbol{y}_{k+1}\,|\,\mathcal{U}_{k+1},\,\mathcal{Y}_{k}) = p(\boldsymbol{y}_{k+1}\,|\,\mathcal{U}_{k},\,\mathcal{Y}_{k}) \tag{B.15}
$$

  The first identity holds because the output $\boldsymbol{y}_{k+1}$ only depends on the state and the input at the same sampling instant. The second instead holds because the state $\boldsymbol{x}_{k+1}$ only depends on the inputs and outputs up to the sampling instant k. As for the third, it is a consequence of the fact that in (B.1) the output $\boldsymbol{y}_k$ does not directly depends on the input at the same sampling instant (i.e. $\boldsymbol{u}_k$). Equation (B.11) therefore reduces to:

$$
p(\boldsymbol{x}_{k+1}\,|\,\mathcal{U}_{k+1},\,\mathcal{Y}_{k+1}) = p(\boldsymbol{x}_{k+1}\,|\,\mathcal{U}_{k},\,\mathcal{Y}_{k})\,\frac{p(\boldsymbol{y}_{k+1}\,|\,\boldsymbol{x}_{k+1},\,\boldsymbol{u}_{k+1})}{p(\boldsymbol{y}_{k+1}\,|\,\mathcal{U}_{k},\,\mathcal{Y}_{k})} \tag{B.16}
$$

The *posterior* PDF $p(\boldsymbol{x}_k\,|\,\mathcal{U}_k,\,\mathcal{Y}_k)$ embodies the entire knowledge about the state $\boldsymbol{x}_k$ at time k after having received the entire measurement set $\{\mathcal{U}_k,\,\mathcal{Y}_k\}$. Therefore, it contains all the information required for state estimation. Depending on the chosen criterion, different (point) state estimators are obtained:

- *Minimum Mean-Square Error* (MMSE) state estimator:

$$\hat{\boldsymbol{x}}_{k|k}^{MMSE} = \underset{\hat{\boldsymbol{x}}_{k|k}}{\operatorname{argmin}}\ \mathbb{E}\left[(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k})^T(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k}) \,|\, \mathcal{U}_k,\, \mathcal{Y}_k\right] \tag{B.17}$$

  It is equivalent to minimize the trace (sum of diagonal elements) of the estimation error covariance matrix. It can be proven that the MMSE estimator is the *conditional mean* of $\boldsymbol{x}_k$:

$$\hat{\boldsymbol{x}}_{k|k}^{MMSE} = \mathbb{E}\left[\boldsymbol{x}_k \,|\, \mathcal{U}_k,\, \mathcal{Y}_k\right] = \int \boldsymbol{x}_k\, p(\boldsymbol{x}_k \,|\, \mathcal{U}_k,\, \mathcal{Y}_k)\, d\boldsymbol{x}_k \tag{B.18}$$

- *Maximum A Posteriori* (MAP) state estimator: is the *mode* of the posterior PDF:

$$\hat{\boldsymbol{x}}_{k|k}^{MAP} = \underset{\hat{\boldsymbol{x}}_{k|k}}{\operatorname{argmax}}\ p(\hat{\boldsymbol{x}}_{k|k} \,|\, \mathcal{U}_k,\, \mathcal{Y}_k) \tag{B.19}$$

  The MAP estimator is used to take into account prior knowledge about the quantity being estimated.

- *Maximum Likelihood* (ML) estimator: is the *mode* of the likelihood function:

$$\hat{\boldsymbol{x}}_{k|k}^{ML} = \underset{\hat{\boldsymbol{x}}_{k|k}}{\operatorname{argmax}}\ p(\boldsymbol{y}_k \,|\, \hat{\boldsymbol{x}}_{k|k},\, \boldsymbol{u}_k) \tag{B.20}$$

  The ML estimator does not account for prior knowledge about the quantity being estimated.

The MAP estimate is closely related to the method of maximum likelihood (ML) estimation, but employs an augmented optimization objective which incorporates a prior distribution (that quantifies the additional information available through prior knowledge of a related event) over the quantity one wants to estimate. MAP estimation can therefore be seen as a regularization of maximum likelihood estimation.

B.2 Maximum A Posteriori (MAP) State Estimation

The MAP state estimate is obtained as the mode of the posterior PDF:

$$\ell_k^A(\boldsymbol{x}_k, \mathcal{U}_k, \mathcal{Y}_k) \triangleq p(\boldsymbol{x}_k \,|\, \mathcal{U}_k, \mathcal{Y}_k) \tag{B.21}$$

or, equivalently, the natural logarithm of that function:

$$L_k^A(\boldsymbol{x}_k, \mathcal{U}_k, \mathcal{Y}_k) \triangleq \ln \ell_k^A(\boldsymbol{x}_k, \mathcal{U}_k, \mathcal{Y}_k) = \ln p(\boldsymbol{x}_k \,|\, \mathcal{U}_k, \mathcal{Y}_k) \tag{B.22}$$

From (B.11) it follows that:

$$\begin{aligned} L_k^A(\boldsymbol{x}_k, \mathcal{U}_k, \mathcal{Y}_k) = {} & \ln p(\boldsymbol{x}_k \,|\, \mathcal{U}_{k-1},\, \mathcal{Y}_{k-1}) + \ln p(\mathcal{Y}_k \,|\, \boldsymbol{x}_k,\, \boldsymbol{u}_k) - \cdots \\ & \cdots - \ln p(\mathcal{Y}_k \,|\, \mathcal{U}_k,\, \mathcal{Y}_{k-1}) \end{aligned} \tag{B.23}$$

Let $\hat{\boldsymbol{x}}_{k|k-1}$ be the *state prediction*, and

$$\tilde{\boldsymbol{x}}_{k|k-1} \triangleq \boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1} \tag{B.24}$$

the *state prediction error*. The *covariance matrix of the prediction error* is defined as:

$$\boldsymbol{P}_{k|k-1} = \mathbb{E}\left[\tilde{\boldsymbol{x}}_{k|k-1}\,\tilde{\boldsymbol{x}}_{k|k-1}^T \,\middle|\, \mathcal{U}_{k-1},\, \mathcal{Y}_{k-1}\right] \tag{B.25}$$

It holds that:

$$\boldsymbol{x}_k \sim \mathcal{N}(\hat{\boldsymbol{x}}_{k|k-1},\, \boldsymbol{P}_{k|k-1})\,, \qquad \boldsymbol{y}_k \sim \mathcal{N}(\boldsymbol{H}_k\boldsymbol{x}_k,\, \boldsymbol{R}_k) \tag{B.26}$$

namely:

$$p(\boldsymbol{x}_k \,|\, \mathcal{U}_{k-1},\, \mathcal{Y}_{k-1}) = \frac{1}{\sqrt{(2\pi)^n \det \boldsymbol{P}_{k|k-1}}} \exp\left[-\tfrac{1}{2}(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1})^T \boldsymbol{P}_{k|k-1}^{-1}(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1})\right] \tag{B.27}$$

$$p(\mathcal{Y}_k \,|\, \boldsymbol{x}_k,\, \boldsymbol{u}_k) = \frac{1}{\sqrt{(2\pi)^p \det \boldsymbol{R}_k}} \exp\left[-\tfrac{1}{2}(\boldsymbol{y}_k - \boldsymbol{H}_k\boldsymbol{x}_k)^T \boldsymbol{R}_k^{-1}\,(\boldsymbol{y}_k - \boldsymbol{H}_k\boldsymbol{x}_k)\right] \tag{B.28}$$

The log-value of the posterior PDF becomes:

$$\begin{aligned} L_k^A(\boldsymbol{x}_k,\mathcal{U}_k,\mathcal{Y}_k) &= -\tfrac{1}{2}(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1})^T \boldsymbol{P}_{k|k-1}^{-1}(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1}) + \cdots \\ &\cdots - \tfrac{1}{2}(\boldsymbol{y}_k - \boldsymbol{H}_k\boldsymbol{x}_k)^T \boldsymbol{R}_k^{-1}\,(\boldsymbol{y}_k - \boldsymbol{H}_k\boldsymbol{x}_k) + \text{const} \end{aligned} \tag{B.29}$$

The MAP state estimate $\hat{\boldsymbol{x}}_{k|k}$ is obtained as the stationary point of the posterior PDF, i.e. the value that makes the gradient equal to zero:

$$\frac{\partial L_k^A(\hat{\boldsymbol{x}}_{k|k},\mathcal{U}_k,\mathcal{Y}_k)}{\partial \boldsymbol{x}_k} = -\boldsymbol{P}_{k|k-1}^{-1}(\hat{\boldsymbol{x}}_{k|k} - \hat{\boldsymbol{x}}_{k|k-1}) + \boldsymbol{H}_k^T \boldsymbol{R}_k^{-1}(\boldsymbol{y}_k - \boldsymbol{H}_k\hat{\boldsymbol{x}}_{k|k}) = \boldsymbol{0} \tag{B.30}$$

The previous expression has been derived by using the following identities (see (84) and (85) in [106]):

$$\frac{\partial}{\partial \boldsymbol{s}}(\boldsymbol{x} - \boldsymbol{A}\boldsymbol{s})^T \boldsymbol{W}(\boldsymbol{x} - \boldsymbol{A}\boldsymbol{s}) = -2\boldsymbol{A}^T\boldsymbol{W}(\boldsymbol{x} - \boldsymbol{A}\boldsymbol{s}) \tag{B.31}$$

$$\frac{\partial}{\partial \boldsymbol{x}}(\boldsymbol{x} - \boldsymbol{s})^T \boldsymbol{W}(\boldsymbol{x} - \boldsymbol{s}) = 2\boldsymbol{W}(\boldsymbol{x} - \boldsymbol{s}) \tag{B.32}$$

After some manipulations it follows that:

$$\hat{\boldsymbol{x}}_{k|k} = \left(\boldsymbol{I} + \boldsymbol{P}_{k|k-1}\,\boldsymbol{H}_k^\top\,\boldsymbol{R}_k\,\boldsymbol{H}_k\right)^{-1}\left(\hat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{P}_{k|k-1}\boldsymbol{H}_k^\top\,\boldsymbol{R}_k^{-1}\boldsymbol{y}_k\right) \tag{B.33}$$

By using the *matrix inversion lemma* (see (157) in [106]):

$$(\boldsymbol{A} + \boldsymbol{U}\boldsymbol{B}\boldsymbol{V})^{-1} = \boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}\boldsymbol{U}\left(\boldsymbol{B}^{-1} + \boldsymbol{V}\boldsymbol{A}^{-1}\boldsymbol{U}\right)^{-1}\boldsymbol{V}\boldsymbol{A}^{-1} \tag{B.34}$$

it follows that:

$$\begin{aligned} \left(\boldsymbol{I} + \boldsymbol{P}_{k|k-1}\,\boldsymbol{H}_k^\top\,\boldsymbol{R}_k\,\boldsymbol{H}_k\right)^{-1} &= \boldsymbol{I} - \boldsymbol{P}_{k|k-1}\boldsymbol{H}_k^\top\left(\boldsymbol{R}_k + \boldsymbol{H}_k\boldsymbol{P}_{k|k-1}\boldsymbol{H}_k^\top\right)^{-1}\boldsymbol{H}_k \\ &= \boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H}_k \end{aligned} \tag{B.35}$$

where:

$$\boldsymbol{K}_k \triangleq \boldsymbol{P}_{k|k-1}\boldsymbol{H}_k^\top \left(\boldsymbol{R}_k + \boldsymbol{H}_k\boldsymbol{P}_{k|k-1}\boldsymbol{H}_k^\top\right)^{-1} \tag{B.36}$$

Therefore, equation (B.33) becomes:

$$\begin{aligned}
\hat{\boldsymbol{x}}_{k|k} &= (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H}_k)\left(\hat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{P}_{k|k-1}\boldsymbol{H}_k^\top\boldsymbol{R}_k^{-1}\boldsymbol{y}_k\right)\\
&= \hat{\boldsymbol{x}}_{k|k-1} - \boldsymbol{K}_k\boldsymbol{H}_k\hat{\boldsymbol{x}}_{k|k-1} + (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H}_k)\boldsymbol{P}_{k|k-1}\boldsymbol{H}_k^\top\boldsymbol{R}_k^{-1}\boldsymbol{y}_k\\
&= \hat{\boldsymbol{x}}_{k|k-1} - \boldsymbol{K}_k\boldsymbol{H}_k\hat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}_k\,\boldsymbol{y}_k\\
&= \hat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}_k\left(\boldsymbol{y}_k - \boldsymbol{H}_k\hat{\boldsymbol{x}}_{k|k-1}\right)
\end{aligned} \tag{B.37}$$

where the following result:

$$\begin{aligned}
&(\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H}_k)\boldsymbol{P}_{k|k-1}\boldsymbol{H}_k^\top\boldsymbol{R}_k^{-1}\\
&\cdots = \boldsymbol{P}_{k|k-1}\boldsymbol{H}_k^\top\boldsymbol{R}_k^{-1} - \boldsymbol{K}_k\boldsymbol{H}_k\boldsymbol{P}_{k|k-1}\boldsymbol{H}_k^\top\boldsymbol{R}_k^{-1} = \cdots\\
&\cdots = \boldsymbol{K}_k\left(\boldsymbol{R}_k + \cancel{\boldsymbol{H}_k\boldsymbol{P}_{k|k-1}\boldsymbol{H}_k^\top} - \cancel{\boldsymbol{H}_k\boldsymbol{P}_{k|k-1}\boldsymbol{H}_k^\top}\right)\boldsymbol{R}_k^{-1} = \boldsymbol{K}_k
\end{aligned} \tag{B.38}$$

has been used to obtain the final expression. Equation (B.37) is the *state update* equation of the filter, and $\boldsymbol{K}_k$ in (B.36) is the *filter gain*. The quantity:

$$\tilde{\boldsymbol{y}}_{k|k-1} \triangleq \boldsymbol{y}_k - \boldsymbol{H}_k\,\hat{\boldsymbol{x}}_{k|k-1} \tag{B.39}$$

is the *innovation process*. The *state estimation error* is equal to:

$$\begin{aligned}
\tilde{\boldsymbol{x}}_{k|k} &\triangleq \boldsymbol{x}_{k|k} - \hat{\boldsymbol{x}}_{k|k}\\
&= \boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1} - \boldsymbol{K}_k\left(\boldsymbol{y}_k - \boldsymbol{H}_k\hat{\boldsymbol{x}}_{k|k-1}\right)\\
&= \boldsymbol{x}_k - (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H}_k)\hat{\boldsymbol{x}}_{k|k-1} - \boldsymbol{K}_k\boldsymbol{y}_k\\
&= (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H}_k)(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1}) + \cancel{\boldsymbol{K}_k\boldsymbol{H}_k\boldsymbol{x}_k} - \boldsymbol{K}_k(\cancel{\boldsymbol{H}_k\boldsymbol{x}_k} + \boldsymbol{v}_k)\\
&= (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H}_k)(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1}) - \boldsymbol{K}_k\boldsymbol{v}_k\\
&= (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H}_k)\tilde{\boldsymbol{x}}_{k|k-1} - \boldsymbol{K}_k\boldsymbol{v}_k
\end{aligned} \tag{B.40}$$

where

$$\tilde{\boldsymbol{x}}_{k|k-1} \triangleq \boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1} \tag{B.41}$$

is the *state prediction error*. The *covariance matrix of the state estimation error* is:

$$\begin{aligned}
\boldsymbol{P}_{k|k} &= \mathbb{E}\left[\tilde{\boldsymbol{x}}_{k|k}\,\tilde{\boldsymbol{x}}_{k|k}^T \,\middle|\, \mathcal{U}_k, \mathcal{Y}_k\right] =\\
&= \mathbb{E}\left[(\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H}_k)\tilde{\boldsymbol{x}}_{k|k-1}\tilde{\boldsymbol{x}}_{k|k-1}^T(\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H}_k)^T\right] + \mathbb{E}\left[\boldsymbol{K}_k\boldsymbol{v}_k\boldsymbol{v}_k^T\boldsymbol{K}_k^T\right]\\
&= (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H}_k)\boldsymbol{P}_{k|k-1}(\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H}_k)^T + \boldsymbol{K}_k\boldsymbol{R}_k\boldsymbol{K}_k^T
\end{aligned} \tag{B.42}$$

which has been obtained by taking into account that $\hat{\boldsymbol{x}}_{k|k-1}$ and $\boldsymbol{v}_k$ are mutually uncorrelated, i.e. $\mathbb{E}[\hat{\boldsymbol{x}}_{k|k-1}\boldsymbol{v}_k^T] = \boldsymbol{0}$. Note that similarly to the first of (B.26), it is also possible to write that:

$$\boldsymbol{x}_k \sim \mathcal{N}(\hat{\boldsymbol{x}}_{k|k}, \boldsymbol{P}_{k|k}) \tag{B.43}$$

namely

$$p(\boldsymbol{x}_k \,|\, \mathcal{U}_k,\, \mathcal{Y}_k) \;=\; \frac{1}{\sqrt{(2\pi)^n \det \boldsymbol{P}_{k|k}}} \exp\left[-\tfrac{1}{2}(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k})^T \boldsymbol{P}_{k|k}^{-1}(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k})\right] \tag{B.44}$$

Moreover, since the innovation process is equal to:

$$\begin{aligned} \tilde{\boldsymbol{y}}_{k|k-1} &= \boldsymbol{y}_k \;-\; \hat{\boldsymbol{y}}_{k|k-1} \\ &= \boldsymbol{H}_k\boldsymbol{x}_k + \boldsymbol{v}_k - \boldsymbol{H}_k\hat{\boldsymbol{x}}_{k|k-1} \\ &= \boldsymbol{H}_k(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1}) + \boldsymbol{v}_k \\ &= \boldsymbol{H}_k\tilde{\boldsymbol{x}}_{k|k-1} + \boldsymbol{v}_k \end{aligned} \tag{B.45}$$

its covariance matrix has the following expression:

$$\begin{aligned} \boldsymbol{\Lambda}_k &\triangleq \mathbb{E}\left[\tilde{\boldsymbol{y}}_{k|k-1}\tilde{\boldsymbol{y}}_{k|k-1}^T \,\middle|\, \mathcal{U}_k,\, \mathcal{Y}_k\right] \\ &= \mathbb{E}\left[\boldsymbol{H}_k\tilde{\boldsymbol{x}}_{k|k-1}\tilde{\boldsymbol{x}}_{k|k-1}^T\boldsymbol{H}_k^T\right] \;+\; \mathbb{E}\left[\boldsymbol{v}_k\boldsymbol{v}_k^T\right] \\ &= \boldsymbol{H}_k\boldsymbol{P}_{k|k-1}\boldsymbol{H}_k^T + \boldsymbol{R}_k \end{aligned} \tag{B.46}$$

Therefore, it is possible to write that:

$$\boldsymbol{y}_k \sim \mathcal{N}\left(\boldsymbol{H}_k\hat{\boldsymbol{x}}_{k|k-1},\, \boldsymbol{\Lambda}_k\right) \tag{B.47}$$

namely:

$$p(\boldsymbol{y}_k \,|\, \mathcal{U}_{k-1},\, \mathcal{Y}_{k-1}) \;=\; \frac{1}{\sqrt{(2\pi)^p \det \boldsymbol{\Lambda}_k}} \exp\left[-\tfrac{1}{2}(\boldsymbol{y}_k - \boldsymbol{H}_k\hat{\boldsymbol{x}}_{k|k-1})^T\boldsymbol{\Lambda}_k^{-1}(\boldsymbol{y}_k - \boldsymbol{H}_k\hat{\boldsymbol{x}}_{k|k-1})\right] \tag{B.48}$$

In (B.37), the state prediction $\hat{\boldsymbol{x}}_{k|k-1}$ is obtained from the state estimate at the previous sampling instant by using the following *state prediction* equation:

$$\hat{\boldsymbol{x}}_{k|k-1} \;=\; \boldsymbol{\Phi}_{k-1}\hat{\boldsymbol{x}}_{k-1|k-1} \;+\; \boldsymbol{\Gamma}_{k-1}\boldsymbol{u}_{k-1} \tag{B.49}$$

The covariance matrix of the prediction error used in (B.36) and (B.42) is obtained from the covariance matrix of the state estimate at the previous sampling instant by using the following equation:

$$\begin{aligned} \boldsymbol{P}_{k|k-1} &= \mathbb{E}\left[\tilde{\boldsymbol{x}}_{k|k-1}\,\tilde{\boldsymbol{x}}_{k|k-1}^T \,\middle|\, \mathcal{U}_k,\, \mathcal{Y}_k\right] \\ &= \mathbb{E}\left[\left(\boldsymbol{x}_k - \boldsymbol{\Phi}_{k-1}\hat{\boldsymbol{x}}_{k-1|k-1} - \boldsymbol{\Gamma}_{k-1}\boldsymbol{u}_{k-1}\right)\left(\boldsymbol{x}_k - \boldsymbol{\Phi}_{k-1}\hat{\boldsymbol{x}}_{k-1|k-1} - \boldsymbol{\Gamma}_{k-1}\boldsymbol{u}_{k-1}\right)^T\right] \\ &= \mathbb{E}\Big[\big(\boldsymbol{\Phi}_{k-1}\boldsymbol{x}_{k-1} \;+\; \cancel{\boldsymbol{\Gamma}_{k-1}\boldsymbol{u}_{k-1}} + \boldsymbol{w}_k - \boldsymbol{\Phi}_{k-1}\hat{\boldsymbol{x}}_{k-1|k-1} - \cancel{\boldsymbol{\Gamma}_{k-1}\boldsymbol{u}_{k-1}}\big)\cdots \\ &\qquad\cdots\big(\boldsymbol{\Phi}_{k-1}\boldsymbol{x}_{k-1} \;+\; \cancel{\boldsymbol{\Gamma}_{k-1}\boldsymbol{u}_{k-1}} + \boldsymbol{w}_k - \boldsymbol{\Phi}_{k-1}\hat{\boldsymbol{x}}_{k-1|k-1} - \cancel{\boldsymbol{\Gamma}_{k-1}\boldsymbol{u}_{k-1}}\big)^T\Big] \\ &= \mathbb{E}\left[\boldsymbol{\Phi}_{k-1}\tilde{x}_{k-1|k-1}\tilde{x}_{k-1|k-1}^T\boldsymbol{\Phi}_{k-1}^T\right] \;+\; \mathbb{E}\left[\boldsymbol{w}_k\boldsymbol{w}_k^T\right] \\ &= \boldsymbol{\Phi}_{k-1}\boldsymbol{P}_{k-1|k-1}\boldsymbol{\Phi}_{k-1}^T \;+\; \boldsymbol{Q}_k \end{aligned} \tag{B.50}$$

It can be noticed that the equations (B.37), (B.36), (B.42), (B.49) and (B.50) are identical to those obtained by Kalman using the method of orthogonal projections.

As a matter of fact, given (B.46), the filter gain (B.36) can be rewritten as:

$$\boldsymbol{K}_k = \boldsymbol{P}_{k|k-1}\boldsymbol{H}_k^T\boldsymbol{\Lambda}_k^{-1} \tag{B.51}$$

The covariance matrix (B.42) of the state estimation error can be rewritten as:

$$\begin{aligned}
\boldsymbol{P}_{k|k} &= (\boldsymbol{I}-\boldsymbol{K}_k\boldsymbol{H}_k)\boldsymbol{P}_{k|k-1}(\boldsymbol{I}-\boldsymbol{K}_k\boldsymbol{H}_k)^T + \boldsymbol{K}_k\boldsymbol{R}_k\boldsymbol{K}_k^T \\
&= \boldsymbol{P}_{k|k-1} - 2\boldsymbol{K}_k\boldsymbol{H}_k\boldsymbol{P}_{k|k-1} + \boldsymbol{K}_k\boldsymbol{H}_k\boldsymbol{P}_{k|k-1}\boldsymbol{H}_k^T\boldsymbol{K}_k^T + \boldsymbol{K}_k\boldsymbol{R}\boldsymbol{K}_k^T \\
&= \boldsymbol{P}_{k|k-1} - 2\boldsymbol{K}_k\boldsymbol{H}_k\boldsymbol{P}_{k|k-1} + \boldsymbol{K}_k\left(\boldsymbol{H}_k\boldsymbol{P}_{k|k-1}\boldsymbol{H}_k^T + \boldsymbol{R}_k\right)\boldsymbol{K}_k^T \\
&= \boldsymbol{P}_{k|k-1} - 2\boldsymbol{K}_k\boldsymbol{H}_k\boldsymbol{P}_{k|k-1} + \boldsymbol{K}_k\boldsymbol{\Lambda}_k\boldsymbol{K}_k \\
&= \boldsymbol{P}_{k|k-1} - 2\boldsymbol{P}_{k|k-1}\boldsymbol{H}_k^T\boldsymbol{\Lambda}_k^{-1}\boldsymbol{H}_k\boldsymbol{P}_{k|k-1} + \boldsymbol{P}_{k|k-1}\boldsymbol{H}_k^T\boldsymbol{\Lambda}_k^{-1}\,\boldsymbol{\Lambda}_k\,\boldsymbol{\Lambda}_k^{-1}\boldsymbol{H}_k\boldsymbol{P}_{k|k-1} \\
&= \boldsymbol{P}_{k|k-1} - \boldsymbol{P}_{k|k-1}\boldsymbol{H}_k^T\boldsymbol{\Lambda}_k^{-1}\boldsymbol{H}_k\boldsymbol{P}_{k|k-1} \\
&= \boldsymbol{P}_{k|k-1} - \boldsymbol{K}_k\,\boldsymbol{\Lambda}_k\,\boldsymbol{K}_k^T
\end{aligned} \tag{B.52}$$

Moreover:

$$\begin{aligned}
\boldsymbol{P}_{k|k} &= \boldsymbol{P}_{k|k-1} - \boldsymbol{P}_{k|k-1}\boldsymbol{H}_k^T\boldsymbol{\Lambda}_k^{-1}\boldsymbol{H}_k\boldsymbol{P}_{k|k-1} \\
&= \boldsymbol{P}_{k|k-1} - \boldsymbol{P}_{k|k-1}\boldsymbol{H}_k^T\left(\boldsymbol{R} + \boldsymbol{H}_k\boldsymbol{P}_{k|k-1}\boldsymbol{H}_k^T\right)^{-1}\boldsymbol{H}_k\boldsymbol{P}_{k|k-1} \\
&= \left(\boldsymbol{P}_{k|k-1}^{-1} + \boldsymbol{H}_k^T\boldsymbol{R}^{-1}\boldsymbol{H}_k\right)^{-1}
\end{aligned} \tag{B.53}$$

where the last expression has been obtained by using the matrix inversion lemma (B.34).

B.3 Problem formulation

The problem considered in this section is to obtain a simultaneous (optimal) estimation of the system state and the covariance matrices of the process and measurement noises, when the latter are assumed to be not precisely known a priori. In the following, they are assumed *diagonal*:

$$\boldsymbol{Q}_k = \begin{bmatrix} q_{11,k} & & 0 \\ & \ddots & \\ 0 & & q_{nn,k} \end{bmatrix}, \qquad \boldsymbol{R}_k = \begin{bmatrix} r_{11,k} & & 0 \\ & \ddots & \\ 0 & & r_{pp,k} \end{bmatrix} \tag{B.54}$$

The parameters of the process and measurement noise covariance matrices are grouped into the following vector of *noise covariance parameters*:

$$\boldsymbol{\xi}^T \triangleq [q_{11,k}\,,\ \ldots\,,\ q_{nn,k}\,,\ r_{11,k}\,,\ \ldots\,,\ r_{pp,k}]^T \tag{B.55}$$

Similarly to Sect. B.2, suppose to adopt a MAP criterion to obtain the joint estimate of the system state and the noise covariance parameters. This corresponds to maximize the joint posterior PDF:

$$\begin{aligned}\ell_k^A(\boldsymbol{x}_k,\,\boldsymbol{\xi},\,\mathcal{U}_k,\,\mathcal{Y}_k) &\triangleq p(\boldsymbol{x}_k,\,\boldsymbol{\xi}\,|\,\mathcal{U}_k,\,\mathcal{Y}_k)\\ &= p(\boldsymbol{x}_k\,|\,\boldsymbol{\xi},\,\mathcal{U}_k,\mathcal{Y}_k)\;p(\boldsymbol{\xi}\,|\,\mathcal{U}_k,\mathcal{Y}_k)\\ &= p(\boldsymbol{x}_k\,|\,\boldsymbol{\xi},\,\mathcal{U}_k,\mathcal{Y}_k)\;p(\mathcal{Y}_k\,|\,\boldsymbol{\xi},\,\mathcal{U}_k)\,\frac{p(\boldsymbol{\xi}\,|\,\mathcal{U}_k)}{p(\mathcal{Y}_k\,|\,\mathcal{U}_k)}\end{aligned} \tag{B.56}$$

where:

$$\begin{aligned}p(\mathcal{Y}_k\,|\,\boldsymbol{\xi},\,\mathcal{U}_k) &= p(\boldsymbol{y}_k,\,\mathcal{Y}_{k-1}\,|\,\boldsymbol{\xi},\,\mathcal{U}_k)\\ &= p(\boldsymbol{y}_k\,|\,\boldsymbol{\xi},\,\mathcal{U}_k,\,\mathcal{Y}_{k-1})\;p(\mathcal{Y}_{k-1}\,|\,\boldsymbol{\xi},\,\mathcal{U}_{k-1})\\ &= p(\boldsymbol{y}_k\,|\,\boldsymbol{\xi},\,\mathcal{U}_{k-1},\,\mathcal{Y}_{k-1})\;p(\mathcal{Y}_{k-1}\,|\,\boldsymbol{\xi},\,\mathcal{U}_{k-1})\\ &\;\;\vdots\\ &= \prod_{i=1}^{k} p(\boldsymbol{y}_i\,|\,\boldsymbol{\xi},\,\mathcal{U}_{i-1},\,\mathcal{Y}_{i-1})\end{aligned} \tag{B.57}$$

or, equivalently, the natural logarithm of that function:

$$\begin{aligned}L_k^A(\boldsymbol{x}_k,\,\boldsymbol{\xi},\,\mathcal{U}_k,\,\mathcal{Y}_k) &\triangleq \ln \ell_k^A(\boldsymbol{x}_k,\boldsymbol{\xi},\mathcal{U}_k,\mathcal{Y}_k)\\ &= \ln p(\boldsymbol{x}_k\,|\,\boldsymbol{\xi},\,\mathcal{U}_k,\,\mathcal{Y}_k) + \ln p(\mathcal{Y}_k\,|\,\boldsymbol{\xi},\,\mathcal{U}_k) + \ln p(\boldsymbol{\xi}\,|\,\mathcal{U}_k) - \ln p(\mathcal{Y}_k\,|\,\mathcal{U}_k)\\ &= \ln p(\boldsymbol{x}_k\,|\,\boldsymbol{\xi},\,\mathcal{U}_k,\,\mathcal{Y}_k) + \sum_{i=1}^{k}\ln p(\boldsymbol{y}_i\,|\,\boldsymbol{\xi},\,\mathcal{U}_{i-1},\,\mathcal{Y}_{i-1}) + \cdots\\ &\qquad\cdots + \ln p(\boldsymbol{\xi}\,|\,\mathcal{U}_k) - \ln p(\mathcal{Y}_k\,|\,\mathcal{U}_k)\end{aligned} \tag{B.58}$$

By using (B.44) and (B.48), the log-value of the joint conditional PDF becomes:

$$\begin{aligned}L_k^A(\boldsymbol{x}_k,\,\boldsymbol{\xi},\,\mathcal{U}_k,\,\mathcal{Y}_k) &= -\tfrac{1}{2}\ln\det \boldsymbol{P}_{k|k} - \tfrac{1}{2}\tilde{\boldsymbol{x}}_{k|k}^T\,\boldsymbol{P}_{k|k}^{-1}\,\tilde{\boldsymbol{x}}_{k|k} + \cdots\\ &\qquad\cdots + \sum_{i=1}^{k}\left[-\tfrac{1}{2}\ln\det\boldsymbol{\Lambda}_i - \tfrac{1}{2}\,\tilde{\boldsymbol{y}}_{i|i-1}^T\,\boldsymbol{\Lambda}_i^{-1}\,\tilde{\boldsymbol{y}}_{i|i-1}\right] + \cdots\\ &\qquad\cdots + \ln p(\boldsymbol{\xi}\,|\,\mathcal{Y}_k) + \text{const}\end{aligned} \tag{B.59}$$

The last term $\ln p(\boldsymbol{\xi}\,|\,\mathcal{Y}_k)$ in (B.59) represents the *a priori* information (i.e. PDF) about the unknown parameter $\boldsymbol{\xi}$ (see also equation (3.3.38) in [99], and equation (15) in [76]). If such information is not available, then the term can be dropped from (B.59). This is equivalent to maximize the log-value of the conditional PDF:

$$\begin{aligned}\ell_k(\boldsymbol{x}_k,\,\boldsymbol{\xi},\,\mathcal{U}_k,\,\mathcal{Y}_k) &\triangleq p(\boldsymbol{x}_k,\,\mathcal{Y}_k\,|\,\boldsymbol{\xi},\,\mathcal{U}_k)\\ &= p(\boldsymbol{x}_k\,|\,\boldsymbol{\xi},\,\mathcal{U}_k,\mathcal{Y}_k)\;p(\mathcal{Y}_k\,|\,\boldsymbol{\xi},\,\mathcal{U}_k)\end{aligned} \tag{B.60}$$

Similarly to (B.60), all the measurements starting from the initial time must be accounted for in the maximization process, in particular because of the structure of (B.57). This would yield a filter with *growing-length memory*, which is unrealistic for online estimation [68]. To enable the implementation of estimators with *fixed-length memory*, the conditional PDF (B.60) (or (B.56) in case the a priori information about the unknown parameter is available) is modified as follows to account for only a limited number of measurements (see also equation (10-19) and the discussion at pp. 76-77 of [70]):

$$\begin{aligned}
\ell_k(\boldsymbol{x}_k,\, \boldsymbol{\xi},\, \mathcal{U}_k^N,\, \mathcal{Y}_k^N) &\triangleq p(\boldsymbol{x}_k,\, \mathcal{Y}_k^N \,|\, \boldsymbol{\xi},\, \mathcal{U}_k^N,\, \mathcal{Y}_{k-N}) \\
&= p(\boldsymbol{x}_k \,|\, \boldsymbol{\xi},\, \mathcal{U}_k, \mathcal{Y}_k^N,\, \mathcal{Y}_{k-N})\ p(\mathcal{Y}_k^N \,|\, \boldsymbol{\xi},\, \mathcal{U}_k,\, \mathcal{Y}_{k-N}) \\
&= p(\boldsymbol{x}_k \,|\, \boldsymbol{\xi},\, \mathcal{U}_k, \mathcal{Y}_k)\ \frac{p(\mathcal{Y}_k^N,\, \mathcal{Y}_{k-N} \,|\, \boldsymbol{\xi},\, \mathcal{U}_k)}{p(\mathcal{Y}_{k-N} \,|\, \boldsymbol{\xi},\, \mathcal{U}_k)} \\
&= p(\boldsymbol{x}_k \,|\, \boldsymbol{\xi},\, \mathcal{U}_k, \mathcal{Y}_k)\ \frac{p(\mathcal{Y}_k \,|\, \boldsymbol{\xi},\, \mathcal{U}_k)}{p(\mathcal{Y}_{k-N} \,|\, \boldsymbol{\xi},\, \mathcal{U}_k)} \\
&= p(\boldsymbol{x}_k \,|\, \boldsymbol{\xi},\, \mathcal{U}_k, \mathcal{Y}_k)\ \frac{\prod_{i=1}^{k}\ p(\boldsymbol{y}_i \,|\, \boldsymbol{\xi},\, \mathcal{U}_{i-1},\, \mathcal{Y}_{i-1})}{\prod_{i=1}^{k-N}\ p(\boldsymbol{y}_i \,|\, \boldsymbol{\xi},\, \mathcal{U}_{i-1},\, \mathcal{Y}_{i-1})} \\
&= p(\boldsymbol{x}_k \,|\, \boldsymbol{\xi},\, \mathcal{U}_k, \mathcal{Y}_k) \prod_{i=k-N+1}^{k} p(\boldsymbol{y}_i \,|\, \boldsymbol{\xi},\, \mathcal{U}_{i-1},\, \mathcal{Y}_{i-1})
\end{aligned} \tag{B.61}$$

where

$$\mathcal{U}_k^N \triangleq \{u_i \,:\, i = k-N+1,\, \ldots,\, k\}\ , \qquad \mathcal{Y}_k^N \triangleq \{y_i \,:\, i = k-N+1,\, \ldots,\, k\} \tag{B.62}$$

The corresponding log-value of this latter conditional PDF is:

$$\begin{aligned}
L_k(\boldsymbol{x}_k,\, \boldsymbol{\xi},\, \mathcal{U}_k,\, \mathcal{Y}_k) \;=\; & -\tfrac{1}{2} \ln\det \boldsymbol{P}_{k|k} \;-\; \tfrac{1}{2} \tilde{\boldsymbol{x}}_{k|k}^T\, \boldsymbol{P}_{k|k}^{-1}\, \tilde{\boldsymbol{x}}_{k|k} \;+\; \cdots \\
& \cdots + \sum_{i=k-N+1}^{k} \left[-\tfrac{1}{2} \ln\det \boldsymbol{\Lambda}_i - \tfrac{1}{2}\, \tilde{\boldsymbol{y}}_{i|i-1}^T\, \boldsymbol{\Lambda}_i^{-1}\, \tilde{\boldsymbol{y}}_{i|i-1} \right] \;+\; \cdots \\
& \cdots + \text{const}
\end{aligned} \tag{B.63}$$

The maximization of (B.63) (or (B.61)) yields a growing-length-memory state estimator, and a fixed-length-memory parameter estimator. It is also immediate to notice that the value $(\hat{\boldsymbol{x}}_{k|k},\, \hat{\boldsymbol{\xi}}_k)$ maximizing (B.63) corresponds to a joint MAP estimator for the system state, and a ML estimator for the unknown parameter (noise covariance parameters). However, in [70] (and all the derived literature) no distinction is made between MAP and ML estimators, and all the objective functions like (B.56), (B.60) or (B.61) are collectively referred as *likelihood functions*. Correspondingly, the estimates obtained from their maximization are simply referred as *maximum likelihood* estimates, regardless of the exact PDF involved in the optimization process.

B.3.1 Determination of the maximum likelihood solution

In the following we consider the maximization of the "likelihood function" (B.63). The value $(\hat{\boldsymbol{x}}_{k|k},\, \hat{\boldsymbol{\xi}}_k)$ that maximizes the function is determined by equating its derivatives with respect to $\boldsymbol{x}_k$ and $\boldsymbol{\xi}$ to zero. The derivative with respect to $\boldsymbol{x}_k$ is equal to[1]:

$$\frac{\partial L_k(\boldsymbol{x}_k,\boldsymbol{\xi})}{\partial \boldsymbol{x}_k} = -\boldsymbol{P}_{k|k}^{-1}\left(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k}\right) \tag{B.64}$$

For computing the derivative with respect to the j^{th} component of $\boldsymbol{\xi}$, denoted as ξ_j, the following intermediate results are exploited (see also (46) in [106]):

$$\frac{\partial}{\partial \xi_j}\ln\det\boldsymbol{A} = \frac{1}{\det\boldsymbol{A}}\frac{\partial\det\boldsymbol{A}}{\partial\xi_j} = \operatorname{Tr}\left[\boldsymbol{A}^{-1}\frac{\partial\boldsymbol{A}}{\partial\xi_j}\right] \tag{B.65}$$

$$\begin{aligned}\frac{\partial}{\partial \xi_j}\left(\boldsymbol{x}^T\boldsymbol{A}^{-1}\boldsymbol{x}\right) &= \frac{\partial\boldsymbol{x}^T}{\partial\xi_j}\boldsymbol{A}^{-1}\boldsymbol{x} + \boldsymbol{x}^T\frac{\partial\boldsymbol{A}^{-1}}{\partial\xi_j}\boldsymbol{x} + \boldsymbol{x}^T\boldsymbol{A}^{-1}\frac{\partial\boldsymbol{x}}{\partial\xi_j}\\ &= 2\,\boldsymbol{x}^T\boldsymbol{A}^{-1}\frac{\partial\boldsymbol{x}}{\partial\xi_j} + \boldsymbol{x}^T\frac{\partial\boldsymbol{A}^{-1}}{\partial\xi_j}\boldsymbol{x}\\ &= 2\,\boldsymbol{x}^T\boldsymbol{A}^{-1}\frac{\partial\boldsymbol{x}}{\partial\xi_j} - \boldsymbol{x}^T\boldsymbol{A}^{-1}\frac{\partial\boldsymbol{A}}{\partial\xi_j}\boldsymbol{A}^{-1}\boldsymbol{x}\\ &= 2\,\boldsymbol{x}^T\boldsymbol{A}^{-1}\frac{\partial\boldsymbol{x}}{\partial\xi_j} - \operatorname{Tr}\left[\boldsymbol{A}^{-1}\boldsymbol{x}\boldsymbol{x}^T\boldsymbol{A}^{-1}\frac{\partial\boldsymbol{A}}{\partial\xi_j}\right]\end{aligned} \tag{B.66}$$

In (B.66) the following identities have been used (see also (59) and (17) in [106]):

$$\frac{\partial\boldsymbol{A}^{-1}}{\partial\xi_j} = -\boldsymbol{A}^{-1}\frac{\partial\boldsymbol{A}}{\partial\xi_j}\boldsymbol{A}^{-1} \tag{B.67}$$

$$\boldsymbol{a}^T\boldsymbol{b} = \operatorname{Tr}\left[\boldsymbol{a}\boldsymbol{b}^T\right] = \operatorname{Tr}\left[\boldsymbol{b}\boldsymbol{a}^T\right] \tag{B.68}$$

By using (B.65) and (B.66) it is immediate to verify that:

$$\begin{aligned}\frac{\partial L_k(\boldsymbol{x}_k,\boldsymbol{\xi})}{\partial\xi_j} &= -\frac{1}{2}\left\{\frac{\partial}{\partial\xi_j}\ln\det\boldsymbol{P}_{k|k} + \frac{\partial}{\partial\xi_j}\left(\tilde{\boldsymbol{x}}_{k|k}^T\boldsymbol{P}_{k|k}^{-1}\tilde{\boldsymbol{x}}_{k|k}\right) + \cdots\right.\\ &\qquad\left.\cdots + \sum_{i=k-N+1}^{k}\frac{\partial}{\partial\xi_j}\ln\det\boldsymbol{\Lambda}_i + \frac{\partial}{\partial\xi_j}\left(\tilde{\boldsymbol{y}}_{i|i-1}^T\boldsymbol{\Lambda}_i^{-1}\tilde{\boldsymbol{y}}_{i|i-1}\right)\right\}\\ &= -\frac{1}{2}\left\{\operatorname{Tr}\left[\left(\boldsymbol{P}_{k|k}^{-1} - \boldsymbol{P}_{k|k}^{-1}\tilde{\boldsymbol{x}}_{k|k}\tilde{\boldsymbol{x}}_{k|k}^T\boldsymbol{P}_{k|k}^{-1}\right)\frac{\partial\boldsymbol{P}_{k|k}}{\partial\xi_j}\right] + 2\,\tilde{\boldsymbol{x}}_{k|k}^T\boldsymbol{P}_{k|k}^{-1}\frac{\partial\tilde{\boldsymbol{x}}_{k|k}}{\partial\xi_j} + \cdots\right.\\ &\left.\cdots + \sum_{i=k-N+1}^{k}\operatorname{Tr}\left[\left(\boldsymbol{\Lambda}_i^{-1} - \boldsymbol{\Lambda}_i^{-1}\tilde{\boldsymbol{y}}_{i|i-1}\tilde{\boldsymbol{y}}_{i|i-1}^T\boldsymbol{\Lambda}_i^{-1}\right)\frac{\partial\boldsymbol{\Lambda}_i}{\partial\xi_j}\right] + 2\,\tilde{\boldsymbol{y}}_{i|i-1}^T\boldsymbol{\Lambda}_i^{-1}\frac{\partial\tilde{\boldsymbol{y}}_{i|i-1}}{\partial\xi_j}\right\}\end{aligned} \tag{B.69}$$

But:

$$\frac{\partial\tilde{\boldsymbol{x}}_{k|k}}{\partial\xi_j} = \frac{\partial}{\partial\xi_j}(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k}) = -\frac{\partial\hat{\boldsymbol{x}}_{k|k}}{\partial\xi_j} \tag{B.70}$$

$$\frac{\partial\tilde{\boldsymbol{y}}_{i|i-1}}{\partial\xi_j} = \frac{\partial}{\partial\xi_j}(\boldsymbol{y}_i - \boldsymbol{H}_i\hat{\boldsymbol{x}}_{i|i-1}) = -\boldsymbol{H}_i\frac{\partial\hat{\boldsymbol{x}}_{i|i-1}}{\partial\xi_j} \tag{B.71}$$

[1] In the following, the notation $L_k(\boldsymbol{x}_k,\boldsymbol{\xi},\mathcal{U}_k,\mathcal{Y}_k)$ is shortened as $L_k(\boldsymbol{x}_k,\boldsymbol{\xi})$.

so that:

$$\begin{aligned}\frac{\partial L_k}{\partial \xi_j} = -\frac{1}{2}\Bigg\{ &\mathrm{Tr}\left[\left(\boldsymbol{P}_{k|k}^{-1} - \boldsymbol{P}_{k|k}^{-1}\tilde{\boldsymbol{x}}_{k|k}\tilde{\boldsymbol{x}}_{k|k}^T\boldsymbol{P}_{k|k}^{-1}\right)\frac{\partial \boldsymbol{P}_{k|k}}{\partial \xi_j}\right] - 2\,\tilde{\boldsymbol{x}}_{k|k}^T\,\boldsymbol{P}_{k|k}^{-1}\,\frac{\partial \hat{\boldsymbol{x}}_{k|k}}{\partial \xi_j} + \cdots \\ &\cdots + \sum_{i=k-N+1}^{k}\mathrm{Tr}\left[\left(\boldsymbol{\Lambda}_i^{-1} - \boldsymbol{\Lambda}_i^{-1}\tilde{\boldsymbol{y}}_{i|i-1}\tilde{\boldsymbol{y}}_{i|i-1}^T\boldsymbol{\Lambda}_i^{-1}\right)\frac{\partial \boldsymbol{\Lambda}_i}{\partial \xi_j}\right] - 2\,\tilde{\boldsymbol{y}}_{i|i-1}^T\,\boldsymbol{\Lambda}_i^{-1}\,\boldsymbol{H}_i\,\frac{\partial \hat{\boldsymbol{x}}_{i|i-1}}{\partial \xi_j}\Bigg\}\end{aligned} \tag{B.72}$$

Equating (B.64) to zero yields:

$$\frac{\partial L_k(\boldsymbol{x}_k, \boldsymbol{\xi})}{\partial \boldsymbol{x}_k} = 0 \qquad \Rightarrow \qquad \boldsymbol{x}_k = \hat{\boldsymbol{x}}_{k|k} \tag{B.73}$$

By considering the previous results, after equating (B.69) to zero it follows that:

$$\begin{aligned}\frac{\partial L_k(\boldsymbol{x}_k, \boldsymbol{\xi})}{\partial \xi_j} = -\frac{1}{2}\Bigg\{ &\mathrm{Tr}\left[\boldsymbol{P}_{k|k}^{-1}\frac{\partial \boldsymbol{P}_{k|k}}{\partial \xi_j}\right] + \sum_{i=k-N+1}^{k}\mathrm{Tr}\left[\left(\boldsymbol{\Lambda}_i^{-1} - \boldsymbol{\Lambda}_i^{-1}\tilde{\boldsymbol{y}}_{i|i-1}\tilde{\boldsymbol{y}}_{i|i-1}^T\boldsymbol{\Lambda}_i^{-1}\right)\frac{\partial \boldsymbol{\Lambda}_i}{\partial \xi_j}\right] - \cdots \\ &\cdots - 2\,\tilde{\boldsymbol{y}}_{i|i-1}^T\,\boldsymbol{\Lambda}_i^{-1}\,\boldsymbol{H}_i\,\frac{\partial \hat{\boldsymbol{x}}_{i|i-1}}{\partial \xi_j}\Bigg\} = 0\end{aligned} \tag{B.74}$$

Equation (B.74) can only be solved using numerical techniques, which, however, demand computational times incompatible with online estimation; or by introducing some simplifications [98]. In the next section, an explicit suboptimal solution is proposed to estimate online process and measurement noise covariance matrices.

B.3.2 Explicit suboptimal solution

In this approach, the equation (B.74) is simplified by assuming that:

$$\mathrm{Tr}\left[\boldsymbol{P}_{k|k}^{-1}\frac{\partial \boldsymbol{P}_{k|k}}{\partial \xi_j}\right] = 0\,, \qquad\qquad \frac{\partial \hat{\boldsymbol{x}}_{i|i-1}}{\partial \xi_j} = 0 \tag{B.75}$$

namely the 1st term in (B.74) is dominated by the 2nd, and the state prediction $\hat{\boldsymbol{x}}_{i|i-1}$ is almost unaffected by the parameter ξ_j. The approximated equation is:

$$\begin{aligned}\frac{\partial \tilde{L}_k(\boldsymbol{x}_k, \boldsymbol{\xi})}{\partial \xi_j} &= -\frac{1}{2}\sum_{i=k-N+1}^{k}\mathrm{Tr}\left[\left(\boldsymbol{\Lambda}_i^{-1} - \boldsymbol{\Lambda}_i^{-1}\tilde{\boldsymbol{y}}_{i|i-1}\tilde{\boldsymbol{y}}_{i|i-1}^T\boldsymbol{\Lambda}_i^{-1}\right)\frac{\partial \boldsymbol{\Lambda}_i}{\partial \xi_j}\right] \\ &= -\frac{1}{2}\sum_{i=k-N+1}^{k}\mathrm{Tr}\left[\Delta\boldsymbol{\Lambda}_i^{-1}\frac{\partial \boldsymbol{\Lambda}_i}{\partial \xi_j}\right] = 0\end{aligned} \tag{B.76}$$

where:

$$\Delta\boldsymbol{\Lambda}_i^{-1} \triangleq \boldsymbol{\Lambda}_i^{-1} - \boldsymbol{\Lambda}_i^{-1}\tilde{\boldsymbol{y}}_{i|i-1}\tilde{\boldsymbol{y}}_{i|i-1}^T\boldsymbol{\Lambda}_i^{-1} \tag{B.77}$$

It holds that:

$$
\begin{aligned}
\frac{\partial \boldsymbol{\Lambda}_i}{\partial r_{j,k}} &= \frac{\partial}{\partial r_{j,k}} \left[\boldsymbol{H}_i \boldsymbol{P}_{i|i-1} \boldsymbol{H}_i^T + \boldsymbol{R}_i \right] \\
&= \boldsymbol{H}_i \frac{\partial \boldsymbol{P}_{i|i-1}}{\partial r_{j,k}} \boldsymbol{H}_i^T + \frac{\partial \boldsymbol{R}_i}{\partial r_{j,k}} \\
&= \boldsymbol{H}_i \frac{\partial \boldsymbol{P}_{i|i-1}}{\partial r_{j,k}} \boldsymbol{H}_i^T + \boldsymbol{J}^{jj}
\end{aligned}
\tag{B.78}
$$

$$
\begin{aligned}
\frac{\partial \boldsymbol{\Lambda}_i}{\partial q_{j,k}} &= \frac{\partial}{\partial q_{j,k}} \left[\boldsymbol{H}_i \boldsymbol{P}_{i|i-1} \boldsymbol{H}_i^T + \boldsymbol{R}_i \right] \\
&= \boldsymbol{H}_i \frac{\partial \boldsymbol{P}_{i|i-1}}{\partial q_{j,k}} \boldsymbol{H}_i^T
\end{aligned}
\tag{B.79}
$$

where $q_{j,k}$ and $r_{j,k}$ are the generic elements of the $\boldsymbol{Q}_k$ and $\boldsymbol{R}_k$ matrices, and $\boldsymbol{J}^{jj}$ is the *single entry* matrix with 1 at position (j, j), and zero otherwise (see p.5 of [106]). By using the above expressions, equation (B.76) reduces to these two conditions for the derivatives with respect to $r_{j,k}$ and $q_{j,k}$:

$$
\begin{aligned}
\frac{\partial \tilde{L}_k}{\partial r_{j,k}} &= -\frac{1}{2} \sum_{i=k-N+1}^{k} \operatorname{Tr} \left[\Delta \boldsymbol{\Lambda}_i^{-1} \frac{\partial \boldsymbol{\Lambda}_i}{\partial r_{j,k}} \right] \\
&= -\frac{1}{2} \sum_{i=k-N+1}^{k} \operatorname{Tr} \left[(\Delta \boldsymbol{\Lambda}_i)_{jj}^{-1} + \Delta \boldsymbol{\Lambda}_i^{-1} \boldsymbol{H}_i \frac{\partial \boldsymbol{P}_{i|i-1}}{\partial r_{j,k}} \boldsymbol{H}_i^T \right] = 0
\end{aligned}
\tag{B.80}
$$

$$
\begin{aligned}
\frac{\partial \tilde{L}_k}{\partial q_{j,k}} &= -\frac{1}{2} \sum_{i=k-N+1}^{k} \operatorname{Tr} \left[\Delta \boldsymbol{\Lambda}_i^{-1} \frac{\partial \boldsymbol{\Lambda}_i}{\partial q_{j,k}} \right] \\
&= -\frac{1}{2} \sum_{i=k-N+1}^{k} \operatorname{Tr} \left[\Delta \boldsymbol{\Lambda}_i^{-1} \boldsymbol{H}_i \frac{\partial \boldsymbol{P}_{i|i-1}}{\partial q_{j,k}} \boldsymbol{H}_i^T \right] = 0
\end{aligned}
\tag{B.81}
$$

where the notation $(\boldsymbol{A})_{jj}$ is used to specify the element at position (j, j) of the matrix $\boldsymbol{A}$ (see p.5 of [106]). Consider equation (B.80) first. In many applications it can be shown that:

$$
\boldsymbol{H}_i \frac{\partial \boldsymbol{P}_{i|i-1}}{\partial r_{j,k}} \boldsymbol{H}_i^T \ll \boldsymbol{I}
\tag{B.82}
$$

so that:

$$
\begin{aligned}
\frac{\partial \tilde{L}_k}{\partial r_{j,k}} &\approx -\frac{1}{2} \sum_{i=k-N+1}^{k} \operatorname{Tr}\left[(\Delta \boldsymbol{\Lambda}_i)^{-1}_{jj}\right] \\
&= -\frac{1}{2} \sum_{i=k-N+1}^{k} (\Delta \boldsymbol{\Lambda}_i)^{-1}_{jj} \\
&= -\frac{1}{2} \sum_{i=k-N+1}^{k} \left(\boldsymbol{\Lambda}_i^{-1} - \boldsymbol{\Lambda}_i^{-1} \tilde{\boldsymbol{y}}_{i|i-1} \tilde{\boldsymbol{y}}_{i|i-1}^T \boldsymbol{\Lambda}_i^{-1}\right)_{jj} = 0
\end{aligned}
\tag{B.83}
$$

It can be proved that:

$$
\boldsymbol{\Lambda}_i^{-1} \tilde{\boldsymbol{y}}_{i|i-1} = \boldsymbol{R}_i^{-1} \tilde{\boldsymbol{y}}_{i|i} \tag{B.84}
$$

where:

$$
\tilde{\boldsymbol{y}}_{i|i} = \boldsymbol{y}_i - \boldsymbol{H}_i \hat{\boldsymbol{x}}_{i|i} \tag{B.85}
$$

In fact:

$$
\begin{aligned}
\tilde{\boldsymbol{y}}_{i|i} &= \boldsymbol{y}_i - \boldsymbol{H}_i \hat{\boldsymbol{x}}_{i|i} \\
&= \boldsymbol{y}_i - \boldsymbol{H}_i \left[\hat{\boldsymbol{x}}_{i|i-1} + \boldsymbol{K}_i \left(\boldsymbol{y}_i - \boldsymbol{H}_i \hat{\boldsymbol{x}}_{i|i-1}\right)\right] \\
&= (\boldsymbol{I} - \boldsymbol{H}_i \boldsymbol{K}_i) \left(\boldsymbol{y}_i - \boldsymbol{H}_i \hat{\boldsymbol{x}}_{i|i-1}\right) \\
&= \left(\boldsymbol{I} - \boldsymbol{H}_i \boldsymbol{P}_{i|i-1}^{-1} \boldsymbol{H}_i^T \boldsymbol{\Lambda}_i^{-1}\right) \tilde{\boldsymbol{y}}_{i|i-1} \\
&= \left(\boldsymbol{\Lambda}_i - \boldsymbol{H}_i \boldsymbol{P}_{i|i-1}^{-1} \boldsymbol{H}_i^T\right) \boldsymbol{\Lambda}_i^{-1} \tilde{\boldsymbol{y}}_{i|i-1} \\
&= \left(\cancel{\boldsymbol{H}_i \boldsymbol{P}_{i|i}^{-1} \boldsymbol{H}_i^T} + \boldsymbol{R}_i - \cancel{\boldsymbol{H}_i \boldsymbol{P}_{i|i-1}^{-1} \boldsymbol{H}_i^T}\right) \boldsymbol{\Lambda}_i^{-1} \tilde{\boldsymbol{y}}_{i|i-1} \\
&= \boldsymbol{R}_i \boldsymbol{\Lambda}_i^{-1} \tilde{\boldsymbol{y}}_{i|i-1}
\end{aligned}
\tag{B.86}
$$

Moreover, by using the matrix inversion lemma (B.34) and the identity (B.53) it follows that:

$$
\begin{aligned}
\boldsymbol{\Lambda}_i^{-1} &= \left(\boldsymbol{H}_i \boldsymbol{P}_{i|i-1} \boldsymbol{H}_i^T + \boldsymbol{R}_i\right)^{-1} \\
&= \boldsymbol{R}_i^{-1} - \boldsymbol{R}_i^{-1} \boldsymbol{H}_i \left(\boldsymbol{P}_{i|i-1}^{-1} + \boldsymbol{H}_i^T \boldsymbol{R}_i^{-1} \boldsymbol{H}_i\right)^{-1} \boldsymbol{H}_i^T \boldsymbol{R}_i^{-1} \\
&= \boldsymbol{R}_i^{-1} - \boldsymbol{R}_i^{-1} \boldsymbol{H}_i \boldsymbol{P}_{i|i} \boldsymbol{H}_i^T \boldsymbol{R}_i^{-1}
\end{aligned}
\tag{B.87}
$$

By using (B.84) and (B.87), equation (B.83) reduces to:

$$
\begin{aligned}
\frac{\partial \tilde{L}_k}{\partial r_{j,k}} &\approx -\frac{1}{2}\sum_{i=k-N+1}^{k}\left(\boldsymbol{\Lambda}_i^{-1} - \boldsymbol{\Lambda}_i^{-1}\tilde{\boldsymbol{y}}_{i|i-1}\,\tilde{\boldsymbol{y}}_{i|i-1}^T\boldsymbol{\Lambda}_i^{-1}\right)_{jj} \\
&= -\frac{1}{2}\sum_{i=k-N+1}^{k}\left(\boldsymbol{\Lambda}_i^{-1} - \boldsymbol{R}_i^{-1}\tilde{\boldsymbol{y}}_{i|i}\,\tilde{\boldsymbol{y}}_{i|i}^T\boldsymbol{R}_i^{-1}\right)_{jj} \\
&= -\frac{1}{2}\sum_{i=k-N+1}^{k}\left(\boldsymbol{R}_i^{-1} - \boldsymbol{R}_i^{-1}\boldsymbol{H}_i\boldsymbol{P}_{i|i}\boldsymbol{H}_i^T\boldsymbol{R}_i^{-1} - \boldsymbol{R}_i^{-1}\tilde{\boldsymbol{y}}_{i|i}\,\tilde{\boldsymbol{y}}_{i|i}^T\boldsymbol{R}_i^{-1}\right)_{jj} \qquad \text{(B.88)} \\
&= -\frac{1}{2}\sum_{i=k-N+1}^{k}\left[\boldsymbol{R}_i^{-1}\left(\boldsymbol{R}_i - \boldsymbol{H}_i\boldsymbol{P}_{i|i}\boldsymbol{H}_i^T - \tilde{\boldsymbol{y}}_{i|i}\,\tilde{\boldsymbol{y}}_{i|i}^T\right)\boldsymbol{R}_i^{-1}\right]_{jj} \\
&= -\frac{1}{2}\sum_{i=k-N+1}^{k} r_{j,i}^{-1}\left(\boldsymbol{R}_i - \boldsymbol{H}_i\boldsymbol{P}_{i|i}\boldsymbol{H}_i^T - \tilde{\boldsymbol{y}}_{i|i}\,\tilde{\boldsymbol{y}}_{i|i}^T\right)_{jj} r_{j,i}^{-1} \;=\; 0
\end{aligned}
$$

where the last identity holds because $\boldsymbol{R}_i$ is diagonal. If the estimation process is essentially *time invariant over the most recent N steps* (see p. 122 of [70]), namely:

$$
\boldsymbol{R}_i \approx \boldsymbol{R}_k\,, \qquad \boldsymbol{r}_{j,i} \approx \boldsymbol{r}_{j,k} \qquad k-N+1 \le i \le k \tag{B.89}
$$

then the previous equation is equivalent to:

$$
\begin{aligned}
&\sum_{i=k-N+1}^{k}\left(\boldsymbol{R}_i - \boldsymbol{H}_i\boldsymbol{P}_{i|i}\boldsymbol{H}_i^T - \tilde{\boldsymbol{y}}_{i|i}\,\tilde{\boldsymbol{y}}_{i|i}^T\right)_{jj} = \cdots \\
&\cdots = \sum_{i=k-N+1}^{k}\left[r_{j,i} - \left(\boldsymbol{H}_i\boldsymbol{P}_{i|i}\boldsymbol{H}_i^T + \tilde{\boldsymbol{y}}_{i|i}\,\tilde{\boldsymbol{y}}_{i|i}^T\right)_{jj}\right] \qquad \text{(B.90)} \\
&\cdots = N\,r_{j,k} - \sum_{i=k-N+1}^{k}\left(\boldsymbol{H}_i\boldsymbol{P}_{i|i}\boldsymbol{H}_i^T + \tilde{\boldsymbol{y}}_{i|i}\,\tilde{\boldsymbol{y}}_{i|i}^T\right)_{jj} = 0
\end{aligned}
$$

which yields the following expression for the estimate of the parameter $r_{j,k}$:

$$
\hat{r}_{j,k} = \frac{1}{N}\sum_{i=k-N+1}^{k}\left(\boldsymbol{H}_i\boldsymbol{P}_{i|i}\boldsymbol{H}_i^T + \tilde{\boldsymbol{y}}_{i|i}\,\tilde{\boldsymbol{y}}_{i|i}^T\right)_{jj} \tag{B.91}
$$

Consider next equation (B.81). By using the following property of the matrix trace (see (16) in [106]):

$$
\mathrm{Tr}\left[\boldsymbol{ABC}\right] = \mathrm{Tr}\left[\boldsymbol{BCA}\right] = \mathrm{Tr}\left[\boldsymbol{CAB}\right] \tag{B.92}
$$

the equation reduces to:

$$
\begin{aligned}
\frac{\partial \tilde{L}_k}{\partial q_{j,k}} &= -\frac{1}{2}\sum_{i=k-N+1}^{k}\mathrm{Tr}\left[\Delta\boldsymbol{\Lambda}_i^{-1}\boldsymbol{H}_i\frac{\partial \boldsymbol{P}_{i|i-1}}{\partial q_{j,k}}\boldsymbol{H}_i^T\right] \\
&= -\frac{1}{2}\sum_{i=k-N+1}^{k}\mathrm{Tr}\left[\boldsymbol{H}_i^T\Delta\boldsymbol{\Lambda}_i^{-1}\boldsymbol{H}_i\frac{\partial \boldsymbol{P}_{i|i-1}}{\partial q_{j,k}}\right] = 0
\end{aligned} \tag{B.93}
$$

From (B.50) it follows that:

$$\begin{aligned}\frac{\partial \boldsymbol{P}_{i|i-1}}{\partial q_{j,k}} &= \boldsymbol{\Phi}_{i-1}\frac{\partial \boldsymbol{P}_{i-1|i-1}}{\partial q_{j,k}}\boldsymbol{\Phi}_{i-1}^{T} + \frac{\partial \boldsymbol{Q}_i}{\partial q_{j,k}} \\ &\approx \frac{\partial \boldsymbol{Q}_i}{\partial q_{j,k}} = \boldsymbol{J}^{jj}\end{aligned} \tag{B.94}$$

where it has assumed that (see p. 123 of [70]):

$$\boldsymbol{\Phi}_{i-1}\frac{\partial \boldsymbol{P}_{i-1|i-1}}{\partial q_{j,k}}\boldsymbol{\Phi}_{i-1}^{T} \approx 0 \tag{B.95}$$

Moreover:

$$\begin{aligned}\boldsymbol{H}_i^T\Delta\boldsymbol{\Lambda}_{\boldsymbol{i}}^{-1}\boldsymbol{H}_i &= \boldsymbol{H}_i^T\left(\boldsymbol{\Lambda}_i^{-1} - \boldsymbol{\Lambda}_i^{-1}\tilde{\boldsymbol{y}}_{i|i-1}\tilde{\boldsymbol{y}}_{i|i-1}^T\boldsymbol{\Lambda}_i^{-1}\right)\boldsymbol{H}_i \\ &= \boldsymbol{H}_i^T\boldsymbol{\Lambda}_i^{-1}\boldsymbol{H}_i^T - \boldsymbol{H}_i^T\boldsymbol{\Lambda}_i^{-1}\tilde{\boldsymbol{y}}_{i|i-1}\tilde{\boldsymbol{y}}_{i|i-1}^T\boldsymbol{\Lambda}_i^{-1}\boldsymbol{H}_i \\ &= \boldsymbol{P}_{i|i-1}^{-1}\left(\boldsymbol{P}_{i|i-1}\boldsymbol{H}_i^T\boldsymbol{\Lambda}_i^{-1}\boldsymbol{H}_i\boldsymbol{P}_{i|i-1} - \cdots\right. \\ &\qquad \left.\cdots - \boldsymbol{P}_{i|i-1}\boldsymbol{H}_i^T\boldsymbol{\Lambda}_i^{-1}\tilde{\boldsymbol{y}}_{i|i-1}\tilde{\boldsymbol{y}}_{i|i-1}^T\boldsymbol{\Lambda}_i^{-1}\boldsymbol{H}_i\boldsymbol{P}_{i|i-1}\right)\boldsymbol{P}_{i|i-1}^{-1} \\ &= \boldsymbol{P}_{i|i-1}^{-1}\left(\boldsymbol{K}_i\boldsymbol{\Lambda}_i\boldsymbol{K}_i^T - \boldsymbol{K}_i\tilde{\boldsymbol{y}}_{i|i-1}\tilde{\boldsymbol{y}}_{i|i-1}^T\boldsymbol{K}_i^T\right)\boldsymbol{P}_{i|i-1}^{-1} \\ &= \boldsymbol{P}_{i|i-1}^{-1}\left(\boldsymbol{P}_{i|i-1} - \boldsymbol{P}_{i|i} - \boldsymbol{K}_i\tilde{\boldsymbol{y}}_{i|i-1}\tilde{\boldsymbol{y}}_{i|i-1}^T\boldsymbol{K}_i^T\right)\boldsymbol{P}_{i|i-1}^{-1} \\ &= \boldsymbol{P}_{i|i-1}^{-1}\left(\boldsymbol{\Phi}_{i-1}\boldsymbol{P}_{i-1|i-1}\boldsymbol{\Phi}_{i-1}^T + \boldsymbol{Q}_i - \boldsymbol{P}_{i|i} - \boldsymbol{K}_i\tilde{\boldsymbol{y}}_{i|i-1}\tilde{\boldsymbol{y}}_{i|i-1}^T\boldsymbol{K}_i^T\right)\boldsymbol{P}_{i|i-1}^{-1}\end{aligned} \tag{B.96}$$

where the last two identities are obtained after (B.52) and (B.50). Therefore (B.93) reduces to:

$$\begin{aligned}\frac{\partial \tilde{L}_k}{\partial q_{j,k}} &= -\frac{1}{2}\sum_{i=k-N+1}^{k}\left(\boldsymbol{H}_i^T\Delta\boldsymbol{\Lambda}_i^{-1}\boldsymbol{H}_i\right)_{jj} \\ &= -\frac{1}{2}\sum_{i=k-N+1}^{k}\left[\boldsymbol{P}_{i|i-1}^{-1}\left(\boldsymbol{U}_i + \boldsymbol{Q}_i - \boldsymbol{P}_{i|i} - \Delta\boldsymbol{x}_i\Delta\boldsymbol{x}_i^T\right)\boldsymbol{P}_{i|i-1}^{-1}\right]_{jj} = 0\end{aligned} \tag{B.97}$$

where it has been set:

$$\boldsymbol{U}_i \triangleq \boldsymbol{\Phi}_{i-1}\boldsymbol{P}_{i-1|i-1}\boldsymbol{\Phi}_{i-1}^T \tag{B.98}$$

$$\Delta\boldsymbol{x}_i \triangleq \boldsymbol{K}_i\tilde{\boldsymbol{y}}_{i|i-1} \tag{B.99}$$

Assuming again that the estimation process is *time invariant over the most recent N steps* (see p. 123 of [70]), namely:

$$\boldsymbol{P}_{i|i-1} = \boldsymbol{P}_{k|k-1}\,, \qquad \boldsymbol{Q}_i = \boldsymbol{Q}_k \qquad k-N+1 \le i \le k \tag{B.100}$$

then (B.97) reduces to:

$$\frac{\partial \tilde{L}_k}{\partial q_{j,k}} = -\frac{1}{2} \sum_{i=k-N+1}^{k} \left[\boldsymbol{P}_{i|i-1}^{-1} \left(\boldsymbol{U}_i + \boldsymbol{Q}_i - \boldsymbol{P}_{i|i} - \Delta \boldsymbol{x}_i \Delta \boldsymbol{x}_i^T \right) \boldsymbol{P}_{i|i-1}^{-1} \right]_{jj}$$

$$= -\frac{1}{2} \left[\boldsymbol{P}_{k|k-1} \left(\sum_{i=k-N+1}^{k} \boldsymbol{U}_i + \boldsymbol{Q}_k - \boldsymbol{P}_{i|i} - \Delta \boldsymbol{x}_i \Delta \boldsymbol{x}_i^T \right) \boldsymbol{P}_{k|k-1} \right]_{jj} = 0 \tag{B.101}$$

which is satisfied if and only if:

$$\sum_{i=k-N+1}^{k} \boldsymbol{U}_i + \boldsymbol{Q}_k - \boldsymbol{P}_{i|i} - \Delta \boldsymbol{x}_i \Delta \boldsymbol{x}_i^T = \boldsymbol{0} \tag{B.102}$$

or, alternatively:

$$\boldsymbol{Q}_k = \frac{1}{N} \sum_{i=k-N+1}^{k} \boldsymbol{P}_{i|i} - \boldsymbol{U}_i + \Delta \boldsymbol{x}_i \Delta \boldsymbol{x}_i^T \tag{B.103}$$

The estimate of the j^{th} element on the diagonal of $\boldsymbol{Q}$ is therefore:

$$\hat{q}_{j,k} = \frac{1}{N} \sum_{i=k-N+1}^{k} \left(\boldsymbol{P}_{i|i} - \boldsymbol{U}_i + \Delta \boldsymbol{x}_i \Delta \boldsymbol{x}_i^T \right)_{jj} \tag{B.104}$$

Bibliography

[1] "World energy outlook 2024", *International Energy Agency*, 2024. [Online]. Available: https://iea.blob.core.windows.net/assets/5e9122fc-9d5b-4f18-8438-dac8b39b702a/WorldEnergyOutlook2024.pdf.

[2] "Commission regulation (eu) 2019/1781", *Official Journal of the European Union*, vol. L 272, pp. 74–94, 2019-01-10. [Online]. Available: https://www.iec.ch/government-regulators/electric-motors.

[3] M. Hirsh. "Not so green technology: The complicated legacy of rare earth mining". [Online]. Available: https://hir.harvard.edu/not-so-green-technology-the-complicated-legacy-of-rare-earth-mining/.

[4] R. U. S. Institute. "Japan's responses to china's supply chain dominance". [Online]. Available: https://rusi.org/explore-our-research/publications/commentary/japans-responses-chinas-supply-chain-dominance#:~:text=Following%20a%20territorial%20dispute%20in,dependent%20industries%20and%20supply%20chains..

[5] U. G. Survey. "Rare earths statistics and information". [Online]. Available: https://www.usgs.gov/centers/national-minerals-information-center/rare-earths-statistics-and-information.

[6] U. G. Survey, *Mineral commodity summaries 2024*, Accessed: 2024-12-30, 2024. [Online]. Available: https://pubs.usgs.gov/periodicals/mcs2024/mcs2024.pdf.

[7] "Mapping the impact and conflicts of rare earth elements". [Online]. Available: https://ips-dc.org/mapping-the-impact-and-conflicts-of-rare-earth-elements/.

[8] N. Haque, A. Hughes, S. Lim, and C. Vernon, "Rare earth elements: Overview of mining, mineralogy, uses, sustainability and environmental impact", *Resources*, vol. 3, no. 4, pp. 614–635, 2014, ISSN: 2079-9276. DOI: 10.3390/resources3040614. [Online]. Available: https://www.mdpi.com/2079-9276/3/4/614.

[9] C. Oprea, A. Dziechciarz, and C. Martis, "Comparative analysis of different synchronous reluctance motor topologies", in *2015 IEEE 15th International Conference on Environment and Electrical Engineering (EEEIC)*, 2015, pp. 1904–1909. DOI: 10.1109/EEEIC.2015.7165463.

[10] T. Senjyu, K. Kinjo, N. Urasaki, and K. Uezato, "High efficiency control of synchronous reluctance motors using extended kalman filter", *IEEE Transactions on Industrial Electronics*, vol. 50, no. 4, pp. 726–732, 2003, ISSN: 1557-9948. DOI: 10.1109/TIE.2003.814998.

[11] I. Boldea, L. N. Tutelea, L. Parsa, and D. Dorrell, "Automotive electric propulsion systems with reduced or no permanent magnets: An overview", *IEEE Transactions on Industrial Electronics*, vol. 61, no. 10, pp. 5696–5711, 2014. DOI: 10.1109/TIE.2014.2301754.

[12] V. Manzolini and S. Bolognani, "On the rotor position self-sensing capability of reluctance and ipm synchronous motors", *IEEE Transactions on Industry Applications*, vol. 56, no. 4, pp. 3755–3766, 2020, ISSN: 1939-9367. DOI: 10.1109/TIA.2020.2984406.

[13] J. Antons and T. Rösmann, "Self-sensing control of a synchronous reluctance machine using an extended kalman filter", in *2015 IEEE International Conference on Industrial Technology (ICIT)*, 2015, pp. 831–839. DOI: 10.1109/ICIT.2015.7125201.

[14] S.-C. Agarliţă, M. Fătu, L. N. Tutelea, F. Blaabjerg, and I. Boldea, "I-f starting and active flux based sensorless vector control of reluctance synchronous motors, with experiments", in *2010 12th International Conference on Optimization of Electrical and Electronic Equipment*, 2010, pp. 337–342. DOI: 10.1109/OPTIM.2010.5510564.

[15] A. Ghaderi and T. Hanamoto, "Wide-speed-range sensorless vector control of synchronous reluctance motors based on extended programmable cascaded low-pass filters", *IEEE Transactions on Industrial Electronics*, vol. 58, no. 6, pp. 2322–2333, Jun. 2011, ISSN: 1557-9948. DOI: 10.1109/TIE.2010.2066537.

[16] Z. Zhao, P. Davari, W. Lu, H. Wang, and F. Blaabjerg, "An overview of condition monitoring techniques for capacitors in dc-link applications", *IEEE Transactions on Power Electronics*, vol. 36, no. 4, pp. 3692–3716, 2021. DOI: 10.1109/TPEL.2020.3023469.

[17] H. Wang, M. Liserre, and F. Blaabjerg, "Toward reliable power electronics: Challenges, design tools, and opportunities", *IEEE Industrial Electronics Magazine*, vol. 7, no. 2, pp. 17–26, 2013. DOI: 10.1109/MIE.2013.2252958.

[18] H. Wen, W. Xiao, X. Wen, and P. Armstrong, "Analysis and evaluation of DC-link capacitors for high-power-density electric vehicle drive systems", *IEEE Transactions on Vehicular Technology*, vol. 61, no. 7, pp. 2950–2964, 2012. DOI: 10.1109/TVT.2012.2206082.

[19] M. Hinkkanen and J. Luomi, "Induction motor drives equipped with diode rectifier and small dc-link capacitance", *IEEE Transactions on Industrial Electronics*, vol. 55, no. 1, pp. 312–320, 2008. DOI: 10.1109/TIE.2007.903959.

[20] L. Mathe, L. Török, D. Wang, and D. Sera, “Resonance reduction for AC drives with small capacitance in the DC link”, *IEEE Transactions on Industry Applications*, vol. 53, no. 4, pp. 3814–3820, 2017. DOI: 10.1109/TIA.2017.2688980.

[21] W.-J. Lee and S.-K. Sul, “DC-link voltage stabilization for reduced DC-link capacitor inverter”, *IEEE Transactions on Industry Applications*, vol. 50, no. 1, pp. 404–414, 2014. DOI: 10.1109/TIA.2013.2268733.

[22] D. Ding, G. Wang, N. Zhao, G. Zhang, and D. Xu, “Enhanced flux-weakening control method for reduced DC-link capacitance IPMSM drives”, *IEEE Transactions on Power Electronics*, vol. 34, no. 8, pp. 7788–7799, 2019. DOI: 10.1109/TPEL.2018.2878877.

[23] G. Wang et al., “Overmodulation strategy for electrolytic capacitorless PMSM drives: Voltage distortion analysis and boundary optimization”, *IEEE Transactions on Power Electronics*, vol. 35, no. 9, pp. 9574–9585, 2020. DOI: 10.1109/TPEL.2020.2971084.

[24] D. Ding et al., “Impedance reshaping for inherent harmonics in PMSM drives with small DC-link capacitor”, *IEEE Transactions on Power Electronics*, vol. 37, no. 12, pp. 14 265–14 279, 2022. DOI: 10.1109/TPEL.2022.3192052.

[25] R. Maheshwari, S. Munk-Nielsen, and K. Lu, “An active damping technique for small DC-link capacitor based drive system”, *IEEE Transactions on Industrial Informatics*, vol. 9, no. 2, pp. 848–858, 2013. DOI: 10.1109/TII.2012.2222036.

[26] D. Wang et al., “Voltage modulation using virtual positive impedance concept for active damping of small DC-link drive system”, *IEEE Transactions on Power Electronics*, vol. 33, no. 12, pp. 10 611–10 621, 2018. DOI: 10.1109/TPEL.2018.2799988.

[27] K. Li et al., “Grid current quality improvement for three-phase diode rectifier-fed small DC-link capacitance IPMSM drives”, *IEEE Transactions on Energy Conversion*, vol. 37, no. 4, pp. 2310–2320, 2022. DOI: 10.1109/TEC.2022.3176143.

[28] H. Shin, Y. Son, and J.-I. Ha, “Grid current shaping method with DC-link shunt compensator for three-phase diode rectifier-fed motor drive system”, *IEEE Transactions on Power Electronics*, vol. 32, no. 2, pp. 1279–1288, 2017. DOI: 10.1109/TPEL.2016.2540651.

[29] D. Ding, G. Zhang, G. Wang, and D. Xu, “Dual antiovervoltage control scheme for electrolytic capacitorless IPMSM drives with coefficient autoregulation”, *IEEE Transactions on Power Electronics*, vol. 35, no. 3, pp. 2895–2907, 2020. DOI: 10.1109/TPEL.2019.2932030.

[30] D. Ding, G. Wang, N. Zhao, G. Zhang, and D. Xu, “An antiovervoltage control scheme for electrolytic capacitorless IPMSM drives based on stator current vector orientation”, *IEEE Transactions on Industrial Electronics*, vol. 67, no. 5, pp. 3517–3527, 2020. DOI: 10.1109/TIE.2019.2918497.

[31] J. Huo et al., "Torque ripple compensation with anti-overvoltage for electrolytic capacitorless PMSM compressor drives", *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol. 10, no. 5, pp. 6148–6159, 2022. DOI: 10.1109/JESTPE.2022.3175897.

[32] D. Wang, K. Lu, P. O. Rasmussen, L. Mathe, and Y. Feng, "Analysis of voltage modulation based active damping techniques for small dc-link drive system", in *2015 IEEE Energy Conversion Congress and Exposition (ECCE)*, 2015, pp. 2927–2934. DOI: 10.1109/ECCE.2015.7310070.

[33] P. C. Krause, O. Wasynczuk, S. D. Sudhoff, and S. Pekarek, Eds., *Analysis of Electric Machinery and Drive Systems* (IEEE Press Series on Power Engineering 37), 3. ed. Hoboken, NJ: Wiley, 2013, ISBN: 978-1-118-02429-4.

[34] S. Skogestad and I. Postlethwaite, *Multivariable feedback control: Analysis and Design*. John Wiley, 2005.

[35] "Tekneidos". Accessed: 2024-11-04. [Online]. Available: https://www.tekneidos.com/.

[36] R. W. Erickson and D. Maksimovic, *Fundamentals of Power Electronics*. Springer, 2001.

[37] H. W. Sorenson, "Least-squares estimation: From Gauss to Kalman", *IEEE Spectrum*, vol. 7, no. 7, pp. 63–68, 1970. DOI: 10.1109/MSPEC.1970.5213471.

[38] S. F. Schmidt, "Application of state-space methods to navigation problems", in ser. Advances in Control Systems, C. Leondes, Ed., vol. 3, Elsevier, 1966, pp. 293–340. DOI: https://doi.org/10.1016/B978-1-4831-6716-9.50011-4. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9781483167169500114.

[39] J. K. Uhlmann, "Dynamic map building and localization: New theoretical foundations", Ph.D. dissertation, University of Oxford, 1995. [Online]. Available: https://books.google.it/books?id=eJb4MgEACAAJ.

[40] S. Bolognani, R. Oboe, and M. Zigliotto, "Sensorless full-digital PMSM drive with EKF estimation of speed and rotor position", *IEEE Transactions on Industrial Electronics*, vol. 46, no. 1, pp. 184–191, 1999. DOI: 10.1109/41.744410.

[41] S. Bolognani, M. Zigliotto, and M. Zordan, "Extended-range PMSM sensorless speed drive based on stochastic filtering", *IEEE Transactions on Power Electronics*, vol. 16, no. 1, pp. 110–117, 2001. DOI: 10.1109/63.903995.

[42] D. Pasqualotto, S. Rigon, and M. Zigliotto, "Sensorless speed control of synchronous reluctance motor drives based on extended Kalman filter and neural magnetic model", *IEEE Transactions on Industrial Electronics*, vol. 70, no. 2, pp. 1321–1330, 2023. DOI: 10.1109/TIE.2022.3159962.

[43] S. Julier and J. Uhlmann, "Unscented filtering and nonlinear estimation", in *Proceedings of the IEEE*, vol. 92, 2004, pp. 401–422. DOI: 10.1109/JPROC.2003.823141.

[44] G. Welch and G. Bishop, "An introduction to the Kalman filter", University of North Carolina at Chapel Hill, USA, Tech. Rep., 1995.

[45] D. Simon, *Optimal State Estimation*. Wiley, Jan. 2006. DOI: 10.1002/0470045345. [Online]. Available: https://doi.org/10.1002/0470045345.

[46] J. Uhlmann, *What's with the name "unscented"?* [Online]. Available: https://ethw.org/First-Hand:The_Unscented_Transform#What%E2%80%99s_with_the_Name_%E2%80%9CUnscented%E2%80%9D?.

[47] S. J. Julier and J. K. Uhlmann, "New extension of the Kalman filter to nonlinear systems", in *SPIE Proceedings*, I. Kadar, Ed., SPIE, Jul. 1997. DOI: 10.1117/12.280797. [Online]. Available: https://doi.org/10.1117/12.280797.

[48] S. Särkkä, *Bayesian Filtering and Smoothing*. Cambridge University Press, Sep. 2013. DOI: 10.1017/cbo9781139344203. [Online]. Available: https://doi.org/10.1017/cbo9781139344203.

[49] R. van der Merwe, "Sigma-point Kalman filters for probabilistic inference in dynamic state-space models", Ph.D. dissertation, OGI School of Science & Engineering at Oregon Health & Science University, 2004.

[50] A. J. Haug, *Bayesian Estimation and Tracking*. Wiley, Jun. 2012. DOI: 10.1002/9781118287798. [Online]. Available: https://doi.org/10.1002/9781118287798.

[51] P. Borsje, T. Chan, Y. Wong, and S. Ho, "A comparative study of Kalman filtering for sensorless control of a permanent-magnet synchronous motor drive", in *IEEE International Conference on Electric Machines and Drives, 2005.*, 2005, pp. 815–822. DOI: 10.1109/IEMDC.2005.195816.

[52] K. Horváth, "Load torque and permanent magnet flux linkage estimation of surface-mounted PMSM by using unscented Kalman filter", in *2023 International Conference on Electrical Drives and Power Electronics (EDPE)*, 2023, pp. 1–8. DOI: 10.1109/EDPE58625.2023.10274042.

[53] A. Darba and K. Salahshoor, "EKF and UKF-based estimation of a sensorless axial flux PM machine under an internal model control scheme using a SVPWM inverter", in *Proceedings of the 29th Chinese Control Conference*, 2010, pp. 5676–5681.

[54] L. An and K. Hameyer, "Rotor position and speed estimation of interior permanent magnet synchronous motor using unscented Kalman filter", in *2014 17th International Conference on Electrical Machines and Systems (ICEMS)*, 2014, pp. 727–732. DOI: 10.1109/ICEMS.2014.7013577.

[55] M. K. Hussain, B. M. Alshadeedi, and R. Hejeejo, "Optimized Kalman filters for sensorless vector control induction motor drives", *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 13, no. 1, pp. 17–27, 2023. DOI: 10.11591/ijece.v13i1.

[56] S. Kumar, J. Prakash, and P. Kanagasabapathy, "A critical evaluation and experimental verification of extended Kalman filter, unscented Kalman filter and neural state filter for state estimation of three phase induction motor", *Applied Soft Computing, Elsevier*, vol. 11, no. 3, pp. 3199–3208, 2011. DOI: 10.11591/ijece.v13i1.

[57] R. Yildiz, M. Barut, and E. Zerdali, "A comprehensive comparison of extended and unscented Kalman filters for speed-sensorless control applications of induction motors", *IEEE Transactions on Industrial Informatics*, vol. 16, no. 10, pp. 6423–6432, 2020. DOI: 10.1109/TII.2020.2964876.

[58] Y. Zhang and X.-F. Cheng, "Sensorless Control of Permanent Magnet Synchronous Motors and EKF Parameter Tuning Research", en, *Mathematical Problems in Engineering*, vol. 2016, pp. 1–12, 2016, ISSN: 1024-123X, 1563-5147. DOI: 10.1155/2016/3916231.

[59] A. H. Mohamed and K. P. Schwarz, "Adaptive Kalman Filtering for INS/GPS", en, *Journal of Geodesy*, vol. 73, no. 4, pp. 193–203, May 1999, ISSN: 0949-7714, 1432-1394. DOI: 10.1007/s001900050236.

[60] S. Bolognani, R. Oboe, and M. Zigliotto, "Sensorless full-digital PMSM drive with EKF estimation of speed and rotor position", *IEEE Transactions on Industrial Electronics*, vol. 46, no. 1, pp. 184–191, Feb. 1999, ISSN: 1557-9948. DOI: 10.1109/41.744410. Accessed: Mar. 4, 2024.

[61] H.-G. Yeh, "Real-time implementation of a narrow-band Kalman filter with a floating-point processor DSP32", en, *IEEE Transactions on Industrial Electronics*, vol. 37, no. 1, pp. 13–18, Feb. 1990, ISSN: 02780046. DOI: 10.1109/41.45838.

[62] L. Ortombina, D. Pasqualotto, F. Tinazzi, and M. Zigliotto, "Automatic Tuning Procedure at Standstill for Extended Kalman Filter in Sensorless Control of Permanent Magnet Synchronous Motors", en, in *2019 IEEE 10th International Symposium on Sensorless Control for Electrical Drives (SLED)*, IEEE, Sep. 2019, pp. 1–6, ISBN: 978-1-72810-766-0. DOI: 10.1109/SLED.2019.8896350.

[63] S. Bolognani, L. Tubiana, and M. Zigliotto, "Extended kalman filter tuning in sensorless PMSM drives", en, *IEEE Trans. on Ind. Applicat.*, vol. 39, no. 6, pp. 1741–1747, Nov. 2003, ISSN: 0093-9994.

[64] K. Shi, T. Chan, Y. Wong, and S. Ho, "Speed estimation of an induction motor drive using an optimized extended Kalman filter", *IEEE Transactions on Industrial Electronics*, vol. 49, no. 1, pp. 124–133, Feb. 2002, ISSN: 1557-9948. DOI: 10.1109/41.982256.

[65] E. Zerdali and M. Barut, "The Comparisons of Optimized Extended Kalman Filters for Speed-Sensorless Control of Induction Motors", en, *IEEE Transactions on Industrial Electronics*, vol. 64, no. 6, pp. 4340–4351, Jun. 2017, ISSN: 0278-0046, 1557-9948. DOI: 10.1109/TIE.2017.2674579.

[66] Y. Laamari, K. Chafaa, and B. Athamena, "Particle swarm optimization of an extended Kalman filter for speed and rotor flux estimation of an induction motor drive", en, *Electrical Engineering*, vol. 97, no. 2, pp. 129–138, Jun. 2015, ISSN: 0948-7921, 1432-0487. DOI: 10.1007/s00202-014-0322-1.

[67] E. Zerdali, "A Comparative Study on Adaptive EKF Observers for State and Parameter Estimation of Induction Motor", en, *IEEE Transactions on Energy Conversion*, vol. 35, no. 3, pp. 1443–1452, Sep. 2020, ISSN: 0885-8969, 1558-0059. DOI: 10.1109/TEC.2020.2979850.

[68] C. T. Fraser and S. Ulrich, "Adaptive extended Kalman filtering strategies for spacecraft formation relative navigation", *Acta Astronautica*, vol. 178, pp. 700–721, Jan. 2021.

[69] R. F. Stengel, *Optimal control and estimation*, eng. Dover Publ, 1994, ISBN: 978-0-486-68200-6.

[70] P. S. Maybeck, *Stochastic models, estimation and control - Part 2*. New York: Academic Press, 1982, vol. 141, ISBN: 978-0-12-480702-0.

[71] Z.-g. Yin, C. Zhao, Y.-R. Zhong, and J. Liu, "Research on Robust Performance of Speed-Sensorless Vector Control for the Induction Motor Using an Interfacing Multiple-Model Extended Kalman Filter", en, *IEEE Transactions on Power Electronics*, vol. 29, no. 6, pp. 3011–3019, Jun. 2014, ISSN: 0885-8993, 1941-0107. DOI: 10.1109/TPEL.2013.2272091.

[72] Z. Yin et al., "A Speed and Flux Observer of Induction Motor Based on Extended Kalman Filter and Markov Chain", en, *IEEE Transactions on Power Electronics*, vol. 32, no. 9, pp. 7096–7117, Sep. 2017, ISSN: 0885-8993, 1941-0107. DOI: 10.1109/TPEL.2016.2623806.

[73] Q. Xia, M. Rao, Y. Ying, and X. Shen, "Adaptive Fading Kalman Filter with an Application", en, *Automatica*, vol. 30, no. 8, pp. 1333–1338, Aug. 1994, ISSN: 00051098. DOI: 10.1016/0005-1098(94)90112-0.

[74] E. Zerdali, R. Yildiz, R. Inan, R. Demir, and M. Barut, "Adaptive Fading Extended Kalman Filter Based Speed-Sensorless Induction Motor Drive", en, in *2018 XIII International Conference on Electrical Machines (ICEM)*, Alexandroupoli: IEEE, Sep. 2018, pp. 1367–1373, ISBN: 978-1-5386-2477-7. DOI: 10.1109/ICELMACH.2018.8507168.

[75] E. Zerdali, R. Yildiz, R. Inan, R. Demir, and M. Barut, "Improved speed and load torque estimations with adaptive fading extended Kalman filter", en, *International Transactions on Electrical Energy Systems*, vol. 31, no. 1, Jan. 2021, ISSN: 2050-7038, 2050-7038. DOI: 10.1002/2050-7038.12684.

[76] R. Mehra, "Approaches to adaptive filtering", en, *IEEE Trans. Automat. Contr.*, vol. 17, no. 5, pp. 693–698, Oct. 1972, ISSN: 0018-9286. DOI: 10.1109/TAC.1972.1100100.

[77] K. Myers and B. Tapley, "Adaptive sequential estimation with unknown noise statistics", en, *IEEE Transactions on Automatic Control*, vol. 21, no. 4, pp. 520–523, Aug. 1976, ISSN: 0018-9286. DOI: 10.1109/TAC.1976.1101260.

[78] E. Zerdali, "Adaptive Extended Kalman Filter for Speed-Sensorless Control of Induction Motors", en, *IEEE Transactions on Energy Conversion*, vol. 34, no. 2, pp. 789–800, Jun. 2019, ISSN: 0885-8969, 1558-0059. DOI: 10.1109/TEC.2018.2866383.

[79] G. Ozkurt and E. Zerdali, "Design and Implementation of Hybrid Adaptive Extended Kalman Filter for State Estimation of Induction Motor", en, *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–12, 2022, ISSN: 0018-9456, 1557-9662. DOI: 10.1109/TIM.2022.3144729.

[80] J. Han, Q. Song, and Y. He, "Adaptive Unscented Kalman Filter and Its Applications in Nonlinear Control", in *Kalman Filter Recent Advances and Applications*, V. M. and A. Pigazo, Eds., InTech, Apr. 2009, ISBN: 978-953-307-000-1. DOI: 10.5772/6799. Accessed: Jul. 23, 2024.

[81] B. Ritter, "Nonlinear State Estimation and Noise Adaptive Kalman Filter Design for Wind Turbines", Ph.D. dissertation, 2020.

[82] Juntong Qi, Jianda Han, and Zhenwei Wu, "Rotorcraft UAV actuator failure estimation with KF-based adaptive UKF algorithm", in *2008 American Control Conference*, Seattle, WA: IEEE, Jun. 2008, pp. 1618–1623, ISBN: 978-1-4244-2078-0. DOI: 10.1109/ACC.2008.4586723. Accessed: Nov. 28, 2023.

[83] M. Mohamed and I. Kar, "Aerodynamic parameter estimation using adaptive unscented Kalman filter", *Aircraft Engineering and Aerospace Technology: An International Journal*, vol. 85, Jun. 2013. DOI: 10.1108/AEAT-Mar-2011-0038.

[84] S. Morimoto, M. Sanada, and Y. Takeda, "High-performance current-sensorless drive for pmsm and synrm with only low-resolution position sensor", *IEEE Transactions on Industry Applications*, vol. 39, no. 3, pp. 792–801, 2003. DOI: 10.1109/TIA.2003.811782.

[85] R. Morales-Caporal and M. Pacas, "Encoderless predictive direct torque control for synchronous reluctance machines at very low and zero speed", *IEEE Transactions on Industrial Electronics*, vol. 55, no. 12, pp. 4408–4416, 2008. DOI: 10.1109/TIE.2008.2007025.

[86] S. Ichikawa, M. Tomita, S. Doki, and S. Okuma, "Sensorless control of synchronous reluctance motors based on extended emf models considering magnetic saturation with online parameter identification", *IEEE Transactions on Industry Applications*, vol. 42, no. 5, pp. 1264–1274, 2006. DOI: 10.1109/TIA.2006.880848.

[87] C. Li, G. Wang, G. Zhang, D. Xu, and D. Xiao, "Saliency-based sensorless control for synrm drives with suppression of position estimation error", *IEEE Transactions on Industrial Electronics*, vol. 66, no. 8, pp. 5839–5849, 2019. DOI: 10.1109/TIE.2018.2874585.

[88] E. Armando, R. I. Bojoi, P. Guglielmi, G. Pellegrino, and M. Pastorelli, "Experimental identification of the magnetic model of synchronous machines", *IEEE Transactions on Industry Applications*, vol. 49, no. 5, pp. 2116–2125, 2013. DOI: 10.1109/TIA.2013.2258876.

[89] R. Hermann and A. Krener, "Nonlinear controllability and observability", *IEEE Transactions on Automatic Control*, vol. 22, no. 5, pp. 728–740, Oct. 1977, ISSN: 1558-2523. DOI: 10.1109/TAC.1977.1101601. Accessed: Nov. 18, 2024.

[90] R. Caro, C. A. Silva, R. Perez, and J. I. Yuz, "Sensorless control of a SynRM for the whole speed range based on a nonlinear observability analysis", in *2017 IEEE International Conference on Industrial Technology (ICIT)*, Toronto, ON: IEEE, Mar. 2017, pp. 336–341, ISBN: 978-1-5090-5320-9. DOI: 10.1109/ICIT.2017.7913253. Accessed: Jan. 8, 2024.

[91] M. Koteich, A. Maloum, G. Duc, and G. Sandou, "Observability analysis of sensorless synchronous machine drives", in *2015 European Control Conference (ECC)*, Jul. 2015, pp. 3560–3565. DOI: 10.1109/ECC.2015.7331085. Accessed: Nov. 18, 2024.

[92] P. Vaclavek, P. Blaha, and I. Herman, "AC Drive Observability Analysis", *IEEE Transactions on Industrial Electronics*, vol. 60, no. 8, pp. 3047–3059, Aug. 2013, ISSN: 1557-9948. DOI: 10.1109/TIE.2012.2203775. Accessed: Jan. 8, 2024.

[93] S. F. Schmidt, "Application of State-Space Methods to Navigation Problems", in *Advances in Control Systems*, C. T. Leondes, Ed., vol. 3, Elsevier, Jan. 1966, pp. 293–340. DOI: 10.1016/B978-1-4831-6716-9.50011-4. Accessed: Nov. 20, 2024.

[94] E. Wan and R. Van Der Merwe, "The unscented Kalman filter for nonlinear estimation", in *Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium (Cat. No.00EX373)*, 2000, pp. 153–158. DOI: 10.1109/ASSPCC.2000.882463.

[95] R. Antonello, L. Ortombina, F. Tinazzi, and M. Zigliotto, "Advanced current control of synchronous reluctance motors", in *2017 IEEE 12th International Conference on Power Electronics and Drive Systems (PEDS)*, 2017, pp. 1, 037–1, 042. DOI: 10.1109/PEDS.2017.8289150.

[96] A. Brosch, F. Tinazzi, O. Wallscheid, M. Zigliotto, and J. Böcker, "Finite set sensorless control with minimum a priori knowledge and tuning effort for interior permanent-magnet synchronous motors", *IEEE Transactions on Power Electronics*, vol. 38, no. 10, pp. 12 508–12 519, 2023. DOI: 10.1109/TPEL.2023.3294557.

[97] S. Bolognani, L. Ortombina, F. Tinazzi, and M. Zigliotto, "Model sensitivity of fundamental-frequency-based position estimators for sensorless PM and reluctance synchronous motor drives", *IEEE Transactions on Industrial Electronics*, vol. 65, no. 1, pp. 77–85, 2018. DOI: 10.1109/TIE.2017.2716902.

[98] P. S. Maybeck, *Stochastic models, estimation and control - Part 3* (Mathematics in science and engineering 141), en. New York: Academic Press, 1982, ISBN: 978-0-12-480701-3 978-0-12-480702-0 978-0-12-480703-7.

[99] P. D. Abramson, "Simultaneous estimation of the state and noise statistics in linear dynamical systems", NTRS Author Affiliations: NASA Electronics Research Center NTRS Document ID: 19700013435 NTRS Research Center: Legacy CDMS (CDMS), Ph.D. dissertation, Mar. 1970.

[100] K. Reif, S. Gunther, E. Yaz, and R. Unbehauen, "Stochastic stability of the discrete-time extended Kalman filter", *IEEE Transactions on Automatic Control*, vol. 44, no. 4, pp. 714–728, Apr. 1999, ISSN: 1558-2523. DOI: 10.1109/9.754809. Accessed: Aug. 26, 2024.

[101] F. Alonge, T. Cangemi, F. D'Ippolito, A. Fagiolini, and A. Sferlazza, "Convergence Analysis of Extended Kalman Filter for Sensorless Control of Induction Motor", *IEEE Transactions on Industrial Electronics*, vol. 62, no. 4, pp. 2341–2352, Apr. 2015, ISSN: 1557-9948. DOI: 10.1109/TIE.2014.2355133. Accessed: Aug. 27, 2024.

[102] Q. Song, Z. Jiang, and J. Han, "Noise Covariance Identification Based Adaptive UKF with Application to Mobile Robot Systems", in *Proceedings 2007 IEEE International Conference on Robotics and Automation*, Rome, Italy: IEEE, Apr. 2007, pp. 4164–4169, ISBN: 978-1-4244-0602-9 978-1-4244-0601-2. DOI: 10.1109/ROBOT.2007.364119. Accessed: Jul. 23, 2024.

[103] Q. Song, J. Qi, and J. Han, "An Adaptive UKF Algorithm and Its Application in Mobile Robot Control", in *2006 IEEE International Conference on Robotics and Biomimetics*, Kunming, China: IEEE, 2006, pp. 1117–1122, ISBN: 978-1-4244-0570-1. DOI: 10.1109/ROBIO.2006.340085. Accessed: Nov. 28, 2023.

[104] Q. Song and Y. He, "Adaptive unscented Kalman filter for estimation of modelling errors for helicopter", in *2009 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, Guilin, China: IEEE, Dec. 2009, pp. 2463–2467, ISBN: 978-1-4244-4774-9. DOI: 10.1109/ROBIO.2009.5420406. Accessed: Nov. 28, 2023.

[105] B. Ritter, E. Mora, T. Schlicht, A. Schild, and U. Konigorski, "Adaptive Sigma-Point Kalman Filtering for Wind Turbine State and Process Noise Estimation", *Journal of Physics: Conference Series*, vol. 1037, p. 032 003, Jun. 2018, ISSN: 1742-6588, 1742-6596. DOI: 10.1088/1742-6596/1037/3/032003. Accessed: Jul. 23, 2024.

[106] K. B. Petersen and M. S. Pedersen, "Matrix - Cookbook", Tech. Rep.

Publications

- D. Pasqualotto, S. Rigon and M. Zigliotto, "Sensorless Speed Control of Synchronous Reluctance Motor Drives Based on Extended Kalman Filter and Neural Magnetic Model," in IEEE Transactions on Industrial Electronics, vol. 70, no. 2, pp. 1321-1330, Feb. 2023, doi: 10.1109/TIE.2022.3159962.
- R. Antonello, S. Rigon, F. Tinazzi and M. Zigliotto, "Capacitorless Synchronous Reluctance Motor Drives for Pump Applications With Resonance Damping and Line Current Harmonics Reduction," in IEEE Journal of Emerging and Selected Topics in Power Electronics, vol. 11, no. 5, pp. 5279-5289, Oct. 2023, doi: 10.1109/JESTPE.2023.3310530.
- S. Rigon, B. Haus, P. Mercorelli and M. Zigliotto, "Comparison Between UKF and EKF in Sensorless Synchronous Reluctance Motor Drives," in IEEE Open Journal of Power Electronics, vol. 5, pp. 1562-1572, 2024, doi: 10.1109/OJPEL.2024.3469533.
- S. Rigon, R. Antonello, P. Mercorelli and M. Zigliotto, "Sensorless Control of AC Motor Drives with Adaptive Extended Kalman Filter," 2024 IEEE Energy Conversion Congress and Exposition (ECCE), Phoenix, AZ, USA, 2024, pp. 5991-5998, doi: 10.1109/ECCE55643.2024.10860865.
- S. Rigon, B. Haus, P. Mercorelli and M. Zigliotto, "Automatic Tuning of Extended Kalman Filter in Synchronous Reluctance Motor Drives with a Master-Slave Configuration," in IEEE Open Journal of Power Electronics, doi: 10.1109/OJPEL.2025.3546752.

The PhD scholarship was co-funded with resources from the PON REACT-EU program.