POLITECNICO
MILANO 1863

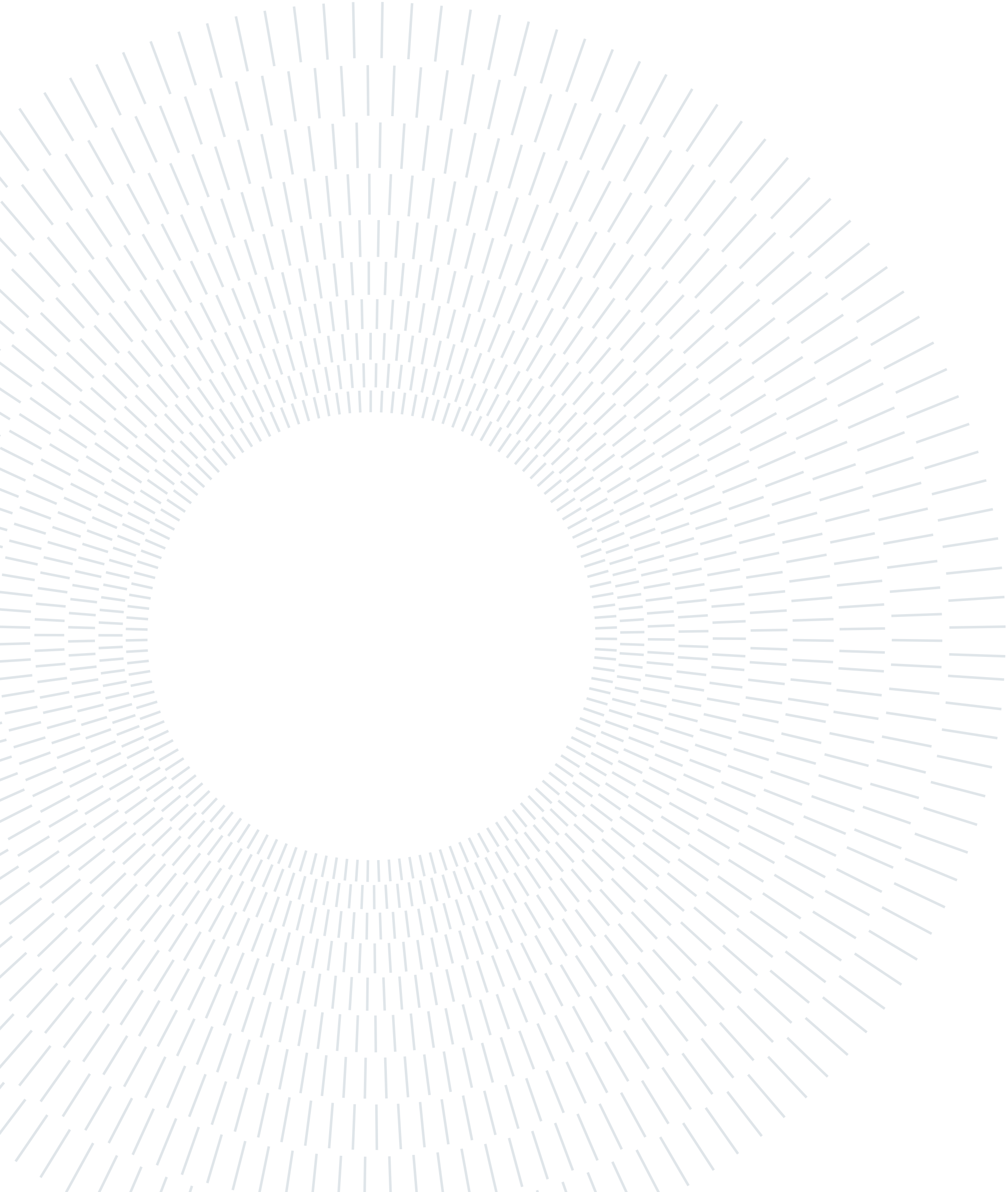
SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

Continuous-time system identification with functional basis expansions

Tesi di Laurea Magistrale in
Space Engineering - Ingegneria Spaziale

Author: **José Antonio Rebollo**

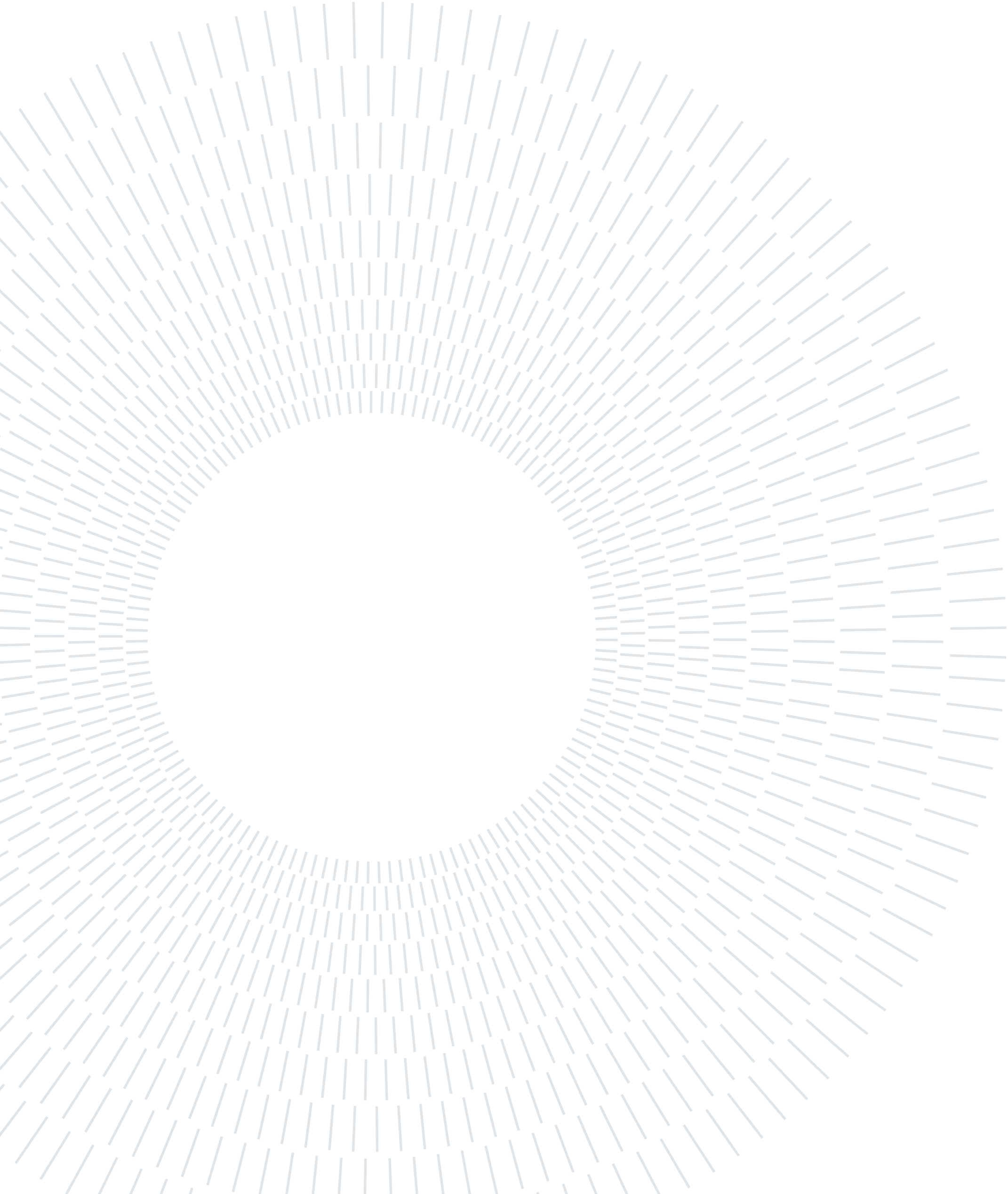
Student ID: 242930
Advisor: Prof. Marco Lovera
Co-advisor: Enrico Barbiero
Academic Year: 2024-25



Abstract

System identification is a fundamental problem in engineering, especially for modeling complex dynamical systems from experimental data. While discrete-time system identification methods are well-developed, the identification of continuous-time systems remains a challenge due to issues like noise, non-uniform sampling, and the need for numerical differentiation. This thesis explores a novel approach to continuous-time system identification using functional basis expansions. Specifically, it leverages orthonormal Schauder bases, such as Hermite and Laguerre functions, to rigorously construct a transformation from continuous-time to discrete-time systems. The methodology extends the well-established Predictor-Based Subspace Identification (PBSID) framework by incorporating Hilbert space theory and spectral analysis techniques. A new method, the Hermite Domain PBSID (HD-PBSID), is introduced, drawing inspiration from quantum mechanics and the quantum harmonic oscillator representation. Theoretical derivations are supported by numerical simulations, comparing the proposed algorithm to the existing Continuous Time PBSID (CT-PBSID), and experimental validation using quadrotor flight data. Results demonstrate the advantages of the proposed approach in achieving accurate system representations while maintaining computational efficiency and robustness to noise. These findings contribute to a deeper understanding of functional basis methods in system identification and open new possibilities for high-fidelity modeling of continuous-time dynamical systems.

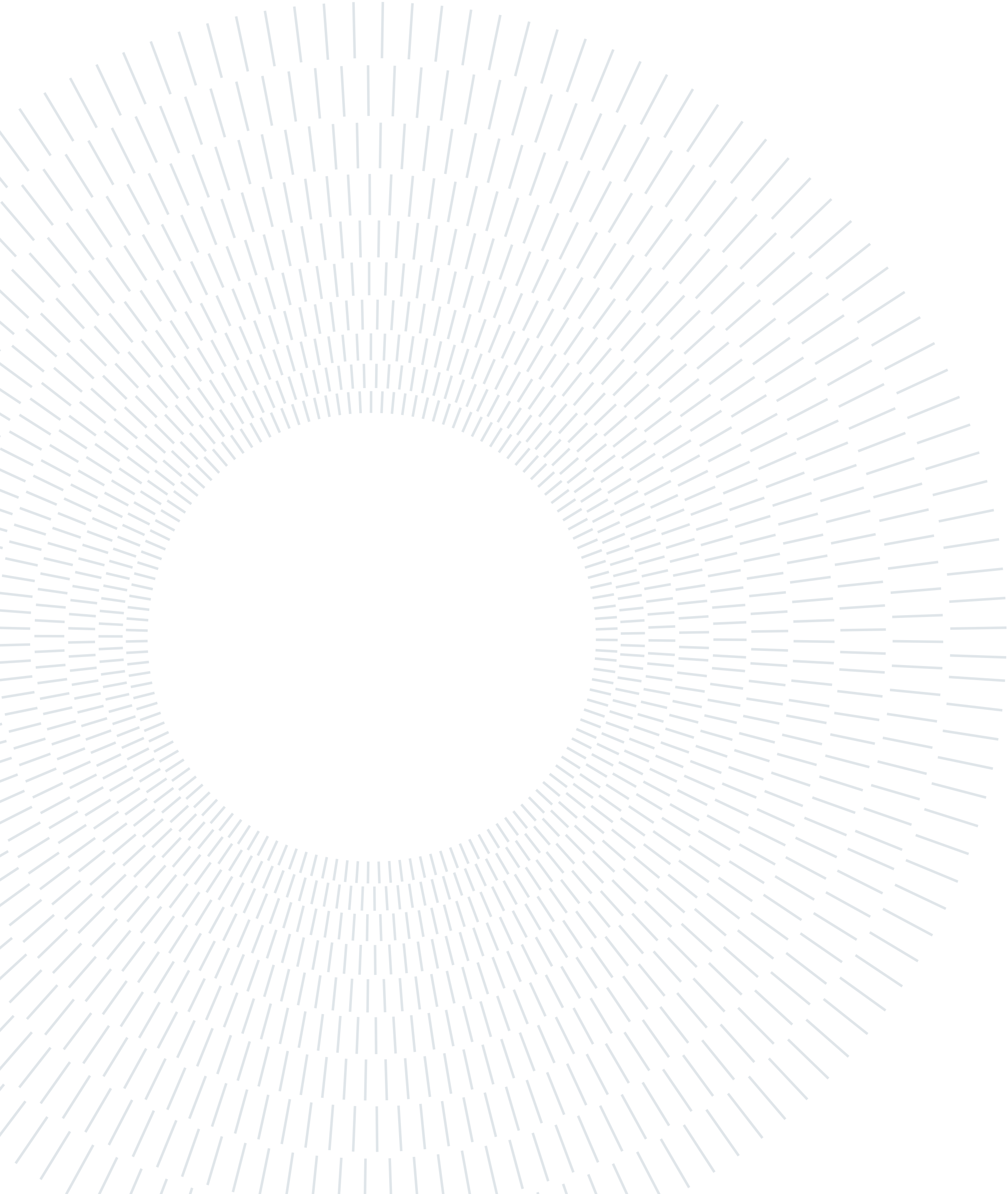
Keywords: Continuous-time system identification, Hilbert spaces, Hermite functions, Laguerre functions, subspace identification methods, quadrotor modeling.



Abstract in lingua italiana

L'identificazione dei sistemi rappresenta un problema fondamentale in ingegneria, soprattutto per la modellazione di sistemi dinamici complessi a partire da dati sperimentali. Mentre i metodi per l'identificazione di sistemi in tempo discreto sono ormai ben sviluppati, l'identificazione di sistemi in tempo continuo continua a presentare sfide, dovute a problematiche quali il rumore, il campionamento non uniforme e la necessità di effettuare differenziazioni numeriche. Questa tesi propone un approccio innovativo all'identificazione di sistemi in tempo continuo basato sull'espansione in basi funzionali. In particolare, si fa ricorso a basi di Schauder ortonormali, come quelle costituite dalle funzioni di Hermite e di Laguerre, per costruire in modo rigoroso una trasformazione che converte i sistemi in tempo continuo in sistemi in tempo discreto. La metodologia sviluppata amplia il consolidato framework di Predictor Based Subspace IDentification (PBSID), integrando la teoria degli spazi di Hilbert e tecniche di analisi spettrale. In questo contesto viene introdotto un nuovo metodo, il PBSID in dominio di Hermite (HD-PBSID), ispirato alla meccanica quantistica e alla rappresentazione dell'oscillatore armonico quantistico. Le derivazioni teoriche sono supportate da simulazioni numeriche, che mettono a confronto l'algoritmo proposto con il Continuous Time PBSID (CT-PBSID) attualmente in uso, e da una validazione sperimentale basata su dati di volo di un quadrirotore. I risultati evidenziano i vantaggi dell'approccio proposto, capace di ottenere rappresentazioni accurate dei sistemi mantenendo al contempo elevata efficienza computazionale e robustezza rispetto al rumore. Questi risultati contribuiscono a una comprensione più approfondita dei metodi basati su basi funzionali nell'identificazione dei sistemi e aprono nuove prospettive per la modellazione ad alta fedeltà di sistemi dinamici in tempo continuo.

Parole chiave: Identificazione dei sistemi a tempo continuo, spazi di Hilbert, funzioni di Hermite, funzioni di Laguerre, metodi di identificazione subspaziale, modellazione di quadrirotori.



Contents

Abstract **i**

Abstract in lingua italiana **iii**

Contents **v**

1 Introduction **1**
- 1.1 Review on system identification for rotorcraft and linear systems 1
- 1.2 Scope of this work 3
- 1.3 General structure 4

2 Preliminaries and Theoretical Background **7**
- 2.1 Fundamentals on linear systems 7
  - 2.1.1 Continuous-time linear systems 8
  - 2.1.2 Discrete-time linear systems 8
  - 2.1.3 Process and measurement noise 8
  - 2.1.4 Representation systems and non-uniqueness of solutions 9
- 2.2 System identification and the PBSID method 10
- 2.3 Hilbert spaces 13
  - 2.3.1 Introduction 13
  - 2.3.2 The $\mathcal{L}^2$ space 14
  - 2.3.3 The ℓ^2 space 14
  - 2.3.4 Projections on Hilbert spaces 15
  - 2.3.5 Orthonormal bases 16
  - 2.3.6 Spectrum of a linear operator 17
  - 2.3.7 Self-adjoint operators 18

3 CT-DT transformation via basis expansions **19**
- 3.1 Linear systems as operators on Hilbert spaces 19

3.2 Schauder bases of $\mathcal{L}^2$ 20
  3.2.1 Laguerre functions 21
  3.2.2 Hermite functions 22
  3.2.3 Construction of CT-DT transformations 23
3.3 Revision of the CT-PBSID 24
3.4 The quantum oscillator approach to CT-DT transformation 27
  3.4.1 Theoretical framework 28
  3.4.2 Numerical examples 32
3.5 The HD-PBSID method 35
  3.5.1 Effective support and scaling 36
  3.5.2 Modified differentiation operator 37
  3.5.3 Integrator form 39
  3.5.4 Final algorithm 40
  3.5.5 Experiment design 42
  3.5.6 Effective algorithm to obtain the Hermite basis 43

4 Generalization through Sturm-Liouville theory 45
4.1 Sturm-Liouville problems and orthogonal bases 45
4.2 Recasting second-order eigenvalue problems 47
4.3 Rodrigues' Formula and Operator Derivation 49
4.4 General Derivative Operator for Orthogonal Functions 52
4.5 Particular Cases 54
  4.5.1 Hermite Functions 54
  4.5.2 Laguerre Functions 55
  4.5.3 Legendre Functions 56
  4.5.4 General approach 57
4.6 Classification of solutions 57

5 Simulation results and analysis 59
5.1 Computational verification 59
5.2 Comparison with the CT-PBSID 62

6 Experimental implementation 73
6.1 Experimental setup 73
  6.1.1 Drone dynamics 73
  6.1.2 System control and excitation 76
  6.1.3 Experiments procedure 77
6.2 Identification and results 78

6.2.1 Identification setup 78
6.2.2 Identification results 78

7 Conclusions and future work **85**
7.1 Conclusions 85
7.2 Future Work 86

Bibliography **87**

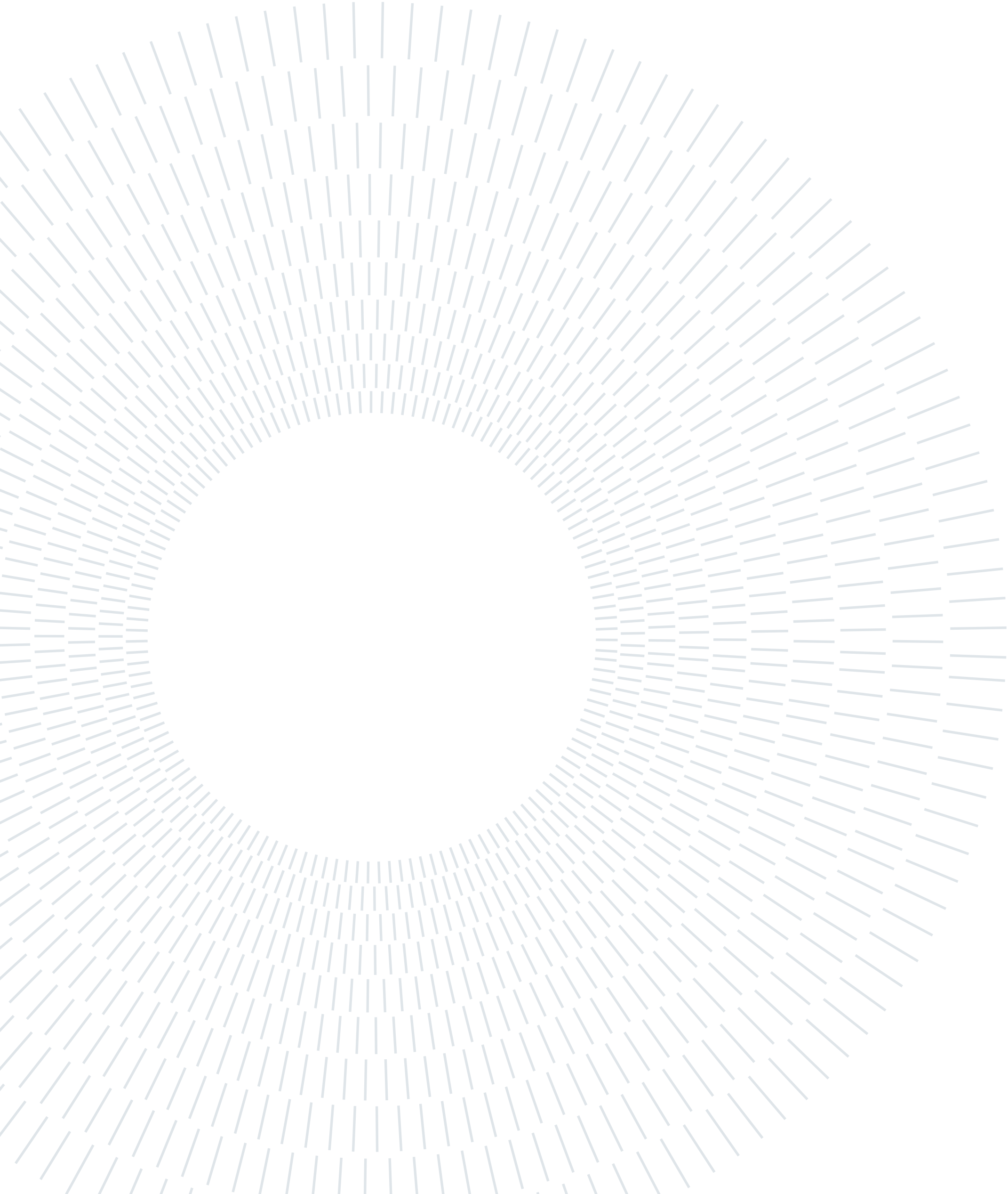
List of Figures **89**

List of Tables **91**

List of Symbols **93**

Acknowledgements **95**

Agradecimientos **97**



1 | Introduction

1.1. Review on system identification for rotorcraft and linear systems

The field of system identification is a well-established area of research, seeking to build mathematical models from experimental data. It has seen major advances over the past decades, driven by the increasing demand for accurate and reliable models in various engineering disciplines. Particularly, in rotorcraft flight dynamics, a detailed knowledge of linear and non-linear dynamics is critical for safe and effective operation, control design, and performance analysis. This presents a series of challenges, as the collected data are often noisy and of limited duration or bandwidth, which makes a robust system identification framework a necessary tool to extract meaningful insights from limited experiments. As highlighted by Bergamasco and Lovera [1], a combination of techniques is often needed to tackle the challenging issues that arise in rotorcraft identification.

The central challenge in system identification is to build accurate models from real experimental data. Furthermore, the measurements are sometimes collected in closed-loop operation, which further complicates the problem. This introduces correlations between the input and noise signals, potentially leading to inconsistent and biased estimates if not handled appropriately [2, 3]. To overcome these challenges, a wide variety of techniques have been developed over the years, ranging from frequency-domain approaches to time-domain methods, as detailed in [4, 5]. Frequency-domain techniques are widely employed for their ability to provide non-parametric estimates of the system's transfer function. As explained by [1], these methods, based on frequency sweeps, result in accurate frequency response functions, which are used as benchmarks for comparison with identified parametric models. While these approaches are useful for describing the system in terms of frequency-dependent behavior, they may not provide direct information on the state-space representations commonly used in control design. Moreover, generating reliable estimates using frequency sweeps is time-consuming and requires carefully designed experiments.

Time-domain methods, on the other hand, use transient inputs, such as multi-step and impulse-like excitation signals. These methods typically focus on identifying state-space representations, which can be directly used for control design purposes. Subspace identification techniques have become prominent in this context due to their ability to estimate MIMO (Multiple Input Multiple Output) models without resorting to iterative optimization procedures. This class of methods relies on numerical linear algebra tools to obtain models directly from data [6, 7]. In particular, the survey by Van der Veen, van Wingerden, Lovera, and Verhaegen [8] provides a comprehensive overview of several subspace identification methods that were proposed to tackle the closed loop identification challenge. For instance, the Multivariable Output-Error State-space (MOESP) method estimates system dynamics by projecting the input-output data onto the orthogonal complement of the input and innovations subspaces. This projection is achieved through an RQ factorization of the input-output block Hankel matrices, followed by a singular value decomposition (SVD) to determine the system order and extract the state-space matrices. The N4SID (Numerical Subspace State-Space Identification) method uses a different approach, directly projecting the input-output data onto subspaces associated with non-steady-state Kalman filter banks to compute the state sequences. These state sequences, representing optimal predictions of the system states, are then used to construct the system matrices. One of the most widely used subspace identification techniques in closed-loop applications is the Predictor Based Subspace IDentification (PBSID) algorithm and its variants [9, 10]. This method, in contrast, combines subspace projections with prediction error minimization. It identifies the system by first estimating prediction errors using oblique projections of the data and then solving a least-squares problem to determine the state-space matrices. This hybrid approach improves model accuracy, especially in closed-loop settings, by explicitly accounting for the feedback structure during the estimation process. Other approaches include using specific excitations or instrumental variables (IV) [11, 12], also detailed in [8], or applying methods that require controller information.

While the vast majority of identification methods focus on discrete-time systems, there is a need for directly identifying continuous-time models, especially when dealing with stiff systems or non-uniformly sampled data [13, 14]. Several approaches for continuous-time identification have been proposed, as detailed in [15]. These often rely on basis functions to avoid the numerical issues related to high order derivative computation. This includes transformations based on Laguerre basis expansions [16] or generalized orthonormal bases, as proposed by [17]. In addition, some methods directly operate in the frequency domain [18], which can be advantageous when an explicit frequency response function is desired.

One major issue that arises is the choice of the model structure to capture the true system behaviour with a reasonable number of parameters. A practical approach, proposed in [1], is to integrate the information from time and frequency domain, to create a unified methodology that provides robust model estimates that can be used with well established control theory results. This implies combining the ability of frequency response methods to provide accurate estimates of the systems response in the desired bandwidth and of time domain models to provide reliable state-space models suitable for control design. This is achieved through a time/frequency approach that combines the most valuable advantages of both techniques, exploiting non-parametric estimates to benchmark parametric identification methods.

Other techniques exist which are capable of achieving similar results to those of the subspace methods presented so far, such as the well established prediction-error methods, or a series of model reduction techniques that can be used once a black-box identification has been performed [7, 19]. These are not discussed in detail in this work for the sake of brevity.

Although the PBSID technique offers a powerful approach to system identification, without the need of an initial estimate nor an iterative approach, and while providing good statistical guarantees, the extension to continuous time systems is an open area of research. The analysis performed in [15] led to two different state-of-the-art strategies: the application of Laguerre filters and the projection into a basis of Laguerre functions, with the posterior providing better results. Interestingly, these two frameworks were found by tackling the problem from two different approaches, namely, all-pass filtering [20] and system lifting [17], respectively. This apparent discrepancy in methodologies, even though both methods ultimately leverage projections onto the Laguerre basis, is an interesting matter of research. Furthermore, it remains unexplored whether other families of functions can be used to construct similar identification approaches. Finally, if possible, a unified framework including the families of functions or filters which can be used for the system discretization would be an interesting tool for both a theoretical understanding and a practical improvement.

1.2. Scope of this work

The main scope of this work is to construct a novel framework for the transformation of linear continuous-time (CT) systems into linear discrete-time (DT) systems by using the properties of orthonormal bases within a functional analysis framework. This transformation is achieved by projecting the system's signals and operators onto the basis and by exploiting the existing relationships between the original system and the transformed one. In doing so, a mathematical framework that allows building a direct link between the original CT system and a corresponding DT system is developed. A particularization for Hermite functions is explored, leading to a novel method denoted as Hermite Domain PBSID (HD-PBSID). This method, thus, represents a valuable contribution to the field of system identification and control.

Furthermore, the proposed framework is not specific to any particular basis. A detailed analysis of the underlying mathematical structure is provided, leveraging results from Sturm-Liouville theory to understand which orthonormal bases can be obtained by an eigenproblem defined by second order linear operators. Finally, the practical aspects of the transformation are explored, by analyzing practical cases through both simulation of realistic scenarios and experimental implementation for the identification of a quadrotor. This allows to validate the theoretical results obtained in this work and verify the performance of the proposed methods for system identification.

This work aims to answer three main research questions:

- What is the relation of the Laguerre filtering and projection methods for continuous-time system identification?
- To what extent can the prior methods be extended to other families of functions?
- What is the common framework for the continuous-time system identification strategies given different function families?

1.3. General structure

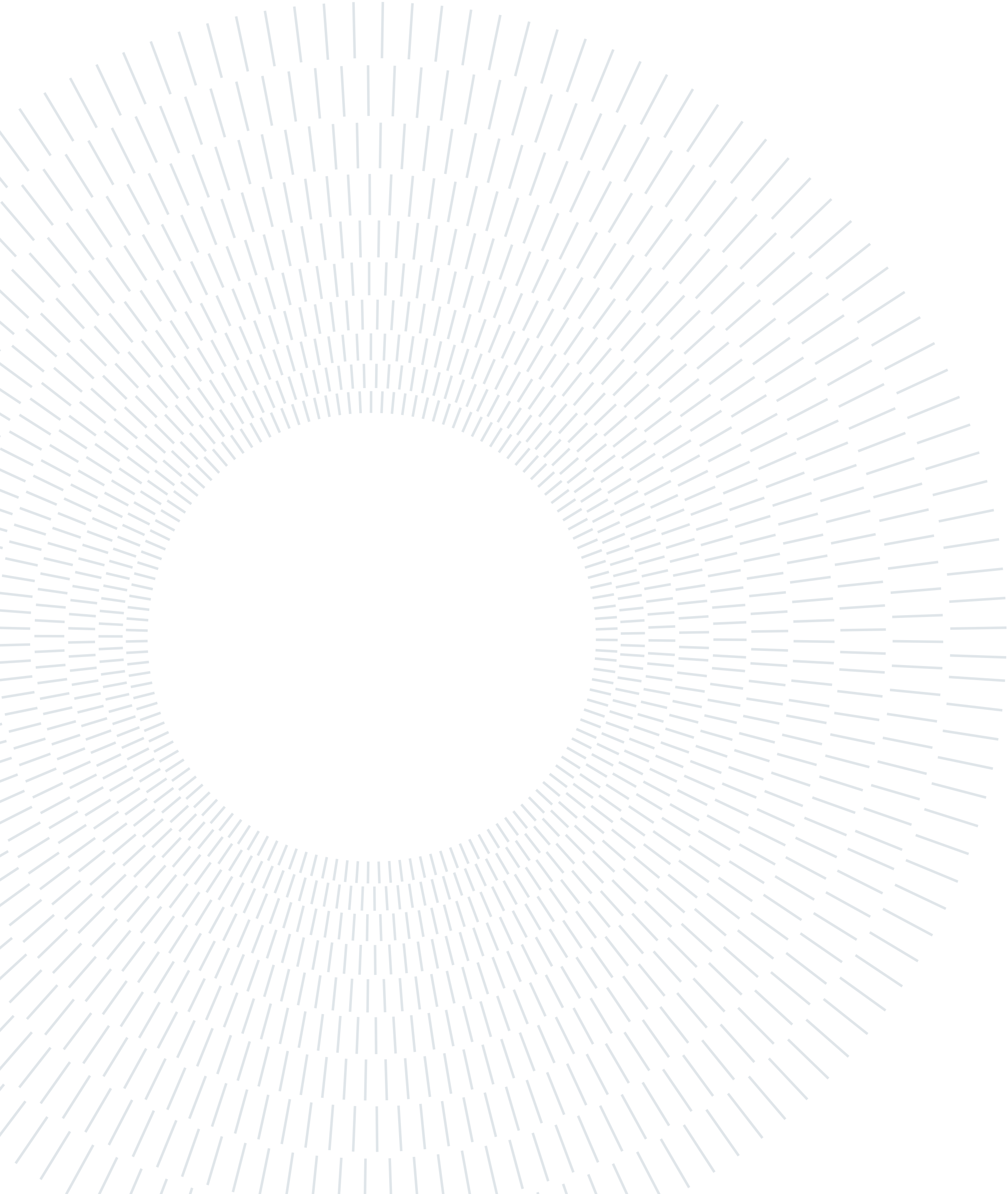
The document is organized as follows. Chapter 2 revisits all the mathematical tools that will be needed for this work, focusing on the properties of linear continuous-time and discrete-time systems, on the core concepts of system identification, and finally, on the fundamental results on functional analysis and Hilbert spaces. This chapter is used to provide the reader with a background needed to fully understand the theoretical contributions of this thesis.

Chapter 3 lays out the first theoretical framework for this work. In this chapter, the core concepts of the transformation of systems by means of linear bases are presented. A description on how signals and operators are transformed to a corresponding orthogonal basis is provided, leveraging on the properties of the inner product. Next, the concepts on functional analysis previously described are applied to the CT-DT transformation. This methodology is first applied to the existing CT-PBSID framework, evaluating its consistency by using the properties of Laguerre bases. Additionally, a novel CT-DT transformation is derived using the Hermite basis, noting that their properties correspond to the ones of the quantum harmonic oscillator. By exploiting this framework, the time-derivative operator is written as a combination of the creation and annihilation operators, well known in Quantum Mechanics. Finally, this chapter provides a clear description of the novel algorithm implementation in a practical scenario.

Chapter 4 generalizes the results achieved by exploring all the available orthonormal bases for linear systems, by leveraging Sturm-Liouville theory. In this chapter, it will be shown that only a few bases arise from this theory, which are suitable for linear system analysis. This mostly theoretical chapter gives a common framework to Laguerre-based and Hermite-based CT-DT strategies.

Chapters 5 and 6 are devoted to the practical aspects of the proposed method. In the former, the proposed framework is validated by applying it to the transformation of systems, and by showcasing its performance in a few representative numerical cases. In the latter, the proposed method is tested in a practical scenario by performing the identification of a quadrotor from experimental measurements.

Finally, Chapter 7 revisits the contributions of this work and sets the basis for future developments.



2 | Preliminaries and Theoretical Background

In this Chapter, the mathematical preliminaries and theoretical background used throughout this work are introduced.

2.1. Fundamentals on linear systems

A dynamical system S is a map transforming an input or control signal u to an output or measurement y,

$$\begin{aligned} S &: U \to Y \\ u &\to y, \end{aligned} \tag{2.1}$$

where U and Y are the sets in which u and y are defined, respectively. In this work, continuous-time (CT) and discrete-time (DT) multidimensional dynamical systems are considered. For CT systems, u and y are vector functions of time,

$$\begin{aligned} u &: \mathbb{T} \to \mathbb{R}^{n_u} \\ y &: \mathbb{T} \to \mathbb{R}^{n_y}. \end{aligned} \tag{2.2}$$

where $\mathbb{T}$ is the time interval in which the system is defined, which may be as $[t_0, t_f]$, $\mathbb{R}$ or $\mathbb{R}^+$. For DT systems, u and y are sequences of vectors,

$$\begin{aligned} u &\in (\mathbb{R}^{n_u})^{\mathbb{M}} \\ y &\in (\mathbb{R}^{n_y})^{\mathbb{M}}, \end{aligned} \tag{2.3}$$

where $\mathbb{M}$ is the sequence cardinal space, which may be as $\mathbb{N} \cap [0, N_{max}]$, $\mathbb{Z}$ or $\mathbb{N}$. Additional conditions might be imposed to the sets U, Y for the systems to be physically feasible. In general, dynamical systems are constructed by the definition of an inner state $x \in X$, which is propagated as given by the input. The output then maps the state and input. Next, the particular structure of CT and DT linear systems is described.

2.1.1. Continuous-time linear systems

A CT linear system is defined by a linear Ordinary Differential Equation (ODE) plus a linear measurement map in the form

$$\begin{aligned}\frac{d}{dt}x &= Ax + Bu \\ y &= Cx + Du.\end{aligned} \tag{2.4}$$

Thus, the system S is completely characterized by matrices $A \in \mathbb{R}^{n_x \times n_x}$, $B \in \mathbb{R}^{n_x \times n_u}$, $C \in \mathbb{R}^{n_y \times n_x}$ and $D \in \mathbb{R}^{n_y \times n_u}$. The uniqueness of signal $x(t)$ and hence $y(t)$ given $u(t)$ requires an initial condition $x(t_0)$ or initial state to be given.

2.1.2. Discrete-time linear systems

A DT linear system is defined by a linear difference equation plus a linear measurement map in the form

$$\begin{aligned}x_{k+1} &= Ax_k + Bu_k \\ y_k &= Cx_k + Du_k.\end{aligned} \tag{2.5}$$

Thus, the system S is completely characterized by matrices $A \in \mathbb{R}^{n_x \times n_x}$, $B \in \mathbb{R}^{n_x \times n_u}$, $C \in \mathbb{R}^{n_y \times n_x}$ and $D \in \mathbb{R}^{n_y \times n_u}$. The uniqueness of the sequence x_k and hence y_k given u_k requires an initial sample x_{k_0} or initial state to be given.

2.1.3. Process and measurement noise

In realistic scenarios, both propagation and measurement equations are corrupted by noise, which is defined in a slightly different manner depending on the structure of the system. For instance, for CT systems, noise is considered by means of a Wiener process and the corresponding Stochastic Differential Equation (SDE),

$$\begin{aligned}dx &= Axdt + Budt + dw \\ dz &= Cxdt + Dudt + dv \\ ydt &= dz.\end{aligned} \tag{2.6}$$

Signals dw and dv are sampled from a multivariate normal distribution of mean zero and covariance

$$\mathrm{Cov}\begin{pmatrix} dw \\ dv \end{pmatrix} = \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} dt. \tag{2.7}$$

Hence, the process noise dw introduces a random walk in the dynamics while the measurement noise dv directly corrupts the measurement with gaussian noise. Typically dw and dv are assumed as uncorrelated, leading to

$$\begin{aligned} dx &= Axdt + Budt + dw \\ y &= Cx + Du + v, \end{aligned} \tag{2.8}$$

with dw and v sampled from normal distributions of mean zero and covariances Qdt and R, respectively.

For DT systems, the addition of noise is analogous, such that

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + w_k \\ y_k &= Cx_k + Du_k + v_k, \end{aligned} \tag{2.9}$$

with

$$\operatorname{Cov}\begin{pmatrix} w_k \\ v_k \end{pmatrix} = \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix}. \tag{2.10}$$

2.1.4. Representation systems and non-uniqueness of solutions

Let a DT linear system given by (2.5). Note that the following analysis equally applies to CT systems. Let the state transformation $Tx'_k = x_k$, for T an invertible square matrix. Thus,

$$\begin{aligned} Tx'_{k+1} &= ATx'_k + Bu_k \\ y_k &= CTx'_k + Du_k. \end{aligned} \tag{2.11}$$

As T is invertible, the following system is equivalent to (2.5),

$$\begin{aligned} x'_{k+1} &= A'x'_k + B'u_k \\ y_k &= C'x'_k + Du_k, \end{aligned} \tag{2.12}$$

for $A' = T^{-1}AT$, $B' = T^{-1}B$ and $C' = CT$. Therefore, given a set of inputs and measurements, any state and system matrices that satisfy the described relations is a correct representation of the original system. Consequently, even for completely accurate signals u_k and y_k, an infinite set of matrices can be obtained, that is, the System Identification problem has a non-unique solution. The desired representation might be obtained by imposing additional constraints based on known physical properties of x. Alternatively, one of the possible representations might be obtained as a solution.

2.2. System identification and the PBSID method

The System Identification problem aims to characterize S given u and y. The presence of noise requires a statistical approach. There exists a broad set of techniques to identify S from the available data. Some of them parameterize the matrices and minimize a cost function or maximize a likelihood [19, 21]. Other techniques operate on the frequency domain, fitting the estimated transfer function from measurements with a model of a given structure [22]. A particularly interesting set of methods are the so called Subspace Identification (SID), which leverage the properties of linear systems and tackle the identification algebraically. These methods are well suited for multidimensional systems, do not require an initial guess for the system nor a parametrization of its matrices, and have good properties regarding the effect of noise.

For DT systems, the Predictor-Based Subspace Identification (PBSID) method is a well known approach that provides excellent results in practical scenarios. It leverages the innovation form of the system to identify. Given the structure in (2.9), let the predictor

$$\hat{x}_{k+1} = A\hat{x}_k + Bu_k + Ke_k, \tag{2.13}$$

where K is the predictor's Kalman gain [23] and the innovation e_k is defined as

$$e_k = y_k - C\hat{x}_k - Du_k. \tag{2.14}$$

Assuming that the Kalman gain is in steady state, the innovation form of the predictor's dynamics is given by

$$\begin{aligned} \hat{x}_{k+1} &= A\hat{x}_k + Bu_k + Ke_k \\ y_k &= C\hat{x}_k + Du_k + e_k. \end{aligned} \tag{2.15}$$

Substituting the measurement equation in the update equation leads to

$$\hat{x}_{k+1} = A\hat{x}_k + Bu_k + Ky_k - KC\hat{x}_k - KDu_k = \bar{A}\hat{x}_k + \bar{B}\begin{pmatrix} u_k \\ y_k \end{pmatrix}, \tag{2.16}$$

with $\bar{A} = A - KC$ and $\bar{B} = \begin{bmatrix} B - KD & K \end{bmatrix}$. Let the extended measurement $z_k = \begin{pmatrix} u_k & y_k \end{pmatrix}^T$. Note that it is possible to recursively compute successive state predictions such as

$$\hat{x}_{k+2} = \bar{A}x_{k+1} + \bar{B}z_{k+1} = \bar{A}^2 x_k + \bar{A}\bar{B}z_k + \bar{B}z_{k+1}. \tag{2.17}$$

Hence, in general,

$$\hat{x}_{k+p} = \bar{A}^p \hat{x}_k + \begin{bmatrix} \bar{B} & \bar{A}\bar{B} & \cdots & \bar{A}^{p-1}\bar{B} \end{bmatrix} \begin{pmatrix} z_{k+p-1} \\ z_{k+p-2} \\ \vdots \\ z_k \end{pmatrix} = \bar{A}^p \hat{x}_k + \bar{\Delta}_p Z^{k+p-1,k}. \tag{2.18}$$

By definition, $\bar{A} = A - KC$ is a stable matrix, so that for large p, $\bar{A}^p \approx 0$ and

$$\hat{x}_{k+p} = \bar{\Delta}_p Z^{k+p-1,k}. \tag{2.19}$$

Substituting this result into the measurements equation,

$$y_{k+p} = C\bar{\Delta}_p Z^{k+p-1,k} + Du_{k+p} + e_{k+p}. \tag{2.20}$$

For a window of size f, a set of equations can be stacked as

$$\begin{aligned} \begin{bmatrix} y_{k+p} & \cdots & y_{k+p+f} \end{bmatrix} &= C\bar{\Delta}_p \begin{bmatrix} Z^{k+p-1,k} & \cdots & Z^{k+p+f-1,k+f} \end{bmatrix} \\ &\quad + D \begin{bmatrix} u_{k+p} & \cdots & u_{k+p+f} \end{bmatrix} + \begin{bmatrix} e_{k+p} & \cdots & e_{k+p+f} \end{bmatrix}. \end{aligned} \tag{2.21}$$

Thus, $C\bar{\Delta}_p$ and D can be obtained as the solution of an unconstrained least squares problem, namely, solving Gauss' normal equations. Let the controllability-like matrix $\bar{\Gamma}_f$ be defined as

$$\bar{\Gamma}_f = \begin{bmatrix} C \\ C\bar{A} \\ \vdots \\ C\bar{A}^{f-1} \end{bmatrix}. \tag{2.22}$$

The Hankel matrix $\bar{\Gamma}_f \bar{\Delta}_d$ is thus given by

$$\bar{\Gamma}_f \bar{\Delta}_d = \begin{bmatrix} C\bar{B} & C\bar{A}\bar{B} & \cdots & C\bar{A}^{p-1}\bar{B} \\ C\bar{A}\bar{B} & C\bar{A}^2\bar{B} & \cdots & C\bar{A}^p\bar{B} \\ \vdots & & \ddots & \\ C\bar{A}^{f-1}\bar{B} & C\bar{A}^f\bar{B} & \cdots & C\bar{A}^{p+f-1}\bar{B} \end{bmatrix}. \tag{2.23}$$

Given the approximation $\bar{A}^k \approx 0, k \geq p$,

$$\bar{\Gamma}_f \bar{\Delta}_d \approx \begin{bmatrix} C\bar{B} & C\bar{A}\bar{B} & \cdots & C\bar{A}^{p-1}\bar{B} \\ C\bar{A}\bar{B} & C\bar{A}^2\bar{B} & \cdots & 0 \\ \vdots & & \ddots & \\ C\bar{A}^{f-1}\bar{B} & C\bar{A}^f\bar{B} & \cdots & 0 \end{bmatrix}, \tag{2.24}$$

which can be constructed by blocks of the already computed matrix $C\bar{\Delta}_p$. Let $Z = \begin{bmatrix} Z^{k+p-1,k} & \ldots & Z^{k+p+f-1,k+f} \end{bmatrix}$. From the structure of Δ_p, it is possible to write

$$\hat{X} \doteq \begin{bmatrix} \hat{x}_{k+p} & \cdots & \hat{x}_{k+p+f} \end{bmatrix} = \bar{\Delta}_p Z. \tag{2.25}$$

Hence,

$$\bar{\Gamma}_f \hat{X} = \bar{\Gamma}_f \bar{\Delta}_p Z. \tag{2.26}$$

Since the right hand side is known, $\bar{\Gamma}_f \hat{X}$ can be computed. An infinite set of representations leads to the same system, as previously discussed. Thus, one might simply determine an $\hat{X}$ by performing a Singular Value Decomposition (SVD) such that

$$USV^T = \bar{\Gamma}_f \hat{X}. \tag{2.27}$$

Only a set of the largest $n_{\hat{x}}$ singular values might be considered, leading to a dimension $x \in \mathbb{R}^{n_{\hat{x}}}$. Thus, for instance, a compatible representation is given by

$$\bar{\Gamma}_f = U\sqrt{S},\ \hat{X} = \sqrt{S}U^T. \tag{2.28}$$

From this point, the computation of A, B and C is straightforward. In particular, the system matrices can be determined by writing the innovation form dynamics in (2.15) for all inputs and measurements, in the same way as described in (2.21), using $\hat{X}$ as a known quantity. Note that K is an output of the PBSID.

This method is a powerful strategy for DT system identification. As the determination of the system's matrices is approached as a LS problem with a white noise source, by leveraging the structure of the innovation form, it has a number of powerful statistical properties. In particular, the PBSID method has been proved to be a consistent estimator, even for closed loop identification, as well as asymptotically efficient. In general, the computational and implementation complexity of the PBSID is small as compared to other identification approaches. Additionally, by choosing the number of singular values in the SVD, the size of the inner state x is a design variable. This is convenient, as it allows to solve the identification problem for various state dimensions, and then select it given some metric. Note that no initial guess is required for the identification of the system matrices A, B and C.

Unfortunately, the PBSID is posed as a DT identification method, which does not have a direct application to CT systems, needing for a previous discretization or CT-DT transformation stage. For this added complexity, which is an open area of study, several methods have been proposed. This additional phase is extremely relevant for the PBSID to provide good results, and is the focus of this work.

2.3. Hilbert spaces

Hilbert spaces provide a rigorous mathematical framework for analyzing functions and sequences, particularly within the context of system identification and signal processing. In this section, only the essentials for this thesis are provided. For more detail, the reader is referred to a manual on functional analysis. A particularly well known reference is [24].

2.3.1. Introduction

A Hilbert space is a vector space $\mathcal{H}$ over the field of real or complex numbers equipped with an inner product $\langle \cdot, \cdot \rangle : \mathcal{H} \times \mathcal{H} \to \mathbb{K}$ (where $\mathbb{K}$ is either $\mathbb{R}$ or $\mathbb{C}$) that satisfies the following properties:

1. Linearity in the first argument: $\langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle$ for all $x, y, z \in \mathcal{H}$ and scalars $\alpha, \beta \in \mathbb{K}$.
2. Conjugate symmetry: $\langle x, y \rangle = \overline{\langle y, x \rangle}$ for all $x, y \in \mathcal{H}$ (where $\overline{z}$ denotes the complex conjugate of z; if $\mathbb{K} = \mathbb{R}$, then $\langle x, y \rangle = \langle y, x \rangle$).
3. Positive definiteness: $\langle x, x \rangle \geq 0$ for all $x \in \mathcal{H}$, with $\langle x, x \rangle = 0$ if and only if $x = 0$.

These properties become useful in various parts of this work. The inner product induces a norm, $\|x\| = \sqrt{\langle x, x \rangle}$, which allows the definition of distance and concepts such as convergence. In particular, a Hilbert space is also required to be complete with respect to the norm induced by the inner product, namely, every Cauchy sequence must converge to an element within the space. This property is fundamental for taking limits and for many approximation techniques. This section introduces the fundamental concepts of Hilbert spaces, emphasizing the $\mathcal{L}^2$ and ℓ^2 spaces, which are very useful frameworks for characterizing continuous-time and discrete-time signals, respectively. These spaces provide the setting for expressing CT and DT systems as operators in a Hilbert space, and allow to construct well-founded transformations between these two formulations.

2.3.2. The $\mathcal{L}^2$ space

The space $\mathcal{L}^2(\mathbb{T})$ is the set of all square-integrable functions defined on a time interval $\mathbb{T} \subseteq \mathbb{R}$. Formally, a function $f : \mathbb{T} \to \mathbb{R}$ belongs to $\mathcal{L}^2(\mathbb{T})$ if

$$\int_{\mathbb{T}} |f(t)|^2 \, dt < \infty. \tag{2.29}$$

The inner product in $\mathcal{L}^2(\mathbb{T})$ is defined as

$$\langle f, g \rangle_{\mathcal{L}^2} = \int_{\mathbb{T}} f(t) g(t) \, dt, \tag{2.30}$$

for $f, g \in \mathcal{L}^2(\mathbb{T})$. The induced norm (or length) of a function $f \in \mathcal{L}^2(\mathbb{T})$ is given by

$$\|f\|_{\mathcal{L}^2} = \sqrt{\langle f, f \rangle_{\mathcal{L}^2}} = \sqrt{\int_{\mathbb{T}} |f(t)|^2 \, dt}. \tag{2.31}$$

The $\mathcal{L}^2$ space is a Hilbert space, which means it is a complete normed vector space. This implies that it is closed under taking limits of Cauchy sequences, ensuring that limits of convergent sequences of functions in this space also belong to the space. This is crucial for dealing with approximations and iterative procedures commonly used in system identification. In the context of system theory, one can consider $\mathcal{L}^2(\mathbb{T}; \mathbb{R}^n)$ as the space of n-dimensional square-integrable vector functions. In that case,

$$\langle f, g \rangle_{\mathcal{L}^2} = \int_{\mathbb{T}} f(t)^T g(t) \, dt, \tag{2.32}$$

for $f, g \in \mathcal{L}^2(\mathbb{T}; \mathbb{R}^n)$.

2.3.3. The ℓ^2 space

The space ℓ^2 is the counterpart of $\mathcal{L}^2$ for discrete-time signals. It is the space of all square-summable sequences. For a sequence $x = (x_k)_{k \in \mathbb{M}}$, where $\mathbb{M} \subseteq \mathbb{Z}$ is the index set, it is said that $x \in \ell^2$ if

$$\sum_{k \in \mathbb{M}} |x_k|^2 < \infty. \tag{2.33}$$

The inner product in ℓ^2 is defined as

$$\langle x, y\rangle_{\ell^2} = \sum_{k\in\mathbb{M}} x_k y_k, \tag{2.34}$$

for $x, y \in \ell^2$. The induced norm of a sequence $x \in \ell^2$ is

$$\|x\|_{\ell^2} = \sqrt{\langle x, x\rangle_{\ell^2}} = \sqrt{\sum_{k\in\mathbb{M}} |x_k|^2}. \tag{2.35}$$

Similar to the $\mathcal{L}^2$ space, the ℓ^2 space is also a Hilbert space, which is essential when dealing with iterative identification algorithms on discrete time data, where the convergence of sequences is critical. In the context of system theory, one can consider $\ell^2(\mathbb{M}; \mathbb{R}^n)$ as the space of n-dimensional square-summable vector sequences. In that case,

$$\langle x, y\rangle_{\ell^2} = \sum_{k\in\mathbb{M}} x_k^T y_k. \tag{2.36}$$

2.3.4. Projections on Hilbert spaces

A fundamental property of Hilbert spaces is the existence of orthogonal projections onto closed subspaces. Let $\mathcal{H}$ be a Hilbert space, and let $\mathcal{M}$ be a closed subspace of $\mathcal{H}$. The Projection Theorem states that for any element $x \in \mathcal{H}$, there exists a unique element $x_{\mathcal{M}} \in \mathcal{M}$, known as the orthogonal projection of x onto $\mathcal{M}$, such that

$$\|x - x_{\mathcal{M}}\| \leq \|x - y\|, \quad \forall y \in \mathcal{M}. \tag{2.37}$$

This element $x_{\mathcal{M}}$ is characterized by the fact that the difference $x - x_{\mathcal{M}}$ is orthogonal to every element in $\mathcal{M}$, i.e.,

$$\langle x - x_{\mathcal{M}}, y\rangle = 0, \quad \forall y \in \mathcal{M}. \tag{2.38}$$

For this work, it is of particular interest to project a given element $x \in \mathcal{H}$ onto the subspace spanned by some $v \in \mathcal{H} - \{0\}$. The resulting vector, denoted as $x_v \in \operatorname{span}(v)$, is such that

$$\langle x - x_v, v\rangle = 0. \tag{2.39}$$

Vector x_v is uniquely given by

$$x_v = \frac{\langle x, v\rangle}{\langle v, v\rangle} v = \langle x, v\rangle \frac{v}{\|v\|}. \tag{2.40}$$

If v has unit norm,

$$x_v = \langle x, v \rangle v. \tag{2.41}$$

The concept of projections provides the best approximation of an element in a subspace in terms of the norm induced by the inner product.

2.3.5. Orthonormal bases

A set of vectors $\{e_i\}_{i \in I}$ in a Hilbert space $\mathcal{H}$ is called an orthogonal basis if:

1. $\langle e_i, e_j \rangle = 0$ for all $i, j \in I$ with $i \neq j$ (orthogonality).
2. The set of all finite linear combinations of the e_i is dense[1] in $\mathcal{H}$.

If, in addition, every vector e_i has unit norm, i.e., $\|e_i\| = \sqrt{\langle e_i, e_i \rangle} = 1$ for all $i \in I$, then the basis is called an orthonormal basis.

Given an orthonormal basis $\{e_i\}_{i \in I}$ and an arbitrary element $x \in \mathcal{H}$, the orthogonal projection of x onto the subspace spanned by a single basis vector e_i is given by

$$x_{e_i} = \langle x, e_i \rangle e_i. \tag{2.43}$$

Since the set of all finite linear combinations of the basis vectors is dense in the Hilbert space $\mathcal{H}$, any element $x \in \mathcal{H}$ can be approximated by a linear combination of the basis vectors:

$$x \approx \sum_{i \in F} c_i e_i, \tag{2.44}$$

where F is a finite subset of I and c_i are the coefficients. In fact, x can be expressed as a convergent series with respect to the norm as

$$x = \sum_{i \in I} \langle x, e_i \rangle e_i, \tag{2.45}$$

where the equality holds in the sense of convergence in the norm induced by the inner

---

[1] The term "dense" means that for any element $x \in \mathcal{H}$ and any arbitrarily small positive number $\epsilon > 0$, there exists a finite linear combination of the basis vectors, say $\sum_{i \in F} c_i e_i$ (where F is a finite subset of I), such that the distance between x and this linear combination is less than ϵ:

$$\left\| x - \sum_{i \in F} c_i e_i \right\| < \epsilon. \tag{2.42}$$

In other words, the set of all finite linear combinations of the basis vectors can get arbitrarily close to any element in the Hilbert space $\mathcal{H}$. This is equivalent to requiring that every element $x \in \mathcal{H}$ can be approximated arbitrarily well by elements in the span of the basis vectors.

product of $\mathcal{H}$. This is called the series expansion of x with respect to the orthonormal basis $\{e_i\}$. The coefficients $c_i = \langle x, e_i \rangle$ are known as the coordinates of x with respect to this basis.

While the complete series expansion in (2.45) gives the exact vector x in terms of elements of the basis, the finite series expansion in (2.44) provides the best approximation of x contained in the span of $\{e_i\}_{i\in F}$. Since I is infinite-dimension for $\mathcal{L}^2$ and ℓ^2 for it to be dense, as given by property 2, only a finite set $F \subset I$ is considered for all series expansions within this work.

2.3.6. Spectrum of a linear operator

Let T be a linear operator over a Hilbert space

$$T : \mathcal{H} \to \mathcal{H} \tag{2.46}$$

The concept of eigenvalue problem can be generalized from linear algebra to any Hilbert space. For this purpose, the *resolvent* operator is defined as

$$R_\lambda(T) = (T - \lambda I)^{-1}. \tag{2.47}$$

A *regular value* is a complex number $\lambda_r \in \mathbb{C}$ such that

1. $R_{\lambda_r}(T)$ exists.
2. $R_{\lambda_r}(T)$ is bounded.
3. $R_{\lambda_r}(T)$ is defined over a dense subset $E \subset \mathcal{H}$.

The set of all regular values is the *resolvent set*, $\rho(T)$. The complementary of the resolvent set is the *spectrum* of the operator, $\sigma(T) = \mathbb{C} - \rho(T)$. In particular, all values λ that do not verify condition 1 are labeled by *point spectrum* $\sigma_p(T)$, those verifying condition 1 but not condition 2 are the *continuous spectrum* $\sigma_c(T)$, and those satisfying conditions 1 and 2 but not condition 3 belong to the *residual spectrum* $\sigma_r(T)$. Note that $\sigma(T) = \sigma_p(T) \cup \sigma_c(T) \cup \sigma_r(T)$, and $\sigma_p \cap \sigma_c = \sigma_p \cap \sigma_r = \sigma_c \cap \sigma_r = \varnothing$. The definition of eigenvalue is only possible for the point spectrum, and can be relaxed to also consider the continuous spectrum. The residual spectrum is not of interest for this work, and is not considered in what follows.

For the point spectrum, by imposing that $T - \lambda I$ is not invertible, the usual definition of eigenvalue is recovered, namely,

$$Tv = \lambda v,\ \ v \in \mathcal{H} - \{0\}. \tag{2.48}$$

Vectors v satisfying the above condition are the eigenvectors of T associated to λ. For the continuous spectrum, an equivalent form known as a Weyl sequence is used.

2.3.7. Self-adjoint operators

Let T be a bounded[2] linear operator over a Hilbert space. The adjoint operator $T^\dagger$ is the only linear operator such that, for any $u, v \in \mathcal{H}$,

$$\langle u, Tv\rangle = \langle T^\dagger u, v\rangle. \tag{2.49}$$

An operator is *self-adjoint* if it satisfies $T = T^\dagger$. Let two eigenvectors u and v be associated to different eigenvalues $\lambda_1 \neq \lambda_2$ of a self-adjoint operator. Then,

$$\langle u, Tv\rangle = \langle Tu, v\rangle = \lambda_1\langle u, v\rangle = \lambda_2\langle u, v\rangle. \tag{2.50}$$

Thus, it must hold that

$$\langle u, v\rangle = 0. \tag{2.51}$$

This is a general property of self-adjoint operators: the eigenvectors associated to two different eigenvalues are orthogonal. Besides, in consequence of the conjugate symmetry property of the inner product,

$$\langle u, Tu\rangle = \overline{\langle Tu, u\rangle} = \lambda_1\|u\|^2 = \overline{\lambda_1}\|u\|^2. \tag{2.52}$$

Consequently, all eigenvalues of a self-adjoint operator are real. Furthermore, the residual spectrum of self-adjoint operators is empty, that is, $\sigma(T) = \sigma_p(T) \cup \sigma_c(T)$.

A particularly interesting property exists for self-adjoint operators having only point spectrum, that is, $\sigma(T) = \sigma_p(T)$. In that case, the set of all eigenvectors of T span the complete space $\mathcal{H}$,

$$\sigma(T) = \sigma_p(T) \to \mathcal{H} = \bigoplus_{\lambda\in\sigma_p(T)} \ker\left(T - \lambda I\right). \tag{2.53}$$

In other words, the eigenvectors of a self-adjoint operator with only point spectrum form an orthogonal basis of $\mathcal{H}$.

---

[2] An operator T is bounded if $\sup_{u\in\mathcal{H}} \|Tu\| < M\|u\|$, for some $M \in \mathbb{R}^+$.

3 | CT-DT transformation via basis expansions

This Chapter explores the transformation of CT linear systems to DT representations using basis expansions in Hilbert spaces. This transformation is essential for many applications in system identification and control, where algorithms are often designed for DT data. During the development of this thesis, it was found that literature on CT-DT transformation for system identification did not approach the problem from a basis transformation perspective. This motivated research on this theoretical framework, which leverages the algebraic properties of Hilbert spaces and the existence of discrete bases for purely continuous spaces.

The chapter begins by establishing a framework where linear systems are viewed as operators acting on Hilbert spaces of signals. It then explores the concept of Schauder bases, with a focus on Laguerre and Hermite functions, which provide a way to represent signals as sequences of coefficients. The construction of CT-DT transformations is then discussed, showing how operators can be discretized through the use of these basis expansions. Furthermore, a revision of previous CT-PBSID methods based on Laguerre functions is provided, highlighting the link between projection and convolution approaches, and examining the validity of its foundation from a Hilbert spaces perspective. Finally, a novel approach based on an analogy with the well known quantum oscillator is introduced, leading to an elegant and direct CT-DT transformation using Hermite functions..

3.1. Linear systems as operators on Hilbert spaces

A CT Linear Time-Invariant (LTI) system can be viewed as a combination of signals and dynamics. Signals are elements of a continuous Hilbert space, namely $\mathcal{L}^2(\mathbb{R})$ or $\mathcal{L}^2(\mathbb{R}^+)$. Let $x(t), u(t), y(t), v(t), w(t) \in \mathcal{L}^2$, where $x(t)$ is the system state, $u(t)$ the input, $y(t)$ the output, and $v(t)$, $w(t)$ some noise processes. The dynamics are linear operators over the Hilbert space, given by

$$T : \mathcal{L}^2 \to \mathcal{L}^2. \tag{3.1}$$

Specifically, considering the structure in (2.8), operators such as the derivative d/dt, and the system matrices A, B, C, and D are considered.

The goal of this work is to find a bijective transformation of such a system to a DT LTI one, sharing the structure with (2.9). In the DT framework, signals are elements of a discrete Hilbert space, namely $\ell^2(\mathbb{Z})$ or $\ell^2(\mathbb{N})$. Let $x_n, u_n, y_n, v_n, w_n \in \ell^2$. The dynamics are (linear) operators over the Hilbert space, given by

$$T_d : \ell^2 \rightarrow \ell^2. \tag{3.2}$$

Specifically, the operators of interest are the index increase operator (z), and the system matrices A_d, B_d, C_d, and D_d.

Additionally, a very important requirement for the CT-DT transformation is that it must preserve the mean and covariance of Wiener processes. This last property appears naturally by using orthonormal bases, as proved later on.

3.2. Schauder bases of $\mathcal{L}^2$

A Schauder basis, or countable basis, for a Hilbert space $\mathcal{H}$ is a sequence $\{b_i\}_{i\in\mathbb{N}}$ of elements of the space, $b_i \in \mathcal{H}$, that satisfies

$$v = \sum_{i\in\mathbb{N}} \alpha_i b_i, \quad \forall v \in \mathcal{H}, \tag{3.3}$$

and for which the coefficients α_i are unique. A Hilbert space that admits a Schauder basis is called separable. Thus, a discrete basis can be defined in a Hilbert space if and only if the space is separable. Conveniently, both $\mathcal{L}^2(\mathbb{R})$ and $\mathcal{L}^2(\mathbb{R}^+)$ are separable spaces. Furthermore, any separable infinite-dimensional Hilbert space is isometrically isomorphic to $\ell^2(\mathbb{N})$. For instance, let two spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ be isometrically isomorphic. Thus, there exists a transformation $T : \mathcal{H}_1 \rightarrow \mathcal{H}_2$ which verifies the following:

- Isometric: T preserves the norm of vectors, namely, $\|v\|_{\mathcal{H}_1} = \|T(v)\|_{\mathcal{H}_2}, \forall v \in \mathcal{H}_1$. For Hilbert spaces, this implies that T is unitary.
- Isomorphic: T is bijective and the algebraic structure of the spaces is preserved. For Hilbert spaces, this implies that T is linear invertible.

It is of interest to find Schauder bases that are orthonormal, as they provide a particularly straightforward transformation strategy based on projections. The standard orthonormal countable bases of $\mathcal{L}^2(\mathbb{R})$ and $\mathcal{L}^2(\mathbb{R}^+)$ are the Hermite functions and the Laguerre functions, respectively. Laguerre functions have been well studied for CT-DT transformation, although their application is not derived from the theory of Hilbert spaces and Schauder bases but by using the methods of system lifting and Laguerre filtering, whereas to the author knowledge, Hermite functions have not been considered for the CT system identification problem.

3.2.1. Laguerre functions

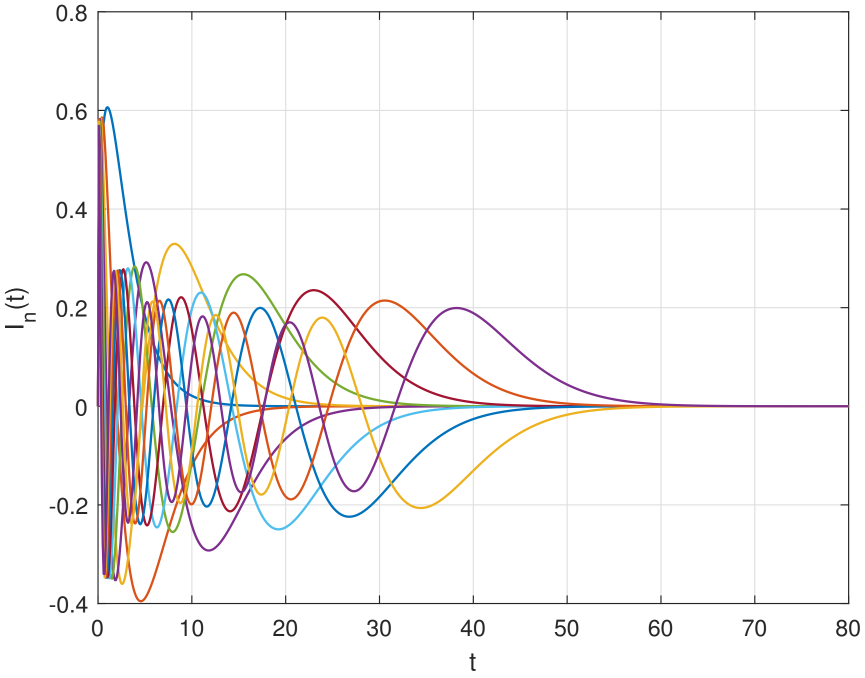


Figure 3.1: Laguerre basis up to order 10, with pole $a = 1$.

The Laguerre basis is a set of functions $\{l_n(t)\} \in \mathcal{L}(\mathbb{R}^+)$, constructed on top of the well known Laguerre polynomials, by means of a normalization. This is a Shauder basis of $\mathcal{L}(\mathbb{R}^+)$ which is also orthonormal, namely,

$$\langle l_n, l_m \rangle = \int_0^\infty l_n(t) l_m(t) dt = \delta_{nm}, \forall n, m \in \mathbb{N}. \tag{3.4}$$

The Laguerre functions of order n can be written as the convolution of a low-pass filter $l_0(t)$, which is the 0-th order Laguerre function, with an n-th order stable all-pass filter, i.e.,

$$\mathcal{A}^n l_0(t) = l_n(t). \tag{3.5}$$

A relevant property of these functions, is their simple expression in the Fourier domain.

Indeed, the Fourier transform of an n-th order all pass filter is given by

$$\mathcal{F}\left\{\mathcal{A}^n\right\} = \left(\frac{s-a}{s+a}\right)^n, \tag{3.6}$$

where a is a constant $a \in \mathbb{R}^+$ known as the filter's pole that characterizes the filter. The Fourier transform of the low-pass filter l_0 is

$$\mathcal{F}\left\{l_0\right\} = \frac{\sqrt{2a}}{s+a}. \tag{3.7}$$

In practice, Laguerre functions can be efficiently computed by means of recurrence relations, which arise from the normalized polynomial nature of this basis.

3.2.2. Hermite functions

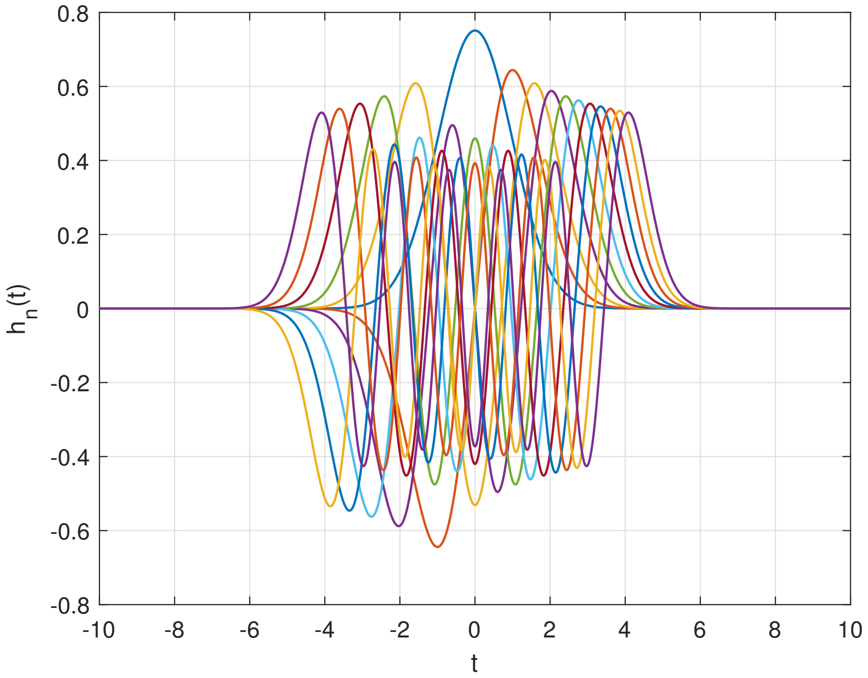


Figure 3.2: Hermite basis up to order 10.

The normalized Hermite functions, $h_n(t)$, are defined using the Hermite polynomials[1]

---

[1] In the literature, the polynomials used in this thesis are sometimes referred to as the *physicist's Hermite polynomials*. An explicit expression for the n-th Hermite polynomial is the particularized Rodrigues' formula, $\mathrm{H}_n(t) = (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2}$, although they are typically generated using recurrence relations, as shown later in this chapter.

$\mathrm{H}_n(t)$ as

$$h_n(t) = e^{-t^2/2} \frac{\mathrm{H}_n(t)}{\sqrt{2^n \sqrt{\pi} n!}}. \tag{3.8}$$

The sequence $\{h_n(t)\}$ is an orthonormal Schauder basis of $\mathcal{L}^2(\mathbb{R})$, that is,

$$\langle h_n, h_m \rangle = \int_{-\infty}^{\infty} h_n(t) h_m(t) dt = \delta_{nm}, \forall n, m \in \mathbb{N}. \tag{3.9}$$

These functions are of particular interest because of their representation as eigenfunctions of the Fourier transform operator and the interesting relations of the derivative operator, as it will be shown next. Additionally, in the same way as Laguerre functions, Hermite functions satisfy recurrence relations, which are fundamental for practical applications.

3.2.3. Construction of CT-DT transformations

Given an operator $T : \mathcal{L}^2 \to \mathcal{L}^2$, it is possible to discretize it using an orthonormal Schauder basis $\{b_i\}$. For instance, let the vectors $v, v' \in \mathcal{L}^2$ satisfying

$$v' = Tv. \tag{3.10}$$

By expanding the vectors into their series, it follows that

$$\sum_{i \in \mathbb{N}} \langle b_i, v' \rangle b_i = T \sum_{i \in \mathbb{N}} \langle b_i, v \rangle b_i = \sum_{i \in \mathbb{N}} \langle b_i, v \rangle T b_i. \tag{3.11}$$

By projecting both sides on the basis element b_j, it follows that

$$\langle b_j, v' \rangle = \sum_{i \in \mathbb{N}} \langle b_i, v \rangle \langle b_j, T b_i \rangle. \tag{3.12}$$

Consequently,

$$\tilde{v}'_j = \sum_{i \in \mathbb{N}} \tilde{T}_{ji} \tilde{v}_i, \tag{3.13}$$

where $\tilde{v}_i = \langle b_i, v \rangle$ and $\tilde{T}_{ji} = \langle b_j, T b_i \rangle$. Therefore, this procedure transforms an operator in the $\mathcal{L}^2$ into an operator in ℓ^2 in the coordinates of the given basis. Note that as per the definition of Schauder bases, any transformed vector $\{\tilde{v}_n\}$ and operator $\tilde{T}_{ji}$ are unique, that is, there is only one representation of a vector v and operator T as a linear combination of the elements of a basis. This uniqueness property is proved to be of interest later on. For instance, let B_1 and B_2 be two different Schauder bases. It can be guaranteed that

$$v_j = \sum_{u_i \in B_1} \tilde{v}_{ji} u_i, \forall v_j \in B_2. \tag{3.14}$$

Thus, let a basis be given by appending B_1 and B_2, $B = \{v \in B_1, B_2\} = \{u_0, u_1, \ldots, v_0, v_1, \ldots\}$. As shown in (3.14), all elements of B_2 are a linear combination of elements of B_1, in other words, the basis B is composed of linearly dependent vectors and, for instance, can never be orthogonal. Therefore, no additional information can be obtained by appending a Schauder basis, but redundant (linearly-dependent) data. All Schauder bases of $\mathcal{L}^2$ are (possibly infinite-terms) linear combinations of each other. The applicability of this result on CT-DT transformation is discussed next.

The structure of $\tilde{T}_{ji}$ is of particular interest. Writing it in matrix form,

$$\tilde{T} = \begin{bmatrix} \langle b_1, Tb_1 \rangle & \langle b_1, Tb_2 \rangle & \cdots \\ \langle b_2, Tb_1 \rangle & \langle b_2, Tb_2 \rangle & \cdots \\ \vdots & \ddots & \ddots \end{bmatrix}. \tag{3.15}$$

The term $\tilde{T}_{ji}$ is nonzero only if Tb_i has some component in the direction of b_j. An operator $\tilde{T}$ with $\tilde{T}_{ji} \neq 0$ is said to *connect* states i and j. This intuition is of interest later on.

In some cases, explicitly computing $\tilde{T}_{ji} = \langle b_j, Tb_i \rangle$ can be a complex task, as it is the case for the derivative operator. As $\tilde{T}_{ji}$ is unique, it might be determined by exploiting the properties of the chosen Schauder basis.

3.3. Revision of the CT-PBSID

Two methods for system discretization based on Laguerre functions were proposed in [15]: a projection method and a convolution method. This section aims to study these methods under the light of functional analysis in Hilbert spaces, not providing a complete description or discussion of their properties, but a characterization in terms of the proposed theoretical framework.

The projection method is based on the methodology explained in the previous section. It defines the following projection for the i-th Laguerre function,

$$\tilde{u}_i(k) = \int_0^\infty (\mathcal{A}_i^k l_0(\tau)) u(t_i + \tau) d\tau, \tag{3.16}$$

which is by definition equivalent to

$$\tilde{u}_i(k) = \langle l_k(t), u(t_i + t)\rangle. \tag{3.17}$$

Therefore, the CT-PBSID expands a given vector $u(t_i+t) \in \mathcal{L}(R^+)$ as a linear combination of the Schauder basis, and then computes the corresponding coefficients by exploiting the fact that the basis is orthonormal. The signal can be completely characterized for any $t_i \in \mathbb{R}^+$ as

$$u(t + t_i) = \sum_{k\in\mathbb{N}} \tilde{u}_i(k) l_k(t). \tag{3.18}$$

For a different sampling time t_j, one can write

$$u(t + t_j) = \sum_{k\in\mathbb{N}} \tilde{u}_j(k) l_k(t + (t_j - t_i)). \tag{3.19}$$

As discussed in the previous section, $\{l_k(t)\}$ and $\{l_k(t + (t_j - t_i))\}$ are two linearly-dependent bases. Therefore, for a sufficient order of expansion, these vectors are linear combinations of each other, and coefficients $\tilde{u}_j$ do not add information to $\tilde{u}_i$ about the projected function $u(t)$.

The original CT-PBSID performs a low order Laguerre expansion of signals, but resamples it at every sampling instant t_i. This approach does not conserve the properties derived from orthogonality and completeness of the Schauder basis. In particular, from this approach, considering n samples and order N expansions, the generated basis is given by

$$B = \{l_0(t), \dots, l_N(t), l_0(t + \Delta t), \dots, l_N(t + \Delta t), \dots, l_0(t + n\Delta t), \dots, l_N(t + n\Delta t)\}. \tag{3.20}$$

Even for $n \gg 1$, this basis is nor orthogonal, nor linearly independent, and certainly not a Schauder basis, lacking the properties shown throughout this text. In contrast, a method that is closer to the theoretical results is a high order (N) expansion of the signal for a single given sampling point t_0, so that $\langle u(t + t_0), l_m\rangle \approx 0, m > N$.

A convolution method is also proposed in [15], based on the framework developed by [20], to benefit from the simple Fourier transform of Laguerre functions. The structure of convolutions is very similar to that of the projections. To illustrate this, let two functions $f, g : \mathbb{R}^+ \to \mathbb{R}$. By definition,

$$(f * g)(t) = \int_{\mathbb{R}^+} f(\tau)g(t - \tau)d\tau = \int_0^t f(\tau)g(t - \tau)d\tau, \tag{3.21}$$

as f, g are defined only for positive real values. Consequently

$$(f * g)(t) = \langle f(\tau), g(t-\tau) \rangle_{[0,t]}, \tag{3.22}$$

that is, the convolution of functions in $\mathbb{R}^+$ evaluated at t is equivalent to an inner product in $[0, t]$. Only if $t \gg 1$, the convolution has the structure of the inner product of $\mathcal{L}^2(\mathbb{R}^+)$. By using convolutions, a discretization of the linear system is obtained. However, for the same basis of Laguerre functions, the coefficients of the expansion are different. Additionally, the convolution-based method performs worse than the projection-based approach.

If both discretizations were correct, this result seems as a contradiction. As previously shown, for a given basis of functions, the projection of a linear operator T over the span of the basis, $\tilde{T}_{ji}$, is unique. This apparent inconsistency of theory and practice disappears by examining the original proof of the projection-based CT-DT transformation [20]. In particular, it assumes that the convolution-based transformation conserves the cross covariance of Wiener processes. For this to be true, the following equality must hold,

$$\int_0^\infty l_q(t-\tau) S l_r(t-\tau) d\tau \overset{?}{=} S\delta_{qr}. \tag{3.23}$$

This is correct only if evaluated at $t \to \infty$, as given by the orthonormality condition of Laguerre functions in $\mathcal{L}(\mathbb{R}^+)$. In other words, Laguerre functions are not orthonormal in any interval $[0, t]$, and the theoretical results previously discussed are only valid for large t. As just shown, convolutions for $t \to \infty$ behave as projections, which leads to the aforementioned projection method. Hence, the convolution-based approach only approximately obtains a DT expansion, even in theory, and does not conserve the mean and covariance of noise signals, among other properties.

Examining (3.23), a set of functions for which a convolution may be used, just as a projection, is the Hermite basis. Indeed, for Wiener processes $v(t)$, $w(t)$ defined over $\mathbb{R}$

with cross-covariance S, it holds that

$$\begin{aligned} E\left\{\left(h_n(t) * v(t)\right)\left(h_m(t) * w(t)\right)\right\} &= E\left\{\left(\int_{-\infty}^{\infty} h_n(t-\tau)dv(\tau)\right)\left(\int_{-\infty}^{\infty} h_m(t-\phi)dw(\phi)\right)\right\} \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_n(t-\tau)h_m(t-\phi)E\{dv(\tau)dw(\phi)\} \\ &= \int_{-\infty}^{\infty} h_n(t-\tau)h_m(t-\tau)Sd\tau = S\delta_{nm}. \end{aligned} \tag{3.24}$$

This is a convenient property, as it ensures that the covariance of random processes is conserved, not only under projection, but also under convolution. This arises an interest on Hermite functions as an alternative to the Laguerre basis for CT-DT projections. Nonetheless, transforming the d/dt operator and hence the dynamics requires a careful analysis and is not straightforward. A similar problem is found in Quantum Mechanics when computing the momentum operator acting on the eigenfunctions of the harmonic oscillator. This framework is adopted to the problem at hand, leading to an elegant method, as follows.

3.4. The quantum oscillator approach to CT-DT transformation

Functional analysis is one of the most relevant mathematical tools in many areas of science and engineering. In particular, Quantum Mechanics built on top of Hilbert spaces theory as a consequence of its First Postulate[2] (see, for instance, [25]). The expansion of functions using Schauder bases is a particularly important technique, as it is the foundation of quantum states, energy levels and state transitions.

This motivates searching for connections between the approaches followed in Quantum Mechanics and those of CT-DT transformations. In this work, one of these connections was found by observing the nature of Hermite functions and the well known quantum harmonic oscillator problem, which is of paramount importance in many applications such as nuclear or molecular physics. This interesting framework is explored in this section, firstly from a theoretical view and then with numerical examples verifying the predicted properties. These techniques are later used to construct a novel CT system identification method.

---

[2] The First Postulate of Quantum Mechanics states that any physical system is completely defined by a function $\Psi(\boldsymbol{r}, t)$, the wavefunction, which belongs to $\mathcal{L}^2$.

3.4.1. Theoretical framework

Hermite functions can be defined as the solutions of the Hermite equation, namely,[3]

$$\left(-\frac{1}{2}\frac{d^2}{dt^2}+\frac{1}{2}t^2\right)h_n(t)=\left(n+\frac{1}{2}\right)h_n(t). \tag{3.25}$$

This is the exact structure of the time-invariant Schrödinger's one-dimensional equation for a quantum harmonic oscillator. The operator $1/2(d^2/dt^2+t^2)$ is defined as the *Hamiltonian*, H. Note that problem (3.25) is also an eigenvalue problem for the operator H, which can be proved to be self-adjoint and to have only point spectrum. The following Hamiltonian decomposition is a common tool for simplification,

$$H=\frac{1}{2}\left(-\frac{d^2}{dt^2}+t^2\right)=\frac{1}{\sqrt{2}}\left(-\frac{d}{dt}+t\right)\frac{1}{\sqrt{2}}\left(\frac{d}{dt}+t\right)+\frac{1}{2}\doteq a^\dagger a+\frac{1}{2}, \tag{3.26}$$

where the creation and annihilation operators, $a^\dagger$ and a, are defined by

$$a^\dagger=\frac{1}{\sqrt{2}}\left(-\frac{d}{dt}+t\right) \tag{3.27}$$

$$a=\frac{1}{\sqrt{2}}\left(\frac{d}{dt}+t\right). \tag{3.28}$$

An important technique for operator manipulation in Quantum Mechanics is the so-called commutator $[\cdot,\cdot]$, defined for two operators $A, B:\mathcal{L}^2\to\mathcal{L}^2$ such that

$$[A,B]\doteq AB-BA. \tag{3.29}$$

Note that, by definition,

$$Ha^\dagger h_n(t)=[H,a^\dagger]h_n(t)+a^\dagger Hh_n(t). \tag{3.30}$$

The second term is simple to compute because, as given in (3.25),

$$Hh_n(t)=\left(n+\frac{1}{2}\right)h_n(t). \tag{3.31}$$

---

[3] The equation is written in a slightly non-conventional manner for a clearer analysis.

The first term is slightly more involved. In particular, the following commutation relationships can be proved

$$[a, a^\dagger] = 1 \tag{3.32}$$

$$[H, a^\dagger] = a^\dagger \tag{3.33}$$

$$[H, a] = -a. \tag{3.34}$$

Thus,

$$Ha^\dagger h_n(t) = a^\dagger h_n(t) + a^\dagger \left(n + \frac{1}{2}\right) h_n(t) = \left(n + 1 + \frac{1}{2}\right) a^\dagger h_n(t). \tag{3.35}$$

Comparing with (3.25), it is seen that $a^\dagger h_n(t)$ is an eigenfunction of H with eigenvalue $(n+1) + \frac{1}{2}$, or that

$$a^\dagger h_n(t) \propto h_{n+1}(t). \tag{3.36}$$

It is equally proved that

$$ah_n(t) \propto h_{n-1}(t). \tag{3.37}$$

By inspection of the commutation relationships, it can be proved that

$$a^\dagger h_n(t) = \sqrt{n+1} h_{n+1}(t) \tag{3.38}$$

$$ah_n(t) = \sqrt{n} h_{n-1}(t). \tag{3.39}$$

Therefore, the derivative operator can be written by combining (3.27) and (3.28),

$$\frac{d}{dt} h_n(t) = \frac{1}{\sqrt{2}} \left(a - a^\dagger\right) h_n(t). \tag{3.40}$$

Note that, for Hermite functions, the derivative operator connects the state n with states $n-1$ and $n+1$ only. Obtaining the derivative operator for the coefficients of a Hermite series expansion is not trivial. Let a function $f(t)$ be written as a series expansion

$$f(t) = \sum_{n=0}^{\infty} f_n h_n(t), \tag{3.41}$$

where

$$f_n = \langle h_n(t), f(t) \rangle. \tag{3.42}$$

Differentiating this function allows to extend the derivative operator to the discrete domain so that

$$\frac{d}{dt} f(t) = \sum_{n=0}^{\infty} f_n \left(\frac{1}{\sqrt{2}} \left(a - a^\dagger\right) h_n(t) \right) \doteq \sum_{n=0}^{\infty} \left(\frac{d}{dt} f_n \right) h_n(t), \tag{3.43}$$

where coefficients df_n/dt are defined so that the previous equality holds. To determine df_n/dt from f_n, it is useful to project the prior equation over $h_m(t)$ and leveraging orthonormality, that is,

$$\left\langle h_m(t), \frac{d}{dt}f(t)\right\rangle = \sum_{n=0}^{\infty}\left\langle h_m(t), \left(\frac{1}{\sqrt{2}}\left(a - a^\dagger\right)\right)h_n(t)\right\rangle f_n = \frac{d}{dt}f_m. \tag{3.44}$$

Applying the adjoint definition within the inner product,

$$\sum_{n=0}^{\infty}\left\langle \left(\frac{1}{\sqrt{2}}\left(a - a^\dagger\right)\right)^\dagger h_m(t), h_n(t)\right\rangle f_n = \left(\frac{1}{\sqrt{2}}\left(a - a^\dagger\right)\right)^\dagger f_m = \frac{d}{dt}f_m. \tag{3.45}$$

Noting that $a^\dagger$ and a are adjoints of each other, one finally gets

$$\frac{d}{dt}f_n = \left(\frac{1}{\sqrt{2}}\left(a - a^\dagger\right)\right)^\dagger f_n = \frac{1}{\sqrt{2}}\left(a^\dagger - a\right)f_n. \tag{3.46}$$

This expression can be written as a matrix operator. In particular, let $\underline{\mathrm{f}} = \begin{pmatrix} f_0 & f_1 & \cdots \end{pmatrix}$.[4] It follows that

$$\frac{d}{dt}\underline{\mathrm{f}} = \underline{\mathrm{f}}\begin{bmatrix} 0 & \sqrt{1/2} & 0 & 0 & \cdots \\ -\sqrt{1/2} & 0 & \sqrt{2/2} & 0 & \cdots \\ 0 & -\sqrt{2/2} & 0 & \sqrt{3/2} & \cdots \\ 0 & 0 & -\sqrt{3/2} & 0 & \ddots \\ \vdots & \vdots & \vdots & \ddots & \ddots \end{bmatrix} = \underline{\mathrm{fD}}. \tag{3.47}$$

When considering a vector function, the definition is identical. Indeed, let

$$\underline{\mathrm{f}} = \begin{pmatrix} f_{x,0} & f_{x,1} & \cdots \\ f_{y,0} & f_{y,1} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \tag{3.48}$$

It holds that

$$\frac{d}{dt}\underline{\mathrm{f}} = \underline{\mathrm{fD}}. \tag{3.49}$$

---

[4] This definition is more convenient than using column vectors, as shown next.

Hence, the transformation of a linear system defined by (2.4) is

$$\frac{d}{dt}x = Ax + Bu \rightarrow \underline{x}D = A\underline{x} + B\underline{u} \tag{3.50}$$

$$y = Cx + Du \rightarrow \underline{y} = C\underline{x} + D\underline{u} \tag{3.51}$$

For the series expansion definition (3.3), let $f(t) : \mathbb{R} \rightarrow \mathbb{R}^n$ be a vector function and $\underline{f}$ its Hermite representation. Besides, let the discretized Hermite basis matrix be

$$\underline{H} = \begin{bmatrix} h_0(t_0) & h_0(t_1) & \dots & h_0(t_N) \\ h_1(t_0) & h_1(t_1) & \dots & h_1(t_N) \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}. \tag{3.52}$$

The sampled time-domain signal $f(t)$ can be recovered at the sampling times $t_0, t_1, \dots, t_N$ so that

$$\begin{pmatrix} f(t_0) & f(t_1) & \cdots & f(t_N) \end{pmatrix} = \underline{f}\underline{H}. \tag{3.53}$$

In practice, a maximum order n_{max} is considered for the Hermite expansion. Let $m = n_{max} + 1$, accounting for the order 0 term. Thus,

$$\underline{f} \in \mathbb{R}^{n\times m}, \underline{D} \in \mathbb{R}^{m\times m}, \underline{H} \in \mathbb{R}^{m\times N}. \tag{3.54}$$

The use of this procedure leads to a much simpler bijective transformation of the system matrices. Furthermore, as seen in (3.50), the introduction of Hermite functions leads to a second-order discrete system, due to the presence of terms of order $n+1$, n, and $n-1$. This is a consequence of the fact that the operator $\underline{D}$, which is the projection of the derivative on the Hermite domain, connects the element n of the basis with $n+1$ and $n-1$. As all three terms $n-1$, n and $n+1$ are present on the discrete dynamics, they are indeed second order.

The projection with Hermite functions leads to a simple matrix-like derivative operator, conveniently for a DT system identification. For completeness, next the convolution method is explored. For a function $f(t) \in \mathcal{L}^2(\mathbb{R})$, let

$$\left[h_n(t) * f(t)\right](t) = \int_{-\infty}^{\infty} h_n(t-\tau) f(\tau) d\tau. \tag{3.55}$$

Considering the symmetry property $h_n(-t) = (-1)^n h_n(t)$ for Hermite functions, the following can be written,

$$[h_n(t) * f(t)]\,(t_i) = (-1)^n \int_{-\infty}^{\infty} h_n(\tau - t_i) f(\tau) d\tau = (-1)^n \langle h_n(t - t_i), f(t) \rangle. \tag{3.56}$$

This demonstrates that convolutions are just projections with alternating sign, leading to an equivalent coefficients expansion. In consistence with theory, and in particular due to the uniqueness of coefficients for an expansion of a given basis, convolutions lead to a set of coefficients only if they coincide with the ones provided by projections. Hermite functions thus provide a framework in which projections or convolutions can be equally used to obtain the coefficients of a series expansion of functions in $\mathcal{L}^2(\mathbb{R})$ over $\ell^2(\mathbb{N})$.

3.4.2. Numerical examples

Various numerical examples of interest are provided to support and clarify the previous theoretical results. In particular, three main constructions are considered:

- Projection: given a function $f(t) \in \mathcal{L}^2(\mathbb{R})$, the expansion on the Hermite basis $\{f_n\}$ is defined as $f_n = \langle h_n(t), f(t) \rangle = \int_{\mathbb{R}} h_n(t) f(t) dt$. In practice, a finite support and order of the expansion is considered.
- Reconstruction: given a Hermite representation $\{f_n\}$ of a function, for a maximum order n_{max}, the reconstructed time-domain signal is defined as $f_h(t) = \sum_{n=0}^{n_{max}} f_n h_n(t)$. This process is equivalent to equation (3.53).
- Hermite-domain differentiation: as described in (3.49), the quantum oscillator approach leads to an effective procedure to differentiate signals on the Hermite domain. The obtained coefficients may be used to reconstruct the time-domain derivative of the original function.

Firstly, the observed results on white noise transformation can be clearly seen in Figure 3.3. For a time-domain signal (top), the Hermite representation (bottom) is provided. The coefficients of an infinite support white noise process follow a white noise process, as expected. Note that, for a finite support noise signal, the resulting coefficients do not follow the same behavior. Thus, the support of signals and the used basis should be adequate. This is discussed in the next section.

In Figure 3.4, a signal and its derivative are compared with the Hermite reconstructed expansions. The following procedure is used to construct the time derivative in the Hermite

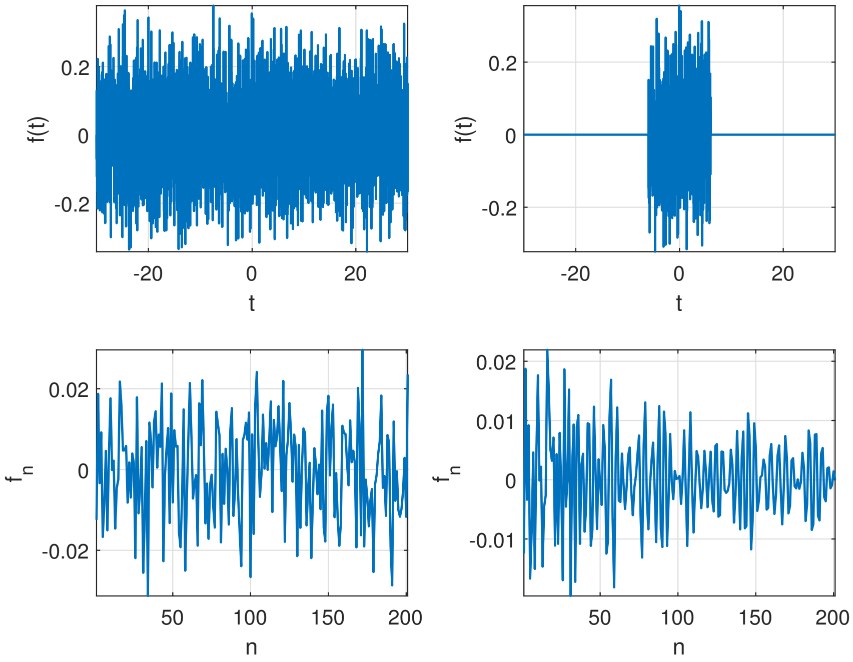


Figure 3.3: Projection of white noise signals of infinite and finite support over the Hermite domain. Each column corresponds to a different signal.

domain:

$$x(t) \to \text{Projection} \to x_n \to \text{Differentiation} \to \frac{dx_n}{dt} \to \text{Reconstruction} \to \frac{dx_h(t)}{dt}.$$

The Hermite expansions effectively allow to compute the signal time evolution and its derivative. Note that, although an order 200 expansion was used, only smaller order coefficients would be required, as a rapid decay in magnitude is observed. This property is a consequence of the fact that Hermite coefficients belong to $\ell^2(\mathbb{N})$, that is, they should decay at $n \to \infty$ for the series to be square-summable.

Figure 3.5 depicts an interesting experiment combining both random processes and the derivative operator. For a white noise signal $w(t)$, a Brownian motion is given by[5]

$$\frac{d}{dt}f(t) = w(t). \tag{3.57}$$

---

[5] A slight abuse of notation was used. More rigorously, given a white noise signal dw with variance $\sigma^2 dt$, a Brownian motion or Wiener process is given by $df = dw$.

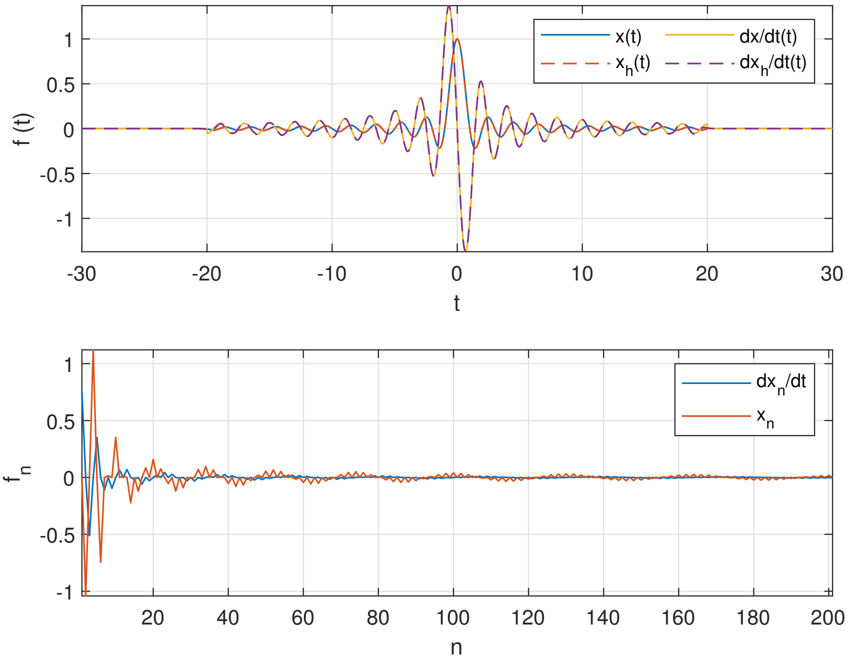


Figure 3.4: Differentiation of signals on the Hermite domain and reconstruction on the time domain.

On the Hermite domain, this expression translates to

$$\underline{\mathrm{f}}\underline{\mathrm{D}} = \underline{\mathrm{w}}. \tag{3.58}$$

The derivative operator $\underline{\mathrm{D}}$ can be inverted[6] to obtain $\underline{\mathrm{f}}$ from $\underline{\mathrm{w}}$. Reconstructing the time domain signals as defined in (3.53),

$$f(t) = \underline{\mathrm{f}}\underline{\mathrm{H}} = \underline{\mathrm{w}}\underline{\mathrm{D}}^{-1}\underline{\mathrm{H}}. \tag{3.59}$$

This is shown in Figure 3.5. A white noise signal is transformed to the Hermite domain, multiplied by $\underline{\mathrm{D}}^{-1}$ and projected to the time domain, obtaining a Brownian motion as expected.

The last numerical example is given by the simulation of a linear system in the Hermite

---

[6] The fact that the order expansion is finite makes $\underline{\mathrm{D}}^{-1}$ singular. This issue can be tackled by using a high order expansion and switching the last element of the last column of $\underline{\mathrm{D}}$ to a small positive number. Alternatively, a modified differentiation operator can be used, as described later in this chapter.

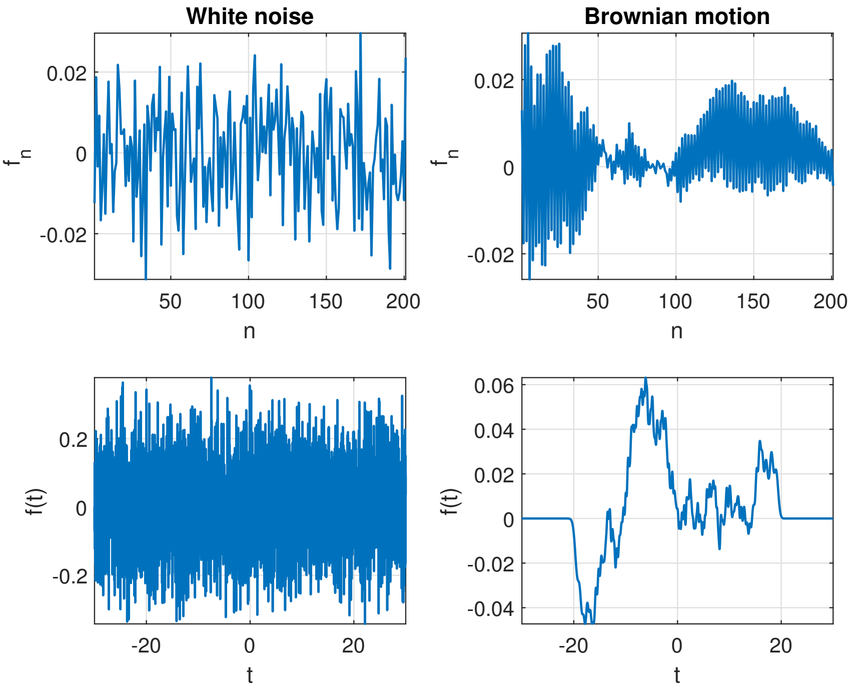


Figure 3.5: Construction of a Brownian motion process on the Hermite domian.

domain. Let a noiseless linear system given by

$$A = \begin{bmatrix} -2 & 1 \\ 0 & -4 \end{bmatrix}, B = \begin{bmatrix} 1 & 0.5 \\ 0 & 1 \end{bmatrix}, C = \begin{bmatrix} 1 & 0 \\ 4 & 1 \end{bmatrix}, D = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Just as for the Brownian motion, a system given by a linear equation can be integrated in the Hermite domain. In particular, note that

$$\underline{\mathrm{xD}} = A\underline{\mathrm{x}} + B\underline{\mathrm{u}} \rightarrow \underline{\mathrm{x}} = (A\underline{\mathrm{x}} + B\underline{\mathrm{u}})\,\underline{\mathrm{D}}^{-1}. \tag{3.60}$$

Given $\underline{\mathrm{u}}$, this problem can be solved in a recursive manner. This is shown in Figure 3.6. Both the input and output signal are well characterized, reconstructing the system's dynamics in the Hermite domain.

3.5. The HD-PBSID method

The previous discussion leads to the construction of a CT-DT transformation compatible with the PBSID algorithm. However, some issues have to be assessed for this framework to behave appropriately. The final result, which is one of the main contributions of this work, is the Hermite Domain PBSID (HD-PBSID).

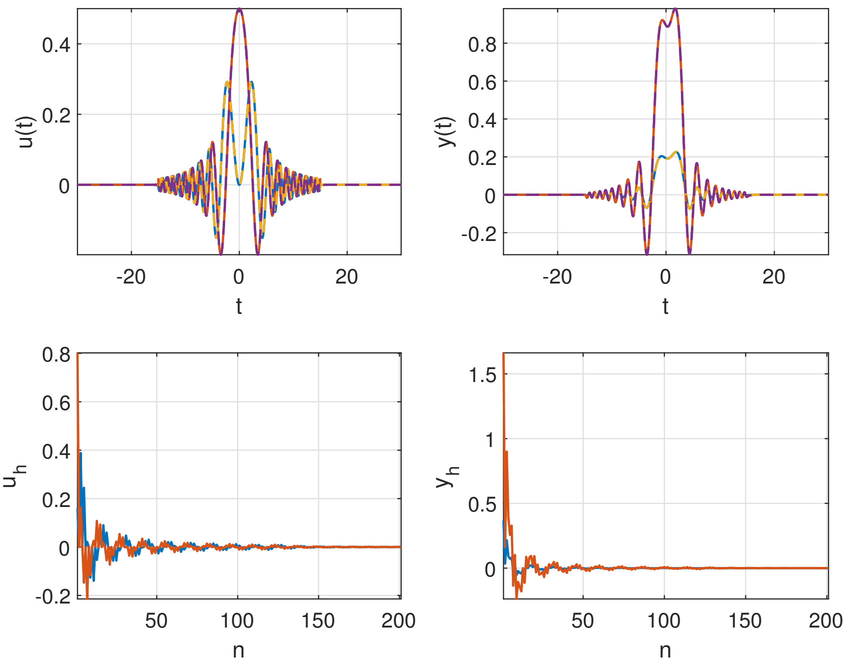


Figure 3.6: Simulation of a linear system with two inputs and two outputs on the Hermite domain, with the time domain dynamics in continuous line and the Hermite reconstructed signals in dashed line. The original and reconstructed signals coincide.

3.5.1. Effective support and scaling

The algorithm begins with a projection onto a Hermite subspace. In particular, a basis of $n_{max}+1$ Hermite functions is considered, in what follows denoted by $\mathcal{B} = \{h_n\}_0^{n_{max}}$. Although an infinite Hermite basis allows to write any function in $\mathcal{L}^2(\mathbb{R})$, a truncated one such as $\mathcal{B}$ is limited to the so-called *effective support*. In Figure 3.7, a base of Hermite function up to order 200 is displayed. As it can be seen, functions rapidly decay outside of the domain $\left[-\sqrt{2n_{max}+1}, \sqrt{2n_{max}+1}\right]$. This is a general property of Hermite functions.

Let an experiment take place in $t \in [-\tau, \tau]$. It is convenient that any function $f(t)$ to be projected is contained within the effective support of $\mathcal{B}$. For that purpose, let the time change of variable be

$$t' = t\frac{\sqrt{2n_{max}+1}}{\tau} = \frac{t}{\alpha}, \tag{3.61}$$

so that $f(t) \to f(t')$. This transformed function perfectly fits the effective support of the

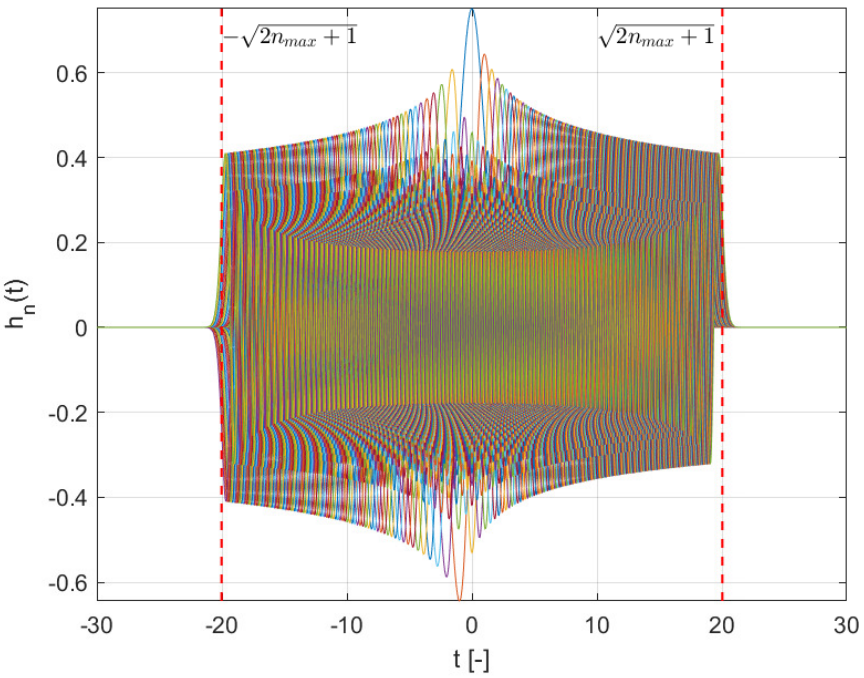


Figure 3.7: Hermite basis of order 200, with effective support.

considered basis. Note that this affects the dynamic equations, so that

$$\frac{d}{dt}x = Ax + Bu \rightarrow \frac{d}{dt'}x = \alpha Ax + \alpha Bu. \tag{3.62}$$

3.5.2. Modified differentiation operator

Let the innovation form dynamics in (2.15) for a CT system, be

$$\begin{aligned} \frac{d}{dt}\hat{x} &= A\hat{x} + Bu + Ke \\ y &= C\hat{x} + Du + e. \end{aligned} \tag{3.63}$$

Therefore,

$$\frac{d}{dt}\hat{x} = (A - KC)\,\hat{x} + \begin{bmatrix} B - KD & K \end{bmatrix} \begin{pmatrix} u \\ y \end{pmatrix} = \bar{A}\hat{x} + \bar{B}z. \tag{3.64}$$

Note that $\bar{A}$ is stable in the CT sense, that is, it has only eigenvalues with negative real part. By performing a CT-DT transformation such as (3.49), accounting for the time scaling described in (3.62), one obtains

$$\underline{\mathrm{xD}} = \alpha\bar{A}\underline{\mathrm{x}} + \alpha\bar{B}\underline{\mathrm{z}}. \tag{3.65}$$

This equation is not well-suited for direct application of the PBSID algorithm, as it requires that $\alpha \bar{A}$ is stable in the DT sense, that is, with eigenvalues within the unit circle, which is not granted. This can be effectively handled by introducing a modified differentiation operator, namely,

$$\underline{\mathrm{D}}' = \left(\frac{\mathrm{D}}{\alpha} + \beta I \right) \frac{1}{\gamma}, \tag{3.66}$$

for $\beta \geq 0$ and $\gamma > 0$. This way, dynamics is written as follows,

$$\underline{\mathrm{x}\mathrm{D}'} = \frac{\bar{A} + \beta I}{\gamma} \underline{\mathrm{x}} + \frac{\bar{B}}{\gamma} \underline{\mathrm{z}} = \bar{A}' \underline{\mathrm{x}} + \bar{B}' \underline{\mathrm{z}}. \tag{3.67}$$

If the eigenvalues of $\bar{A}$ are contained in a ball $B_{\beta,\gamma}$ with center $-\beta$ on the real axis and radius γ, as depicted in Figure 3.8, the resulting transformation guarantees the properties required for the PBSID. In practice, it is of interest to set γ as a representative value for the expected eigenvalues of the system, and β as a large positive number. Indeed, It is clear that eigenvalues λ that are close to the boundary of $B_{\beta,\gamma}$ lead to an increased bias in the solution with respect to those near the center of the stability region. For this reason, it is convenient to use enlarged stability regions to improve the estimation and reduce bias. A very high value of β, together with a high order of differentiation,

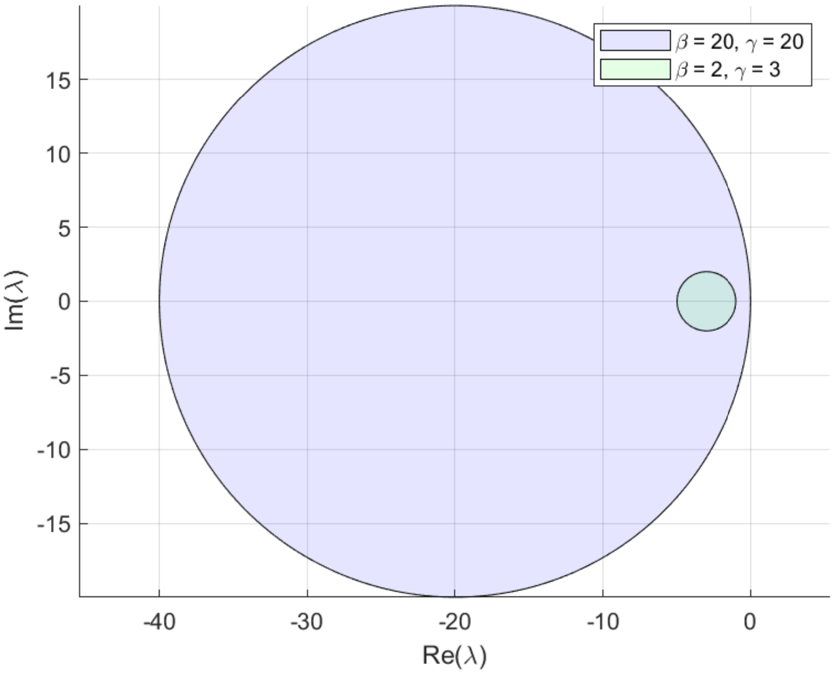


Figure 3.8: Stability region in the complex plane for the modified differentiation operator.

might induce ill-posedness in the PBSID Least Squares problem. For this reason, if very sparse eigenvalues are expected, a partial identification of fast and slow modes separately, together with adequate experiment duration and value of α, might be needed. For double precision solvers (such as `MATLAB`), γ values of 30 still lead to numerical accuracy.

3.5.3. Integrator form

The PBSID can be applied by first differentiating (3.67) multiple times such that

$$\underline{\mathrm{xD}}'^{n} = \bar{A}'^{n}\underline{\mathrm{x}} + \bar{B}'\underline{\mathrm{zD}}'^{n-1} + \bar{A}'\bar{B}'\underline{\mathrm{zD}}'^{n-2} + \cdots + \bar{A}'^{n-1}\bar{B}'\underline{\mathrm{z}}. \tag{3.68}$$

Substituting $\bar{A}'^{n} \approx 0$, the following is obtained

$$\underline{\mathrm{xD}}'^{n} \approx \begin{bmatrix} \bar{B}' & \bar{A}'\bar{B}' & \cdots & \bar{A}'^{n-1}\bar{B}' \end{bmatrix} \begin{bmatrix} \underline{\mathrm{zD}}'^{n-1} \\ \underline{\mathrm{zD}}'^{n-2} \\ \vdots \\ \underline{\mathrm{z}} \end{bmatrix}. \tag{3.69}$$

Therefore, the operator $\underline{\mathrm{D}}'$ has to be applied to the input-output signal $\underline{\mathrm{z}}$ multiple times. It is important to recall that this matrix operator is the projection of a differentiation operator on the time domain. In presence of noise, a differentiation leads to its amplification, which is detrimental to the performance of the algorithm. Conversely, the inverse of $\underline{\mathrm{D}}'$ is an integration-like operator and, as such, acts as a filter. By multiplying the previous equation by $\underline{\mathrm{D}}'^{-n}$, or what is the same, by integrating it n times, the following expression is obtained,

$$\underline{\mathrm{x}} = \begin{bmatrix} \bar{B}' & \bar{A}'\bar{B}' & \cdots & \bar{A}'^{n-1}\bar{B}' \end{bmatrix} \begin{bmatrix} \underline{\mathrm{zD}}'^{-1} \\ \underline{\mathrm{zD}}'^{-2} \\ \vdots \\ \underline{\mathrm{zD}}'^{-n} \end{bmatrix}. \tag{3.70}$$

In this case, $\underline{\mathrm{zD}}'^{-m}$ is a filtered quantity, which does behave adequately in presence of noise. This equation, denoted by integrator form, has proved to improve significantly the identification results, and is used hereafter.

3.5.4. Final algorithm

From the previously described modifications, a powerful system identification algorithm based on Hermite projections can be obtained. This section aims to provide a straightforward reference to implement the HD-PBSID in practice. In particular, the following function is described:

$$\left\{u(t_1^u), u(t_2^u), \dots, u(t_{n_u}^u), y(t_1^y), \dots, y(t_{n_y}^y)\right\} \rightarrow \text{HD-PBSID} \rightarrow \left\{\hat{A}, \hat{B}, \hat{C}, \hat{D}, \hat{K}\right\}.$$

Note that the sampling times t^u and t^y for u and y do not have to coincide. Firstly, the input and output signals are scaled and projected into a Hermite subspace of dimension $n_{max} + 1$. Defining $\alpha = \tau/\sqrt{2n_{max}+1}$, for τ one half of the total experiment time, the Hermite domain projection of the m-th dimension of each signal is given by

$$\underline{u}_{mn} = \int_{t^u/\alpha} u_m\left(t/\alpha\right) h_n(t')dt' \tag{3.71}$$

$$\underline{y}_{mn} = \int_{t^y/\alpha} y_m\left(t/\alpha\right) h_n(t')dt' \tag{3.72}$$

for $n = 0, 1, \dots, n_{max}$. The integration should be performed with a noise-robust quadrature scheme, such as a trapezoid rule. This step is perhaps the most important of the whole identification. Adequately choosing n_{max} is paramount to the algorithm. Additionally, signals u and y should be time shifted so that their time vector is symmetric, that is, so that both signals are defined in a support $[-\tau, \tau]$. In order to correctly select n_{max}, it is very useful to compute the DT-CT transformation matrix $\underline{H}$ such as (3.52) and recover a CT signal from $\underline{u}$, $\underline{y}$. After undoing the time scaling, the recovered signals should match the initial ones and fit their frequency spectrum. A proper experiment design is also very relevant for this step to be well posed, as discussed later.

Once $\underline{u}$ and $\underline{y}$ are computed, the modified differentiation operator is to be constructed. For this purpose, both β and γ need to be defined. β should be an average of the real parts expected for the system under study. γ should be as high as possible, with appropriate values around 30. The definition of these parameters has proved not to be critical for performance. Let

$$\underline{\mathrm{D}}' = \begin{bmatrix} 0 & \frac{\sqrt{1/2}+\alpha\beta}{\alpha\gamma} & 0 & 0 & \cdots \\ \frac{-\sqrt{1/2}+\alpha\beta}{\alpha\gamma} & 0 & \frac{1+\alpha\beta}{\alpha\gamma} & 0 & \cdots \\ 0 & \frac{-1+\alpha\beta}{\alpha\gamma} & 0 & \frac{\sqrt{3/2}+\alpha\beta}{\alpha\gamma} & \cdots \\ 0 & 0 & \frac{-\sqrt{3/2}+\alpha\beta}{\alpha\gamma} & 0 & \ddots \\ \vdots & \vdots & \vdots & \ddots & \ddots \end{bmatrix}, \tag{3.73}$$

with dimension $(2n_{max}+1) \times (2n_{max}+1)$. Then, the following matrices are constructed,

$$\underline{\mathrm{z}} = \begin{bmatrix} \underline{\mathrm{u}} \\ \underline{\mathrm{y}} \end{bmatrix}, \tag{3.74}$$

and, given a design parameter p, typically in the order of 10,

$$\underline{\mathrm{Z}} = \begin{bmatrix} \underline{\mathrm{z}}\underline{\mathrm{D}}'^{-1} \\ \underline{\mathrm{z}}\underline{\mathrm{D}}'^{-2} \\ \vdots \\ \underline{\mathrm{z}}\underline{\mathrm{D}}'^{-p} \end{bmatrix}. \tag{3.75}$$

As in the conventional PBSID, let the Least Squares (LS) problem

$$\underline{\mathrm{y}} = C\bar{\Delta}_p\underline{\mathrm{Z}} + D\underline{\mathrm{u}} + \underline{\mathrm{e}}. \tag{3.76}$$

This problem is solved for $C\bar{\Delta}_p$ and D, minimizing the norm $\|\underline{\mathrm{y}} - C\bar{\Delta}_p\underline{\mathrm{Z}} - D\underline{\mathrm{u}}\|$. From the obtained matrices, the following is constructed,

$$\bar{\Gamma}_f\bar{\Delta}_d \approx \begin{bmatrix} C\bar{B}' & C\bar{A}'\bar{B}' & \cdots & C\bar{A}'^{p-1}\bar{B}' \\ C\bar{A}'\bar{B}' & C\bar{A}'^{2}\bar{B}' & \cdots & 0 \\ \vdots & & \ddots & \\ C\bar{A}'^{f-1}\bar{B}' & C\bar{A}'^{f}\bar{B}' & \cdots & 0 \end{bmatrix}. \tag{3.77}$$

Following the PBSID, the inner state vector in the Hermite domain is obtained by means of a SVD, namely,

$$USV^T = \bar{\Gamma}_f\bar{\Delta}_d\bar{Z} \rightarrow \underline{\mathrm{x}} = \sqrt{S'}V^T, \tag{3.78}$$

where S' is a submatrix of S with the desired number of singular values. This selection is performed by first sorting the diagonal matrix S in decreasing size of the singular values, and then extracting a square submatrix going from the first row and column to the N-th ones, being N the dimension of the resulting system. Matrix C is the solution of a LS problem,

$$\underline{\mathrm{y}} = C\underline{\mathrm{x}} + D\underline{\mathrm{u}} + \underline{\mathrm{e}}. \tag{3.79}$$

The innovation vector in the Hermite domain is then given by the residual

$$\underline{\mathrm{e}} = \underline{\mathrm{y}} - C\underline{\mathrm{x}} - D\underline{\mathrm{u}}. \tag{3.80}$$

Matrices A and B are the solution of a LS problem given by

$$\gamma\underline{\mathrm{xD}'} = A\underline{\mathrm{x}} + B\underline{\mathrm{u}} + K\underline{\mathrm{e}}. \tag{3.81}$$

As seen in the structure of the algorithm, this method directly provides A, B, C and D as solutions of LS problems, in contrast to the CT-PBSID, where an additional transformation is needed to obtain the system's CT matrices from the DT ones (see [15]).

3.5.5. Experiment design

For an infinite order expansion, any Schauder basis of $\mathcal{L}^2$ can be used to characterize any function $f(t)$ precisely. However, in practice, only finite order expansions can be computed, and thus the fitting of the series expansion and the original function might not be perfect. An adequate selection of a function basis that is similar to the original function provides an improvement in the quality of the approximation. For instance, it is well known that periodic signals are well modeled by a sum of sines and cosines satisfying that periodicity.

For Hermite functions expansions, it is also important to select a suitable signal for projection. Although in system identification the output cannot be directly controlled (and perhaps not even the input, as happens for closed-loop systems), the experiment can be designed in a way that benefits the identification and, in this case, the projections. Hermite functions, as depicted in Figure 3.2, are pulse-like signals with a finite effective support and decaying intensity at both ends. They are also symmetric or anti-symmetric for even and odd indices, respectively. In linear system identification, the frequency components of the input remains in the output, as linear operators can only introduce phase and magnitude changes. Therefore, if the input (or controller reference) is a Hermite-like signal, so should be the output.

Another important consideration for system identification of continuous systems is that the input should excite all frequencies of interest for the system's dynamics. Modes that are not excited are less likely to be identified. Therefore, the Fourier Transform of the used input signal should be rich in the range of interest for the identification. When using Hermite functions, there is not direct control of the frequency spectrum of the resulting signal, so a different procedure is necessary to construct a suitable input.

During the experimental validation of the algorithm, a solution to this problem was found. First, the user must define an initial input $f_0(t)$ signal which excites the frequencies of interest. It does not need to resemble a Hermite function. Then, the signal is projected to a finite Hermite basis, yielding a sequence $\underline{\mathrm{f}}$. The maximum order of the expansion should be lower than the maximum order for identification. Then, the time domain input is reconstructed from the Hermite approximation such that $f(t) = \underline{\mathrm{fH}}$. This signal is then analyzed in terms of frequency spectrum and shape, to verify that it is adequate for the experiment. If not rich enough, a greater order of Hermite expansion is necessary. If suitable, this signal can be used as the input for the identification process. With this method, it is straightforward to generate signals that satisfy the experiment requirements while also providing good properties for the CT-DT process.

3.5.6. Effective algorithm to obtain the Hermite basis

In practical scenarios, the order of the expansions needed for identification may be around 200. For the identification algorithm to be usable, the computation of the Hermite basis must be very efficient. Although analytical expressions do exist for Hermite functions, the memory requirements make this approach not practical. Fortunately, Hermite functions exhibit the so-called recurrence relations, which provide straightforward and computationally efficient methods to evaluate hundreds of functions rapidly. For the normalized functions used in this work, the resulting algorithm can be obtained from basic properties of Hermite polynomials.

Hermite polynomials $\mathrm{H}_n(x)$ satisfy the following recurrence relation

$$\mathrm{H}_{n+1}(x) = 2x\mathrm{H}_n(x) - 2n\mathrm{H}_{n-1}(x). \tag{3.82}$$

The normalized Hermite functions are computed by applying a normalization coefficient,

$$h_n(x) = \mathrm{H}(x)\frac{\exp\left(-\frac{x^2}{2}\right)}{\sqrt{2^n\sqrt{\pi}n!}}. \tag{3.83}$$

Thus, the recurrence relation for the Hermite functions is then found to be

$$h_{n+1}(x) = x\sqrt{\frac{2}{n+1}}h_n(x) - \sqrt{\frac{n}{n+1}}h_{n-1}(x). \tag{3.84}$$

Additionally, the first two Hermite functions are given by

$$h_0(x) = \frac{\exp\left(\frac{-x^2}{2}\right)}{\pi^{1/4}} \tag{3.85}$$

$$h_1(x) = \sqrt{2}x\frac{\exp\left(\frac{-x^2}{2}\right)}{\pi^{1/4}}. \tag{3.86}$$

This gives a closed algorithm to compute the Hermite expansion up to any order with minimal memory and computing requirements. In this thesis, the described algorithm was implemented in `C++`, providing a very fast method to project and reconstruct signals. This method was verified in simulation.

4 | Generalization through Sturm-Liouville theory

In the previous chapter, the problem of CT-DT transformation was observed under the light of functional analysis and Hilbert spaces. This led to a general understanding which covered both the previous Laguerre-based frameworks and a novel methodology based on Hermite functions. The HD-PBSID was developed by leveraging well-known results from quantum mechanics, namely, the connection between the derivative and the creation-annihilation operators in a harmonic quantum oscillator. In contrast, the CT-PBSID uses techniques such as system lifting to derive a CT-DT transformation. Both Hermite and Laguerre functions form Schauder bases of Hilbert spaces, and can be obtained by the normalization of their associated orthogonal polynomials. Additionally, they can both be efficiently computed by means of known recurrence relations, which allow to use large expansions on limited hardware.

After the development of the HD-PBSID, it became of interest whether other CT-DT frameworks based on different bases could be obtained in a systematic way. Such a methodology should then provide the CT-PBSID and the HD-PBSID as particularizations of a more general formulation. Exploring this connection and the nature of Laguerre and Hermite functions naturally led to the Sturm-Liouville theory for differential equations.

This chapter describes the construction of a common framework for CT-DT transformation containing, among others, the Laguerre and Hermite-based methods. For this purpose, the Sturm-Liouville theory is combined with the previous theoretical foundation. By the understanding of CT-DT transformation from a more general perspective, various interesting general results are proved.

4.1. Sturm-Liouville problems and orthogonal bases

The Sturm-Liouville theory is the study of a particular family of ODEs known as the Sturm-Liouville problem. To define this problem precisely, consider real-valued functions

$p(x), q(x), w(x)$ defined on an interval $\mathbb{I} \subseteq \mathbb{R}$. For the theory to be well-behaved, certain conditions are typically imposed on these functions. It is generally assumed that $p(x)$ has a continuous first derivative, and $q(x)$ and $w(x)$ are continuous on the interval $\mathbb{I}$. Furthermore, for many standard results to hold, it is often required that $p(x) > 0$ and $w(x) > 0$ for all x in the interior of $\mathbb{I}$. The function $w(x)$ is specifically known as the weight function. A second-order linear ODE is termed a Sturm-Liouville equation if it can be expressed in the form

$$-\frac{d}{dx}\left(p(x)\frac{d}{dx}f(x)\right) + q(x)f(x) = \lambda w(x)f(x), \tag{4.1}$$

where $f(x)$ is the unknown function, and $\lambda \in \mathbb{R}$ is a parameter, known as the eigenvalue. The objective is to find non-trivial solutions $f(x)$ that satisfy this equation for specific values of λ. Such non-trivial solutions $f(x)$ are called eigenfunctions corresponding to the eigenvalue λ. It is convenient to define Sturm-Liouville operator as

$$L = -\frac{d}{dt}\left(p(x)\frac{d}{dx}\right) + q(x). \tag{4.2}$$

Then, a Sturm-Liouville problem can be simply written as

$$Lf(x) = \lambda w(x)f(x). \tag{4.3}$$

This form emphasizes that the Sturm-Liouville problem is fundamentally about finding the eigenvalues and eigenfunctions of the operator L with respect to the weight function $w(x)$.

The operator L is a linear differential operator acting on the space of twice differentiable functions. The properties of this operator, particularly under appropriate boundary conditions, are crucial for understanding the characteristics of the eigenfunctions and eigenvalues, and their role in forming orthogonal bases. In particular, the operator L can be shown to be formally self-adjoint under suitable conditions, which is a key property that ensures the eigenvalues are real and eigenfunctions corresponding to distinct eigenvalues are orthogonal, as discussed in Chapter 2. This orthogonality is the cornerstone for constructing orthogonal bases, which is of central interest in the context of CT-DT transformations.

To analyze whether the Sturm-Liouville operator is self-adjoint, a standard procedure is constructed through integration by parts and by imposing suitable boundary conditions.

Specifically, given two functions $f(x), g(x) \in \mathcal{L}^2[a,b]$, integrating by parts yields

$$\int_a^b f(x)Lg(x)dx = \int_a^b g(x)Lf(x)dx + \left[p(x)\left(g(x)\frac{d}{dx}f(x) - f(x)\frac{d}{dx}g(x)\right)\right]_a^b. \tag{4.4}$$

For Hermiticity, the boundary term must vanish, which is achieved through appropriate boundary conditions,

$$\left[p(x)\left(g(x)\frac{d}{dx}f(x) - f(x)\frac{d}{dx}g(x)\right)\right]_a^b = 0. \tag{4.5}$$

In that case, it is known that the eigenvalue problem possesses a discrete spectrum. Furthermore, for eigenfunctions $f_1(x)$ and $f_2(x)$ corresponding to eigenvalues $\lambda_1 \neq \lambda_2$, as shown in equation (2.51), their associated eigenfunctions are orthogonal. Eigenfunctions corresponding to the same eigenvalue are not necessarily unique, but orthogonalization procedures can be applied to form an orthonormal set. In general, for the eigenfunctions $\{f_n(x)\}$, the orthogonality condition can be expressed as:

$$\int_a^b f_n(x)f_m(x)w(x)dx = d_n^2\delta_{nm}, \tag{4.6}$$

where d_n^2 is the normalization constant and δ_{nm} is the Kronecker delta.

4.2. Recasting second-order eigenvalue problems

A general second-order eigenvalue problem can be written as

$$p_0(x)f''(x) + p_1(x)f'(x) - q_0(x)f(x) = -\lambda f(x). \tag{4.7}$$

In general, the differential operator in this form is not immediately recognizable as being of the Sturm-Liouville type. However, it can be transformed into the Sturm-Liouville form by introducing a suitable weighting factor $w(x)$. Multiplying equation (4.7) by $w(x)$ and requiring that $w(x)$ satisfies

$$\frac{d}{dx}(w(x)p_0(x)) = w(x)p_1(x), \tag{4.8}$$

leads to

$$w(x)p_0(x)f''(x) + \frac{d}{dx}(w(x)p_0(x))f'(x) - w(x)q_0(x)f(x) = -\lambda w(x)f(x). \tag{4.9}$$

This can be rewritten as

$$\frac{d}{dx}(w(x)p_0(x)f'(x)) - w(x)q_0(x)f(x) = -\lambda w(x)f(x), \tag{4.10}$$

or in the standard Sturm-Liouville form

$$-\frac{d}{dx}(p(x)f'(x)) + q(x)f(x) = \lambda w(x)f(x), \tag{4.11}$$

by defining $p(x) = p_0(x)w(x)$ and $q(x) = q_0(x)w(x)$.

The weighting function $w(x)$ can be derived by solving equation (4.8) as

$$w(x) = \frac{1}{p_0(x)} \exp\left(\int \frac{p_1(y)}{p_0(y)} dy\right). \tag{4.12}$$

In consequence, any equation with the form of (4.7) and satisfying some non-restrictive conditions[1] is a Sturm-Liouville problem. This new formulation, however, is more convenient, as both Laguerre and Hermite equations can be written in the form of (4.7). Note that this construction makes the weight function implicit, and is the one for which the associated basis is orthogonal. Let the differentiation operator

$$L' = p_0(x)\frac{d^2}{dx^2} + p_1(x)\frac{d}{dx} - q_0(x). \tag{4.13}$$

The eigenvalue problem can be written as

$$L'f(x) = -\lambda f(x). \tag{4.14}$$

For the eigenfunctions $f(x)$ to be polynomials, which is often of interest for practical computations, the functions $p_0(x), p_1(x)$, and $q_0(x)$ must also be polynomials. To prove this, let $f(x)$ be an n-th order polynomial,

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_n. \tag{4.15}$$

---

[1] Due to the definition of $w(x)$in (4.8), it is clear that $p_0(x)$ should be differentiable. For $w(x)$ being finite and continuous, $p_1(x)$ and $dp_0(x)/dx$ must be continuous. Requiring $w(x)$ to be strictly positive yields $p_0(x) > 0$. Note that these conditions coincide with the standard requirements introduced at the beginning of the chapter.

Introducing $f(x)$ in equation (4.14),

$$\begin{aligned} & p_0(x)\left(a_0 n(n-1)x^{n-2} + a_1(n-1)(n-2)x^{n-3} + \ldots + a_{n-2}\right) \\ + \; & p_1(x)\left(a_0 n x^{n-1} + a_1(n-1)x^{n-2} + \ldots + a_{n-1}\right) \\ - \; & q_0(x)\left(a_0 x^n + a_1 x^{n-1} + \ldots + a_n\right) \\ = \; & -\lambda\left(a_0 x^n + a_1 x^{n-1} + \ldots + a_n\right). \end{aligned} \tag{4.16}$$

For $f(x)$, $f'(x)$ and $f''(x)$ to be nontrivial, $p_0(x)$, $p_1(x)$ and $q_0(x)$ must be polynomials. Note that any non-polynomial part in the left-hand side must sum up to zero. Removing the homogeneous part of equation (4.14), the remaining particular solution satisfies the polynomial form. Specifically, $p_0(x)$ is at most quadratic, $p_1(x)$ is at most linear, and $q_0(x)$ is a constant. For simplicity in the following derivations, $q_0(x)$ will be set to 0, noting that a constant $q_0(x)$ only introduces a shift in the eigenvalues λ. This is the most general form of a Sturm-Liouville problem returning polynomials.

4.3. Rodrigues' Formula and Operator Derivation

Under the assumption that $p_0(x), p_1(x)$, and $q_0(x)$ satisfy the polynomial conditions mentioned above, the eigenfunctions $f_n(x)$ of the eigenvalue problem can be represented by Rodrigues' formula [26]

$$f_n(x) = \frac{c_n}{w(x)}\frac{d^n}{dx^n}(w(x)p_0^n(x)), \tag{4.17}$$

where c_n is a normalization constant. Differentiating this expression, a recursive relation can be derived. In particular,

$$\frac{d}{dx}\frac{d^n}{dx^n}(w(x)p_0^{n+1}(x)) = \frac{d^n}{dx^n}\left(\frac{d}{dx}(w(x)p_0(x))p_0^n(x) + nw(x)p_0^n(x)p_0'(x)\right). \tag{4.18}$$

Imposing equation (4.8),

$$\frac{d}{dx}\frac{d^n}{dx^n}(w(x)p_0^{n+1}(x)) = \frac{d^n}{dx^n}[wp_0^n(p_1(x) + np_0'(x))]. \tag{4.19}$$

Let $\pi_n(x) = p_1(x) + np_0'(x)$. Note that $\pi_n(x)$ is a first degree polynomial. The following equality was obtained to leverage this structure.

Theorem 4.1. *Let a general class $\mathcal{C}^n$ function $f(x)$. Then, the following holds:*

$$\frac{d^n}{dx^n}[f(x)(ax+b)] = \frac{d^n f(x)}{dx^n}(ax+b) + \frac{d^{n-1}f(x)}{dx^{n-1}}an. \tag{4.20}$$

Proof. This equality, which was found by inspection, can be proved by mathematical induction for the derivative order. The base case is given by the first order derivative,

$$\frac{d}{dt}[f(x)(ax+b)] = \frac{df(x)}{dx}(ax+b) + f(x)a, \tag{4.21}$$

which is consistent with Theorem 4.1. For the induction step, let

$$\frac{d^{n+1}}{dx^{n+1}}[f(x)(ax+b)] = \frac{d}{dx}\frac{d^n}{dx^n}[f(x)(ax+b)] = \frac{d}{dx}\left(\frac{d^n f(x)}{dx}(ax+b) + \frac{d^{n-1}f(x)}{dx^{n-1}}an\right). \tag{4.22}$$

Expanding the derivative and rearranging terms,

$$\frac{d^{n+1}}{dx^{n+1}}[f(x)(ax+b)] = \frac{d^{n+1}f(x)}{dx^{n+1}}(ax+b) + \frac{d^n f(x)}{dx^n}a(n+1), \tag{4.23}$$

which is the result to prove. □

With the previous result, one can expand

$$\frac{d^n}{dx^n}[w(x)p_0^n(x)\pi_n(x)] = \frac{d^n}{dx^n}(w(x)p_0^n(x)) + \frac{d^{n-1}}{dx^{n-1}}(w(x)p_0^n(x))n\pi_n'. \tag{4.24}$$

Function π_n' does not depend on x, as it is constant by construction. For the last term, using the linear eigenproblem (4.14) as an identity and substituting Rodrigues' formula,

$$\begin{aligned} p_0(x)\frac{d^2}{dx^2}\left(\frac{c_n}{w(x)}\frac{d^n}{dx^n}(w(x)p_0(x))\right) + p_1(x)\frac{d}{dx}\left(\frac{c_n}{w(x)}\frac{d^n}{dx^n}(w(x)p_0(x))\right)& \\ = -\lambda_n\left(\frac{c_n}{w(x)}\frac{d^n}{dx^n}(w(x)p_0(x))\right)&. \end{aligned} \tag{4.25}$$

This only holds if the eigenvalues verify the following identity,

$$\lambda_n = -n\left(p_1' + \frac{n-1}{2}p_0''\right). \tag{4.26}$$

Substituting this result back to equation (4.24), together with (4.17), yields

$$\frac{c_n}{c_{n+1}}f_{n+1}(x) = f_n(x)\pi_n(x) - f_n'(x)\frac{n\pi_n'}{\lambda_n}p_0(x). \tag{4.27}$$

This recurrence relation allows for the definition of rising and lowering operators. In particular, let the general rising operator for polynomials,

$$f_{n+1}(x) = \frac{c_{n+1}}{c_n}\frac{n\pi_n'}{\lambda_n}\left(\frac{\lambda_n \pi_n(x)}{n\pi_n'} + p_0(x)\frac{d}{dx}\right) f_n(x). \tag{4.28}$$

For normalized functions $u_n(x) = \frac{1}{d_n}\sqrt{w(x)} f_n(x)$, where d_n is a normalization factor, differentiation leads to

$$d_n u_n'(x) = \sqrt{w(x)} f_n'(x) + \frac{w'(x)}{2\sqrt{w(x)}} f_n(x) = \sqrt{w(x)}\left(f_n'(x) + \frac{p_1(x) - p_0'(x)}{2p_0(x)} f_n(x)\right). \tag{4.29}$$

Consequently, a substitution can be performed such that

$$\sqrt{w(x)p_0(x)} f_n'(x) = d_n u_n'(x) p_0(x) - \frac{p_1(x) - p_0'(x)}{2} d_n u_n(x). \tag{4.30}$$

Substituting back the definition of the normalized basis $u_n(x)$, a rising operator can be defined as

$$\frac{d_{n+1}c_n}{d_n c_{n+1}} u_{n+1}(x) = \frac{n\pi_n'}{\lambda_n}\left(\frac{\lambda_n \pi_n(x)}{n\pi_n'} - \frac{p_1(x) - p_0'(x)}{2} - p_0(x)\frac{d}{dx}\right) u_n(x) \doteq a^\dagger u_n(x). \tag{4.31}$$

Similarly, from the structure of the rising operator, it preserves the inner product,

$$\left\langle \frac{d_n c_{n+1}}{d_{n+1}c_n} a^\dagger u_n, \frac{d_m c_{m+1}}{d_{m+1}c_m} a^\dagger u_m \right\rangle = \langle u_{n+1}, u_{m+1}\rangle = \delta_{nm} = \langle u_n, u_m\rangle, \tag{4.32}$$

where orthogonality was used. Thus, a lowering operator can be constructed as its adjoint, which coincides with its inverse. Doing so[2] yields

$$\begin{aligned} &\frac{(n-1)\pi_{n-1}'}{\lambda_{n-1}}\left(\frac{\lambda_{n-1}\pi_{n-1}(x)}{(n-1)\pi_{n-1}'} - \frac{p_1(x) - p_0'(x)}{2} + \left(p_0'(x) + p_0(x)\frac{d}{dx}\right)\right) u_n(x) \\ &= \frac{d_n c_{n-1}}{d_{n-1}c_n} u_{n-1}(x) \doteq a u_n(x). \end{aligned} \tag{4.33}$$

For orthonormal bases, the coefficients c_n appearing in (4.17) can be obtained from their definition as

$$c_n = \left[\int_a^b \frac{1}{w(x)}\left(\frac{d^n}{dx^n}\left(w(x)p_0^n(x)\right)\right)^2 dx\right]^{-1/2}. \tag{4.34}$$

---

[2] Just by integrating $\langle u_n | a^\dagger | u_m\rangle$ by parts, as done to prove that Sturm-Liouville operators are self-adjoint. In particular, the standard definition $\langle u_n | a^\dagger u_m\rangle \doteq \langle a u_n | u_m\rangle$ is used.

The above results are sufficient if $p_0(x)$ is a constant[3]. Indeed, in that case, $p_0'(x)$ vanishes, and the derivative operator times a constant can directly be computed as a linear combination of $a^\dagger$ and a. For general cases, however, another tool is required.

4.4. General Derivative Operator for Orthogonal Functions

If $p_0(x)$ is a first or second degree polynomial, the derivative operator d/dx cannot obtained explicitly from the generalized creation and annihilation operators, $a^\dagger$ and a. In particular, a linear combination of d/dx, xd/dx and $x^2 d/dx$, given by the structure of $p_0(x)$, is derived. The following recurrence relation holds for orthogonal polynomials (see, e.g. [26]),

$$x f_n(x) = \alpha_n f_{n+1}(x) + \beta_n f_n(x) + \alpha_{n-1} f_{n-1}(x), \tag{4.35}$$

where

$$\alpha_n = \langle x f_n, f_{n+1} \rangle \tag{4.36}$$

$$\beta_n = \langle x f_n, f_n \rangle. \tag{4.37}$$

By applying this relation twice,

$$\begin{aligned} x^2 f_n(x) &= \alpha_n x f_{n+1}(x) + \beta_n x f_n(x) + \alpha_{n-1} x f_{n-1}(x) \\ &= \alpha_n \left(\alpha_{n+1} f_{n+2}(x) + \beta_{n+1} f_{n+1}(x) + a_n f_n(x)\right) \\ &+ \beta_n \left(\alpha_n f_{n+1}(x) + \beta_n f_n(x) + \alpha_{n-1} f_{n-1}(x)\right) \\ &+ \alpha_{n-1} \left(\alpha_{n-1} f_n(x) + \beta_{n-1} f_{n-1}(x) + \alpha_{n-2} f_{n-2}(x)\right) \\ &= \alpha_n \alpha_{n+1} f_{n+2}(x) + (\alpha_n \beta_{n+1} + \alpha_n \beta_n) f_{n+1}(x) \\ &+ (\alpha_n^2 + \beta_n^2 + \alpha_{n-1}^2) f_n(x) + (\alpha_{n-1}\beta_n + \alpha_{n-1}\beta_{n-1}) f_{n-1}(x) \\ &+ \alpha_{n-1}\alpha_{n-2} f_{n-2}(x). \end{aligned} \tag{4.38}$$

The previous results can be conveniently recast in matrix form such that

$$x\underline{\mathrm{f}} = \underline{\mathrm{f}} \begin{bmatrix} \beta_0 & \alpha_0 & 0 & 0 & \cdots \\ \alpha_0 & \beta_1 & \alpha_1 & 0 & \cdots \\ 0 & \alpha_1 & \beta_2 & \alpha_2 & \cdots \\ 0 & 0 & \alpha_2 & \beta_3 & \ddots \\ \vdots & \vdots & \vdots & \ddots & \ddots \end{bmatrix} = \underline{\mathrm{fX}}, \tag{4.39}$$

---

[3] As discussed next, this is what occurs for the Hermite basis.

as well as

$$x^2\underline{\mathrm{f}} = \underline{\mathrm{f}}\begin{bmatrix} \beta_0^2+\alpha_0^2 & \alpha_0\beta_1+\alpha_0\beta_0 & \alpha_0\alpha_1 & 0 & \cdots \\ \alpha_0\beta_1+\alpha_0\beta_0 & \beta_1^2+\alpha_1^2+\alpha_0^2 & \alpha_1\beta_2+\alpha_1\beta_1 & \alpha_1\alpha_2 & \cdots \\ \alpha_0\alpha_1 & \alpha_1\beta_2+\alpha_1\beta_1 & \beta_2^2+\alpha_2^2+\alpha_1^2 & \alpha_2\beta_3+\alpha_2\beta_2 & \cdots \\ 0 & \alpha_1\alpha_2 & \alpha_2\beta_3+\alpha_2\beta_2 & \beta_3^2+\alpha_3^2+\alpha_2^2 & \ddots \\ \vdots & \vdots & \vdots & \ddots & \ddots \end{bmatrix} = \underline{\mathrm{f}}\underline{\mathrm{X}}^2. \quad (4.40)$$

Notably, this result also holds for orthogonal functions $u_n(x)$, with the corresponding normalization coefficients. As $p_0(x)$ is at most second degree in x, the previous expressions characterize its multiplying effect on the basis. Conveniently, the x and x^2 only connect a state to its 1 and 2 neighbors, respectively. Furthermore, note that, defining $\hat{a}^\dagger$ as the index rising and $\hat{a}$ as the index decreasing operators,

$$\frac{c_n d_{n+1}}{c_{n+1}d_n}\frac{\lambda_n}{n\pi_n'}\hat{a}^\dagger - \frac{c_{n-1}d_n}{c_n d_{n-1}}\frac{\lambda_{n-1}}{(n-1)\pi_{n-1}'}\hat{a} = p_0'(x) + 2p_0(x)\frac{d}{dx} + \frac{\lambda_n\pi_n(x)}{n\pi_n'} - \frac{\lambda_{n-1}\pi_{n-1}(x)}{(n-1)\pi_{n-1}'}. \quad (4.41)$$

Therefore, in general,

$$2p_0(x)\frac{d}{dx} = \frac{c_n d_{n+1}}{c_{n+1}d_n}\frac{\lambda_n}{n\pi_n'}\hat{a}^\dagger - \frac{c_{n-1}d_n}{c_n d_{n-1}}\frac{\lambda_{n-1}}{(n-1)\pi_{n-1}'}\hat{a} - \frac{\lambda_n\pi_n(x)}{n\pi_n'} + \frac{\lambda_{n-1}\pi_{n-1}(x)}{(n-1)\pi_{n-1}'} - p_0'(x). \quad (4.42)$$

This is the major result of this chapter. The derivative operator d/dt in a basis space of orthonormal functions is explicitly given by a linear combination of the index rising $\hat{a}^\dagger$ and index lowering $\hat{a}$ operators. If $p_0(x)$ is linear, the quotient $\frac{\lambda_n}{np'}$, together with $\pi_n'(x)$, do not depend on n, allowing to write,

$$2p_0(x)\frac{d}{dx} = \frac{\lambda_n}{n\pi_n'}\left(\frac{c_n d_{n+1}}{c_{n+1}d_n}\hat{a}^\dagger - \frac{c_{n-1}d_n}{c_n d_{n-1}}\hat{a}\right) - p_0'(x). \quad (4.43)$$

Function $p_0'(x)$ is at most linear in x, and can be written in terms of the $\underline{\mathrm{X}}$ operator. Furthermore, $p_0(x)$ can be written, at worst, as a linear combination of $\underline{\mathrm{X}}$ and $\underline{\mathrm{X}}^2$, which can be inverted or multiplied to all the dynamic equations to obtain an equivalent formulation. This expression relates the derivative operator in CT to discrete operators on DT, providing a generalized framework for derivative transformation in orthogonal bases derived from Sturm-Liouville problems.

To better understand the obtained results, they are applied to standard eigenbases. A systematic approach to the construction of bases and the discretization of the d/dt operator is provided.

4.5. Particular Cases

The general framework derived from Sturm-Liouville theory can be particularized to recover well-known orthogonal function systems like Hermite, Laguerre, and Legendre functions by choosing specific forms for $p_0(x)$ and $p_1(x)$.

4.5.1. Hermite Functions

For Hermite polynomials, the generating equation is derived from setting $p_0(x) = 1$ and $p_1(x) = -x$. The eigenvalue problem (4.7) becomes Hermite's equation,

$$\frac{1}{2}f'' - xf' = -\lambda f = -nf. \tag{4.44}$$

where the dependence of eigenvalues with the basis index is as (4.26). Using (4.12), the weight function can be computed as

$$w(x) = \exp\left(\int -y dy\right) = \exp\left(-x^2\right), \tag{4.45}$$

The requirement of Hermiticity fixes the domain at which the basis is defined. In particular, the boundary condition (4.5) must be imposed. Due to the definition of $p(x)$ as $p_0(x)w(x)$, this product must be zero at the boundaries of the interval where the basis is constructed. In this case, as $p_0(x) = 1$ and $w(x)$ is zero only at $x \to \pm\infty$, functions must be defined at $\mathbb{R}$, as expected. Note that the integration constant can be set to any value, as it is compensated by the weight coefficient d_n. In particular, as given in (4.6),

$$d_n = \sqrt{\int_{\mathbb{R}} f_n(x)^2 w(x) dx} = \sqrt{\sqrt{\pi} 2^n n!}. \tag{4.46}$$

The Rodrigues' coefficients are computed as (4.34),

$$c_n = \left[\int_{\mathbb{R}} \exp\left(x^2\right) \left(\frac{d^n}{dx^n} \exp\left(x^2\right)\right)^2 dx\right]^{-1/2} = (-1)^n. \tag{4.47}$$

These are the ingredients necessary to compute the derivative operator in the Hermite domain as given by (4.43). In particular, substituting the previous results,

$$\frac{d}{dx} = \frac{1}{\sqrt{2}}(\sqrt{n}\hat{a} - \sqrt{n+1}\hat{a}^{\dagger}). \tag{4.48}$$

This derivative operator is consistent with the quantum harmonic oscillator analogy, but has been computed from a completely general framework.

4.5.2. Laguerre Functions

For Laguerre polynomials, the generating equation, namely, the Laguerre equation, is obtained by particularizing (4.7) with $p_0(x) = x$ and $p_1(x) = \alpha + 1 - x$,

$$xf'' + (\alpha + 1 - x)f' = -\lambda f = -nf. \tag{4.49}$$

The eigenvalue is related to the basis index through (4.26). The weight function is computed as (4.12), yielding

$$w(x) = \frac{1}{x}\exp\left(\int \frac{\alpha+1-y}{y}dy\right) = x^{\alpha}\exp(-x). \tag{4.50}$$

For the Sturm-Liouville operator to be Hermitian, (4.5) must be satisfied. In this case, the boundaries are clearly $x = 0$ and $x \to +\infty$, as the only zeros of $p(x)$. The weight coefficients are then given by (4.6)

$$d_n = \sqrt{\int_{\mathbb{R}^+} f_n(x)^2 w(x)dx} = \sqrt{\frac{\Gamma(\alpha+n+1)}{n!}} \overset{\alpha\in\mathbb{N}}{=} \sqrt{\frac{(\alpha+n)!}{n!}}. \tag{4.51}$$

Note that the last equality is true only if $\alpha \in \mathbb{N}$, as for that condition the Gamma function behaves as a factorial. The Rodrigues' coefficients can then be computed as (4.34)

$$c_n = \left[\int_{\mathbb{R}^+} x^{-\alpha}\exp(x)\left(\frac{d^n}{dx^n}\left(x^{\alpha+n}\exp(-x)\right)\right)^2 dx\right]^{-1/2} = \frac{1}{n!}. \tag{4.52}$$

Finally, the derivative operator is computed as (4.43), giving

$$x\frac{d}{dx} = \frac{1}{2}\left(\sqrt{\alpha+n+1}\sqrt{n+1}\hat{a}^{\dagger} - \sqrt{\alpha+n}\sqrt{n}\hat{a} - 1\right). \tag{4.53}$$

Note that a multiplying x term is present in the derivative. Observing the structure of (4.39), inverting this term introduces a filter-like behavior, as observed in practice. Furthermore, a constant term $-1/2$ appears in the right-hand side. This explains the fact that an independent term must be added in the CT-PBSID when writing down the discrete dynamics (see [1]). This result is consistent with what is obtained through system lifting, and was verified computationally.

4.5.3. Legendre Functions

For Legendre polynomials, the generating equation, known as Legendre equation, is derived from $p_0(x) = 1 - x^2$ and $p_1(x) = -2x$,

$$(1-x^2)f''(x) - 2xf'(x) = -\lambda f(x) = -n(n+1)f(x). \tag{4.54}$$

Note that, in this case, $p_0(x)$ is not linear but quadratic, forcing to use (4.42) as the differential operator definition. Additionally, this quadratic term modifies the dependence between λ and n, as observed in (4.26). The weight function given by (4.12),

$$w(x) = \frac{1}{1-x^2}\exp\left(\int -\frac{2y}{1-y^2}dy\right) = 1. \tag{4.55}$$

Imposing (4.5) for the Sturm-Liouville to be self-adjoint, $p_0(x)$ must be zero at the boundaries of the interval. Thus, the basis is constructed in $[-1, 1]$. For this domain, the weight coefficients are computed as (4.6)

$$d_n = \sqrt{\int_{-1}^{1} f_n(x)^2 w(x)dx} = \sqrt{\frac{2}{2n+1}}. \tag{4.56}$$

Furthermore, the Rodrigues' coefficients are recovered from (4.34) as

$$c_n = \sqrt{\int_{-1}^{1}\left(\frac{d^n}{dx^n}(1-x^2)^n\right)^2 dx} = \frac{(-1)^n}{2^n n!}. \tag{4.57}$$

From these results, the derivative operator is given by (4.42), that is,

$$2(1-x^2)\frac{d}{dx} = n\sqrt{\frac{2n+1}{2n-1}}\hat{a} - (n+1)\sqrt{\frac{2n+1}{2n+3}}\hat{a}^\dagger + x. \tag{4.58}$$

In this case a second degree term appears multiplying d/dx, behaving like a second order filter. Furthermore, a linear term on x is present in the right-hand side. This result was verified computationally.

4.5.4. General approach

After showing the procedure for obtaining a derivative operator in the bases of Hermite, Laguerre and Legendre functions, a general algorithm can be constructed to extend these results to any basis. In particular, one can proceed as follows:

1. Generating equation. Define the generating ODE as (4.7), with $p_0(x)$ at most quadratic and $p_1(x)$ at most linear polynomials of x.
2. Compute the relationship between the eigenvalue λ and the index n as (4.26).
3. Compute the weight function $w(x)$ from (4.12).
4. Define the basis domain by imposing (4.5).
5. Compute the orthonormality weight coefficients d_n from (4.6).
6. Compute the Rodrigues' coefficients c_n as (4.34).
7. Apply (4.42) to compute the derivative operator in the basis domain. If $p_0(x)$ is at most linear with x, (4.43) provides a simpler method.
8. If x or x^2 appear in the definition of the derivative operator, they can be expressed as linear operators in the basis space as (4.39) and (4.40), respectively.

This method is general, and provides orthogonal bases and their associated differentiation operators systematically.

4.6. Classification of solutions

The possible expressions for $p_0(x)$ and $p_1(x)$, compatible with the boundary conditions in (4.5), allow to introduce an interesting classification of bases. In particular, $p_0(x)$ can be a linear or a quadratic function on x. Let $p_0(x)$ be a linear function. Hence, in general, $p_0(x) = a_0 x + b_0$ and $p_1(x) = a_1 x + b_1$. If $a_0 = 0$, up to a scale factor, $p_0(x) = 1$, and therefore

$$w(x) = \exp\left(\int (a_1 y + b_1) dy\right) = \exp\left(\frac{a_1}{2}x^2 + b_1 x\right). \tag{4.59}$$

Imposing the Hermitian condition on the Sturm-Liouville operator,

$$[p_0(x)w(x)]_a^b = \left[\exp\left(\frac{a_1}{2}x^2 + b_1 x\right)\right]_a^b = 0. \tag{4.60}$$

This condition is satisfied for any $a_1 < 0, b_1$, as long as $a \to -\infty$ and $b \to +\infty$. In this work, this family of solutions are called Hermite-like.

If $a_0 \neq 0$, up to a scale factor, $p_0(x) = x + b_0$, so that

$$w(x) = \frac{1}{x+b_0} \exp\left(\int \frac{a_1 y + b_1}{y + b_0} dy \right) = (x+b_0)^{b_1 - a_1 b_0 - 1} \exp\left(a_1(x+b_0)\right). \tag{4.61}$$

Imposing that the Sturm-Liouville operator is self-adjoint,

$$[p_0(x)w(x)]_a^b = \left[(x+b_0)^{b_1 - a_1 b_0} \exp\left(a_1(x+b_0)\right)\right]_a^b = 0. \tag{4.62}$$

This equality holds as long as $b \to \infty$, for $a_1 < 0$ and $a = -b_0$. These functions are denoted by Laguerre-like.

If $p_0(x)$ is quadratic, standard solutions which can be obtained include the Legendre and Jacobi functions. The analysis of these families of functions, besides obtaining an expression for the derivative operator in the Legendre domain, is left out of the scope of this work. Note that this is an interesting line of future research.

From this Sturm-Liouville theory perspective, several important conclusions can be drawn for CT-DT transformations. Firstly, the terms involving x and x^2 in the differential equation act like inverse derivatives, implying that their presence in the system behaves like first and second-order filters, respectively. This is an interesting behavior, which explains the presence of Laguerre (first order) filters in the CT-PBSID. Within the framework of polynomial Sturm-Liouville eigenproblems, it was proved that Hermite, Laguerre, Legendre, and Jacobi-like solutions can be obtained as orthonormal bases. Hermite-like solutions are defined over the whole real line and do not inherently introduce filtering. Laguerre-like solutions are defined over an interval $[a, +\infty)$ and introduce first-order filtering characteristics. Finally, Legendre and Jacobi-like solutions, defined over $[-1, 1]$, introduce second-order filtering properties. These conclusions provide a theoretical foundation for understanding the filtering properties associated with different orthogonal bases used in CT-DT transformations and offer a systematic way to select and design appropriate bases for specific applications. From this perspective, previous to this work, a good understanding of Laguerre functions was provided through the CT-PBSID. This work introduces the HD-PBSID, which makes use of Hermite functions. An interesting topic of research is the extension of these methods for Legendre-like functions, Jacobi-like functions and other existing families arising from the proposed theoretical methodology.

5 | Simulation results and analysis

This section aims to verify and validate computationally the proposed identification algorithm. For this purpose, various simulation scenarios are used to benchmark the performance of the HD-PBSID with varying conditions, as well as to compare its behavior to the CT-PBSID.

5.1. Computational verification

In this section, the HD-PBSID is tested against a simple second order system, given by

$$A = \begin{bmatrix} -2 & 1 \\ 0 & -4 \end{bmatrix}, B = \begin{bmatrix} 0.5 \\ 1 \end{bmatrix}, C = \begin{bmatrix} 1 & 4 \\ 2 & 1 \end{bmatrix}, D = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

A noise component of varying Signal to Noise Ratio (SNR) is introduced. Note that the system is stable, with eigenvalues $\lambda_A = -2, -4$, and fully observable. The excitation is based on a linear sine sweep,

$$u(t) = \sin\left[2\pi\left(f_1 t + \frac{f_2 - f_1}{2T} t^2\right)\right], \tag{5.1}$$

where $T = 10$ s is the duration of the experiment, and where $f_1 = 0$ Hz and $f_1 = 8$ Hz are the minimum and maximum exciting frequencies, respectively. For a maximum expansion order of the Hermite basis n_{max}, this function is projected onto a basis up to order $n_{max,u} = \lfloor 0.8 n_{max} \rfloor$. Hence, the actual exciting function is given by

$$\hat{u}(t) = \sum_{i=0}^{n_{max,u}} h_n(t)\langle h_n(t), u(t)\rangle. \tag{5.2}$$

As discussed in Section 3.5.5, this improves the quality of the representation of the system in the Hermite domain. For this setup, various identification conditions were tested, in all cases through Monte Carlo simulations with varying seeds for the random number generator in the noise input.

Firstly, the effect of noise is analyzed for a default identification configuration. In particular, 5000 samples were obtained at constant rate during the 10 s of the experiment. A maximum Hermite order of $n_{max} = 300$ was considered. The differentiation coefficients β and γ were set to 3 and 20, respectively. In Figures 5.1–5.4, the bode diagrams of the identified systems, together with the real one, are displayed for varying SNR values. Similarly, the modes of the identified dynamics are compared to the exact ones in Figures 5.5–5.8. For low noise, all simulations converge closely to the real system, with a very good fitting for the bode diagram. As the SNR decreases, the identified systems deviate the actual dynamics, particularly at the extremes of frequency values. Note that high frequencies were not excited, as the sine sweep reaches at most 8 Hz. Longer periods to the duration of the experiment cannot be excited neither. This explains a better fitting for intermediate frequencies. Excluding a small set of experiments for very low SNR, almost all identified dynamics lead to real eigenvalues.

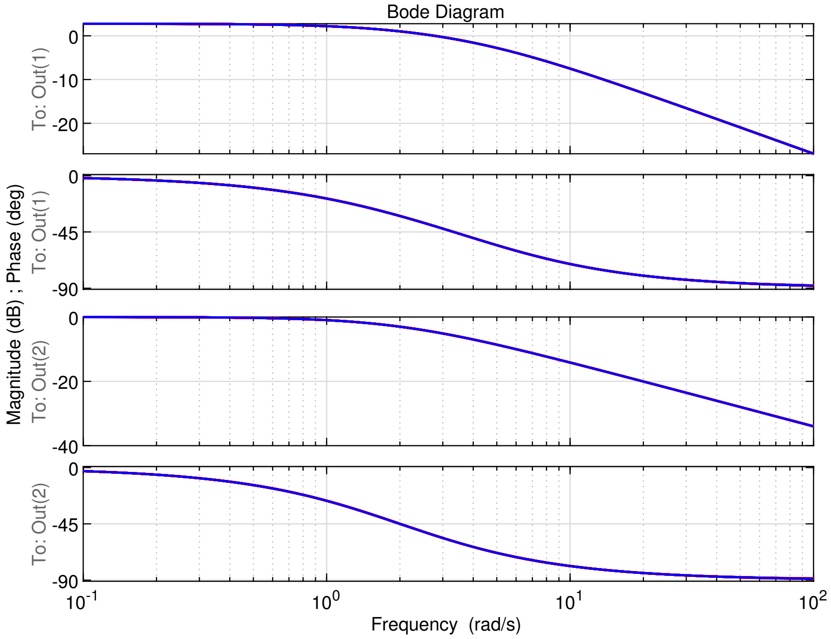


Figure 5.1: Bode diagram of the real system and of the HD-PBSID identified plant for SNR = 50.

Next, the effect of the parameters within the HD-PBSID was evaluated. In particular, four scenarios were considered. In all cases, a moderate-high noise with SNR of 20 was introduced. Firstly, the effect of the number of samples was studied considering values between 100 and 10000 samples, as depicted in Figure 5.9. As expected, more samples provide a better identification of the dynamics. Interestingly, the quality of the identified

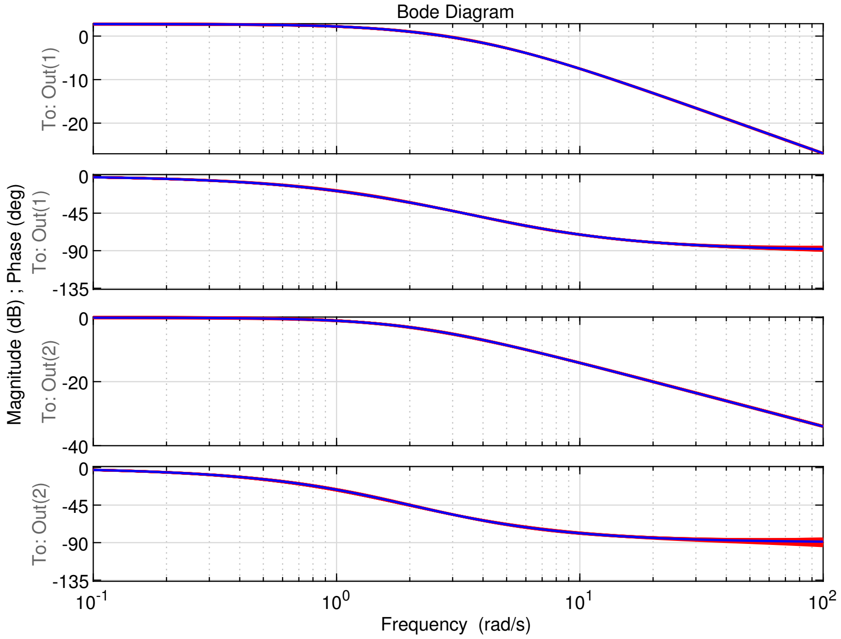


Figure 5.2: Bode diagram of the real system and of the HD-PBSID identified plant for SNR $= 20$.

system does not improve significantly for more than 2000 samples. Note that a very low sampling frequency not only decreases the amount of available information but also reduces the maximum frequency which can be identified.

In second place, the effect of the basis size was analyzed for expansions from order 10 to order 1000. In this case, an expansion order lower than 30 produces a very poor identification accuracy, while increasing the order over 50 does not improve the accuracy. This is what theory predicts: for low order, a significant information is lost when projecting on the Hermite domain. When the order reaches the maximum order excited by the system, increasing the size of the expansion only provides terms of negligible effect. As all functions, when projected to the Hermite domain, belong to ℓ^2, their coefficients must eventually fall to zero, allowing to always define a maximum order.

Finally, the effect of β and γ was studied. As discussed in Chapter 3, these parameters do not have a significant effect in the identification, as long as γ is large enough and does not compromise numerical stability. For varying β and $\gamma = 20$, both $-2/20 = -0.1$ and $-4/20 = -0.2$ are smaller than one, while there is not combination $(\lambda + \beta)/\gamma$ for the proposed values that is outside of the unit circle. For $\beta = 3$ and varying γ, the discretized eigenvalues are both, in absolute value, $1/\gamma$. Hence, the system is stable as long as $\gamma > 1$.

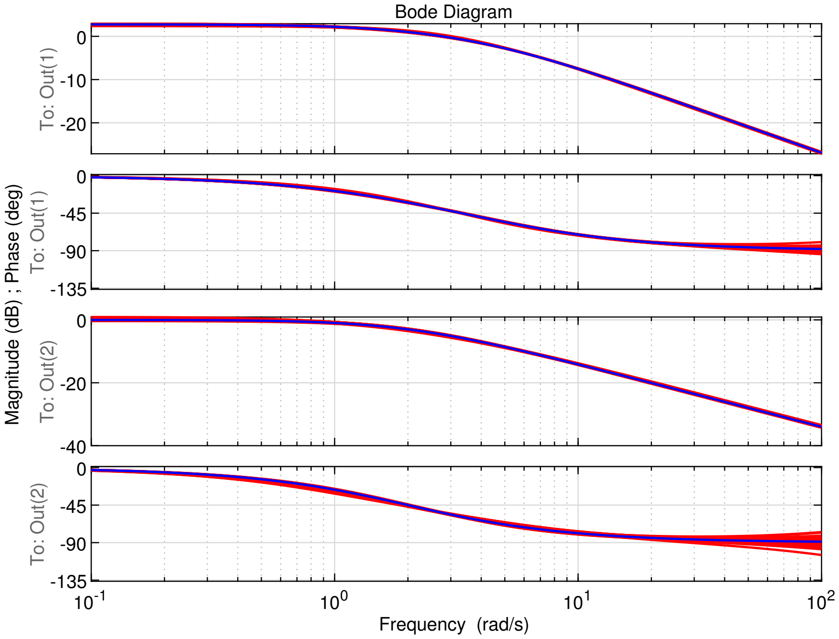


Figure 5.3: Bode diagram of the real system and of the HD-PBSID identified plant for SNR = 10.

Hence, all experiment provide essentially the same errors. A slight improvement in error is observed by increasing γ, as this reduces the bias of the estimator by making A'^n converge faster.

5.2. Comparison with the CT-PBSID

In this section, the controlled linear dynamics of a quadrotor are simulated and identified with both the CT-PBSID and the HD-PBSID for varying noise conditions. In particular, the uncontrolled dynamics are as

$$A = \begin{bmatrix} -0.1068 & 0.1192 & -9.81 \\ -5.9755 & -2.6478 & 0 \\ 0 & 1 & 0 \end{bmatrix}, B = \begin{bmatrix} -10.1647 \\ 450.71 \\ 0 \end{bmatrix},$$

$$C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ -0.1068 & 0.1192 & 0 \end{bmatrix}, D = \begin{bmatrix} 0 \\ 0 \\ 0 \\ -10.1647 \end{bmatrix}.$$

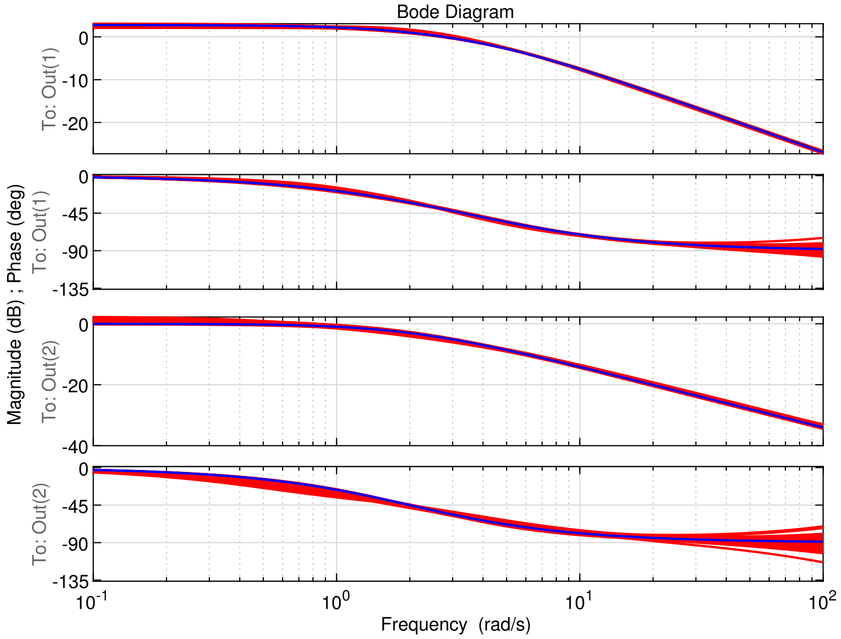


Figure 5.4: Bode diagram of the real system and of the HD-PBSID identified plant for SNR = 6.

A controller is used to stabilize the system. In particular, a constant gain is considered so that $u = Kx + u'$ for u' the exciting signal. The gain is given by

$$K = \begin{bmatrix} 0.0150 & -0.0091 & -0.0231 \end{bmatrix}.$$

The resulting stable plant has poles at -5 and $-1 \pm i$, and is the one to be identified. The excitation signal u' is constructed following the same approach used in the previous section. The HD-PBSID is constructed with a maximum expansion order of 250, for $\beta = 3$ and $\gamma = 30$. The CT-PBSID is characterized by $p = f = 10$ and a pole $a = 25$. A total of 5000 samples are obtained at constant frequency during a 10 s experiment. The results for varying SNR are depicted in Figures 5.13–5.18. Table 5.1 provides a quantitative summary of the comparison. For low noise, both algorithms successfully converge to the real system. However, for higher noise, the CT-PBSID fails to provide good estimations, exhibiting very large standard deviations and a significant bias. In contrast, the HD-PBSID provides accurate and unbiased estimates even for $SNR = 10$. For extremely high noise, the CT-PBSID is not able to provide even a reasonable identification, diverging at multiple cases, while the HD-PBSID still converges to a system with a similar structure to the actual plant. In general, the fact that the HD-PBSID originates from a rigorous non-

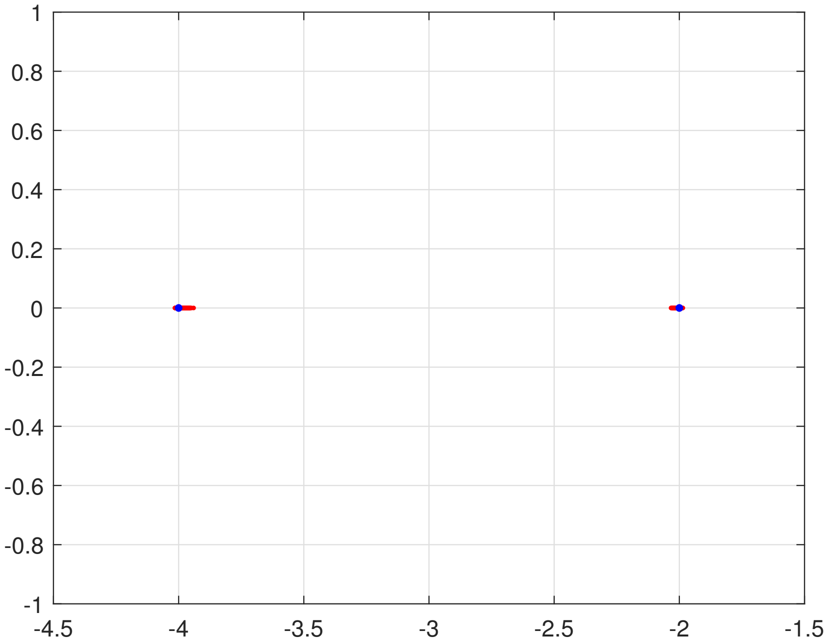


Figure 5.5: Eigenvalues of the real system and of the HD-PBSID identified plant for SNR = 50.

redundant approach to projection of signals and operators, together with the considered input setup based on a lower order Hermite expansion, seems to improve the accuracy of the identification in very noisy scenarios. This is an interesting property which is promising for practical applications.

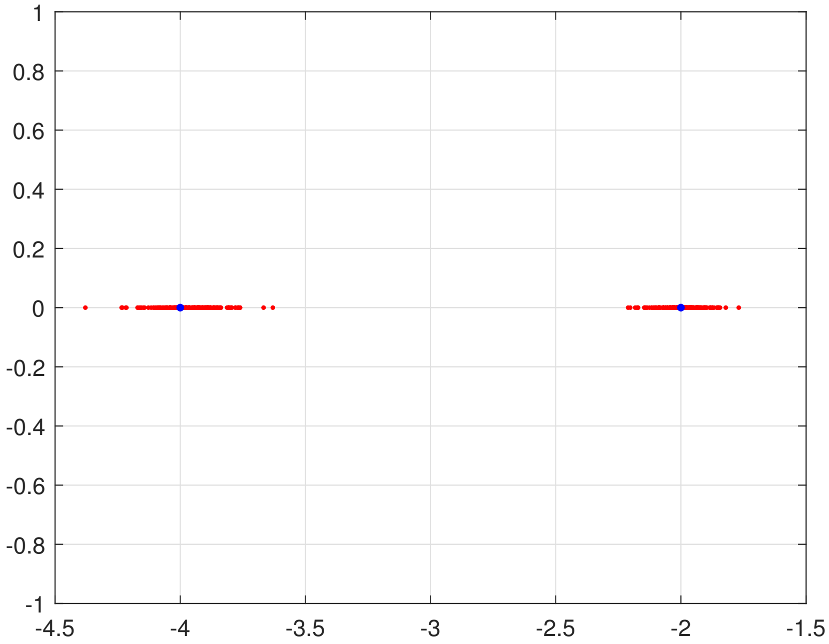


Figure 5.6: Eigenvalues of the real system and of the HD-PBSID identified plant for SNR $= 20$.

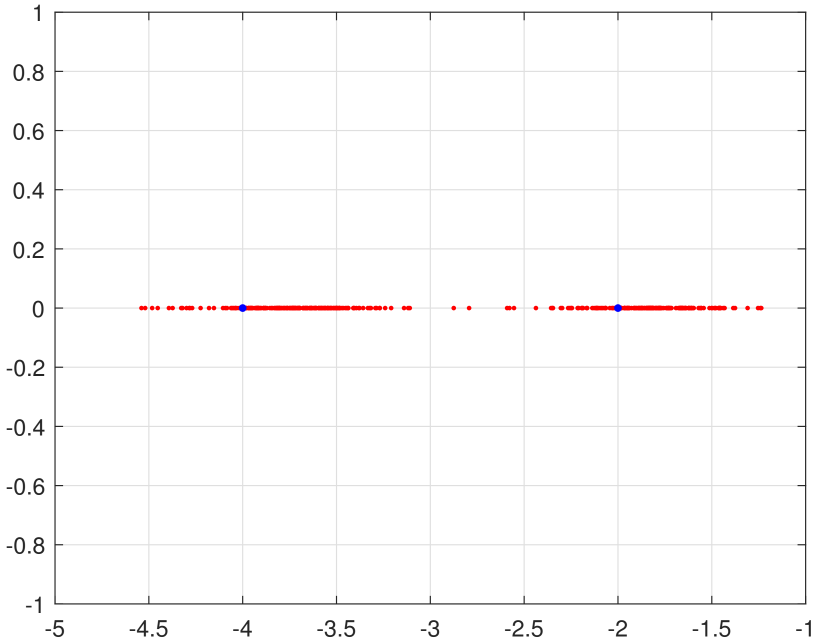


Figure 5.7: Eigenvalues of the real system and of the HD-PBSID identified plant for SNR $= 10$.

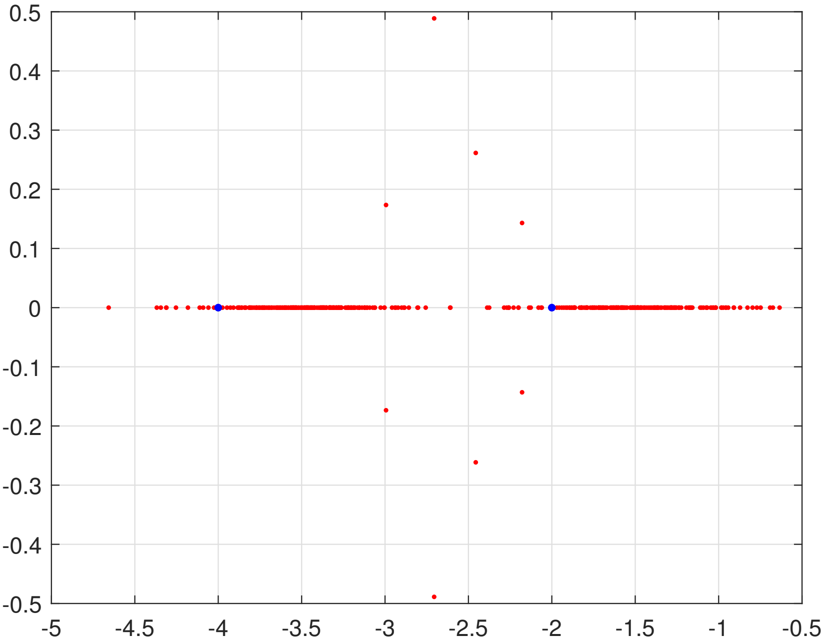


Figure 5.8: Eigenvalues of the real system and of the HD-PBSID identified plant for SNR = 6.

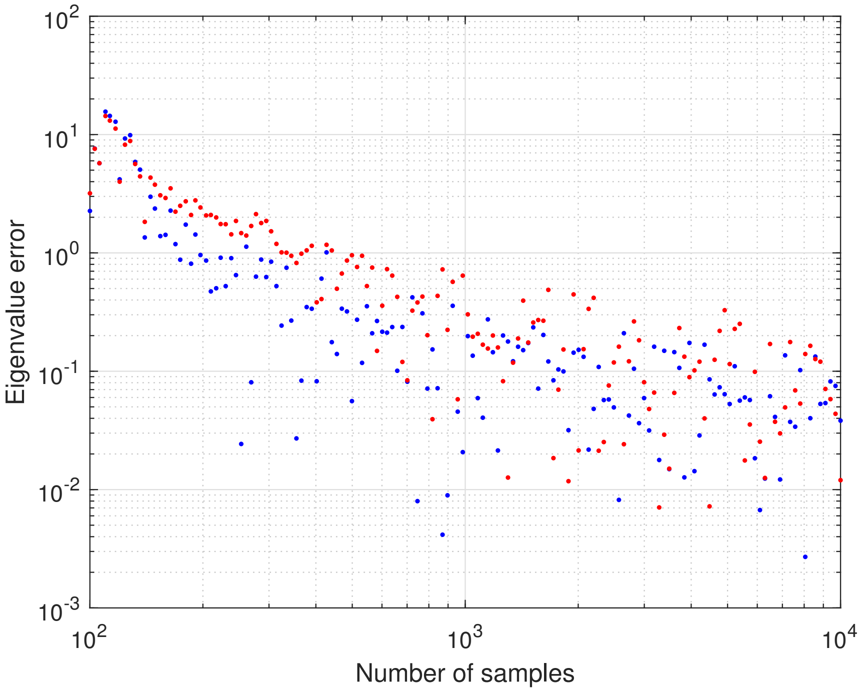


Figure 5.9: Error of the eigenvalues identified by the HD-PBSID for a varying number of samples.

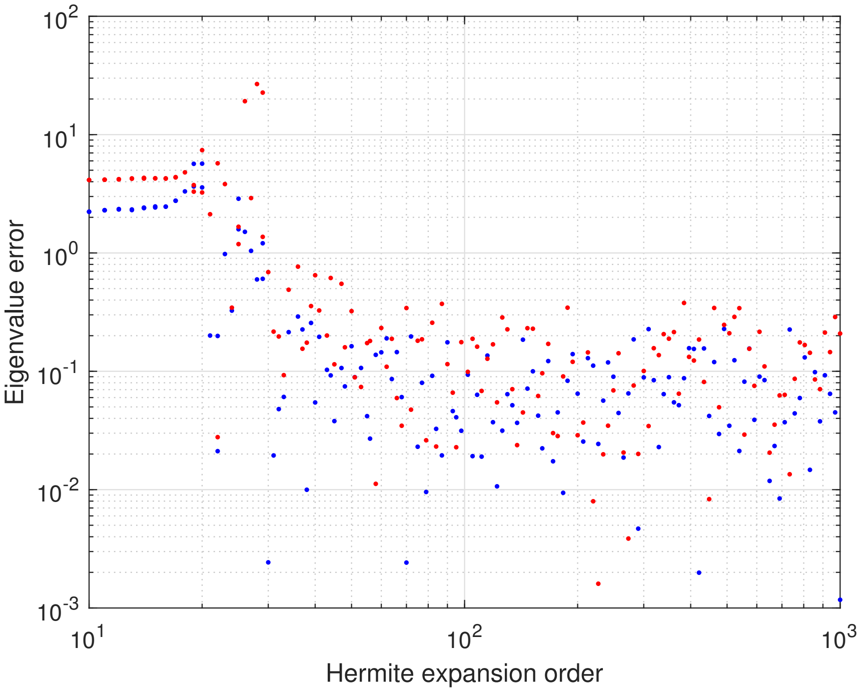


Figure 5.10: Error of the eigenvalues identified by the HD-PBSID for a varying Hermite basis order.

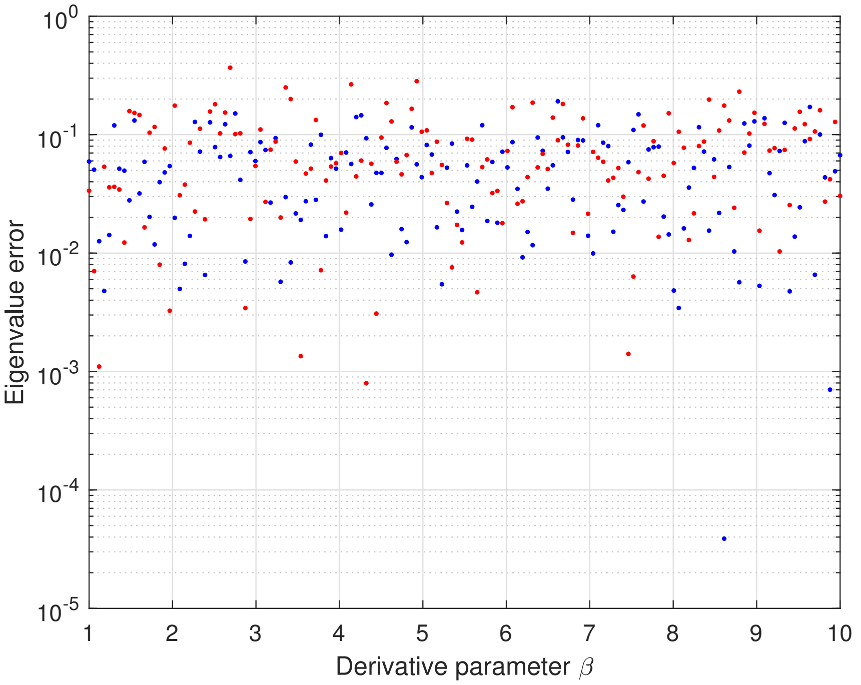


Figure 5.11: Error of the eigenvalues identified by the HD-PBSID for a varying β parameter.

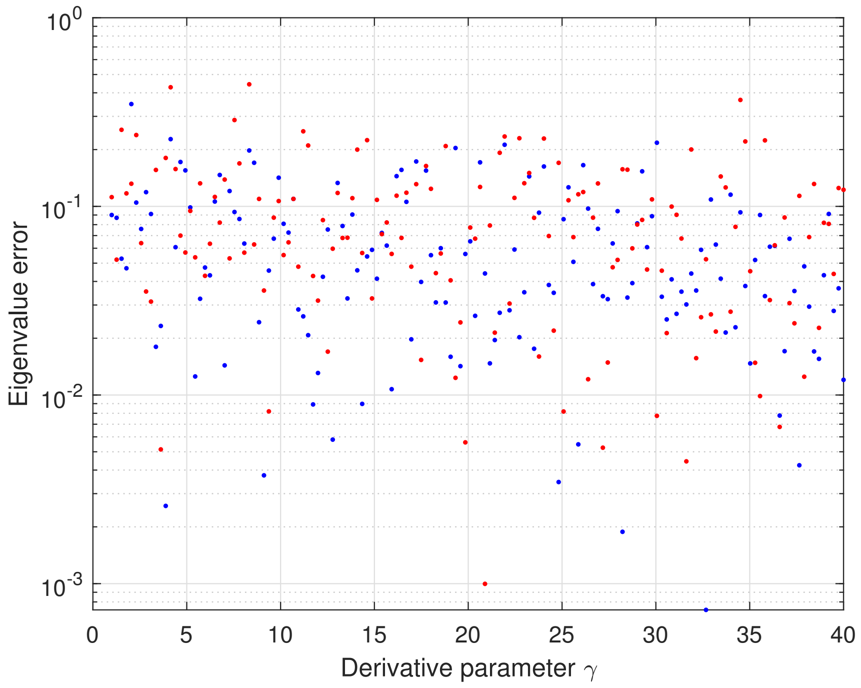


Figure 5.12: Error of the eigenvalues identified by the HD-PBSID for a varying γ parameter.

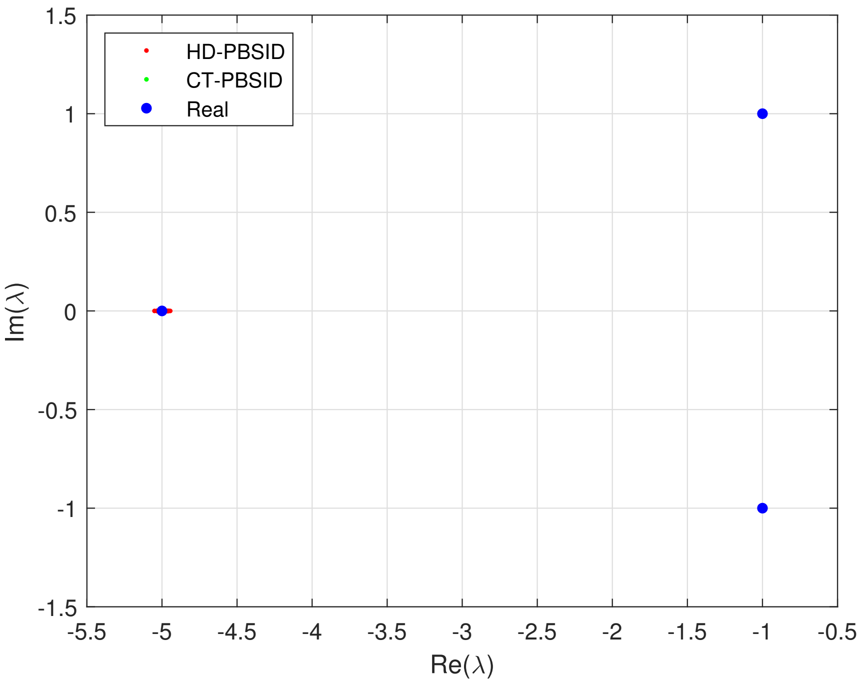


Figure 5.13: Eigenvalues obtained by the CT-PBSID and HD-PBSID for SNR = 50.

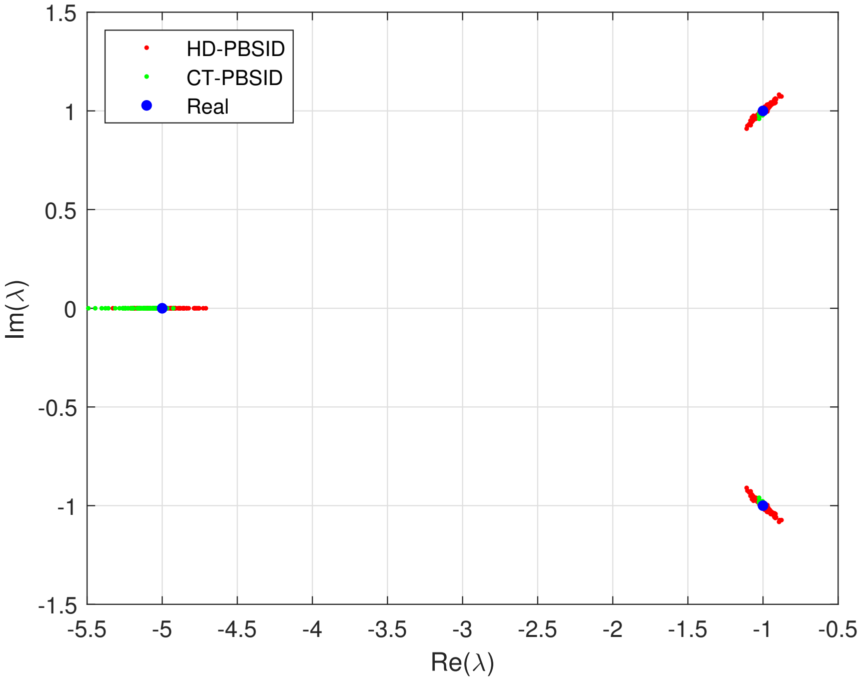


Figure 5.14: Eigenvalues obtained by the CT-PBSID and HD-PBSID for SNR $= 20$.

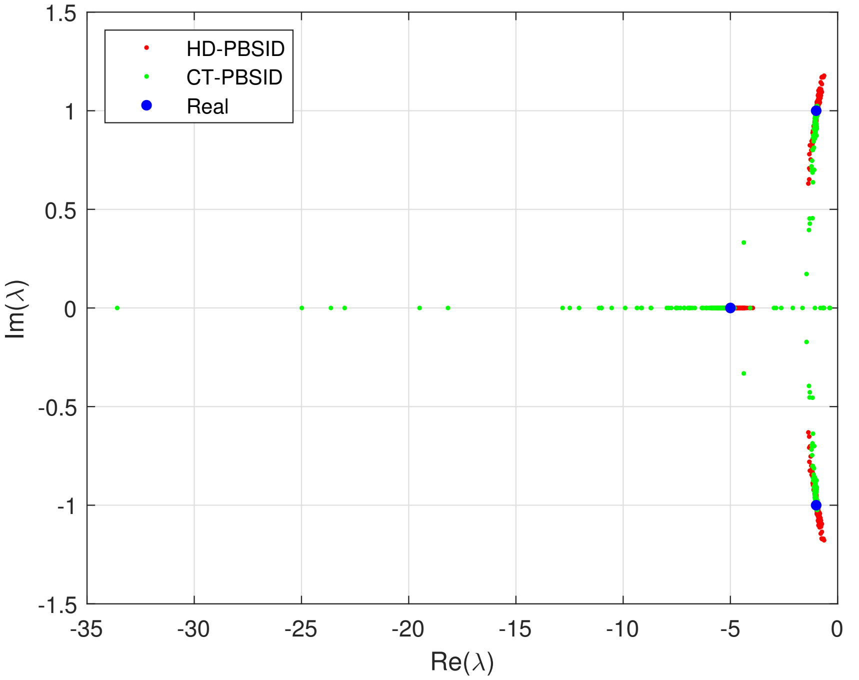


Figure 5.15: Eigenvalues obtained by the CT-PBSID and HD-PBSID for SNR $= 10$.

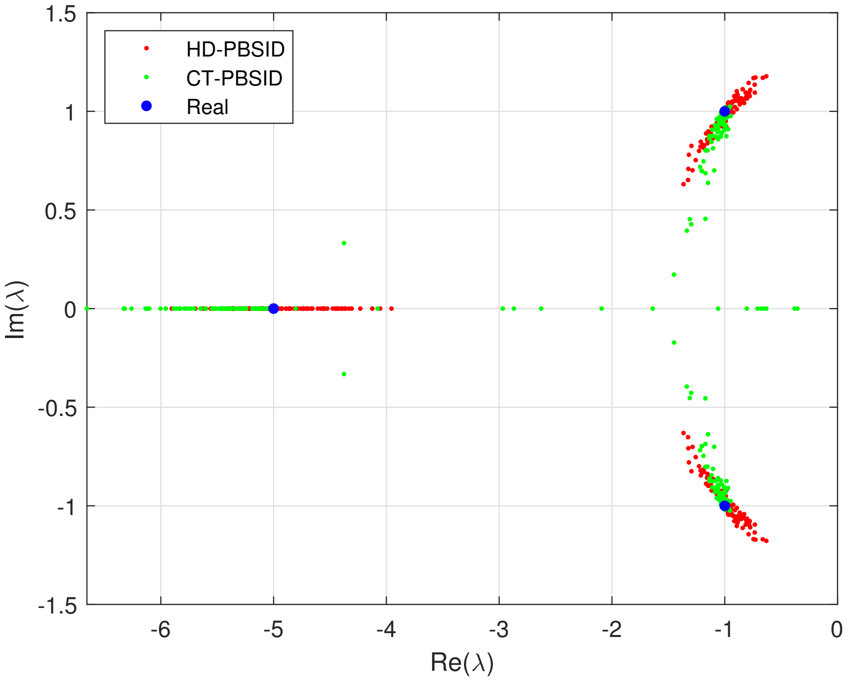


Figure 5.16: Eigenvalues obtained by the CT-PBSID and HD-PBSID for SNR = 10, zoomed for detail.

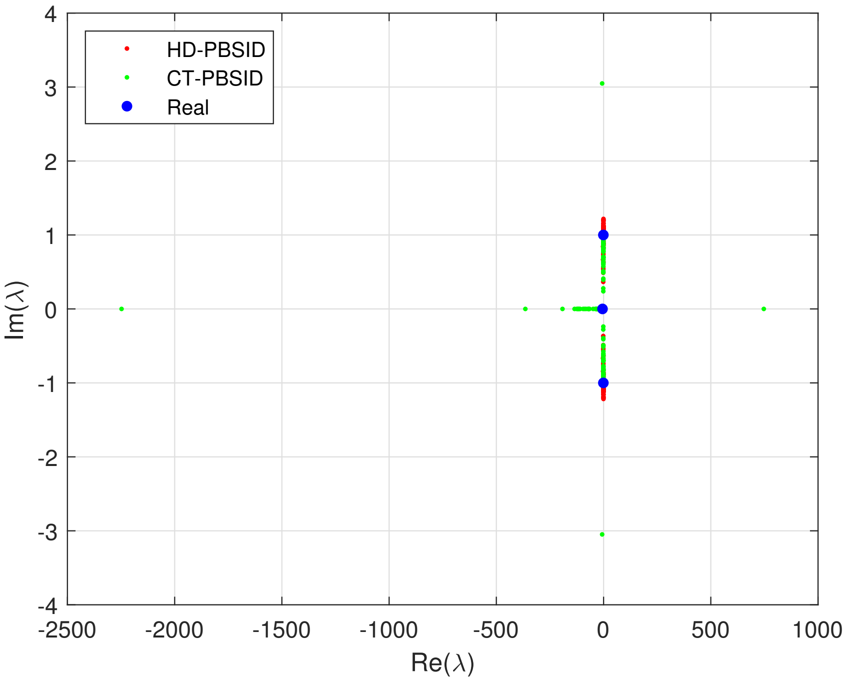


Figure 5.17: Eigenvalues obtained by the CT-PBSID and HD-PBSID for SNR = 6.

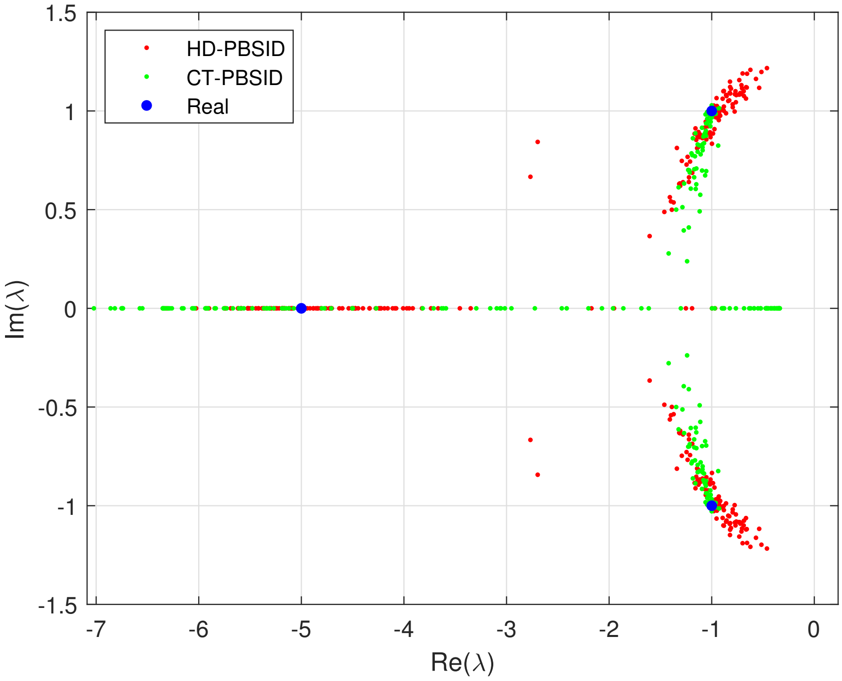
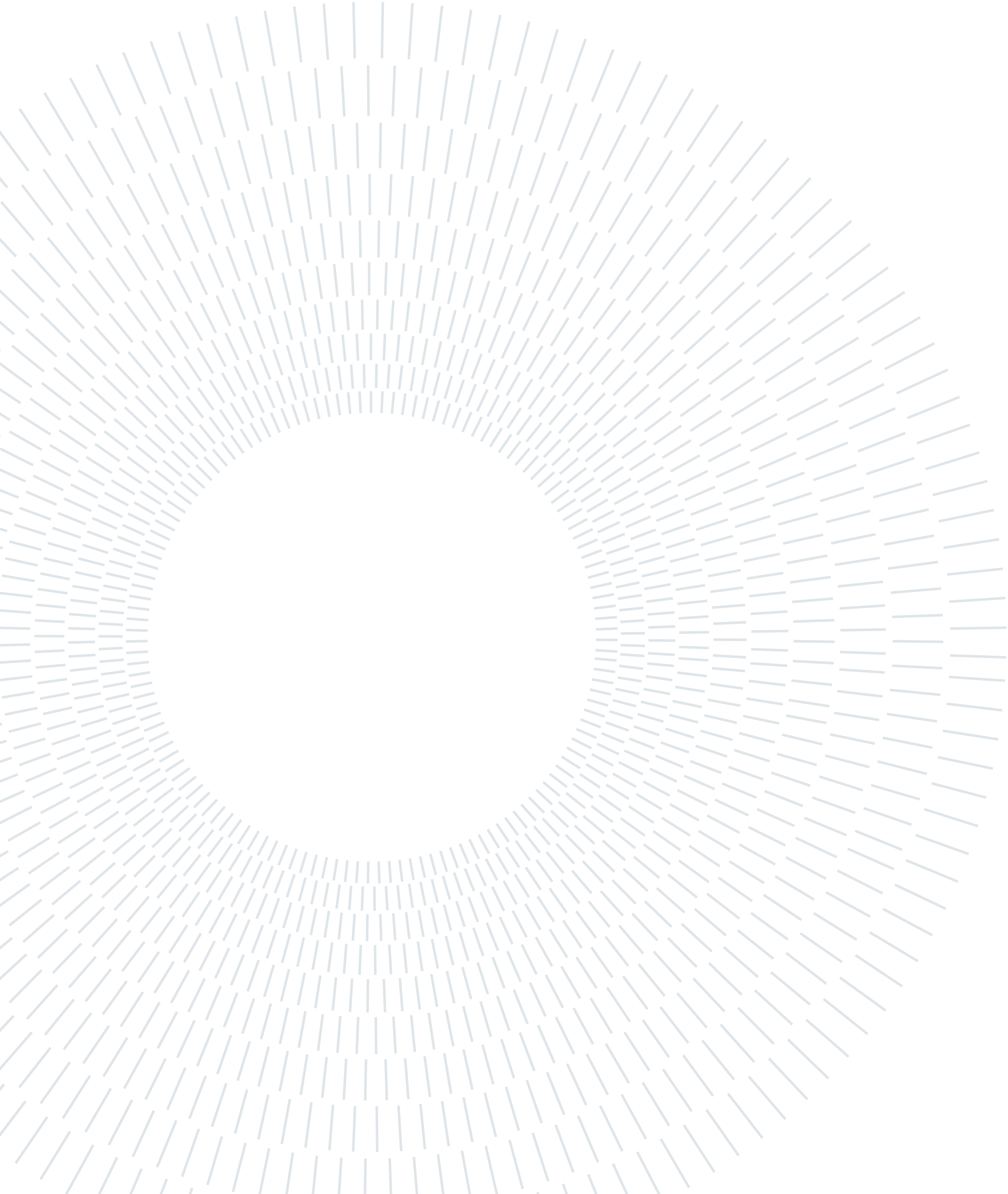


Figure 5.18: Eigenvalues obtained by the CT-PBSID and HD-PBSID for SNR $= 6$, zoomed for detail.

SNR	Method	Bias		STD	
		Dim 1	Dim 2	Dim 1	Dim 2
50	Hermite	0.005854	0.00071335	0.012384	0.0038998
	Laguerre	0.00047555	6.6732e-05	0.00095678	0.00017196
20	Hermite	0.00046385	0.0048368	0.077576	0.035362
	Laguerre	0.10017	0.0061251	0.1052	0.0083464
10	Hermite	0.05526	0.027607	0.23503	0.11277
	Laguerre	2.5129	0.18066	4.6628	0.30477
6	Hermite	0.15335	1.8918	0.45364	0.19165
	Laguerre	45.89	1.8254	227.25	1.0726

Table 5.1: Bias and standard deviation (STD) results for different SNR values for the CT-PBSID and HD-PBSID.



6 | Experimental implementation

In this chapter, the results obtained for the experimental system identification of a quadrotor are provided.

6.1. Experimental setup

The experimental analysis was performed by means of ANT-X Drone Lab [27], which features a quadrotor UAV, as shown in Figure 6.1. Additionally, the Drone Lab makes use of a ground control station, allowing to configure and operate the quadrotor, and a motion capture system, which can be used to produce high frequency samples of kinematic variables. This work focuses on the longitudinal and lateral angular dynamics of the drone, for which measurements on angular velocity and angular acceleration are generated.

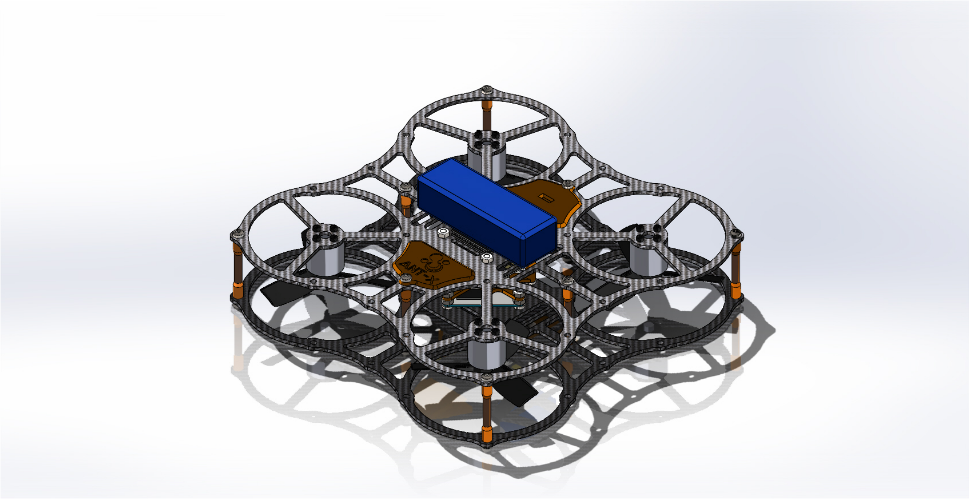


Figure 6.1: CAD model of the used quadrotor [27].

6.1.1. Drone dynamics

The angular dynamics of the drone can be modeled as follows. Firstly, the vehicle's structure can be assumed a rigid body, that is, neglecting the effect of flexibility. Let I an inertial reference frame fixed at ground, and B the body reference frame of the vehicle. Note that B is not centered on the drone's center of mass G, but on the point P where the torque produced by the propellers for a given moment-only control input does not generate any longitudinal thrust[1]. Thus, vector $\boldsymbol{r}_{PG}^{B}$ is the relative position of G with respect to P. Let the velocity of point P, $\boldsymbol{v}_P$, and center of mass, $\boldsymbol{v}_G$, as well as its acceleration $\boldsymbol{a}_G$. Let the attitude director cosines matrix (DCM) from inertial to body axes, which is denoted as C. Let J the drone's inertia tensor in the B frame, which is constant for a rigid body. Let $\boldsymbol{\omega}$ and $\boldsymbol{\alpha}$ be the angular velocity and acceleration of frame B with respect to frame I. Let $\boldsymbol{M}$ the external moments acting on the UAV with respect to point P, which are generated by the propellers, air interaction and gravity[2]. The total angular momentum of the drone around P is given by

$$\boldsymbol{h}_P = J\boldsymbol{\omega} + m(\boldsymbol{r}_{PG} \times \boldsymbol{v}_G). \tag{6.1}$$

This equation can be differentiated at the B frame, yielding

$$\frac{d}{dt}\boldsymbol{h}_P^B = J\dot{\boldsymbol{\omega}} + \boldsymbol{\omega} \times J\boldsymbol{\omega} + m\left((\boldsymbol{v}_G - \boldsymbol{v}_P + \boldsymbol{\omega} \times \boldsymbol{r}_{PG}) \times \boldsymbol{v}_G + \boldsymbol{r}_{PG} \times (\boldsymbol{a}_G + \boldsymbol{\omega} \times \boldsymbol{v}_G)\right). \tag{6.2}$$

Note that, using the Jacobi identity for the triple cross product,

$$(\boldsymbol{\omega} \times \boldsymbol{r}_{PG}) \times \boldsymbol{v}_G + \boldsymbol{r}_{PG} \times (\boldsymbol{\omega} \times \boldsymbol{v}_G) = \boldsymbol{\omega} \times (\boldsymbol{r}_{PG} \times \boldsymbol{v}_G) \tag{6.3}$$

Simplifying this result and applying the Kinetic Moment Theorem,

$$J\dot{\boldsymbol{\omega}} + \boldsymbol{\omega} \times J\boldsymbol{\omega} + m\left(-\boldsymbol{v}_P \times \boldsymbol{v}_G + \boldsymbol{r}_{PG} \times \boldsymbol{a}_G + \boldsymbol{\omega} \times (\boldsymbol{r}_{PG} \times \boldsymbol{v}_G)\right) = \boldsymbol{M} + m(\boldsymbol{v}_G \times \boldsymbol{v}_P). \tag{6.4}$$

This expression can be further reduced to

$$J\dot{\boldsymbol{\omega}} + \boldsymbol{\omega} \times J\boldsymbol{\omega} = \boldsymbol{M} - m\left(\boldsymbol{r}_{PG} \times \boldsymbol{a}_G\right) - m\boldsymbol{\omega} \times \left(\boldsymbol{r}_{PG} \times \boldsymbol{v}_G\right). \tag{6.5}$$

---

[1] In practice, this point coincides with the geometrical center of thrust of the UAV, which typically coincides horizontally with the center of mass, but has a vertical deviation.

[2] Note that gravity exerts torque, as the point at which the dynamics are written is not the center of mass.

The terms $m\left(\boldsymbol{r}_{PG} \times \boldsymbol{a}_G\right)$ and $m\boldsymbol{\omega} \times \left(\boldsymbol{r}_{PG} \times \boldsymbol{v}_G\right)$ are typically called inertial moments. The quadrotor's angular dynamics in the body frame are then given by

$$\frac{d}{dt}C = -\boldsymbol{\omega}^B \times C \tag{6.6}$$

$$\frac{d}{dt}\boldsymbol{\omega}^B = \boldsymbol{\alpha}^B \tag{6.7}$$

$$J\boldsymbol{\alpha}^B + \boldsymbol{\omega}^B \times J\boldsymbol{\omega}^B = \boldsymbol{M}^B - m\left(\boldsymbol{r}_{PG}^B \times \boldsymbol{a}_G^B\right) - m\boldsymbol{\omega}^B \times \left(\boldsymbol{r}_{PG}^B \times \boldsymbol{v}_G^B\right). \tag{6.8}$$

The evolution of $\boldsymbol{v}_G$ is given by Newton's Second Law,

$$m\frac{d}{dt}\boldsymbol{v}_G = \boldsymbol{T}. \tag{6.9}$$

Vector $\boldsymbol{T}$, the drone's total thrust, is a free variable, which is adjusted by the flight control system to maintain a leveled flight, that is, depending on position, typically through the height component, velocity itself and attitude.

For the identification, these dynamics are linearized around steady state conditions, namely, $C = I$ being the identity and for $\boldsymbol{\omega} = \boldsymbol{\alpha} = \boldsymbol{v}_G = 0$. Let the Euler angles $\boldsymbol{\theta} = \begin{pmatrix}\phi & \theta & \psi\end{pmatrix}^T$ characterizing the drone's attitude. The control input $\boldsymbol{u}^B$ can be defined as a signal directly controlling the torque generated by propellers. Therefore, with a slight abuse of notation, and linearizing $\boldsymbol{T}^I$ with altitude and attitude, and $\boldsymbol{M}^B$ as a function of the input signal $\boldsymbol{u}^B$, attitude $\boldsymbol{\theta}$, linear and angular velocity[3],

$$\frac{d}{dt}h = \left(\boldsymbol{u}_h^I\right)^T \boldsymbol{v}^I = v_3^I \tag{6.10}$$

$$m\frac{d}{dt}\boldsymbol{v}^I = T_h h + T_\theta \boldsymbol{\theta} \tag{6.11}$$

$$\frac{d}{dt}\boldsymbol{\theta} = \boldsymbol{\omega}^B \tag{6.12}$$

$$\frac{d}{dt}\boldsymbol{\omega}^B = \boldsymbol{\alpha}^B \tag{6.13}$$

$$J\boldsymbol{\alpha}^B = M_v \boldsymbol{v}^I + M_\theta \boldsymbol{\theta} + M_\omega \boldsymbol{\omega}^B + M_u \boldsymbol{u}^B - \left(\boldsymbol{r}_{PG}^B\right)^\times T_h h - \left(\boldsymbol{r}_{PG}^B\right)^\times T_\theta \boldsymbol{\theta}. \tag{6.14}$$

Note that $\boldsymbol{u}_h^I$ is defined as the unitary vector on the vertical direction. Hence, for the state representation $x = \begin{pmatrix}h & \left(\boldsymbol{v}^I\right)^T & \boldsymbol{\theta}^T & \left(\boldsymbol{\omega}^B\right)^T\end{pmatrix}^T$ and given the input $u = \boldsymbol{u}^B$ and output

---

[3] To account for gravity and aerodynamic torques.

$$y = \left(\left(\boldsymbol{\omega}^B\right)^T \quad \left(\boldsymbol{\alpha}^B\right)^T \right)^T,$$

$$\frac{d}{dt}x = \begin{bmatrix} 0 & \left(\boldsymbol{u}_h^I\right)^T & 0 & 0 \\ T_h/m & T_\theta/m & 0 & 0 \\ 0 & 0 & 0 & I \\ -\left(\boldsymbol{r}_{PG}^B\right)^\times T_h & J^{-1}M_v & J^{-1}M_\theta - \left(\boldsymbol{r}_{PG}^B\right)^\times T_\theta & J^{-1}M_\omega \end{bmatrix} x + \begin{bmatrix} 0 \\ 0 \\ 0 \\ J^{-1}M_u \end{bmatrix} u \quad (6.15)$$

$$y = \begin{bmatrix} 0 & 0 & 0 & I \\ -\left(\boldsymbol{r}_{PG}^B\right)^\times T_h & J^{-1}M_v & J^{-1}M_\theta - \left(\boldsymbol{r}_{PG}^B\right)^\times T_\theta & J^{-1}M_\omega \end{bmatrix} x + \begin{bmatrix} 0 \\ J^{-1}M_u \end{bmatrix} u \quad (6.16)$$

For J diagonal in the body frame, and assuming T_θ, M_θ, M_ω and M_u do not couple different axes, the previous equations can be written separately for each direction. In particular, for $x_i = \left(h \quad v_i^B \quad \theta_i \quad \omega_i^B\right)^T$ and $y_i = \left(\omega_i^B \quad \alpha_i^B\right)^T$, a linear system as follows is written,

$$\frac{d}{dt}x_i = A_i x_i + B_i u_i \quad (6.17)$$

$$y = C_i x_i + D_i u_i. \quad (6.18)$$

In consequence, from a theoretical view, a fourth order system should be identified. Higher order dynamics may arise from a more complicated definition of $\boldsymbol{T}$ with respect to state variables, perhaps from a PID controller.

6.1.2. System control and excitation

The drone's dynamics are unstable, and therefore a closed loop system identification has to be performed. In particular, the input signal is constructed as

$$\boldsymbol{u} = \boldsymbol{u}_c + \boldsymbol{u}_{inj}. \quad (6.19)$$

The first contribution is the feedback input, which is computed by the flight controller during the experiment to stabilize the plant. The second contribution is the one that excites the desired dynamics, commonly referred to as LMNF injection. In particular, only moments (L, M and N) are excited, with no perturbation to the controller total thrust F. The injection is performed one axis at a time, that is, in one experiment only the longitudinal (pitch) or lateral (roll) dynamics are excited. The yaw dynamics are much simpler, typically consisting of just a second or even first order system, and thus are not studied.

For each axis, three experiments were performed using the same input excitation for comparison. The injected moment is the same for both the longitudinal and lateral axis, as they can be expected to behave similarly. In particular, the strategy of signal construction through Hermite projections described in Chapter 5 was followed. The frequencies of interest were chosen between 1 and 4 Hz. A total experiment duration of 10 seconds was considered. An initial sine sweep signal satisfying these conditions was constructed through an order 150 Hermite expansion, as depicted in Figure 6.2. This input signal satisfies both the frequency requirements and the Hermite domain representation condition. Besides, it has near to constant amplitude, producing noticeable effects in the system, and increasing the SNR for all frequencies.

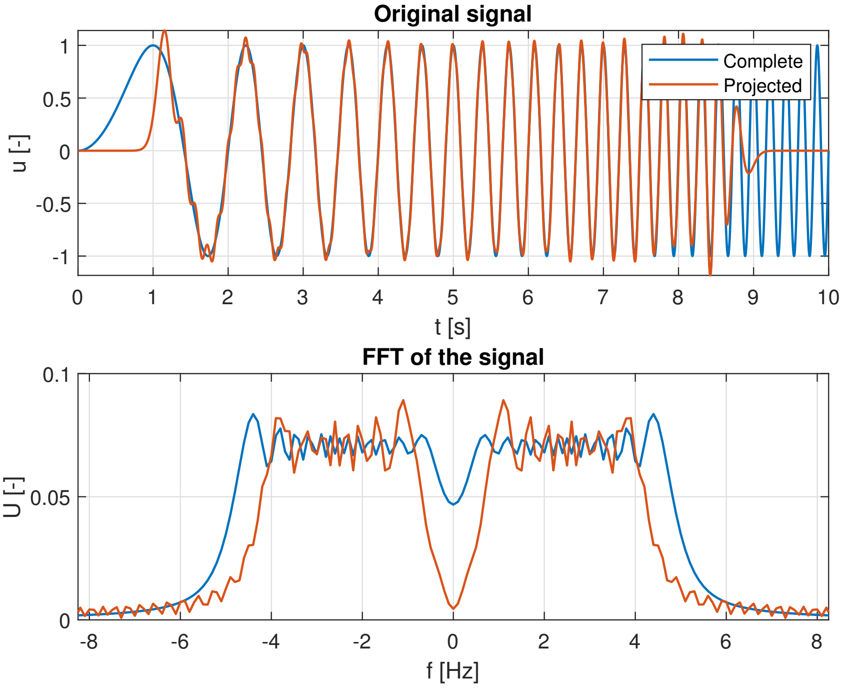


Figure 6.2: Excitation signal for all experiments, constructed via a sine sweep projection, together with its Fourier transform.

6.1.3. Experiments procedure

Experiments were assessed systematically. For each axis, three experiments were performed during the same flight. Using a fully charged battery, the drone was set to steady flight mode. Once the vehicle's position and attitude was stabilized, the corresponding LMNF injection was sent and executed. One experiment was performed once the drone had stabilized from the previous one. During this process, safety measures such as not entering the experiment area while the drone was armed and having a kill-switch ready at all times were taken. After the experiments were executed, the recorded data was recovered from the drone's memory and post-processed to extract, in each case, the signals of interest for the identification.

Input and measurement signals are not sampled homogeneously. Furthermore, a different sampling times vector is used for each variable. This is not an issue for the HD-PBSID, as the first step is a projection, which is compatible with any sampling time. This integration is performed using a trapezoid rule quadrature, as this enhances robustness against noise. The time between two samples is of roughly 4.20 ms. For an average experiment span of 10 s, the available number of samples for each run is of around 2400.

6.2. Identification and results

6.2.1. Identification setup

The same identification algorithm, based on the HD-PBSID, was used for both longitudinal and lateral dynamics. An expansion of 400 terms was used for the Hermite basis, although a lower maximum order does not seem to affect the results substantially, as discussed in Chapter 5. The differential operator parameters were set to $\beta = 6$ and $\gamma = 16$. As for the number of singular values and correspondingly the system state dimensions, motivated by the theoretical model and the results observed in Figures 6.3 and 6.4, a fourth order system was identified. As depicted in both figures, a third order system might be enough to capture the observed experimental behavior. Nonetheless, note that the fourth singular value is in general not much smaller in absolute value than the third one, justifying the selection.

6.2.2. Identification results

The obtained eigenvalues and bode plots for the longitudinal dynamics, considering all three experiments are shown in Figures 6.5 and 6.6. The specific obtained eigenvalues are provided in Tables 6.1 and 6.2. In general, all cases provide rather similar results, mainly in the frequency range excited by the input signal, that is, for $\omega \in [2\pi, 8\pi]$ rad. A good agreement of the three identified systems can be seen for higher frequencies. To obtain a better fitting for a wider frequency range, and in particular for a better characterization of the slow modes, a longer experiment is necessary. An unstable oscillating mode is observed around $0 \pm 5i$. The remaining eigenvalues are slow damping modes that can be associated to dissipation effects. A similar behavior is obtained for the lateral dynamics, as depicted in Figures 6.7 and 6.8.

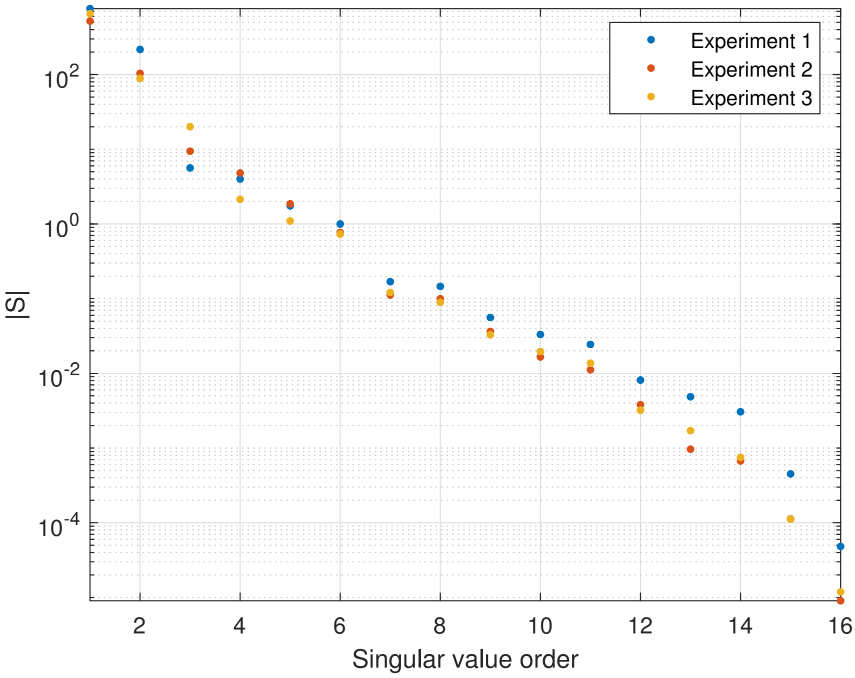


Figure 6.3: Singular values for the longitudinal system identification.

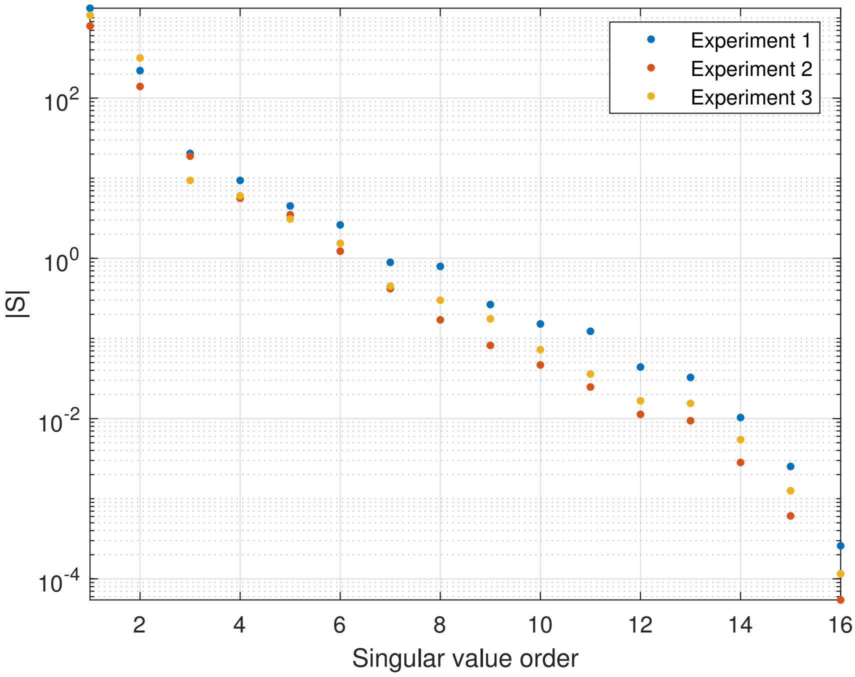


Figure 6.4: Singular values for the lateral system identification.

In order to verify the identified plant, it is of interest to simulate the estimated system under the experimental input. However, as the system is unstable, the output diverges if

Experiment	Eigenvalues			
1	-28.1279	-10.4272	$1.2072 + 5.8205i$	$1.2072 - 5.8205i$
2	-29.6969	-12.2414	$1.5850 + 4.9288i$	$1.5850 - 4.9288i$
3	-37.5718	-6.1102	$0.5082 + 4.3069i$	$0.5082 - 4.3069i$

Table 6.1: Eigenvalues of the identified plant for the longitudinal dynamics.

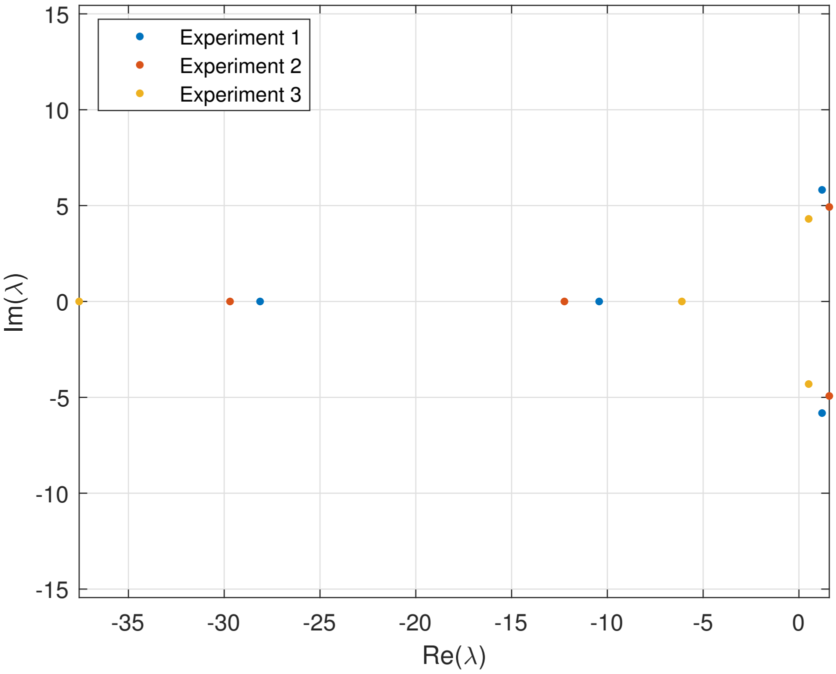


Figure 6.5: Eigenvalues obtained for the longitudinal dynamics.

not controlled in closed loop. Indeed, as matrix A has unstable eigenvalues, any perturbation which is not controlled produces an exponential response. To handle this difficulty, it is useful to recall that the PBSID algorithm also provides a Kalman gain K as part of the identification. Therefore, an estimator for the system state can be constructed as

$$\frac{d}{dt}\hat{x} = A\hat{x} + Bu + K(y - C\hat{x} - Du). \tag{6.20}$$

Defining $z = \begin{pmatrix} u^T & y^T \end{pmatrix}^T$,

$$\frac{d}{dt}\hat{x} = \begin{bmatrix} A - KC \end{bmatrix} \hat{x} + \begin{bmatrix} B - KD & K \end{bmatrix} z. \tag{6.21}$$

This innovation form of the plant is always stable. For this construction, the estimated

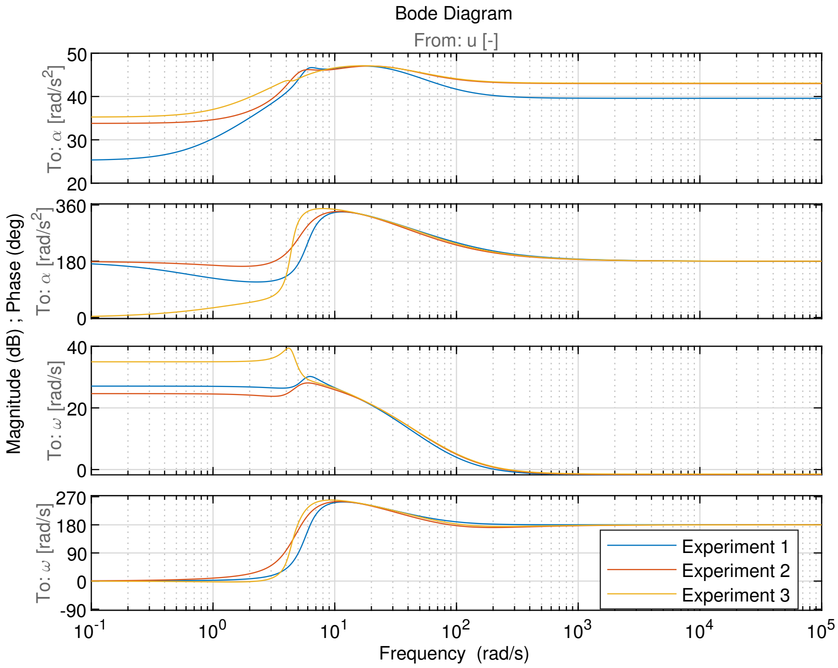


Figure 6.6: Bode diagram obtained for the longitudinal dynamics.

Experiment	Eigenvalues			
1	-28.9986	-11.1375	$2.4045 + 4.9270i$	$2.4045 - 4.9270i$
2	-27.4011	-12.2941	$1.1678 + 5.3542i$	$1.1678 - 5.3542i$
3	-31.7832	-9.7939	$0.8442 + 5.3266i$	$0.8442 - 5.3266i$

Table 6.2: Eigenvalues of the identified plant for the lateral dynamics.

an measured outputs are compared for both the longitudinal and lateral dynamics in Figures 6.9 and 6.10, respectively. In both cases, a good fitting of the identified observer with the measured responses is obtained, perhaps with a slight divergence of ω for the lateral case. The fact that the performed identification provide a good fitting to the experimental data, while giving slightly different eigenvalues, means that the estimated models are adequate for the excited frequency range, but not necessarily for slower or faster excitations. Naturally, if necessary, more experiments at different regimes would provide a more complete picture of the system dynamics. The obtained results provide a promising experimental validation of the HD-PBSID, which appears as an interesting contribution to the problem of CT system identification.

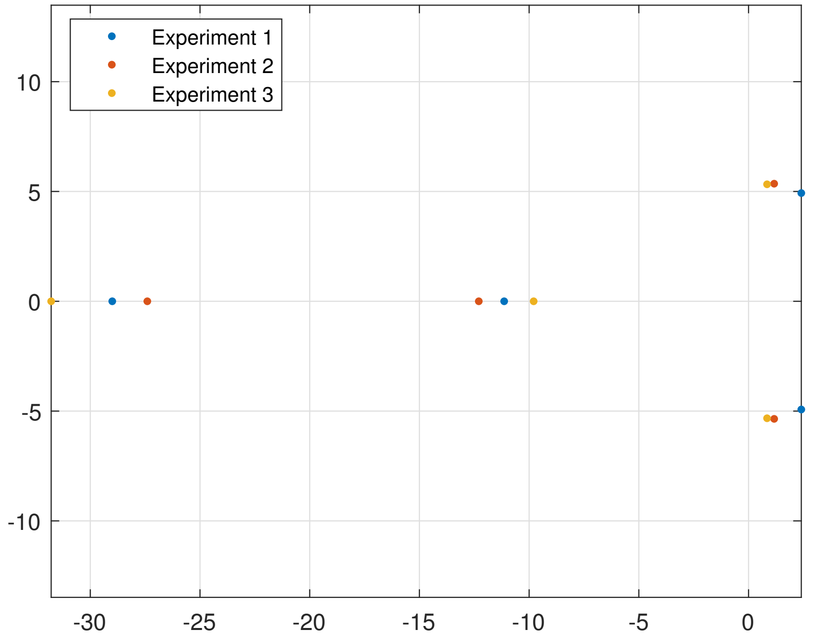


Figure 6.7: Eigenvalues obtained for the lateral dynamics.

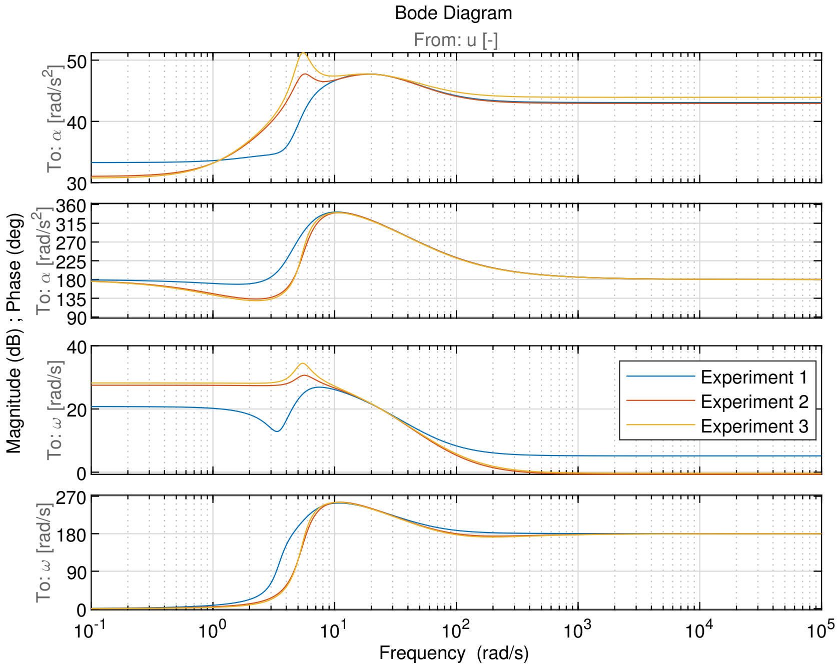


Figure 6.8: Bode diagram obtained for the lateral dynamics.

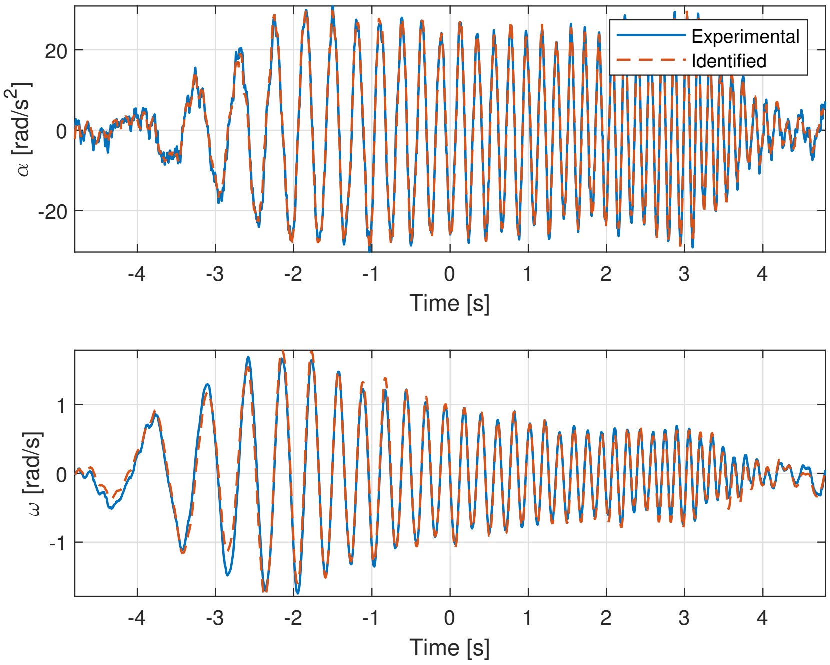


Figure 6.9: Simulated and experimental output for the longitudinal dynamics.

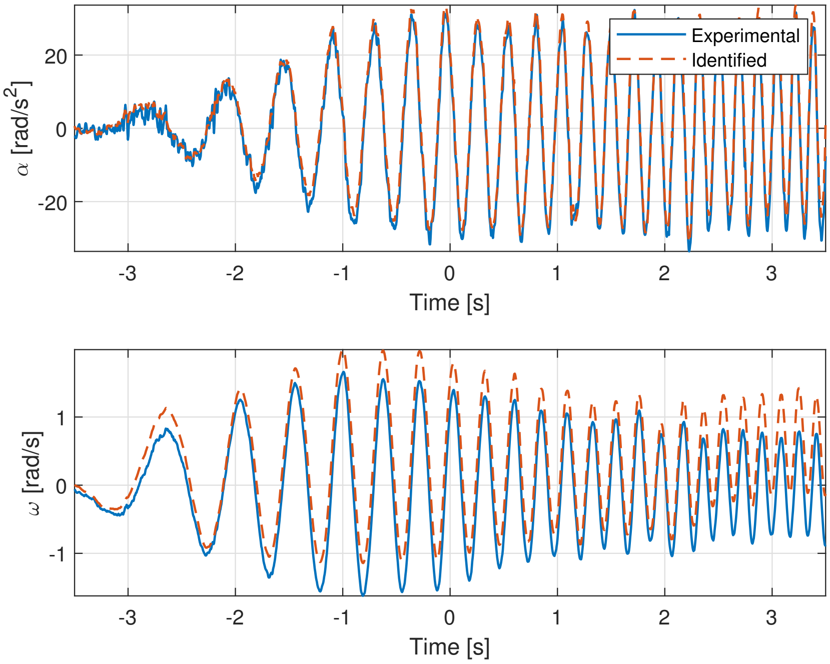
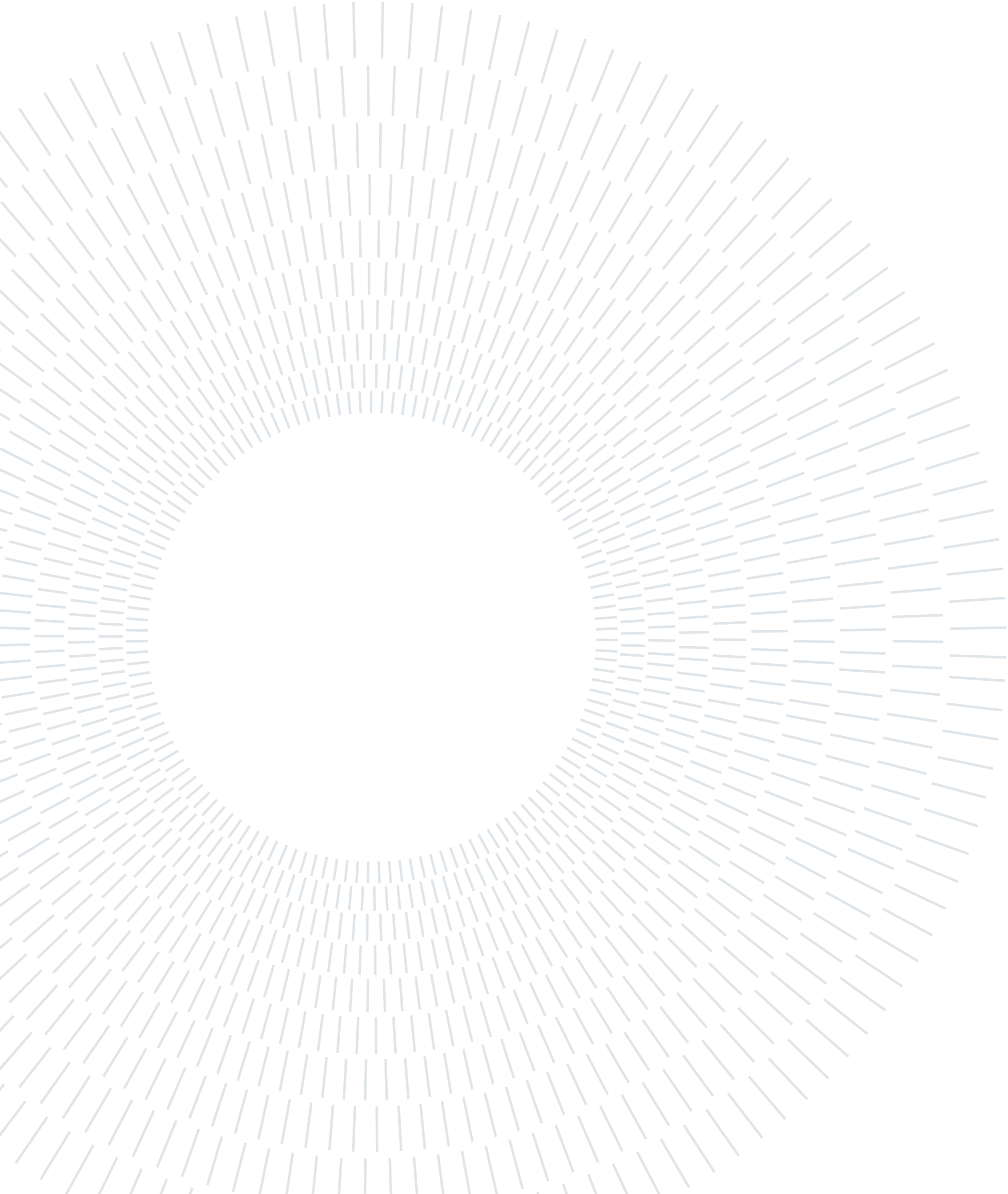


Figure 6.10: Simulated and experimental output for the lateral dynamics.



7 | Conclusions and future work

7.1. Conclusions

This thesis has introduced a novel framework for the identification of continuous-time (CT) systems using orthogonal function bases, with a particular focus on the Hermite domain-based PBSID (HD-PBSID) method. The proposed approach provides an alternative to other proposed CT identification methods such as the CT-PBSID, with particularly interesting properties regarding noisy environments.

One of the key contributions of this work is the rigorous mathematical formulation that connects the choice of orthonormal function bases with the filtering properties and numerical stability of the identification process. While Laguerre-based transformations have been widely studied, this thesis extends the framework by incorporating Hermite functions, exploiting useful properties such as compact support and well-defined spectral characteristics. The mathematical developments presented here lay the foundation for a broader understanding of CT-DT transformations beyond previously used techniques.

The performance of the HD-PBSID method was evaluated through numerical simulations and experimental validation. The simulation results demonstrated the robustness of the approach in handling noise and model uncertainties, while the experimental implementation on a quadrotor UAV confirmed its practical applicability. The method successfully identified both longitudinal and lateral dynamics, showing strong agreement between estimated and measured responses. The results suggest that the use of orthonormal function bases in the way proposed in this thesis can enhance the accuracy and reliability of CT system identification, especially in applications where high-fidelity models are required.

Despite these promising results, certain limitations remain. The computational complexity of the method increases with the order of the selected basis functions, which can become a constraint for real-time applications. Additionally, while the experimental validation demonstrated the feasibility of the method, further investigations on different platforms and operating conditions, as well as validation steps with already identified systems, would be beneficial to assess its broader applicability. These aspects highlight the need for further research and refinement of the proposed methodology.

7.2. Future Work

The findings of this thesis open several avenues for future research. One promising direction is the exploration of alternative function bases beyond Hermite functions. The theoretical framework developed in this work suggests that other orthonormal bases, such as Legendre or Jacobi functions, could be employed to tailor the transformation process to specific system dynamics. Investigating these alternatives may provide further insights into optimal basis selection and its impact on identification accuracy.

The application of the HD-PBSID method in closed-loop identification also represents an interesting research direction. Most of the analyses in this thesis were conducted in open-loop conditions, but many practical systems operate under feedback control. Investigating the robustness of the method in closed-loop scenarios and developing strategies to ensure consistency in the presence of feedback effects would enhance its practical usability.

In addition, efforts should be made to improve the computational efficiency of the proposed approach. While the method provides accurate results, its computational cost increases with the order of the selected basis functions. Exploring strategies such as adaptive order selection, sparsity-based representations, or efficient numerical implementations could improve scalability and enable real-time applications.

Finally, the method should be tested on a wider range of real-world systems. While the experimental validation in this work was performed on a quadrotor UAV, future studies could extend its application to more complex aerospace, robotic, or industrial systems where accurate continuous-time identification is crucial for control design. By addressing these challenges, the HD-PBSID framework can be further refined and extended, contributing to the advancement of system identification methodologies and their applications.

Bibliography

[1] M. Bergamasco and M. Lovera, "Rotorcraft model identification: a black-box time/frequency domain approach," 2013.

[2] S. Bittanti and M. Lovera, "Bootstrap-based estimates of uncertainty in subspace identification methods," *Automatica*, vol. 36, no. 11, pp. 1605–1615, 2000.

[3] A. Chiuso and G. Picci, "Consistency analysis of certain closed-loop subspace identification methods," *Automatica*, vol. 41, no. 3, pp. 377–391, 2005.

[4] M. Tischler and R. Remple, *Aircraft and Rotorcraft System Identification: Engineering Methods with Flight-Test Examples.* AIAA, 2006.

[5] R. Jategaonkar, *Flight Vehicle System Identification.* AIAA, 2006.

[6] P. Van Overschee and B. De Moor, *Subspace identification: theory, implementation, application.* Kluwer Academic Publishers, 1996.

[7] M. Verhaegen and V. Verdult, *Filtering and System Identification: A Least Squares Approach.* Cambridge University Press, 2007.

[8] G. van der Veen, J.-W. van Wingerden, M. Bergamasco, M. Lovera, and M. Verhaegen, "Closed-loop subspace identification methods: an overview," *IET Control Theory and Applications*, vol. 7, no. 10, pp. 1339–1358, 2013.

[9] M. Jansson, "Subspace identification and arx modeling," 2003.

[10] A. Chiuso, "On the relation between cca and predictor based subspace identification," *IEEE Transactions on Automatic Control*, vol. 52, no. 10, pp. 1795–1812, 2007.

[11] C. Chou and M. Verhaegen, "Subspace algorithms for the identification of multivariable dynamic errors-in-variables models," *Automatica*, vol. 33, no. 10, pp. 1857–1869, 1997.

[12] H. Oku and T. Fujii, "Direct subspace model identification of lti systems operating in closed-loop," 2004.

[13] H. Garnier and L. Wang, eds., *Identification of continuous-time models from sampled data.* Springer, 2008.

[14] H. Garnier, T. Soderstrom, and J. Yuz, "Editorial special issue on continuous-time model identification," *Control Theory Appl. IET*, vol. 5, no. 7, pp. 839–841, 2011.

[15] M. Bergamasco, "Continuous-time model identification with applications to rotorcraft dynamics," in *PhD thesis, Politecnico di Milano*, 2012.

[16] M. Bergamasco and M. Lovera, "Continuous-time subspace identification in closed-loop using laguerre filters," *Proc. 49th IEEE Conf. Control Decis. Control. Atlanta*, 2011.

[17] Y. Ohta, "Realization of input-output maps using generalized orthonormal basis functions," *Syst Control Lett*, vol. 22, no. 6, pp. 437–444, 2005.

[18] N. Brady, D. Karlsson, and T. McKelvey, "Frequency-domain continuous-time identification by means of bilinear model reduction," *Control Theory Appl. IET*, vol. 5, no. 7, pp. 909–918, 2011.

[19] L. Ljung, *System identification: theory for the user.* Prentice Hall PTR, 2 ed., 1999.

[20] B. Haverkamp, C. Chou, M. Verhaegen, and R. Johansson, "Identification of continuous-time mimo state space models from sampled data, in the presence of process and measurement noise," in *Proceedings of 35th IEEE Conference on Decision and Control*, vol. 2, pp. 1539–1544, IEEE, 1996.

[21] T. Söderström and P. Stoica, *System Identification.* Prentice Hall, 1989.

[22] R. Pintelon and J. Schoukens, *System Identification: A Frequency Domain Approach.* Wiley, 2nd ed., 2012.

[23] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Transactions of the ASME—Journal of Basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.

[24] E. Kreyszig, *Introductory functional analysis with applications*, vol. 17. John Wiley & Sons, 1991.

[25] D. J. Griffiths and D. F. Schroeter, *Introduction to quantum mechanics.* Cambridge university press, 2019.

[26] G. Szegő, *Orthogonal Polynomials*, vol. 23 of *American Mathematical Society Colloquium Publications.* American Mathematical Society, 4th ed., 1975.

[27] ANT-X Team, *ANT-X Documentation*, 2024. Accessed: 2025-02-09.

List of Figures

3.1 Laguerre basis up to order 10, with pole $a = 1$. 21
3.2 Hermite basis up to order 10. 22
3.3 Projection of white noise signals of infinite and finite support over the Hermite domain. Each column corresponds to a different signal. 33
3.4 Differentiation of signals on the Hermite domain and reconstruction on the time domain. 34
3.5 Construction of a Brownian motion process on the Hermite domian. 35
3.6 Simulation of a linear system with two inputs and two outputs on the Hermite domain, with the time domain dynamics in continuous line and the Hermite reconstructed signals in dashed line. The original and reconstructed signals coincide. 36
3.7 Hermite basis of order 200, with effective support. 37
3.8 Stability region in the complex plane for the modified differentiation operator. 38

5.1 Bode diagram of the real system and of the HD-PBSID identified plant for SNR $= 50$. 60
5.2 Bode diagram of the real system and of the HD-PBSID identified plant for SNR $= 20$. 61
5.3 Bode diagram of the real system and of the HD-PBSID identified plant for SNR $= 10$. 62
5.4 Bode diagram of the real system and of the HD-PBSID identified plant for SNR $= 6$. 63
5.5 Eigenvalues of the real system and of the HD-PBSID identified plant for SNR $= 50$. 64
5.6 Eigenvalues of the real system and of the HD-PBSID identified plant for SNR $= 20$. 65
5.7 Eigenvalues of the real system and of the HD-PBSID identified plant for SNR $= 10$. 65
5.8 Eigenvalues of the real system and of the HD-PBSID identified plant for SNR $= 6$. 66

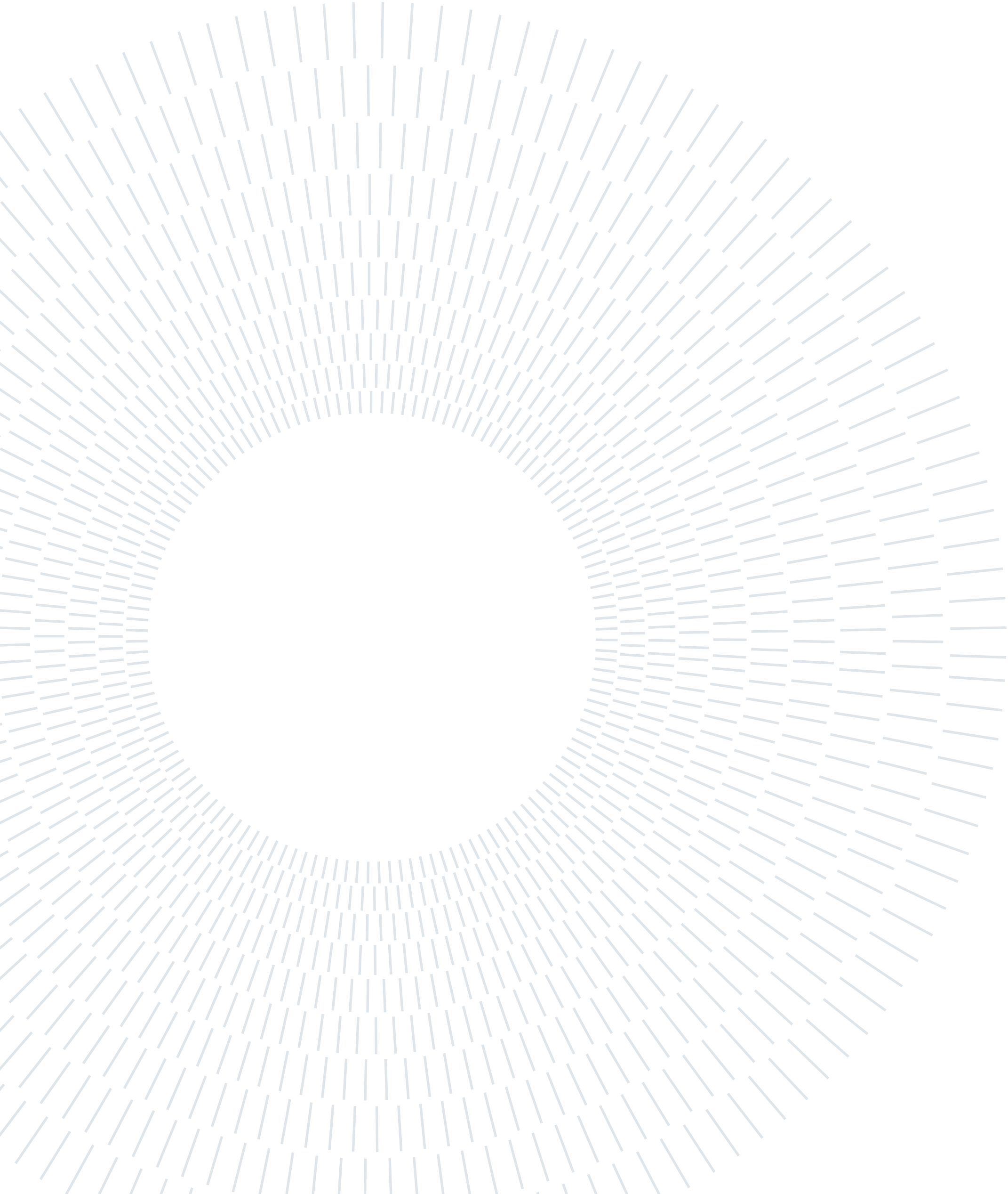
5.9 Error of the eigenvalues identified by the HD-PBSID for a varying number of samples. 66
5.10 Error of the eigenvalues identified by the HD-PBSID for a varying Hermite basis order. 67
5.11 Error of the eigenvalues identified by the HD-PBSID for a varying β parameter. 67
5.12 Error of the eigenvalues identified by the HD-PBSID for a varying γ parameter. 68
5.13 Eigenvalues obtained by the CT-PBSID and HD-PBSID for SNR = 50. . . 68
5.14 Eigenvalues obtained by the CT-PBSID and HD-PBSID for SNR = 20. . . 69
5.15 Eigenvalues obtained by the CT-PBSID and HD-PBSID for SNR = 10. . . 69
5.16 Eigenvalues obtained by the CT-PBSID and HD-PBSID for SNR = 10, zoomed for detail. 70
5.17 Eigenvalues obtained by the CT-PBSID and HD-PBSID for SNR = 6. . . . 70
5.18 Eigenvalues obtained by the CT-PBSID and HD-PBSID for SNR = 6, zoomed for detail. 71

6.1 CAD model of the used quadrotor [27]. 73
6.2 Excitation signal for all experiments, constructed via a sine sweep projection, together with its Fourier transform. 77
6.3 Singular values for the longitudinal system identification. 79
6.4 Singular values for the lateral system identification. 79
6.5 Eigenvalues obtained for the longitudinal dynamics. 80
6.6 Bode diagram obtained for the longitudinal dynamics. 81
6.7 Eigenvalues obtained for the lateral dynamics. 82
6.8 Bode diagram obtained for the lateral dynamics. 82
6.9 Simulated and experimental output for the longitudinal dynamics. 83
6.10 Simulated and experimental output for the lateral dynamics. 83

List of Tables

5.1 Bias and standard deviation (STD) results for different SNR values for the CT-PBSID and HD-PBSID. 71

6.1 Eigenvalues of the identified plant for the longitudinal dynamics. 80
6.2 Eigenvalues of the identified plant for the lateral dynamics. 81



List of Symbols

Symbol	**Description**	**SI Unit**
$\langle a, b \rangle$	Inner product of a and b in Hilbert space	
$\mathcal{H}$	Hilbert space	
$\mathcal{L}^2$	Space of square-integrable functions	
ℓ^2	Space of square-summable sequences	
$\|\cdot\|$	Norm in a Hilbert space	
$\mathbb{E}[\cdot]$	Expectation operator	
δ_{nm}	Kronecker delta function	
$\bar{a}$	Conjugate of $a \in \mathbb{C}$	
$T^\dagger$	Adjoint of an operator T over a Hilbert space	
$f * g$	Convolution of f and g	
$l_n(x)$	Laguerre function of order n	
$\mathrm{H}_n(x)$	Hermite polynomial of degree n	
$h_n(x)$	Hermite function of order n	
$\mathcal{A}^n$	All-pass filter of order n	
$a^\dagger$	Creation operator in the harmonic oscillator	
a	Annihilation operator in the harmonic oscillator	
$\hat{a}^\dagger$	Rising operator	
$\hat{a}$	Lowering operator	
$[A, B]$	Commutator of operators A and B	
λ_n	Eigenvalue associated with index n	
Γ_f	Projection matrix in PBSID method	
x	State vector	
u	Input vector	
y	Output vector	
$\hat{x}$	Estimated state vector	
A, B, C, D	System matrices	
K	Kalman gain	
w	Process noise vector	

Symbol	**Description**	**SI Unit**
v	Measurement noise vector	
SNR	Signal-to-noise ratio	dB
$\underline{\mathrm{f}}$	Hermite-domain matrix for a signal $f(t)$	
$\underline{\mathrm{D}}$	Differentiation matrix	
$\underline{\mathrm{D}}'$	Modified differentiation matrix	
L	Sturm-Liouville operator	
$w(x)$	Orthogonal polynomial weight function	
c_n	Rodrigues' formula coefficient of order n	
d_n	Orthogonal polynomial n-th normalization coefficient	
J	Inertia tensor of the quadrotor	kg m^2
m	Mass of the quadrotor	kg
$\boldsymbol{M}$	External moment acting on the UAV	N m
$\boldsymbol{\omega}$	Angular velocity	rad/s
$\boldsymbol{\alpha}$	Angular acceleration	rad/s^2

Acknowledgements

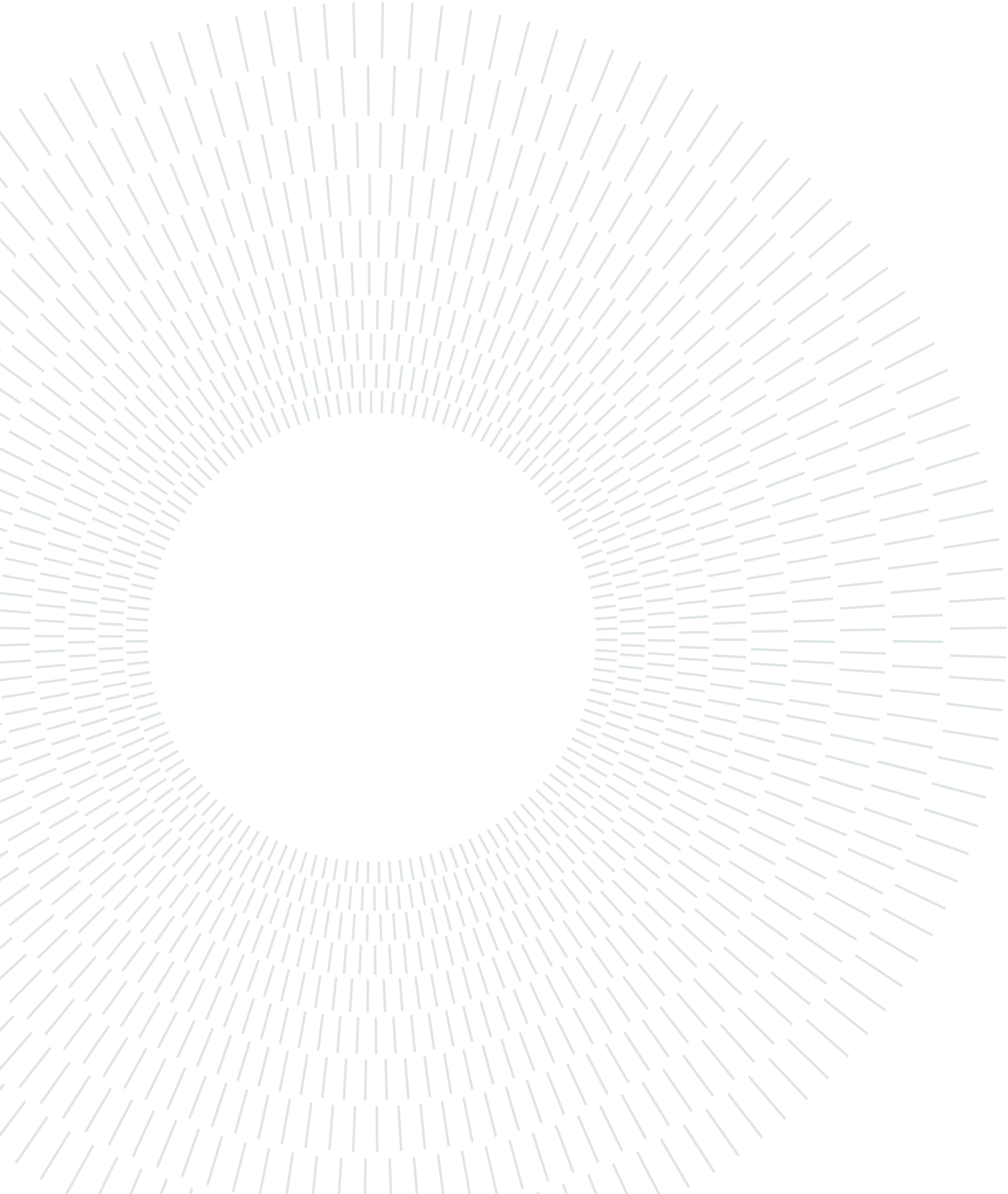
This thesis marks the culmination of two wonderful years at Politecnico di Milano, two years filled with learning, challenges, and countless rewarding experiences. Studying here has been an incredible journey, one that has not only deepened my knowledge but also broadened my perspective, surrounded by inspiring professors, brilliant colleagues, and an environment that fosters curiosity and innovation.

First and foremost, I would like to thank my family, whose tireless support has made this journey possible. Despite the distance from Seville, they have never felt far. To my mom, for always being there to talk, whether it was about my research, my day, or just about anything. To my dad, who has always seen the solution to problems and who has always helped me before I could even ask, no matter how busy he was. And to my sisters, for keeping me grounded, making me laugh, and reminding me that there is life beyond work. This thesis is as main as is yours.

I am deeply grateful to my advisor, Professor Marco Lovera, for giving me the opportunity to work on such an interesting topic and for his guidance along the way. A special thank you to my co-advisor, Enrico Barbiero, whose patience, expertise, and support have been instrumental in shaping this work. His advice and insights have been truly invaluable.

I want to thank Professor Rafael Vázquez and Professor Francisco Gavilán from the University of Seville for introducing me to this opportunity and for their unwavering support. Their insightful guidance, thoughtful perspectives, and generosity have been invaluable throughout this journey. Their willingness to offer encouragement, share their expertise, and welcome me with open arms has made a profound impact, for which I am deeply grateful.

Finally, to you—the reader—thank you for taking the time to go through this document. I hope you find it at least half as interesting as it was difficult to write (and, let's be honest, probably just as difficult to read).



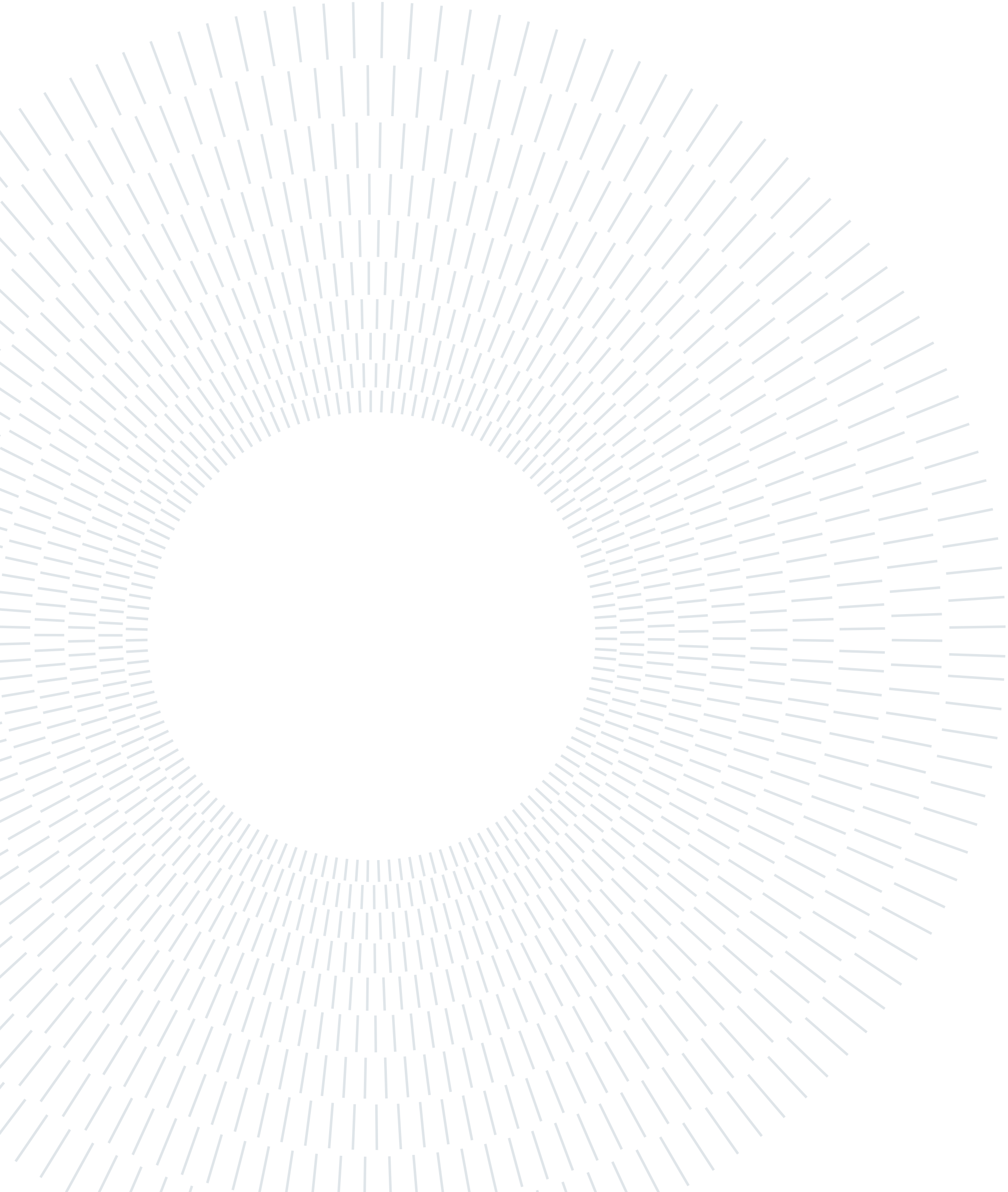
Agradecimientos

Esta tesis marca el broche final de dos maravillosos años en el Politécnico de Milán, dos años llenos de aprendizaje, desafíos y experiencias enormemente gratificantes. Estudiar aquí ha sido un viaje increíble, no solo por la profundidad de conocimientos que he adquirido, sino también por la forma en que ha ampliado mi perspectiva. He tenido la suerte de estar rodeado de profesores inspiradores, compañeros brillantes y un entorno que fomenta la curiosidad y la innovación.

En primer lugar, quiero agradecer a mi familia, cuyo apoyo incondicional ha hecho posible este camino. A pesar de los kilómetros a Sevilla, nunca los he sentido lejos. A mi madre, por estar siempre ahí para hablar, ya fuera sobre mis investigaciones, mi día o cualquier cosa que se nos ocurriera. A mi padre, por su capacidad para encontrar siempre la solución a los problemas y por ayudarme incluso antes de que pudiera pedírselo, sin importar lo ocupado que estuviera. Y a mis hermanas, por mantenerme con los pies en la tierra, hacerme reír y recordarme que hay vida más allá del trabajo. Esta tesis es tan mía como lo es vuestra.

Estoy profundamente agradecido a mi tutor, el profesor Marco Lovera, por darme la oportunidad de trabajar en un tema tan interesante y por su orientación a lo largo del camino. Quiero agradecer especialmente a mi co-tutor, Enrico Barbiero, cuya paciencia, conocimiento y apoyo han sido fundamentales para dar forma a este trabajo. Sus consejos y observaciones han sido realmente valiosos.

También quiero agradecer a los profesores Rafael Vázquez y Francisco Gavilán, de la Universidad de Sevilla, por haberme descubierto esta oportunidad y por su constante apoyo. Su guía, sus perspectivas y su generosidad han sido esenciales durante todo este proceso. Su disposición para motivarme, compartir su experiencia y acogerme con los brazos abiertos ha dejado una huella profunda en mí, por lo que les estaré siempre agradecido.

Por último, a ti, lector, gracias por tomarte el tiempo de recorrer estas páginas. Espero que encuentres este trabajo al menos la mitad de interesante de lo que fue difícil escribirlo (y, siendo honestos, probablemente igual de difícil de leer).