

ADAPTIVE OBSERVER DESIGN FOR PARABOLIC PARTIAL DIFFERENTIAL EQUATIONS

by
PEDRO A. ASCENCIO

A Thesis submitted in fulfilment of requirements for the degree of
Doctor of Philosophy of Imperial College London

Power and Control Research Group
Department of Electrical and Electronic Engineering
Imperial College London
2017

Dedicated to
my wife Mónica, our daughter Elena
pem.

Declaration of Originality

I hereby declare that this is an original thesis and it is entirely based on my own work. I acknowledge the sources in every instance where I have used the ideas of other writers. This thesis was not and will not be submitted to any other university or institution for fulfilling the requirements of a degree.

Copyright Declaration

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.

Pedro A. Ascencio
London, 2017.

Abstract

This thesis addresses the observer design problem, for a class of linear one-dimensional parabolic Partial Differential Equations, considering the simultaneous estimation of states and parameters from boundary measurements.

The design is based on the Backstepping methodology for Partial Differential Equations and extends its central idea, the “Volterra transformation”, to compensate for the parameters uncertainties. The design steps seek to reject time-varying parameter uncertainties setting forth a type of differential boundary value problems (Kernel-PDE/ODEs) to accomplish its objective, the solution of which is computed at every fixed sampling time and constitutes the observer gains for states and parameters. The design does not include any pre-transformation to some canonical form and/or a finite-dimensional formulation, and performs a direct parameter estimation from the original model.

The observer design problem considers two cases of parameter uncertainty, at the “boundary”: control gain coefficient, and “in-domain”: diffusivity and reactivity parameters, respectively. For a Luenberger-type observer structure, the problems associated to one and two points of measurement at the boundary are studied through the application of an intuitive modification of the Volterra-type and Fredholm-type transformations. The resulting Kernel-PDE/ODEs are addressed by means of a novel methodology based on polynomial optimization and Sum-of-Squares decomposition. This approach allows recasting these coupled differential equations as convex optimization problems readily implementable resorting to semidefinite programming, with no restrictions to the spectral characteristics of some integral operators or system’s coefficients.

Additionally, for polynomials Kernels, uniqueness and invertibility of the Fredholm-type transformation are proved in the space of real analytic and continuous functions. The direct and inverse Kernels are approximated as the optimal polynomial solution of a Sum-of-Squares and Moment problem with theoretically arbitrary precision.

Numerical simulations illustrate the effectiveness and potentialities of the methodology proposed to manage a variety of problems with different structures and objectives.

Acknowledgments

I would like to express, on behalf of my family, our deep appreciation to my supervisors Professor Alessandro Astolfi and Professor Thomas Parisini for their invaluable support and guidance. Our feelings of endless gratitude for making possible my research and our family life.

We are indebted to Mr. Pedro Gonzalez, my former Manager in Chile. Many thanks don Pedro to believe in us, helps to our daughter's education, and provide us your generous support.

I would like to thank my colleagues from the Control and Power research group, in particular to Pedro Ramírez and his family for their unique friendship, generosity and innumerable advice. Many thanks to Felipe Tobar for his kind and prompt answers. My gratitude also goes to Mrs. Michelle Hennessy-Hammond for her diligent support. I also thanks to Dr. Riccardo Ferrari for his guidance in my initial steps of research.

From Chile, I want to express my sincere gratitude to Professor Daniel Sbarbaro, Professor José Espinoza and Professor Freddy Paiva for their constant support in many opportunities. I am also grateful to Chile's National Commission for Scientific and Technological Research (CONICYT) for funding my PhD studies at Imperial College.

I would like to thank my examiners, Professor Richard Vinter and Dr. Christophe Prieur, for their valuable feedback on my thesis.

Finally, I would like to thanks my wife Mónica Salazar and our daughter Elena Ascencio. There is no words to express all that they have done throughout these years, their braveness, endurance, love and joy of living.

Pedro A. Ascencio
London, 2017.

Contents

Declaration of Originality	2
Copyright Declaration	2
Abstract	3
Acknowledgements	4
Contents	5
List of Tables	9
List of Figures	10
1 Introduction	11
1.1 Motivation	11
1.2 A brief review of Distributed Parameter Systems theory	12
1.2.1 Control	13
1.2.2 State Estimation	14
1.2.3 Parameter Estimation	16
1.2.3.1 Adaptive Observers	17
1.3 Thesis Outline	18
1.3.1 Objectives	18
1.3.2 Organization and Contributions	19
1.3.2.1 Publications	21
1.4 Preliminary Terminology	21
1.4.1 Notation	21
1.4.2 Definition	22
1.4.3 Abbreviations	22
2 Formulation of Differential Equations as Convex Optimization Problems	23
2.1 Sum-of-Squares Approach for PDEs	24

2.1.1	Compact Basic Semi-algebraic Sets	24
2.1.2	Minimax Approximation	26
2.1.3	Least Squares Approximation	28
2.1.4	Polytopic Domain Decomposition	30
2.2	Moment Approach for PDEs	31
2.3	Computational Examples	32
2.3.1	One-Dimensional Differential Equation-BVP	32
2.3.1.1	Steady Convection-dominated Problem	33
2.3.1.2	Steady Diffusion-Reaction Problem	35
2.3.1.3	Sturm–Liouville Eigenvalue Problem	38
2.3.2	Two-Dimensional PDE-BVP	42
2.3.2.1	Poisson 2-Dimensional Equation	42
2.3.3	Rational Polynomial Functions	42
2.3.3.1	Reciprocal Function Problem	42
2.3.4	Nonlinear Differential Equations	44
2.3.4.1	Algebraic Riccati Differential Equation	44
3	Convex Optimization Approach for Backstepping PDE Design	47
3.1	Introduction	48
3.1.1	Integral Compact Operators	49
3.2	Parabolic PDE and the Volterra-type Operator	51
3.2.1	Problem Setting	51
3.2.2	Kernel-PDE as a Convex Optimization Problem	52
3.2.3	Approximate Inverse Transformation	55
3.3	Hyperbolic PIDE and the Fredholm-type Operator	59
3.3.1	Problem Setting	59
3.3.2	Existence, Uniqueness and Invertibility	61
3.3.3	Kernel-PIDE as a Convex Optimization Problem	66
3.3.4	Approximate Inverse Transformation	70
3.4	Numerical Results	71
3.4.1	Parabolic PDE with constant reactivity term	71
3.4.2	Parabolic PDE with spatially varying reactivity term	73
3.4.3	Hyperbolic PIDE	75
4	Adaptive Observer Design for a Class of Parabolic PDEs	77
4.1	Introduction	78
4.2	Observer Design for one Uncertain Boundary Parameter	79
4.2.1	Problem Setting	79

4.2.2	Adaptive Observer	80
4.2.3	Design via the Volterra-type Transformation	80
4.2.3.1	Kernel-PDE/ODE as a Convex Optimization Problem	84
4.3	Design for an Uncertain Reactivity Parameter	85
4.3.1	Problem Setting	85
4.3.2	Target System	86
4.3.3	Design for one boundary measurement: Volterra-type Transformation	87
4.3.3.1	Kernel-PDE/ODE as a Convex Optimization Problem	91
4.3.4	Design for two boundary measurements: Fredholm-type Transformation	92
4.3.4.1	Kernel-PDE/ODE as a Convex Optimization Problem	95
4.4	Numerical Results	96
4.4.1	System with an Uncertain Boundary Parameter	96
4.4.2	System with an Uncertain Reactivity Parameter	97
5	Adaptive Observer for a Model of Lithium-Ion Batteries	100
5.1	Introduction	101
5.2	Single Particle Model of the Lithium-Ion Batteries	103
5.3	Adaptive Observer Design	106
5.3.1	SPM Formulation for Observer Design	106
5.3.2	Target System	106
5.3.3	Design via the Volterra-type Transformation	108
5.3.4	Coupled PDE-ODE as a Convex Optimization Problem	111
5.3.5	Uncoupled Kernel-PDE/ODE via Convex Optimization	112
5.4	Output Mapping Inversion	114
5.4.1	Inversion via Moment Approach	115
5.5	Numerical Results	116
6	Conclusions	119
6.1	Differential Equations as Convex Optimization Problems	119
6.2	Convex Formulation of Backstepping Design for PDEs	120
6.3	Adaptive Observer Design	121
6.3.1	Adaptive Observer for Lithium-Ion Batteries	122
6.4	Future Research Directions	123
7	Bibliography	124
Appendix A	Polynomial Optimization via Sum-of-Squares	147
A.1	Polynomial Optimization	147
A.1.1	Sum-of-Squares	149

A.1.2	Moments	150
A.1.2.1	Multi-dimensional Notation	153
A.1.3	Primal-Dual Perspective	154
Appendix B	Backstepping Design for PDEs	156
B.1	The Fundamental Idea	156
B.2	Continuous-Time Approach for Observer Design	157
Appendix C	Pseudo Two-Dimensional Lithium-Ion Battery Model	159
C.1	Governing Equations	159
C.1.1	Input-Output Signals	160
C.1.2	Potentials in the Solid Phase	160
C.1.3	Potentials in the Electrolyte	160
C.1.4	Transport in the Solid Phase	160
C.1.5	Transport in the Electrolyte	161
C.1.6	Conservation of Charge	161
C.1.7	Butler-Volmer Kinetics	161
C.1.8	Effective Coefficients	161
C.1.9	Variables and Parameters	162

List of Tables

3.1	Extract of the moment sequence obtained via convex optimization	72
3.2	Maximum bounds of residual functions for a spatially varying reactivity term	73
3.3	Root mean square of residual functions via partition of the domain-based optimization	75
C.1	Variables and constants in the Pseudo Two-Dimensional Lithium-Ion Battery Model	162
C.2	Parameters in the Pseudo Two-Dimensional Lithium-Ion Battery Model	163

List of Figures

2.1.1 Maximum and minimum values of a residual function in a compact domain.	27
2.1.2 Integral of the square of a residual function in a compact domain.	29
2.1.3 Example of a Polytopic Two-Dimensional Domain Decomposition	30
2.3.1 Polynomial approximation of Steady Convection-dominated Problem	34
2.3.2 Polynomial approximation Steady Diffusion-Reaction Problem with homogeneous Dirichlet boundary conditions	37
2.3.3 Polynomial approximation of Steady Diffusion-Reaction Problem with homogeneous Mixed Dirichlet-Neumann boundary conditions	39
2.3.4 Polynomial approximation of the Sturm–Liouville Eigenvalue Problem	41
2.3.5 Polynomial approximation of 2-Dimensional Poisson PDE	43
2.3.6 Reciprocal polynomial approximation	45
2.3.7 Polynomial approximation of the Algebraic Riccati Differential Equation	46
3.4.1 Parabolic PDE. Bounds for the residual functions	72
3.4.2 Parabolic PDE. Approximate Direct and Inverse Kernels	72
3.4.3 Approximate Kernel via partition of the domain-based optimization	75
3.4.4 Hyperbolic PIDE. Approximate Direct and Inverse Kernels	76
4.4.1 Adaptive Observer Design for PDE with Boundary Uncertain Parameter	98
4.4.2 Adaptive Observer Design for PDE with Reactivity Uncertain Parameter	99
5.2.1 Schematic of an Intercalation Cell and its Single Particle Model.	103
5.5.1 Inversion of the Nonlinear Output Mapping in Lithium-Ion Batteries	117
5.5.2 Adaptive Observer Design for Lithium-Ion Batteries: Case 1	118
5.5.3 Adaptive Observer Design for Lithium-Ion Batteries: Case 2	118
C.1.1 Schematic of an Intercalation Cell.	159

Chapter 1

Introduction

—Jean Baptiste Joseph Fourier (1768-1830)—

J. Fourier is best known for his work: “Théorie Analytique de la Chaleur”, where he made a detailed study of series of trigonometric functions, the convergence of which was clarified later on the basis of the work of G.P. Dirichlet and G. Riemann on discontinuous functions, and led to correct the A. Cauchy’s theorem on the sum of a series of continuous functions, result finally achieved by K. Weierstrass. J. Fourier interest in numerical algorithms can be regarded by his own words: “a numerical interpretation ...is however necessary...the true which it is proposed to discover is no less hidden in the formula of analysis than it was in the physical problem itself”. This view is found in a lesser-known facette of his work, addressing various types of problems in which inequalities appears, where he determined region of solutions, such as convex polygons, and proposed methods to determine optimal points. Due to technical difficulties at that time, these ideas remained dormant until the 20th century, being now elements of the subject known as “Linear Programming”.

1.1 Motivation

Control System theory commonly bases its analysis, design and synthesis on mathematical models of processes. These models generally include dynamics *partially known* which are usually formulated via *parameterized* functions. To determine these parameters, System Identification theory considers *off-line* and *on-line* approaches. In contrast with off-line methodologies,¹ on-line schemes perform a dynamical (*adaptive*) parameter estimation, allowing to analyse conditions, take corrective actions and make predictions as the process evolves. In particular, for Diagnosis and Failure Detection, *combined on-line* “state and parameter” estimation based on parameterizations with *physical mean* provides significant advantages to evaluate the process condition and nature of its failures.

In general, the “parameter estimation problem” is particularly complex for Distributed Parameter Systems (DPS). For instance, in a system with an heterogenous and anisotropic medium, this problem could have infinite unknowns since its parameters are dependent on the position and direction in the spatial domain, respectively [1]. A relevant area where problems of this type can be found is the “Damage Detection in Flexible Structures”, via the estimation of parameters

¹In an off-line approach the model parameters are fitted according to historical data as result of an optimization problem defined for a complete set of these data which is not updated over time [2].

representing elastic properties such as mass, stiffness and damping. In this case the model of the system is typically formulated via “hyperbolic” Partial Differential Equations (PDEs) to describe wave dynamics [3, 4]. Other extensive area corresponds to processes involving “phenomena of transport and conversion”, such as chemical reactors, sediment formation, population dispersal, circulatory physiology, to mention a few. Models for these systems commonly are formulated via first-order hyperbolic and “parabolic” PDEs, where the uncertain *functional* parameters are part of mechanisms of diffusion, reaction and advection [4, 5].

An additional complexity for the parameter estimation problem in DPS is the availability for measuring an incomplete number of states of the system, in particular, the more physically realistic and practical case where only *boundary* measurements are feasible.² In these cases the *simultaneous on-line* “state and parameter” estimation, commonly carried out by the so-called *adaptive observers*, is considered as a challenging mathematical problem and a systematical design is only currently available for some systems which can be formulated via certain “canonical” forms [6, 7, 8].

Nowadays, an appealing type of DPS with boundary control/sensing are “secondary electric batteries”, in particular *Lithium-Ion*-based chemistries. Their electrochemical models are mainly based on the “porous electrode theory” and basically involve diffusion dynamics, the parameters of which are partially known and can vary both temporally and spatially. In batteries the only physical “on-line” points of control and measuring are their terminals/collectors, via the variables of current and voltage, being unfeasible to access to the concentration of the ions, which are considered as the main states of these type of models. Thus, the dynamic relation between *parameterized functions* in the model and, for instance, the “ageing” mechanisms of the battery cannot be validated directly, leading to models with limited ability to make predictions (long-term behaviour; SOC: State-Of-Charge) and diagnose conditions (SOH: State-Of-Health) [9, 10].

The above described system is one of the multiple examples which illustrates the fundamental challenge of estimating states and parameters in DPS from boundary measurements. The underlined practical and theoretical relevance of these problems motivates the study of this thesis for which currently there are no clear ideas or methodologies proposed to deal with it [11].

1.2 A brief review of Distributed Parameter Systems theory

DPS, which are characterized by dynamical models where one or more of its independent variables are functions over the spatial dimensions and time [12], arise in numerous application areas and industrial processes (chemical, electrical, civil, mechanical, aeronautical, etc.) (e.g., see [13, 14, 15, 16] and references therein). Although the practical relevance of DPS is known, control, estimation

²This problem is a particular case of the general subject denominated “Boundary/Point Control of PDEs”, where the actuation and sensing are applied either through the boundary or as points at the interior of the spatial domain. Mathematically, these models involve *unbounded* input and output operators which make much more delicate the analysis of the regularity conditions of the PDEs and control/observer schemes.

and identification for these kind of systems have received considerably less attention than for Lumped-Parameter Systems (LPS) (linear/nonlinear) [17]. This is in partly due to the “infinite dimensional” nature³ of DPS which are normally modelled by PDEs, and its higher complexity⁴ in comparison with Finite Dimensional Systems (FDS) modelled by Ordinary Differential/difference Equations (ODEs). Indeed, a common or traditional approach to address DPS is based on the simplified assumption that its variables are spatially uniform [4].

1.2.1 Control

Control of DPS modelled by PDEs has been studied since the 60’s, research that has been characterized for discovering new phenomena with no proper analogues in finite dimensions [23]. In contrast to LPS, there is no a unified methodology of its treatment, and each problem call for different techniques. These can be mainly classified into two categories [24]: (i) *Early Lumping*: a reduction of the dimensionality is implemented as an initial step so that the system can be consequently studied via a lump parameter model. (ii) *Late Lumping*: the infinite dimension nature is kept as long as possible, particularly during the analysis and design, relegating the reduction of dimension to the synthesis stage.

In the context of *Early Lumping* approach, commonly “model reduction” is carried out via a direct discretization of the spatial operators or via spatial/time transformations [4, 18] leading to a model made up of a set of ODEs or difference equations.⁵ These approaches imply loss of PDE structure and relevant information⁶ needed for a suitable analysis and design [32] (see [24, 33] for fundamental disadvantages).

In the frame of *Late Lumping* approach, for distributed control of linear PDEs (actuation and sensing in the domain), Semigroup Theory⁷ provides a unified framework [12, 34, 35, 36, 37, 38].

³In contrast to LPS, differential operators as functions of the spatial coordinate have infinite number of eigenvalues [18]. The analysis of the behaviour of these infinite elements is far more complex and pathological than for finite dimensions (this can be seen by means of the analysis of infinite series and its convergence, keeping the notion of linear combination and linear independence of finite-dimensional vector spaces [19]), for instance: there is no equivalence for all norms; a closed and bounded set in a metric space does not need to be compact; a bijective linear mapping does not need to be invertible; linear operators do not need to be continuous; etc. [20].

⁴As is commented in [21], this complexity turns out “impenetrable” for control engineers commonly with no background in PDEs and functional analysis. An example of such complexity can be seen in the design of optimal controllers for PDEs by the resolution of Riccati equations in terms of continuous operators [22].

⁵Finite Difference Method (FDM) is one of the most popular method that allow a simple representation of PDEs by a direct approximation of its differential operators. A substantive number of control applications have used FDM to extend ideas from LPS to DPS (see discussion and references in [25]).

⁶In general common discretization methods do not “preserve structure” [26]. Structure-preserving formulation of PDEs mimics their essential mathematical-geometrical properties and physical principles (symmetries, divergence free, conservation laws, etc.) [28, 29, 27]. The Discrete Exterior Calculus methodology is a possible general approach for formulating different mimetic numerical methods for PDEs [31, 30].

⁷In the context of Semigroup theory, the Hille-Yosida theorem and the Lumer-Phillips theorem plays a relevant role providing a necessary and sufficient condition and more simple verification of this condition, respectively, for a differential operator to be an infinitesimal generator of semigroup of contractions. This is the condition for a unique solution of the abstract Cauchy problem which, for instance, by means of its spectrum analysis allows determining the stability of its equilibrium points [34, 12, 18, 35].

Indeed, control theory in infinite dimensional spaces owes its initial development to semigroup theory by means of the extension of the spectral analysis and Lyapunov stability theory from FDS to Infinite Dimensional Systems (IDS) [12, 37, 39]. The control action on abstract differential equations (Banach or Hilbert spaces [40]) is commonly formulated in terms of a bounded operator, where semigroup theory is suitable to solve the problem. By contrast, this situation change notably if the actuation/sensing is applied only through the boundary. “Boundary control” of PDEs is physically more realistic (industrial settings) and mathematically much more challenging. In this case, the input operators are unbounded, making difficult their representation and treatment [22, 41, 42]. A particularly important extension, based on semigroup theory, is the classical theory of optimal control to IDS [12, 34, 43, 44, 45], which leads to Operator Riccati Equations, an area of research in its own right (e.g, see [46, 47, 48] and references therein). In the context of “late lumping”, besides semigroup theory, a short variety of methodologies has been developed for extending the control design to IDS, each one with its own limitations (see discussion and references in [49]), where the so-called *Backstepping for PDEs*⁸ emerges as an innovative approach providing a systematical design [21, 50].

1.2.2 State Estimation

The extension of state estimation theory of FDS⁹ to IDS has been mainly carried out on the basis of the *least square* optimization problem formulated in Hilbert spaces [55], for instance: the optimal filtering problem which leads to the Riccati equation and Kalman filter [56]. This problem is related to the “detectability” of the system, a weaker property than observability,¹⁰ which for IDS depends on the location of the measurement points [12].

Essentially in *observer design* the central idea of the Luenberger observer [59] has been extended to IDS (see a survey in [60]). Similar to “controllability”, the concept of “observability” in DPS has been extensively discussed, with no absence of controversies.¹¹ Practical considerations have been discussed due to the complexity of the PDE models [64, 61, 12, 24, 17], leading to exact and approximate definitions¹²[12]. An example of these definitions, which particularly shows the

⁸For a complete exposition of this methodology see [21, 50, 51, 18, 49]. A brief summary related to the observer design problem is presented in Appendix B.

⁹Estimation theory has occupied a prominent place in FDS from the pioneering work of Norbert Wiener [52] and the seminal contributions of Rudolph Kalman [53], in particular for state-spaces models of linear systems, where the method of *Least Squares* (dates back to Karl Gauss around 1795) has been the main theoretical/practical tool [54].

¹⁰A linear FDS is detectable if all unstable modes are observable. This allows “detecting”, via output measurements, a difference in trajectories based on differing their asymptotic behavior [57, 58]. In FDS stability can be analyzed via its spectrum, which is not always feasible in IDS since the corresponding semigroup does not provide a proper lower bound, property denominated “spectrum determined growth assumption” [12].

¹¹In a interesting discussion in [61] the authors argue about the definition of observability based on the uniqueness of solutions instead of recovering of the initial condition [62, 63]. An example includes a PDE model with a unique solution and observable in a specific time-interval domain function of the spatial variable.

¹²As is commented in [12] many conceptions of Controllability/Observability have a limited use in practical control synthesis and design. Definitions of FDS (“exact” concepts) can be very strong in IDS and most of the systems can achieve these in an “approximate” sense. However, in general, approximate notions need not imply stabilizability/detectability [12].

dependence of the observability on the spatial location of the sensors is the case of *Linear Parabolic PDEs*, where by means modal analysis, an approximate observability is always satisfied with one boundary measurement [64, 24].¹³

Observer design using *early lumping* approaches is a straightforward strategy to deal with DPS (e.g., see discussions and references about this in [24, 33, 32, 25]). However, its deficiencies are well-known, as it has been early reported in the control/observer design for flexible structures:

- Based on a discretized model via the Finite Element Method (FEM)[65], it is shown that a second-order observer is not convergent, although the reduced model is observable [66]. On the other hand via modal analysis [4, 67],¹⁴ there is an unreliable estimation of spatially variable medium parameters (such as mass, stiffness and damping) [69, 70].
- With respect to the spectral method, a phenomenon known as *control and observation spillover* can occur (source of closed-loop instability) due to the residual modes (uncontrolled) not considered in FDS representation [71, 72].

In the context of *late lumping approaches*, aside from the extension of linear LPS theory to DPS [12, 55], a few alternatives of observer designs can be found in the literature [60]. These designs can mainly be classified under the Lyapunov stability analysis or energy minimization methodologies [18, 73, 74, 75, 76]. Recently two methodologies of observer design, particularly appealing due to its ability to reduce the Lyapunov analysis (on the full PDE model) to a simple finite-dimensional problem, have been developed. One of these approaches derives Linear Matrix Inequalities (LMIs) conditions for exponential stability directly from the model parameters and their bounds [77, 78]. The second alternative uses positive definite matrices to parameterize Lyapunov functions via Sum-of-Squares (SOS) decomposition and polynomial optimization [79, 80, 81]. On the other hand, Backstepping for PDEs allows a simple observer design where the dynamical properties of the estimate stated are in advance imposed by the selection of a particular stable (target) error system (e.g., see [21, 50, 49, 82, 83, 84, 85, 86]). All these approaches previously mentioned provides a systematical design and effective constructive tools for analysis and control/estimation of DPS, with the main advantage of avoiding abstract formulation in Hilbert spaces and operator equations.

¹³A practical test for approximate contrabability/observability is based on the “spectral analysis” of the differential operators denominated *Modal Controlable/Observable* condition [64]. However, its equivalence is true for “Riesz-spectral systems”, such as linear parabolic PDEs.

¹⁴Spectral/Modal decomposition consists in a time-scale separation and model reduction. Commonly the popular Galerkin Method [68] is used to derive an approximate FDS, where well-known LPS theories can be applied directly. Its application depends on the eigenspectrum of the spatial differential operators. For instance, for first-order hyperbolic PDEs, the eigenvalues are along vertical or nearly vertical asymptotes in the complex plane. It implies a large number of modes to capture the PDE dynamic behaviour and therefore makes prohibitive the use of this method [4].

1.2.3 Parameter Estimation

The parameter estimation problem (inverse problem)¹⁵ is ubiquitous in modeling and particularly complex in DPS, since the non-uniqueness difficulties not only arise from noisy data and insufficient number of observations, but also from sparse/boundary measurements and the objective to determine “functional” model parameters [88, 89, 90]. This problem has received considerable attention, the initial research of which was carried out by means of “early lumping” approaches (for a extensive review of early methodologies see [91]). Similar to filtering and state space estimation in IDS, parameter estimation has been principally addressed using the *Least Squares* method. In particular on the basis of *off-line* schemes considering “abstract evolution” equations in Hilbert spaces, where the quadratic optimization is analysed as an infinite dimensional problem (e.g., see [5, 92, 3, 38] and references therein; substantive body of research of H.T. Banks, K. Ito, F. Kappel, I.G. Roseen, K. Kunish and J.A. Burns). Methodologies based on the variational approach have been less widespread, which involve a numerical solution of Riccati-like PDEs [93].

In parameter estimation a performance criterion is usually defined as a functional of the output estimation error. The question of whether estimated parameters coincide with the parameters of the process when the criterion is minimized (or output error vanishes) is referred to the “parameter identifiability problem” [94, 95]. Commonly, an analysis of *sensitivity* of the performance criterion with respect to the unknown parameters is one of the first steps of analysis, providing insights about the tractability of the estimation and, for instance, its dependence on the spatial location of the measurements. However, this useful tool cannot state “sufficient” conditions to guarantee uniqueness of the solution in the identification [91]. For linear parabolic PDEs, the identifiability problem has received significant attention, with an extensive history of theoretical and practical results, in particular which are framed in the Sturm-Liouville theory [96] (e.g., see [97, 98, 95] and references therein). For instance, some identifiability results show (necessary and sufficient conditions) that it is possible to identify a constant parameter (diffusivity) from only one measurement at the boundary [95, 99, 100, 101].¹⁶ Surprisingly, these studies involve “transformation operators” (deformation formulas [102]) which lead to PDE problems and methodologies similar to those associated with the Backstepping design for PDEs [103, 96].

The parameter estimation problem has had a relevant practical role in “Fault detection and Isolation”,¹⁷ where the identification of spatially varying parameters [5] conducts to detect physical changes that can be evaluated as faults [69, 105, 3, 106]. In general such approaches, by means of

¹⁵An inverse problem considers as unknown the system inputs (reconstruction problem) or the system parameters (identification problem)[1, 87].

¹⁶It is worth noting that approximations of differential operators, such as the “finite difference method”, can cause lost of identifiability of a parameter even though this is mathematically (theoretically) identifiable. See example 2 in [95].

¹⁷Damage identification for mechanical, civil and aerospace systems (flexible structures) has been investigated for many years, including analytical methods based on models of the structures. However, a mayor part of this investigation mainly corresponds to signal processing, where most of the techniques consider the spectral analysis of vibrations measurements (e.g., see [104] and references therein).

discretization methods, transform the IDS problem into a finite dimensional parameter estimation problem. These can be seen as part of more general *Non-Destructive Testing* (NDT) methods, carried out by modal analysis, time or frequency domain, mostly applied in the analysis of flexible structures [107, 108]. In this context a novel approach using spatio-temporal identification based on the *Coupled Map Lattice* (CML) model¹⁸ and orthogonal forward regression least-squares algorithm has been presented by [110]. This work is part of a wide contribution of the authors in the field of systems identification, firstly on LPS and now extending these ideas to DPS (e.g, see [109, 111] and references therein).

1.2.3.1 Adaptive Observers

Parameter estimation has also been addressed as an *on-line* problem of simultaneous estimation of states and parameters, scheme which is commonly known as *adaptive observer*. The main body of research of this approach, for a class of PDEs and time-varying parameters, is based on the semigroup theory according to the initial work of W. Scondo [112, 113] and the extensive research of M.A. Demetriou [114, 115, 116, 117, 118, 119, 120, 7] (for a unified view see [121]).

In the context of *late lumping* a very limited number of alternative approaches can be found in the literature. One of these, which is currently an active line of research, is the adaptive version of Backstepping design for PDEs [50]. In general adaptive methods in Backstepping needs state measurements on the full space domain [50, 122, 123, 124]. In common with adaptive observers based on semigroup theory using boundary measurements, *a feasible design is “only” achievable for particular classes of system structures* (“canonical observer form” [125]) where the unknown parameters multiply the output of the system. For instance, in systems similar to (linearized model of heat conduction in solid propellants [126]):

$$\begin{aligned}\frac{\partial u}{\partial t}(x, t) &= \frac{\partial^2 u}{\partial x^2}(x, t) + \gamma(t)y(t), \\ \frac{\partial u}{\partial x}(0, t) &= -\rho(t)u(0, t), \quad u(1, t) = U(t), \\ y(t) &= u(0, t),\end{aligned}\tag{1.1}$$

where $x \in [0, 1]$ is the spatial variable, u is the dependent variable (temperature), γ is the “in-domain” unknown parameter, ρ is the “boundary” unknown parameter, U is the control signal and y is the measurable output. It is worth noting that this type of structure can be seen as part of the “*Positive Real*” *infinite dimensional systems*¹⁹(e.g., see [6] and references therein).

¹⁸CML can exhibit rich dynamical behaviour and it has been applied to model a vast kind of systems. It can be formulated directly in discrete time or by using discretization on PDEs, which is usually carried out via simple finite differences schemes [109].

¹⁹Similarly to adaptive control, for a class of systems, linear adaptive observers in FDS can be design via the Lyapunov’s stability analysis based on the condition of existence of “Positive Real” (PR) functions, relation stated via the Kalman-Yakubovich-Popov lemma and its variants [127]. This relation has been extended to IDS setting forth the stability condition in terms of the existence of a solution of the Riccati equation or Lur’e equation (operator

In particular, Backstepping-based adaptive observers have been applied to estimate “boundary” unknown parameters [130, 131, 132, 133, 134, 135]. As has been commented in [11], “there is no clear way” to address systems with more general structures and the adaptive observer design remain as a *significant fundamental challenge* in IDS.

1.3 Thesis Outline

1.3.1 Objectives

This research has as main objective the *simultaneous estimation* of “states and uncertain parameters” for a class of 1-dimensional linear parabolic PDEs described by:

$$\begin{aligned} \frac{\partial u}{\partial t}(x, t) &= \epsilon(t) \frac{\partial^2 u}{\partial x^2}(x, t) + \lambda(t)u(x, t), \\ a_1 u(0, t) + b_1 \frac{\partial u}{\partial x}(0, t) &= c_1 \theta(t)U(t), \\ a_2 u(1, t) + b_2 \frac{\partial u}{\partial x}(1, t) &= c_2 \theta(t)U(t), \\ y(t) &= u(0, t) \quad \text{and/or} \quad y(t) = u(1, t), \end{aligned} \tag{1.2}$$

where “ t ” is the time variable, “ x ” is the spatial variable in the unit interval $\Omega = [0, 1]$, “ u ” is the dependent variable which describes the states of the DPS model, “ ϵ ” is the *diffusivity* coefficient and “ λ ” is the *reactivity* coefficient, which are considered as *in-domain time-varying uncertain* parameters. This class of DPS *only* considers available measurements at the boundary (“ y ”: output of the model) and a control action “ U ” applied on a single boundary, where “ θ ” is its *time-varying unknown* gain. The coefficients a_1, a_2, b_1, b_2, c_1 and c_2 allows managing mixed boundary conditions (Dirichlet and/or Newman).

This objective is pursued considering two central conditions:

- Develop a “systematic/constructive” observer design methodology in the context of the *late lumping approaches*, i.e., the analysis and design do not include a model reduction of the PDEs.
- Develop a “computational/numerical method” to determine the observer gains formulated as a *finite dimensional optimization* problem. In particular, obtain “approximated-analytic” functions for these gains.

The objective is separated into three cases of study, of increasing complexity, where the observer design only deals with one unknown parameter. (i) Firstly, the gain of the control action “ θ ” is

equations). Commonly, to obtain feasible equations an “exact observability” (controllability) property is assumed, condition which is very strong requirement in IDS. More practical/simple approaches consider the “exponential detectability” condition and the observer design is carried out so that the compensated error system generates a “exponential stable C_0 semigroup” [128, 129, 12, 6].

considered as the unknown parameter. (ii) Secondly, the reactivity coefficient “ λ ” is assumed as the uncertain parameter. (ii) Finally, the diffusivity coefficient “ ϵ ” is the only uncertain parameter in the model that is considered.

1.3.2 Organization and Contributions

In accordance with the cases/objectives of study mentioned, this thesis is organized as follows:

- **Chapter 2** presents a novel methodology to solve Differential Equations-Boundary Value Problems (DE-BVPs) based on polynomial functions, Sum-of-Squares decomposition and convex optimization. This chapter includes numerical examples to illustrate the performance of the method proposed in comparison with some traditional approaches. This method is used to solve PDEs/ODEs derived from the observer design problem addressed in the subsequent chapters. Its contributions are:
 - The DE-BVPs are recast as a convex optimization problem, readily implementable via semidefinite programming tools, allowing fast computations and accurate results.
 - The method avoids consideration of weak formulations, boundary conditions approximations and build-up of the resulting algebraic relations. On the contrary, this method is applied directly to the original DE-BVP for a selected polynomial degree.
 - The method leads to “approximated-analytic” solutions (polynomial functions). This can also be applied to some classes of differential equations with quadratic terms, multiplicative inverse functions, amongst other classes of problems.
- **Chapter 3** describes the boundary PDE control/observer design under Backstepping methodology formulated as a convex optimization problem. The Volterra-type and Fredholm-type transformations are analysed and the resulting Kernel-PDEs are solved via semidefinite programming. Inverse Kernels are approximated as the optimal solution of a Sum-of-Squares and moment problem. The main contributions in this chapter are:
 - The approach proposed allows solving approximately Kernel-PDEs with sufficient precision to guarantee the stability of the closed-loop system in $\mathcal{L}^2$ -norm topology.
 - The convex formulation proposed and its numerical solution allows applying this methodology to a wide class of problems, without restriction on the spectral characteristic of some integral operators or classes of system’s coefficients.
 - For polynomial Kernels, uniqueness and invertibility of the Fredholm-type transformation is proved in the space of real-analytic functions and continuous functions.
 - The inverse Kernels (for the Volterra-type or Fredholm-type transformations) can be determined approximately using the direct Kernels, avoiding to solve PDE-Kernels using an inverse problem formulation.

- **Chapter 4** presents an adaptive observer design for a linear parabolic PDE under two separated cases of uncertain time-varying parameters:

- ▷ at the boundary: control gain parameter,
- ▷ in-domain: reactivity parameter.

The design proposes a modified integral transformation to compensate for the parameter uncertainty. In addition to an observer based on the Volterra-type transformation, an observer design method with two boundary measures via a Fredholm-type transformation is also formulated. The coupled Kernel-PDEs and ODEs are recast as convex optimization problems and solved via semidefinite programming. The contributions of this chapter are:

- The adaptive observer performs a simultaneous state and parameter estimation considering only measurements at the boundary. The design does not require any pre-transformation to some canonical form and/or finite dimensional formulation, enabling a direct parameter identification from the original PDE model.
 - The modified Volterra-type and Fredholm-type transformations proposed allows compensating for the uncertain time-varying “boundary” or “in-domain” parameter.
 - For two points of measurement, a Fredholm-type integral transformation is proposed which allows managing the resulting non-strict feedback error system.
- **Chapter 5** deals with the adaptive observer design problem for a Single Particle Model of Lithium-Ion batteries where the diffusivity coefficient is considered the uncertain parameter. The design proposes a modification of a Volterra-type transformation to compensate for the parameter uncertainty. The resulting coupled/uncoupled Kernel-PDE/ODE is recast as a convex optimization problem and solved via semidefinite programming. In addition, based on the Moment approach, a novel scheme of inversion of the nonlinear output mapping of the battery is formulated. The contributions of this chapter are:
- The adaptive observer performs a simultaneous estimation of the solid Lithium-Ion concentration and of the diffusivity parameter, considering only boundary measurements of voltage and current. The design does not require any model reduction, taking into account the full nature of the PDE model.
 - The observer design methodology also considers the uncertain gain in the boundary condition which is a reciprocal function of the diffusivity parameter. This boundary uncertainty is dynamically compensated in conjunction with the in-domain uncertainty, so that the adaptation scheme proposed leads to an accurate parameter estimation.
 - The linear feedback problem in the observer is solved via a novel method of inversion which circumvents the limited performance of gradient-based linear approximations com-

monly used. The method proposed achieves a high precision for the whole domain and is able to determine multiple optimal points.

- **Chapter 6** provides concluding remarks of this research and suggestions for future lines of investigation.

1.3.2.1 Publications

The contribution of this research has been disseminated in the following papers:

- P. Ascencio, A. Astolfi and T. Parisini, “Backstepping PDE Design: A Convex Optimization Approach,” Submitted to *IEEE Transactions on Automatic Control*, 2016.
- P. Ascencio, A. Astolfi and T. Parisini, “Backstepping PDE-based adaptive observer for a Single Particle Model of Lithium-Ion Batteries,” *IEEE 56th Conference on Decision and Control*, 2016, pp. 5623-5628.
- P. Ascencio, A. Astolfi and T. Parisini, “An Adaptive Observer for a class of Parabolic PDEs based on a Convex Optimization Approach of Backstepping PDE Design,” *American Control Conference*, 2016, pp. 3429-3434.
- P. Ascencio, A. Astolfi and T. Parisini, “Backstepping PDE Design, Volterra and Fredholm Operators: a Convex Optimization Approach,” *IEEE 55th Conference on Decision and Control*, 2015, pp. 7048-7053.

1.4 Preliminary Terminology

1.4.1 Notation

$\mathcal{A}(\Omega)$, $\mathcal{C}(\Omega)$ and $\mathcal{L}^2(\Omega)$ stand for the space of real-analytic functions, continuous functions and square integrable functions on the domain Ω , respectively. $\mathbb{I}$, $\mathbb{A}$, $\mathbb{V}$ and $\mathbb{F}$ denote the identity, integral, Volterra-type and Fredholm-type operator, respectively.

$\mathbb{R}[\mathbf{x}]$ denotes the ring of real polynomials in n variables $\mathbf{x} = [x_1, x_2, \dots, x_n]^T$ and $\mathbb{P}[\mathbf{x}] = \{p \in \mathbb{R}[\mathbf{x}] : p(x) \geq 0, \forall x \in \mathbb{R}^n\}$ stands for the set of non-negative real polynomials. The notation $\mathbb{R}_{n,r}[\mathbf{x}]$ and $\mathbb{P}_{n,r}[\mathbf{x}]$ explicitly indicates polynomials in n variables with degree at most r , whereas Σ_s represents the subset of polynomials with Sum-of-Squares (SOS) decomposition. In particular $\mathbb{P}(\mathbf{K})$ represents the non-negative polynomials on the set $\mathbf{K}$.

$\Phi_r = [1, x_1, \dots, x_n, x_1^2, x_1x_2, \dots, x_n^r]^T$ is the standard vector basis of $\mathbb{R}_{n,r}[\mathbf{x}]$. Polynomials are expressed by multi-index notation: $p(\mathbf{x}) = \sum_{j=0}^{z(r)-1} p_j \mathbf{x}^{\bar{\alpha}_j} = \langle \bar{p}, \Phi_r \rangle$, where $r \in \mathbb{N}$ is the polynomial degree, $z(r) = \binom{n+r}{r}$ is the number of polynomial coefficients $\bar{p} = [p_0, p_1, \dots, p_{z(r)-1}]^T \in \mathbb{R}^{z(r)}$, $\mathbf{x}^{\bar{\alpha}_j} = x_1^{\alpha_j^1} \cdots x_n^{\alpha_j^n}$ represents the j -th monomial with powers $\bar{\alpha}_j = [\alpha_j^1, \dots, \alpha_j^n]$ such that $|\bar{\alpha}_j| =$

$\sum_{k=1}^n \alpha_j^k \leq r$, $\alpha_j^k \in \mathbb{N}$. For simplicity it is expressed by $p(\mathbf{x}) = \sum_{\alpha \in \mathbb{N}_r^n} p_\alpha \mathbf{x}^\alpha$ with powers $\alpha \in \mathbb{N}_r^n$, where $\mathbb{N}_r^n = \{\bar{\alpha}_j \in \mathbb{N}^n; |\bar{\alpha}_j| \leq r, \forall j = 0, \dots, z(r) - 1\} = \{\bar{\alpha}_0, \dots, \bar{\alpha}_{z(r)-1}\}$.

$\mathcal{M}(\mathbf{K})_+$ stands for the space of finite positive Borel measures on $\mathbf{K}$. $\mathbb{S}^m$ and $\mathbb{S}_+^m$ denote the set of symmetric and semidefinite positive matrices of dimension $m \times m$, respectively.

The expression $\delta = \delta(\mathbf{x}) \Big|_{Q \approx M}^{P \approx N}$ indicates that in the function δ , P and Q have been substituted by the polynomials N and M , respectively.

1.4.2 Definition

For the representation of bivariate polynomial Kernels $P(x, y) = \sum_{k=0}^{z(d)-1} p_k x^{\alpha_k} y^{\beta_k}$ of degree $d \in \mathbb{N}$, coefficients $p_k \in \mathbb{R}$ and powers $\alpha_k \in \mathbb{N}$, $\beta_k \in \mathbb{N}$, the following basis of monomials Φ_d are considered:

$$\begin{aligned} \Phi_d &:= \{1, x, y, x^2, xy, y^2, x^3, x^2y, xy^2, y^3, \dots, \dots, x^d, x^{d-1}y, \dots, xy^{d-1}, y^d\}, \\ \alpha &:= \{0, [1, 0], [2, 1, 0], \dots, [d, d-1, \dots, 1, 0]\}, \\ \beta &:= \{0, [0, 1], [0, 1, 2], \dots, [0, 1, \dots, d-1, d]\}, \end{aligned} \quad (1.3)$$

with $z(d) = (d+1)(d+2)/2$ terms, powers of which are grouped in $d+1$ blocks, where each j -th block contains $j+1$ elements ordered in relation with $[x^{j-k}y^k]$, $k=0, \dots, j$, $j=0, \dots, d$.

Backstepping design for 1-dimensional PDEs with a Volterra-type or Fredholm-type transformation involves the following domains [21, 82]:

$$\Omega := \{0 \leq x \leq 1\}, \quad (1.4)$$

$$\Omega_L := \{0 \leq x \leq 1, 0 \leq y \leq x\}, \quad (1.5)$$

$$\Omega_U := \{0 \leq x \leq 1, x \leq y \leq 1\}, \quad (1.6)$$

$$\Omega_S := \{0 \leq x \leq 1, 0 \leq y \leq 1\}, \quad (1.7)$$

which are the unit interval, the lower and upper triangular subsets sharing a boundary along $x = y$, and the unit square subsets of $\mathbb{R}^2$, respectively.

1.4.3 Abbreviations

BVP	Boundary Value Problem	LPS	Lumped-Parameter Systems
DFN	Doley-Fuller-Newman	OCP	Open Circuit Potential
DPS	Distributed Parameter Systems	ODE	Ordinary Differential Equation
FDS	Finite Dimensional Systems	P2D	Pseudo Two-Dimensional
FDM	Finite Difference Method	PDE	Partial Differential Equation
FEM	Finite Element Method	PDAE	Partial Differential Algebraic Equation
IDS	Infinite Dimensional Systems	PIDE	Partial Integral Differential Equation
IVP	Initial Value Problem	SOS	Sum-of-Squares
LI	Linearly Independent	SPM	Single Particle Model

Chapter 2

Formulation of Differential Equations as Convex Optimization Problems¹

—Pierre de Fermat (1601-1665)—

Beyond his contributions on the “Number Theory” and the celebrated so-called Fermat’s last theorem, P. Fermat is considered father of the “Analytic Geometry”, title shared with René Descartes. Fermat’s work on this area, dealing with the relationship of algebra to geometry, was one of the primary inspirations for the work of Sir Isaac Newton on the subject that today is known as “Differential Calculus”. His idea to obtain a tangent line of a curve, based on simple geometric observations, led to find local maxima and minima of differentiable functions, which is now known as the “interior extremum theorem”.

On the basis of the Weierstrass approximation theorem continuous smooth solutions of differential equations (boundary value problems), in compact sets, can be uniformly approximated by polynomials. In contrast to the traditional approaches such as discrete approximations of differential operators or variational formulations, a simple methodology to approximate these solutions can be formulated via recasting the *equality* condition of the differential problem as *inequalities* written in terms of polynomial coefficients, and relaxing these via their Sums-of-Squares (SOS) decomposition over a compact set. This methodology leads to a convex optimization problem, readily implementable resorting to semidefinite programming tools, the objective of which corresponds to minimize the residual functions with respect to the original equality. This flexible and generic approach is suitable to be applied on a wide class of differential equations, linear or with quadratic terms, amongst other classes of problems with polynomial terms.

¹A brief theoretical background of this chapter is presented in Appendix A. For a detailed exposition see [136, 137, 138] and references therein.

2.1 Sum-of-Squares Approach for PDEs

Let $u = u(\mathbf{x}) \in \mathcal{C}^\infty(\Omega)$ be the solution of a linear PDE-Boundary Value Problem (BVP):

$$\mathbf{P} : \begin{cases} \mathbb{L}[u(\cdot)](\mathbf{x}) = f(\mathbf{x}), \quad \forall \mathbf{x} \in \mathring{\Omega} \\ \mathbb{B}_i[u(\cdot)](\mathbf{x})|_{\mathbf{x} \in \partial\Omega_i} = \mathbf{u}_i, \quad i = 1, \dots, r \in \mathbb{N} \end{cases}, \quad (2.1)$$

in a compact domain $\Omega \subset \mathbb{R}^n$, where $\mathbb{L}$ and $\mathbb{B}_i$ are linear differential operators with polynomials terms, $f \in \mathbb{R}[\mathbf{x}]$, $\partial\Omega \supseteq \bigcup_{i=1}^r \partial\Omega_i$ represents the boundary of $\Omega = \mathring{\Omega} \cup \partial\Omega$, with $\mathring{\Omega}$ interior subset of Ω , and $\mathbf{u}_i \in \mathbb{R}$ is the value of $\mathbb{B}_i[u]$ on the boundary subset $\partial\Omega_i$. Based on the Weierstrass approximation theorem [139], the solution u can be uniformly approximated by polynomials with theoretical arbitrary precision.

Let

$$\delta_d(\mathbf{x}) = \left(\sum_{\alpha \in \mathbb{N}_d^n} p_\alpha \mathbb{L}(\mathbf{x}^\alpha) - f(\mathbf{x}) \right), \quad \forall \mathbf{x} \in \mathring{\Omega} \quad (2.2)$$

be the **residual function** due to the polynomial approximation $u(\mathbf{x}) \approx \mathbf{p}_d(\mathbf{x}) = \sum_{\alpha \in \mathbb{N}_d^n} p_\alpha \mathbf{x}^\alpha \in \mathbb{R}[\mathbf{x}]$ in $\mathbf{P}$ (2.1), of degree d in n variables $\mathbf{x} = [x_1, \dots, x_n]^\top \in \mathbb{R}^n$ and coefficients $\mathbf{p} = [p_{\alpha_0}, \dots, p_{\alpha_{z(n,d)-1}}]^\top \in \mathbb{R}^{z(n,d)}$ with $z(n,d) = \binom{n+d}{d}$.

In contrast to the traditional and well-established methodologies such as for instance: the Finite Difference Method (FDM) or the Finite Element Method (FEM),² a direct methodology to determine the polynomial $\mathbf{p}_d$ consists in minimizing the *residual function* δ_d , by means of formulating it as a **Polynomial Optimization Problem** in terms of the polynomial coefficients p_α . This approach can be seen as part of the *residual methods* to address Boundary Value Problems [147].

2.1.1 Compact Basic Semi-algebraic Sets

The main idea to solve (2.1) as a Polynomial Optimization Problem is based on a notable result from Real Algebraic Geometry: the *Positivstellensatz* [136]. Peculiarly, this result of *positivity certification* does not depend on the characteristics of the polynomials involved in the problem. On the contrary, this result only depends on the kind of *algebraic representation* of the domain Ω . Thus, if Ω can be described by

$$\Omega = \{\mathbf{x} \in \mathbb{R}^n; g_1(\mathbf{x}) \geq 0, \dots, g_m(\mathbf{x}) \geq 0\}, \quad (2.3)$$

²FDM, which is based on the approximation of differential operators by differences (forward, central, backward) on a grid or mesh [140, 141], or FEM, which could be derived from the *weak (variational) formulation* of a differential equation and the *Ritz-Galerkin approximation* [142, 65, 143, 144], can lead to non-convergent solutions. For the FDM, an analysis of convergence can be carried out based on the “maximum principle” [145]. For the case of the FEM, a sufficient condition to obtain a convergent solution consists in to formulate structures in accordance with the Lax-Milgram lemma (bilinear form: $\mathcal{A}(u, v) : V \times V \rightarrow \mathbb{R}$, with V Hilbert space, continuous and coercive), for instance, the particular case of self-adjoint operators [146].

with $m \in \mathbb{N}$ and $g_j \in \mathbb{R}[\mathbf{x}]$, $\forall j = 1, \dots, m$, and this description is a **compact basic semi-algebraic set**, based on the representation theorems of Schmüdgen or Putinar,³ the non-negative function $h(\mathbf{x}) = \pm\delta_d(\mathbf{x}) \mp \gamma$, for some $\gamma \geq 0$, can be formulated as:

$$\pm\delta_d(\mathbf{x}) \mp \gamma = \sum_J s_J(\mathbf{x})G_J(\mathbf{x}), \quad s_J \in \Sigma_s, \quad \forall \mathbf{x} \in \Omega, \quad (2.4)$$

where G_J denotes a particular combination of polynomial constraints g_j 's in accordance with the Schmüdgen or Putinar representation selected ($G_0 = 1$). Moreover, this allows relaxing (2.4) via

$$\left(\pm\delta_d(\mathbf{x}) \mp \gamma - \sum_{J \neq 0} s_J(\mathbf{x})G_J(\mathbf{x}) \right) \in \Sigma_s, \quad s_J \in \Sigma_s, \quad \forall \mathbf{x} \in \Omega, \quad (2.5)$$

which is a SOS decomposition (hierarchy of SOS relaxations), equivalent to a convex optimization problem, numerically implementable via semidefinite programming [136, 138].

Consider the particular domains involved in Backstepping PDE design.

Lemma 1. *The domains (1.4), (1.5), (1.6) and (1.7) have the representations:*

$$\Omega \equiv \{x \in \mathbb{R}, g_1(x) = x(1-x) \geq 0\}, \quad (2.6)$$

$$\Omega_L \equiv \{(x, y) \in \mathbb{R}^2, g_1(x) = x(1-x) \geq 0, g_2(x, y) = y(x-y) \geq 0\} \quad (2.7)$$

$$\Omega_U \equiv \{(x, y) \in \mathbb{R}^2, g_1(x) = x(1-x) \geq 0, g_3(x, y) = (y-x)(1-y) \geq 0\} \quad (2.8)$$

$$\Omega_S \equiv \{(x, y) \in \mathbb{R}^2, g_1(x) = x(1-x) \geq 0, g_4(x, y) = y(1-y) \geq 0\} \quad (2.9)$$

respectively, which are “Compact Basic Semi-Algebraic Sets” and their associated quadratic modules are “Archimedean”.⁴

Proof. By definition (2.6), (2.7), (2.8) and (2.9) are basic semi-algebraic sets [136, 149]. The equivalence of the sets (1.4)-(1.7) and their respective representations (2.6)-(2.9) is immediate, following the feasible solution set of the inequalities involved. Moreover, (1.4) is a closed and bounded subset of $\mathbb{R}$ as well as the sets (1.5)-(1.7) are in $\mathbb{R}^2$, and hence compact [150]. The Archimedean property is verified directly since for each domain representation the set $\{\mathbf{x} \in \mathbb{R}^n : g(\mathbf{x}) \geq 0\}$ is compact for some $j \in \{1, \dots, m\}$ [138]. This property can also be validated since the quadratic modules associated to Ω , Ω_L , Ω_U and Ω_S satisfy the condition [148, 138]:

³These certificates have been considered due to they are computationally practical and simple of implementing. In particular, Putinar's Positivstellensatz only involves a linear number of SOS polynomials with respect to the compact domain representation [148, 136, 138].

⁴Since the domains considered are polytopes, the functions g_i can also be selected as affine inequalities where this condition is always satisfied [138].

$$\exists N \in \mathbb{N} \text{ s.t.: } N - \sum_{i=1}^n x_i^2 \in \left\{ s_0 + \sum_{j=1}^m s_j g_j; (s_j)_{j=0}^m \in \Sigma_s \right\}. \quad (2.10)$$

For instance, for Ω as in (2.6), selecting $N = 1$, $s_1 = 2 \geq 0$ and $s_0(x) = (x-1)^2 \geq 0$. For the lower triangular domain Ω_L as in (2.7), using:

$$\begin{aligned} N_L &= 2 \in \mathbb{N}, s_1 = 4 \geq 0, s_2 = 2 \geq 0, \\ s_0(x, y) &= 3x^2 - 2xy - 4x + y^2 + 2 = [1, x, y]Q_L[1, x, y]^T \in \Sigma_s, \\ Q_L &= \begin{pmatrix} 2 & -2 & 0 \\ -2 & 3 & -1 \\ 0 & -1 & 1 \end{pmatrix} \succeq 0 \Rightarrow N_L - x^2 - y^2 = s_0(x, y) + s_1 g_1(x) + s_2 g_2(x, y). \end{aligned}$$

Likewise, in the case of the upper triangular domain Ω_U as in (2.8), considering:

$$\begin{aligned} N_U &= 3 \in \mathbb{N}, s_1 = 4 \geq 0, s_2 = 2 \geq 0, \\ s_0(x, y) &= 3x^2 - 2x - 2y - 2xy + y^2 + 3 = [1, x, y]Q_U[1, x, y]^T \in \Sigma_s, \\ Q_U &= \begin{pmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 1 \end{pmatrix} \succeq 0 \Rightarrow N_U - x^2 - y^2 = s_0(x, y) + s_1 g_1(x) + s_2 g_3(x, y). \end{aligned}$$

With respect to the unit square domain Ω_S described via (2.9), choosing:

$$\begin{aligned} N_S &= 2 \in \mathbb{N}, s_1 = 2 \geq 0, s_2 = 2 \geq 0, \\ s_0(x, y) &= x^2 - 2x - 2y + y^2 + 2 = [1, x, y]Q_S[1, x, y]^T \in \Sigma_s, \\ Q_S &= \begin{pmatrix} 2 & -1 & -1 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \succeq 0 \Rightarrow N_S - x^2 - y^2 = s_0(x, y) + s_1 g_1(x) + s_2 g_4(x, y). \end{aligned}$$

□

Remark 1. As it has been confirmed in the proof of Lemma 1, due to the algebraic domain representations (2.6)-(2.9) involve quadratic forms g_j 's, the polynomial SOS condition: $s_j \in \Sigma_s, \forall j = 1, \dots, m$ in (2.4) and (2.10) can be simplified to a scalar non-negativity: $s_j \geq 0$. In this particular case the Putinar's certificate is equivalent to the S-procedure [136]. This feature allows reducing significantly computational costs in higher dimensions.

2.1.2 Minimax Approximation

The *axiom of completeness* has a far-outstanding consequence in the generalized *Min-Max theorem*, which sets forth the existence of extreme values of a continuous real-valued function on compact

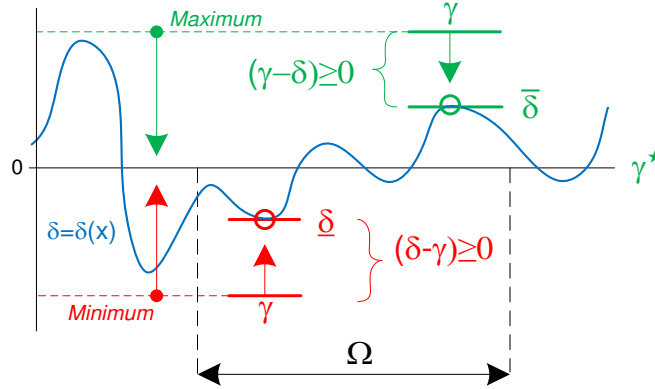


Figure 2.1.1: Maximum and minimum values of a residual function in a compact domain.

domains [19]. Thus, the residual function (2.2) is bounded by $\underline{\delta} \leq \delta_d(\mathbf{x}) \leq \bar{\delta}$, where $\underline{\delta}$ is the *minimum* and $\bar{\delta}$ is the *maximum* of δ_d on Ω compact domain. In addition, the exact solution of (2.1), i.e. $\delta_d = 0$ in (2.2), can be approximated using the simple idea depicted in Figure 2.1.1, imposing $\underline{\delta} \rightarrow 0$ and $\bar{\delta} \rightarrow 0$ ($\delta_d \rightarrow 0$ in Ω) on the extreme values of δ_d .

Intuitively, to achieve this objective, the optimization cost function: $\min \{ \max_{x \in \Omega} \{ |\bar{\delta}|, |\underline{\delta}| \} < \gamma \} \rightarrow 0$ can be formulated, which can be seen as the standard *Uniform Best approximation* approach used to approximate the *null function* (zero value in Ω) by δ_d in terms of $\mathcal{L}^\infty$ -norm (uniform error) in Ω , scheme also denominated *Minimax approximation* [151, 152]. In addition, if Ω has the representation (2.3) and this is a “compact basic semi-algebraic set” (for instance one of the sets considered in Lemma 1), the positivity constraints

$$\begin{aligned} \bar{\delta} - \delta_d(\mathbf{x}) &\geq 0 \quad (\text{maximum}) \\ \delta_d(\mathbf{x}) - \underline{\delta} &\geq 0 \quad (\text{minimum}) \end{aligned}, \quad (2.11)$$

$\forall \mathbf{x} \in \Omega$, can be relaxed and formulated via a SOS decomposition, so that there is an efficient computation of $\underline{\delta}$ and $\bar{\delta}$. Therefore, an approximate solution of u can be found via the convex optimization problem:

$$\mathbf{P}|_{u \approx p_d} \Rightarrow \mathbf{P}_d \left\{ \begin{array}{l} \gamma^* = \inf_{\gamma \in \mathbb{R}, p_\alpha \in \mathbb{R}} \gamma \\ \text{subj. to: } (\bar{\delta} - \delta_d(\mathbf{x}) - \sum_J s_J^1(\mathbf{x}) G_J(\mathbf{x})) \in \sum_s \\ (\delta_d(\mathbf{x}) - \underline{\delta} - \sum_J s_J^2(\mathbf{x}) G_J(\mathbf{x})) \in \sum_s \\ \gamma \geq 0, s_J^1 \in \sum_s, s_J^2 \in \sum_s \\ \begin{bmatrix} \gamma & \underline{\delta} \\ \underline{\delta} & \gamma \end{bmatrix} \succeq 0, \quad \begin{bmatrix} \gamma & \bar{\delta} \\ \bar{\delta} & \gamma \end{bmatrix} \succeq 0 \\ \mathbb{B}_i[p_d(\cdot)](\mathbf{x})|_{\mathbf{x} \in \partial\Omega_i} = \mathbf{u}_i, \quad i = 1, \dots, r \in \mathbb{N} \end{array} \right., \quad (2.12)$$

where G_J stands for combinations of g_j 's in accordance with the type of *positivity certificate* selected (Schmüdgen or Putinar). The problem $\mathbf{P}_d$ imposes $\delta_d(\mathbf{x}) \rightarrow 0, \forall \mathbf{x} \in \Omega$, by means of the

minimization of the global extreme values of δ_d in Ω , in terms of the polynomial coefficients p_α .

Remark 2. *With respect to the generality of the compact domain Ω , under the existence of a non-degenerate coordinate transformation⁵ (typical method to obtain canonical or known PDE forms [155, 156]), this domain can be transformed to a symmetric unit hypercube with respect to the origin $([-1, 1]^n)$, which is useful in particular for orthogonal polynomials.*

2.1.3 Least Squares Approximation

The other standard approach in approximation theory consists in the *Best approximation* in terms of $\mathcal{L}^2(\Omega)$ -norm, scheme also denominated *Least Squares approximation* [152]. This has the main advantage of obtaining a unique solution due to the strict convexity of the problem.

For **P** (2.1) with δ_d as in (2.2), the formulation of a quadratic cost functional: $\min_{p_\alpha} \int_\Omega \delta_d^2(\mathbf{x}) d\mathbf{x}$ leads to a non-convex problem (multiplications of p_α unknown parameters). However, due to the semi-positivity and equality constraints

$$\begin{bmatrix} \left(\frac{1}{\int_\Omega d\mathbf{x}}\right) \gamma - T(\mathbf{x}) & \delta_d(\mathbf{x}) \\ \delta_d(\mathbf{x}) & \gamma \end{bmatrix} \succeq 0, \quad (2.13)$$

$$\int_\Omega T(\mathbf{x}) d\mathbf{x} = 0$$

the Schur's complement provides a sufficient convex condition for this quadratic optimization problem, namely

$$\int_\Omega \left(\frac{1}{\int_\Omega d\mathbf{x}} \gamma^2 - \gamma T(\mathbf{x}) - \delta_d^2(\mathbf{x}) \right) d\mathbf{x} \geq 0 \Rightarrow \int_\Omega \delta_d^2(\mathbf{x}) d\mathbf{x} \leq \gamma^2, \quad \forall \mathbf{x} \in \Omega, \quad (2.14)$$

with $T = T(\mathbf{x}) \in \mathbb{R}[\mathbf{x}]$, and considering that Ω has a representation as a “compact basic semialgebraic set” (such as one of the sets described in in Lemma 1), (2.13) can be relaxed and formulated via a SOS decomposition in its matrix-polynomial version.⁶

Remark 3. *It is worth noting that the formulation (2.13) allows minimizing directly the “root” of $\int_\Omega \delta_d^2(\mathbf{x}) d\mathbf{x}$ instead of the traditional “Least Squares” cost functional. The standard least squares objective can be obtained using*

$$\begin{bmatrix} T(\mathbf{x}) & \delta_d(\mathbf{x}) \\ \delta_d(\mathbf{x}) & 1 \end{bmatrix} \succeq 0, \quad (2.15)$$

$$\gamma = \int_\Omega T(\mathbf{x}) d\mathbf{x}$$

where γ is the cost functional of the optimization problem. On the other hand, due to the quadratic structure of this inequality, if $\deg(T) = 2\deg(\delta_d)$ is selected, a better numerical performance is achieved (this can be rather restrictive for dimensions $n \geq 2$).

⁵Bijjective variable transformation. Jacobian determinant does not vanish on the domain considered [153, 154].

⁶This condition can also be “scalarized”, see details in [136, 138, 157].

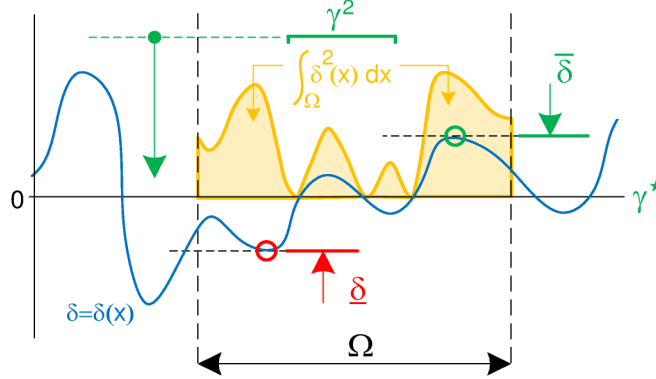


Figure 2.1.2: Integral of the square of a residual function in a compact domain.

Therefore, as Figure 2.1.2 illustrates, an approximate solution of u can be found via the convex optimization problem:

$$\mathbf{P}|_{u \approx \mathbf{p}_d} \Rightarrow \mathbf{P}_d \left\{ \begin{array}{l} \gamma^* = \inf_{\gamma \in \mathbb{R}, \mathbf{p}_\alpha \in \mathbb{R}} \gamma \\ \text{subj. to: } \left(\begin{bmatrix} \left(\frac{1}{\int_\Omega d\mathbf{x}} \right) \gamma - T(\mathbf{x}) & \delta_d(\mathbf{x}) \\ \delta_d(\mathbf{x}) & \gamma \end{bmatrix} - \sum_J S_J(\mathbf{x}) G_J(\mathbf{x}) \right) \in \sum_s^{2 \times 2} \\ \int_\Omega T(\mathbf{x}) d\mathbf{x} = 0 \\ \gamma \geq 0, S_J \in \sum_s^{2 \times 2} \\ \mathbb{B}_i[\mathbf{p}_d(\cdot)](\mathbf{x})|_{\mathbf{x} \in \partial\Omega_i} = \mathbf{u}_i, \quad i = 1, \dots, r \in \mathbb{N} \end{array} \right. , \quad (2.16)$$

where, as in the previous section was indicated, G_J represents a particular combination of g_j 's according to the type of Positivstellensatz (Schmüdgen or Putinar) selected. The problem $\mathbf{P}_d$ (2.16) enforces $\delta_d(\mathbf{x}) \rightarrow 0, \forall \mathbf{x} \in \Omega$, by means of the minimization of the “square root” of the integral of the residual function.

Remark 4. For the particular case of Putinar’s Positivstellensatz with $\Omega = \{\mathbf{x} \in \mathbb{R}^n : g_1(\mathbf{x}) \geq 0, g_2(\mathbf{x}) \geq 0\}$ “compact semi-algebraic set with Archimedean quadratic module”, matrices with structure:

$$S_1(\mathbf{x}) = \begin{bmatrix} s_1(\mathbf{x}) & 0 \\ 0 & 0 \end{bmatrix} \in \sum_s^{2 \times 2}, \quad s_1 \in \sum_s \quad (2.17)$$

$$S_2(\mathbf{x}) = \begin{bmatrix} 0 & 0 \\ 0 & s_2(\mathbf{x}) \end{bmatrix} \in \sum_s^{2 \times 2}, \quad s_2 \in \sum_s \quad (2.18)$$

can be considered in the problem (2.16). For these structures the SOS matrix condition is immediately verified based on the existence of polynomial matrices B_j such that $S_j = B_j^\top(\mathbf{x})B_j(\mathbf{x})$, $j = 1, 2$.

Remark 5. In the one-dimensional case the set of non-negative polynomials ($\mathbb{P}_{n,2d}$) coincides with the set of SOS polynomials ($\sum_s$). However, in general, to the extent the dimension of the space increases, $\sum_s$ is a strictly smaller subset of $\mathbb{P}_{n,2d}$ [136, 284]. Thus, the application of the approach proposed is limited to low dimensional differential BVPs (in practice, $n \leq 3$), and for one dimension is expected to achieve its best numerical performance.

2.1.4 Polytopic Domain Decomposition

The idea of the Finite Element Method of decomposing the whole domain into structural elements and approximating the global solution as the contribution of local polynomial approximations of these sub-domains [142, 65, 143, 144, 158], can directly applied to the formulation of differential equations as convex optimization problems. If the selected sub-domains are polytopes,⁷ such as for instance: intervals in $\mathbb{R}$, polygons in $\mathbb{R}^2$, etc., these can be formulated as *compact basic semi-algebraic sets* which always satisfy the Putinar's Positivstellensatz [138]. Thus, the non-negative representation of polynomials in these sets allows solving *local differential problems* via the convex optimization formulation (2.12) or (2.16). The “global solution” is the contribution of each *local solution*, the continuity of which is imposed by simple optimization constraints on every local boundary (between sub-domains). An example of a two-dimensional polytope decomposition is shown in Figure 2.1.3 in accordance with Lemma 1.

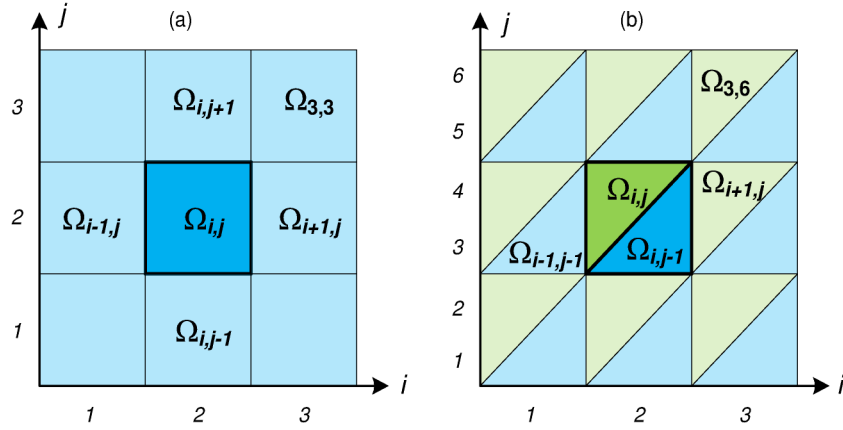


Figure 2.1.3: Example of a Polytopic Two-Dimensional Domain Decomposition. (a) Squares. (b) Triangles (Upper: $\Omega_{i,j}$ and Lower: $\Omega_{i,j-1}$ Triangular Domains).

Let $\Omega = \bigcup_{\beta \in I} \Omega_{\beta} \subset \mathbb{R}^n$ be the domain partitioned into N polytopes Ω_{β} such that $\Omega_{\beta} = \mathring{\Omega}_{\beta} \cup \partial\Omega_{\beta}$ and $\bigcap_{\beta \in I} \mathring{\Omega}_{\beta} = \emptyset$, $\forall \beta \in I$, where $I = \{\beta_1, \dots, \beta_N\}$ with $\beta_k \in \mathbb{N}^n$, $k = 1, \dots, N$.

- For instance for the two-dimensional square-based decomposition shown in Figure 2.1.3(a): $N = 9$, $\beta = (i, j) \in [1, 3] \times [1, 3] \subset \mathbb{N}^2$, $I = \{(1, 1), (2, 1), (3, 1), \dots, (3, 3)\}$ and $\Omega = \bigcup_{(i,j) \in I} \Omega_{(i,j)}$.

Let $\phi_{\beta}(\mathbf{x}) = \sum_{\alpha \in \mathbb{N}_d^n} c_{\alpha}^{\beta} \mathbf{x}^{\alpha}$ be a polynomial of degree d with coefficients $\mathbf{c}^{\beta} \in \mathbb{R}^{z(n,d)}$ associated with each sub-domain Ω_{β} such that δ_{β} is its respective “local” residual function. The global approximation of the solution in (2.1) can be written as $u(\mathbf{x}) \approx \mathbf{p}_d(\mathbf{x}) = \sum_{\beta \in I} \chi_{\Omega_{\beta}}(\mathbf{x}) \phi_{\beta}(\mathbf{x})$, where

⁷Polytopes are bounded (compact and convex) polyhedrons which can be generated by a set of finitely many linear inequalities and equalities [136].

$\chi_{\Omega_\beta}(\mathbf{x}) = \{1, \mathbf{x} \in \Omega_\beta; 0, \mathbf{x} \notin \Omega_\beta\}$ is the indicator function, and it can be determined by the convex optimization problem:

$$\mathbf{P}_{d,N} \left\{ \begin{array}{ll} \gamma^* = \inf_{\gamma \in \mathbb{R}, c_k^\beta \in \mathbb{R}} \gamma \\ \text{subj. to:} \\ \forall \beta \in I = \{\beta_1, \dots, \beta_N\} & : \begin{cases} \left(\bar{\delta}_\beta - \delta_\beta(x) - \sum_J s_{\beta,J}^1(\mathbf{x}) G_J(\mathbf{x}) \right) \in \sum_s \\ \left(\delta_\beta(x) - \underline{\delta}_\beta - \sum_J s_{\beta,J}^2(\mathbf{x}) G_J(\mathbf{x}) \right) \in \sum_s \\ s_{\beta,J}^1 \in \sum_s, \quad s_{\beta,J}^2 \in \sum_s \\ \begin{bmatrix} \gamma & \underline{\delta}_\beta \\ \underline{\delta}_\beta & \gamma \end{bmatrix} \succeq 0, \quad \begin{bmatrix} \gamma & \bar{\delta}_\beta \\ \bar{\delta}_\beta & \gamma \end{bmatrix} \succeq 0 \end{cases} \\ \forall l = 1, \dots, M, \forall m = 1, \dots, M & : \begin{cases} \mathbb{D}_k(\phi_{\beta_l}(\mathbf{x}) - \phi_{\beta_m}(\mathbf{x}))|_{\mathbf{x} \in \Theta_{lm}} = 0, \\ \forall k = 0, \dots, t \end{cases} \\ \text{s.t.: } \Theta_{lm} = \partial\Omega_{\beta_l} \cap \partial\Omega_{\beta_m} \neq \emptyset & : \gamma \geq 0 \\ \forall i = 1, \dots, t \in \mathbb{N}, \forall \sigma \in I & : \mathbb{B}_i[\mathbf{p}_d(\cdot)](\mathbf{x})|_{\mathbf{x} \in \Lambda_{i\sigma}} = \mathbf{u}_i, \\ \text{s.t.: } \Lambda_{i\sigma} = \partial\Omega_i \cap \partial\Omega_\sigma \neq \emptyset \end{array} \right. \quad (2.19)$$

where $M \in \mathbb{N}$ is the total number of shared boundaries of local elements and $\mathbb{D} = \frac{\partial^k}{\partial \mathbf{x}^k}$ is a differential linear operator which enforces continuity ($k = 0$) and a selected degree of smoothness ($k = 1, \dots, t \in \mathbb{N}$) between the local solutions. Similar to this *Minimax* optimization problem, a *Least Squares* minimization can be carried out based on (2.16).

2.2 Moment Approach for PDEs

In the context of the so-called the Generalized Moment Problem (GMP), one of the multiple and useful applications of the duality between SOS and Moments⁸ is the solution of differential equations (ODEs, PDEs) [137].

For linear PDEs with polynomial data, an approximate solution $u \in \mathcal{L}^2(\mathbf{K}, \lambda)$ can be obtained interpreting it as a “density function” with respect to the Lebesgue measure $\lambda = \lambda(\mathbf{x})$, $\forall \mathbf{x} \in \mathbf{K}$. This by means of the computation of finitely many moments $\mathbf{y} = (m_\alpha)_{\alpha \in \mathbb{N}_r^n}$ of the Borel measure μ , $d\mu(\mathbf{x}) = u(\mathbf{x})d\lambda(\mathbf{x})$ (u : Radon-Nikodym derivative of μ with respect to λ) as $m_\alpha = \int_\Omega \mathbf{x}^\alpha d\mu(\mathbf{x})$ [159, 137].

The method of approximation involves two steps:

- (i) Consists of generating a “moment sequence” $\mathbf{y} = (m_\alpha)_{\alpha \in \mathbb{N}_r^n}$ through of solving the adjoint operator equation derived from a weak formulation of the differential equation, namely

⁸For an introduction to this duality, see Appendix A. The proper cone of non-negative polynomials on $\mathbf{K}$ corresponds to the proper cone of *moment sequences* $\mathbf{y} = (m_\alpha)_{\alpha \in \mathbb{N}^n}$ with representing finite Borel measure μ : $\mathbb{P}^*(\mathbf{K}) = \{(m_\alpha) \in \mathbb{R}^{\mathbb{N}^n} : \exists \mu \in \mathcal{M}(\mathbf{K})_+; m_\alpha = \int_{\mathbf{K}} \mathbf{x}^\alpha d\mu(\mathbf{x}), \forall \alpha \in \mathbb{N}^n\}$ [148, 138].

$$\sum_{j=0}^m l_j(\mathbf{x}) \frac{\partial^{|\bar{\alpha}_j|} u(\mathbf{x})}{\partial x_1^{\alpha_j^1} \cdots \partial x_n^{\alpha_j^n}} = f(\mathbf{x}) \Leftrightarrow \mathbb{L}[u(\cdot)](\mathbf{x}) = f(\mathbf{x}), \quad (2.20)$$

$$\langle u, \mathbb{L}^*[\phi_i] \rangle_{\mathcal{L}^2(\Omega)} = \langle f, \phi_i \rangle_{\mathcal{L}^2(\Omega)}, \quad \forall \phi_i \in \Phi,$$

with $\mathbb{L}^*$ adjoint operator, $l_j, f \in \mathbb{R}[\mathbf{x}]$ and $\phi_i = \mathbf{x}^{\bar{\alpha}_i}$ the i -th element of the basis $\Phi = \{\phi_1, \dots, \phi_i, \dots\}$, via a hierarchy of semidefinite programs [161, 160].

- (ii) Corresponds to the “inverse optimization” procedure, which formulates a polynomial approximation for the unknown density u as $\hat{u} = \theta^\top \Phi(\mathbf{x})$ and performs a mean squares minimization:

$$\begin{aligned} \inf_{\hat{u} \in \mathbb{R}_{n,r}[\mathbf{x}]} \|u - \hat{u}\|_{\mathcal{L}^2(\Omega)} &= \inf_{\hat{u} \in \mathbb{R}_{n,r}[\mathbf{x}]} \int_{\Omega} (u(\mathbf{x}) - \hat{u}(\mathbf{x}))^2 d\mathbf{x} \\ &= \inf_{\hat{u} \in \mathbb{R}_{n,r}[\mathbf{x}]} \int_{\Omega} \hat{u}^2(\mathbf{x}) d\mathbf{x} - 2 \int_{\Omega} \hat{u}(\mathbf{x}) \underbrace{u(\mathbf{x}) d\mathbf{x}}_{d\mu} + \underbrace{\int_{\Omega} u^2(\mathbf{x}) d\mathbf{x}}_{\geq 0} \\ &\leq \inf_{\substack{\theta \in \mathbb{R}^{z(r)} \\ \mathbf{x} \in \Omega}} \theta^\top \left(\int_{\Omega} \Phi(\mathbf{x}) \Phi^\top(\mathbf{x}) d\mathbf{x} \right) \theta - 2 \sum_{k=0}^{z(r)-1} \theta_k \underbrace{\int_{\Omega} \phi_k(\mathbf{x}) d\mu}_{m_k} \\ &= \inf_{\substack{\theta \in \mathbb{R}^{z(r)} \\ \mathbf{x} \in \Omega}} \theta^\top \mathbf{M} \theta - 2\theta^\top \mathbf{y}, \end{aligned} \quad (2.21)$$

only based on the knowledge of a finite moment sequence $\mathbf{y}_r = \{m_0, \dots, m_{z(r)-1}\} \in \mathbb{R}^{z(r)}$, equivalent to a semidefinite programming problem [162]. It is worth noting that the non-negativity of $\hat{u}$ can be ensured if this condition (non-negativity) is included as constraint on Ω using positive certificates on compact basic semialgebraic sets.

2.3 Computational Examples

The following examples, one and two-dimensional cases, illustrate the numerical performance of the method proposed for solving differential equations-boundary value problems. These allow evaluating its subsequent implementation in the coupled PDE-ODE problems derived from the design of Backstepping PDE-based adaptive observers (next chapters). Its extension to higher dimensions is straightforward but still limited by the SOS polynomial set itself (see Remark 5 and Appendix A, footnote 11), its positive representation and the current stage of semidefinite programming tools. The numerical solution of these convex optimization problems has been obtained via the Yalmip toolbox for Matlab [163] and SOSTOOLS [164], using the semi-definite programming solver SeDuMi [165] and the SDP package part of the Mosek solver [166].

2.3.1 One-Dimensional Differential Equation-BVP

This section considers *steady diffusion-reaction/advection problems* which arise from several types of systems, for instance: steady heat conduction, longitudinal deformation of an elastic rod, trans-

verse deflection of a cable, etc [167, 89]. These are typical steady BVP used to exemplify the standards one-dimensional numerical methods such as the Shooting Method (BVP as a pair of initial-value problems), Finite Difference Method (approximation of differential operators by discrete differences), Rayleigh-Ritz-Galerkin (variational approach, integral functional minimization) or Spectral-Collocation methods [147, 168, 169, 170, 141, 171, 172, 173, 174, 175]. The nature of its solutions (regular Sturm-Liouville-problem, solutions of which can span $\mathcal{L}^2(0,1)$ [176, 177]) are suitable to test polynomial approximations.

2.3.1.1 Steady Convection-dominated Problem

This example considers the second order linear elliptic two-point BVP

$$\mathbf{P} : \begin{cases} -\epsilon \frac{d^2 u}{dx^2}(x) + b \frac{du}{dx}(x) = 0, \quad \forall x \in (0,1) \\ u(0) = 0 \\ u(1) = 1 \end{cases} \Rightarrow u(x) = \frac{e^{\frac{bx}{\epsilon}} - 1}{e^{\frac{b}{\epsilon}} - 1}, \quad (2.22)$$

with exact analytical solution where $\epsilon > 0$ and $b > 0$. If the convection (transport) term ($b \cdot \nabla u$) dominates the dynamic of the problem (small ϵ/b), this leads to solutions with strong gradients close to its boundaries (boundary layers) [146, 178]. As it is well-known, Galerkin-based numerical approximations exhibit oscillatory behaviour if the local Péclet number: $P_e = bh/(2\epsilon) > 1$. To avoid this, for instance, a small (in certain cases unpractical) grid-size h is required.

The problem (2.22) has been approximately solved for $\epsilon = 1/50$ and $b = 2$ via the convex formulation (2.12) using SOSTOOLS-SeDuMi packages. As Figure 2.3.1 shows the characteristic of the solution at the boundary requires a high polynomial degree ($d \geq 100$) to reach a residual function $\delta_d \leq 10^{-3}$. Figure 2.3.1(a) depicts the exact solution and its polynomial approximation for a degree $d = 100$. A piece-wise linear Centered Finite Difference approximation⁹ is also shown for $N = 101$ nodes (discretization points) in the unit interval. The effect of the strong gradient at the boundary over the precision of the polynomial solution is shown in Figure 2.3.1(b). In this case an approximation error $(u - p) \leq 3.8 \cdot 10^{-5}$ (see y-axis: right-hand side) and a residual function bound $|\gamma| \leq 1.5 \cdot 10^{-3}$ (see y-axis: left-hand side) are achieved. Figure 2.3.1(c) illustrates the oscillation in the Galerkin-based numerical approximation¹⁰ for a grid-size h such that $P_e = bh/(2\epsilon) > 1$. The decreasing characteristic of the approximation error $\|u - p_d\|_\infty$ and of the $\mathcal{L}^\infty$ -norm of the residual function δ_d is depicted in Figure 2.3.1(d). This is compared with the $\mathcal{L}^\infty$ -norm of the approximation error obtained via the *centered* Finite Difference Method with respect to the number of nodes (secondary x-axis, below polynomial degree d).

⁹The system of equations obtained from the application of the *Centered Finite Difference Method* coincides with the system derived from the *Finite Element Method* with piecewise-linear polynomials under a uniform grid space discretization and constant parameters.

¹⁰This oscillation can be overcome considering *Stabilization Methods* [146].

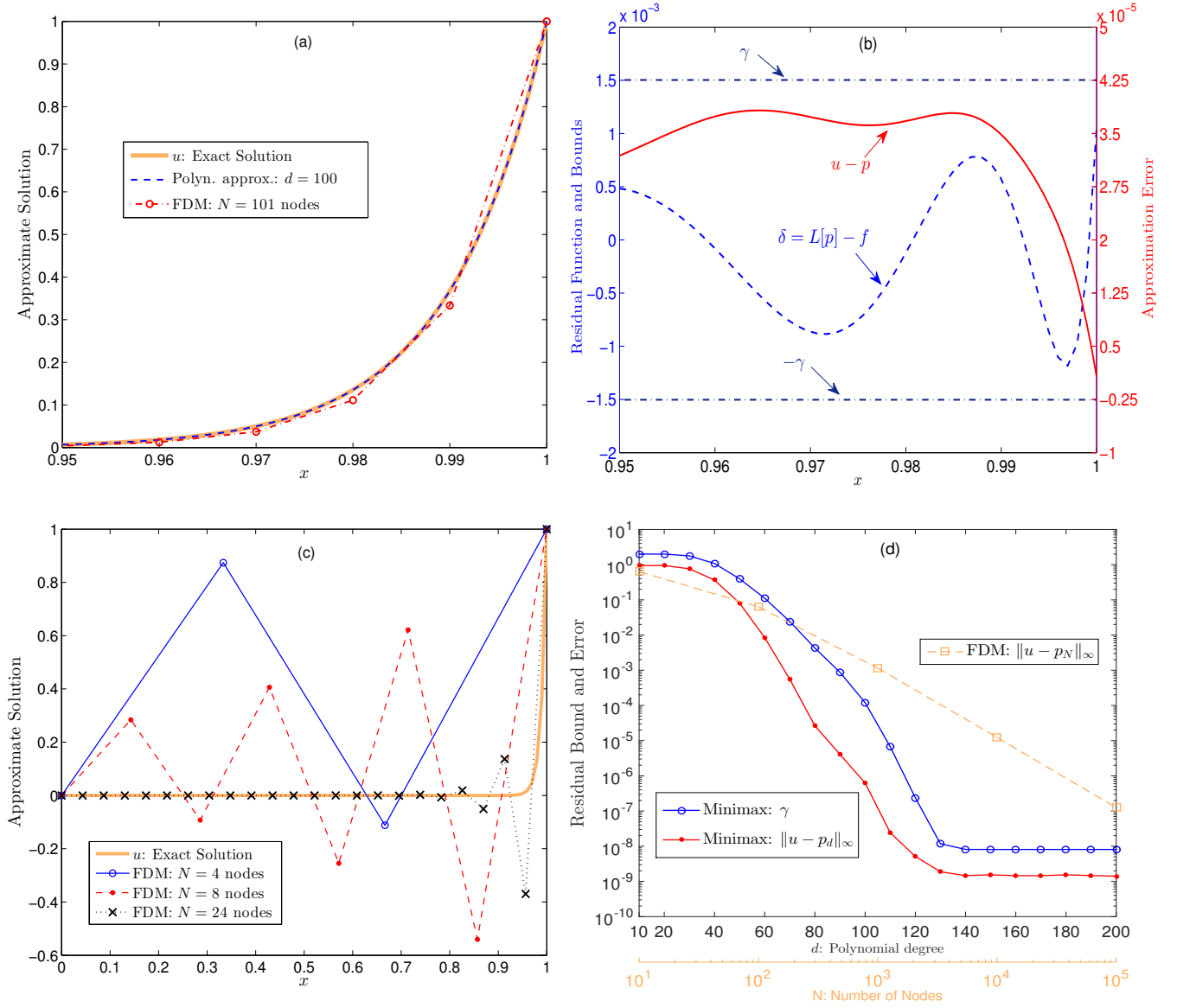


Figure 2.3.1: Polynomial approximation of Steady Convection-dominated Problem. (a) Exact solution $u = u(x)$: solid line, polynomial solution p_d for polynomial degree $d = 100$: dash line, piece-wise linear *Centered Finite Difference* approximation with $N = 101$ nodes: dash-dot line (circles: elements nodes). For clearness, the solutions have been plotted on the interval $[0.95, 1]$ instead of $[0, 1]$. (b) Approximation error $(u - p)$: solid line (y-axis: right-hand side), residual function $\delta = \delta(x)$: dash line (y-axis: left-hand side), optimal bound γ : dash-dot line (y-axis: left-hand side). (c) *Centered Finite Difference* approximation for $N = 4$: solid-circle line, $N = 8$: dash-point line and $N = 24$ dot-x line. (d) Residual function bound and error for polynomial approximation (γ : solid-circle line, $\|u - p_d\|_\infty$: solid-point line) (primary x-axis. d : Polynomial degree) and Finite Difference Method ($\|u - p_N\|_\infty$: dash-square line (secondary x-axis. N Number of Nodes).

2.3.1.2 Steady Diffusion-Reaction Problem

In general, in contrast to the convection-dominated problem, diffusion-reaction differential equations present a well-behaved solution. This is similar to the classic one-dimensional Poisson problem where the ellipticity constant is dominant in comparison with the rest of coefficients in the equation [146, 178].¹¹ To illustrate the flexibility of the approach proposed, numerical examples with two types of boundary conditions are considered.

- **Case 1:** Dirichlet Homogeneous Boundary conditions.

This case considers the second order linear elliptic two-point BVP with exact analytical solution

$$\mathbf{P} : \begin{cases} -\frac{d^2 u}{dx^2}(x) + u(x) = x, & \forall \mathbf{x} \in (0, 1) \\ u(0) = 0 \\ u(1) = 0 \end{cases} \Rightarrow u(x) = x - \frac{\sinh(x)}{\sinh(1)}, \quad (2.23)$$

with homogeneous Dirichlet boundary conditions [179].

The problem (2.23) has been approximately solved via the *Minimax* convex formulation (2.12) as well as *root Least Squares* (2.16) using Yalmip-SeDuMi packages. As Figure 2.3.2 shows, the approximate polynomial solution reaches an accurate result for low polynomial degrees. Figure 2.3.2(a) depicts approximations for polynomials degrees of $d = 2$ and $d = 4$ obtained via the *Minimax* approach (2.12):

$$\begin{aligned} u &\approx u_2 = -0.2500x^2 + 0.2500x - 1.494 \cdot 10^{-19}, \\ u &\approx u_4 = -0.01896x^4 - 0.1270x^3 - 0.002579x^2 + 0.1486x - 5.72 \cdot 10^{-18}, \end{aligned}$$

illustrating a comparison with a piece-wise linear *Finite Element Method* approximation of 6 elements. Figure 2.3.2(b) shows a residual function bound and approximation error (in norm) with monotonic decreasing tendency for $d \leq 10$. For higher polynomial degrees, the semidefinite package run into numerical problems (optimum around to the maximum threshold precision). Figure 2.3.2(c) shows an approximate solution using the Finite Element Method with piece-wise quadratic Lagrange polynomials (3 basis per element), namely

$$\begin{aligned} u &\approx u_{\text{FEM}_2}(x) = \sum_{i=0}^N c_i \phi_i(x) + \sum_{j=1}^N d_j \psi_j(x), \\ \phi_i(x) &= \begin{cases} (2/h^2)(x - x_{i-1})(x - (x_{i-1} + h/2)), & \forall x \in [x_{i-1}, x_i] \\ (2/h^2)(x - x_{i+1})(x - (x_{i+1} - h/2)), & \forall x \in [x_i, x_{i+1}] \end{cases}, \\ \psi_j(x) &= -(4/h^2)(x - x_{j-1})(x - x_j) \quad \forall x \in [x_{j-1}, x_j], \end{aligned} \quad (2.24)$$

¹¹ Similar to the advection dominated case, diffusion-reaction differential problems can also present boundary layers under very small diffusion coefficients with respect to the reaction ones [178].

where $h = (x_j - x_{j-1})$ is the uniform discretization step, with $N = 3$ elements and parameters: $c_0 = c_3 = 0$, $c_1 = 0.0444$, $c_2 = 0.0564$, $d_1 = 0.02419$, $d_2 = 0.05659$ and $d_3 = 0.03927$. Figure 2.3.2(d) depicts the $\mathcal{L}^\infty$ -norm approximation error using the FEM with piece-wise linear (FEM₁), quadratic (FEM₂) and quartic (FEM₄) polynomials as functions of the number of elements N . It is worth noting that for the FEM approximate solution around 250 quartic polynomials (~ 50 Elements, 5 basis per element) are needed to achieve an approximation error of 10^{-11} . This accuracy can be obtained using a polynomial approximation of degree $d = 10$ via the convex optimization approach proposed.

• **Case 2:** Mixed Dirichlet-Neumann Homogeneous Boundary conditions

This case considers the second order linear elliptic two-point BVP with exact analytical solution

$$\mathbf{P} : \begin{cases} -\frac{d^2 u}{dx^2}(x) + u(x) = 2e^{-x}, & \forall \mathbf{x} \in (0, 1) \\ u(0) = 0 \\ \frac{du}{dx}(1) = 0 \end{cases} \Rightarrow u(x) = xe^{-x}, \quad (2.25)$$

with homogeneous mixed Dirichlet-Neumann boundary conditions [169].

The problem (2.25) has been approximately solved via both the variant *Least Squares* (2.15) in (2.16) and the *Minimax* (2.12) optimization problem, using Yalmip-Mosek packages. As Figure 2.3.3 shows, due to the zero gradient condition at one boundary, this problem is slight harder than (2.23). Figure 2.3.2(a) depicts approximate solutions for polynomials degrees of $d = 2$ and $d = 4$ obtained via the standard *Least Squares* (2.15) in (2.16):

$$\begin{aligned} u \approx u_2 &= -0.4534x^2 + 0.9067x, \\ u \approx u_4 &= -0.09095x^4 + 0.4460x^3 - 0.9871x^2 + 1.0000x. \end{aligned}$$

Figure 2.3.3(b) shows a monotonic decreasing residual function bound and approximation error (in norms) for $d \leq 12$. In this case the combination of Yalmip-Mosek packages obtained lower residual bounds ($\gamma \sim 10^{-14}$) than the previous case via Yalmip-SeDuMi packages. This optimization also ran into numerical problems for higher polynomial degrees (around to the maximum threshold precision). Figure 2.3.3(c) shows an approximate solution using the Finite Element method with piece-wise quartic Lagrange polynomials (5 basis per element), namely

$$\begin{aligned} u \approx u_{\text{FEM}_4}(x) &= \sum_{i=0}^N c_i \phi_i(x) + \sum_{k=1}^3 \sum_{j=1}^N d_{k,j} \psi_{k,j}(x), \\ \phi_i(x) &= \begin{cases} \left(\frac{32}{3h^4}\right)(x - x_{i-1})(x - (x_{i-1} + \frac{h}{4}))(x - (x_{i-1} + \frac{h}{2}))(x - x_{i-1} + \frac{3h}{4}), & \forall x \in [x_{i-1}, x_i] \\ \left(\frac{32}{3h^4}\right)(x - x_{i+1})(x - (x_{i+1} - \frac{h}{4}))(x - (x_{i+1} - \frac{h}{2}))(x - (x_{i+1} - \frac{3h}{4})), & \forall x \in [x_i, x_{i+1}] \end{cases} \end{aligned} \quad (2.26)$$

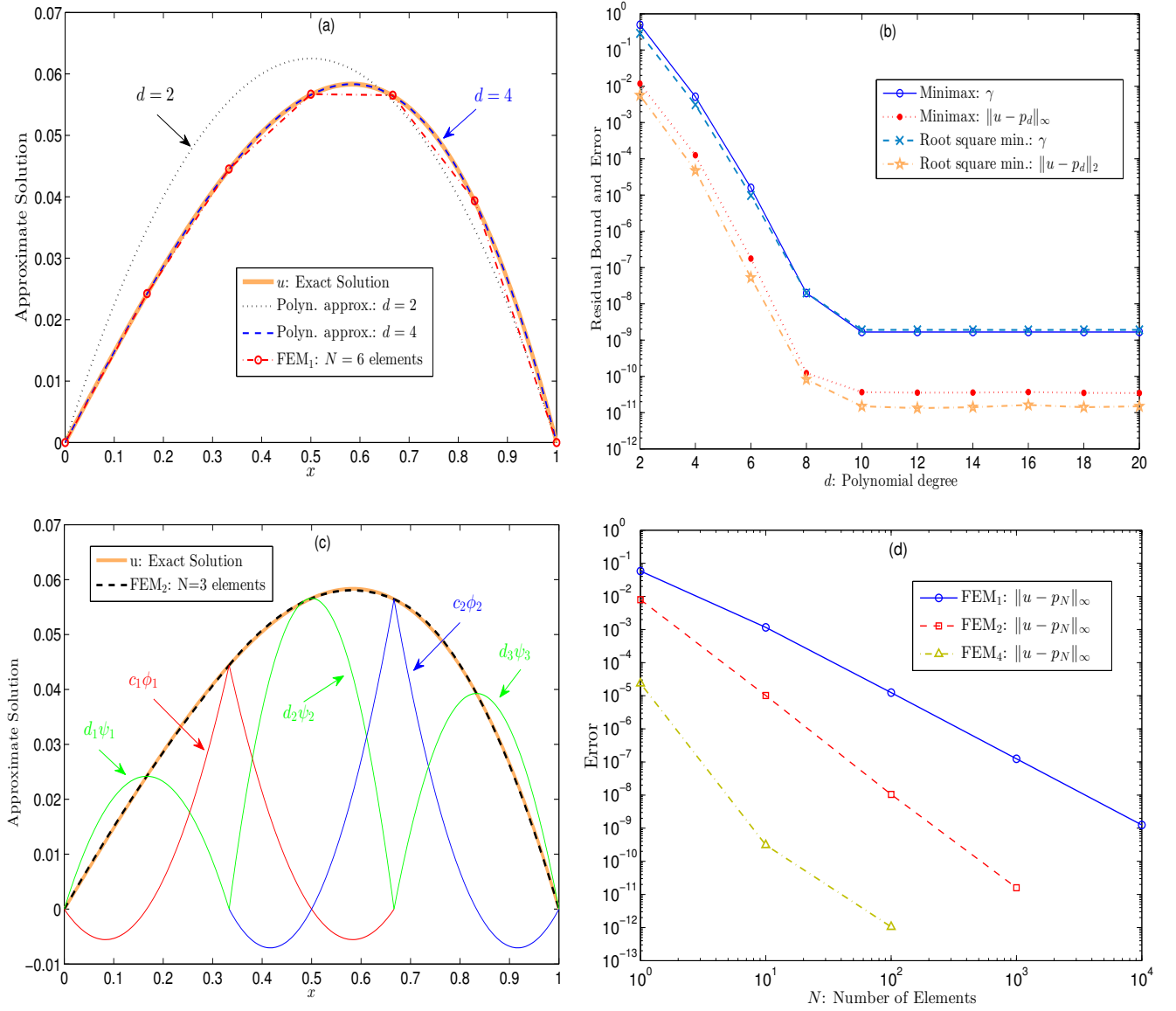


Figure 2.3.2: Polynomial approximation of **Case 1**: Steady Diffusion-Reaction Problem with homogeneous Dirichlet boundary conditions. (a) Exact solution $u = u(x)$: solid line, polynomial approximation $p = p(x)$ for polynomial degree $d = 2$: dot line and $d = 4$: dash line, piece-wise linear Finite Element (FEM₁) approximation with $N = 6$ elements: dash-dot line (circles: elements nodes). (b) Optimal bound for the residual function and error. *Minimax* optimization: γ : solid-circle line, $\|u - p\|_\infty$: dot-point line. *Root Least Squares* optimization: γ : dash-x line, $\|u - p\|_2$: dash-dot-start line. (c) Approximation using the Finite Element method with piece-wise quadratic polynomials (FEM₂). Exact solution $u = u(x)$: solid line, FEM₂ approximation: dash line. (d) $\mathcal{L}^\infty$ -norm error approximation using FEM₁: solid-circle line, FEM₂: dash-square line and FEM₄: dash-dot-triangle line.

$$\begin{aligned}\psi_{1,j}(x) &= -\left(\frac{128}{3h^4}\right)(x-x_{j-1})(x-(x_{i-j}+\frac{h}{2}))(x-x_{j-1}+\frac{3h}{4})(x-x_j) \\ \psi_{2,j}(x) &= \left(\frac{64}{h^4}\right)(x-x_{j-1})(x-(x_{i-j}+\frac{h}{4}))(x-x_{j-1}+\frac{3h}{4})(x-x_j) \quad , \quad \forall x \in [x_{j-1}, x_j], \\ \psi_{3,j}(x) &= -\left(\frac{128}{3h^4}\right)(x-x_{j-1})(x-(x_{i-j}+\frac{h}{4}))(x-x_{j-1}+\frac{h}{2})(x-x_j)\end{aligned}$$

where $h = (x_j - x_{j-1})$ is the uniform discretization step, with $N = 1$ elements and parameters: $c_0 = 0$, $c_1 = 0.36906$, $d_{1,1} = 0.1950$, $d_{2,1} = 0.3038$ and $d_{3,1} = 0.3551$. Figure 2.3.3(d) depicts the $\mathcal{L}^2$ -norm approximation error using the Finite Element method with piece-wise linear (FEM₁), quadratic (FEM₂) and quartic (FEM₄) polynomials as functions of the number of elements N . Similar to the previous case, the FEM needs around 250 quartic polynomials (~ 50 Elements, 5 basis per element) to achieve an approximation error of 10^{-9} . This accuracy can be obtained using a polynomial approximation of degree $d = 10$ via the convex optimization approach proposed.

2.3.1.3 Sturm–Liouville Eigenvalue Problem

This example considers the second order two-point BVP with non-homogeneous Dirichlet boundary conditions and exact analytical solution

$$\mathbf{P} : \begin{cases} \frac{d^2 u}{dx^2}(x) + \omega^2 u(x) = 0, \quad \forall x \in (0, 1) \\ u(0) = 1 \\ u(1) = 0 \end{cases} \Rightarrow u(x) = \cos(\omega x), \quad (2.27)$$

for an angular frequency of $\omega = 5\pi/2$, which is a particular case of the classical Sturm–Liouville Eigenvalue problem [177].

The problem (2.27) has been approximately solved via the variant *Least Squares* (2.15) in (2.16) using Yalmip-Mosek packages. In comparison with the former problems (2.23) and (2.25), as it is shown in Figure 2.3.4(b), if an unique polynomial is considered as approximate solution, to obtain a residual function bound $\gamma \sim 10^{-11}$ a higher polynomial degree ($d \geq 20$) is needed. To illustrate a reduction of the polynomial degree, a polytopic decomposition of the domain is alternatively formulated according to Section 2.1.4 using the *Least Squares* variant (2.15) in (2.19).

Let $\Omega = \bigcup_{j=1}^N \Omega_j = [0, 1]$ be the domain partitioned into N uniform intervals $\Omega_j = [x_j, x_{j+1}]$, over the set of equidistant nodes $\{x_1, \dots, x_j, x_{j+1}, \dots, x_{N+1}\}$. Let $\phi_j = \sum_{k=0}^d c_k^j x^k$ be a polynomial of degree d , with coefficients $c_k^j \in \mathbb{R}$, associated to each sub-domain Ω_j such that $\delta_j(x) = \frac{d^2 \phi_j}{dx^2}(x) + \omega^2 \phi_j(x)$ is its respective “local” residual function. The global approximate solution of (2.27) can be written as $u \approx \mathbf{p}_d = \sum_{j=1}^N \chi_{[x_j, x_{j+1}]}(x) \phi_j(x)$, where $\chi_{[a,b]}(x) = \{1, x \in [a, b]; 0, x \notin [a, b]\}$ is the indicator function, which can be determined by the convex optimization problem:

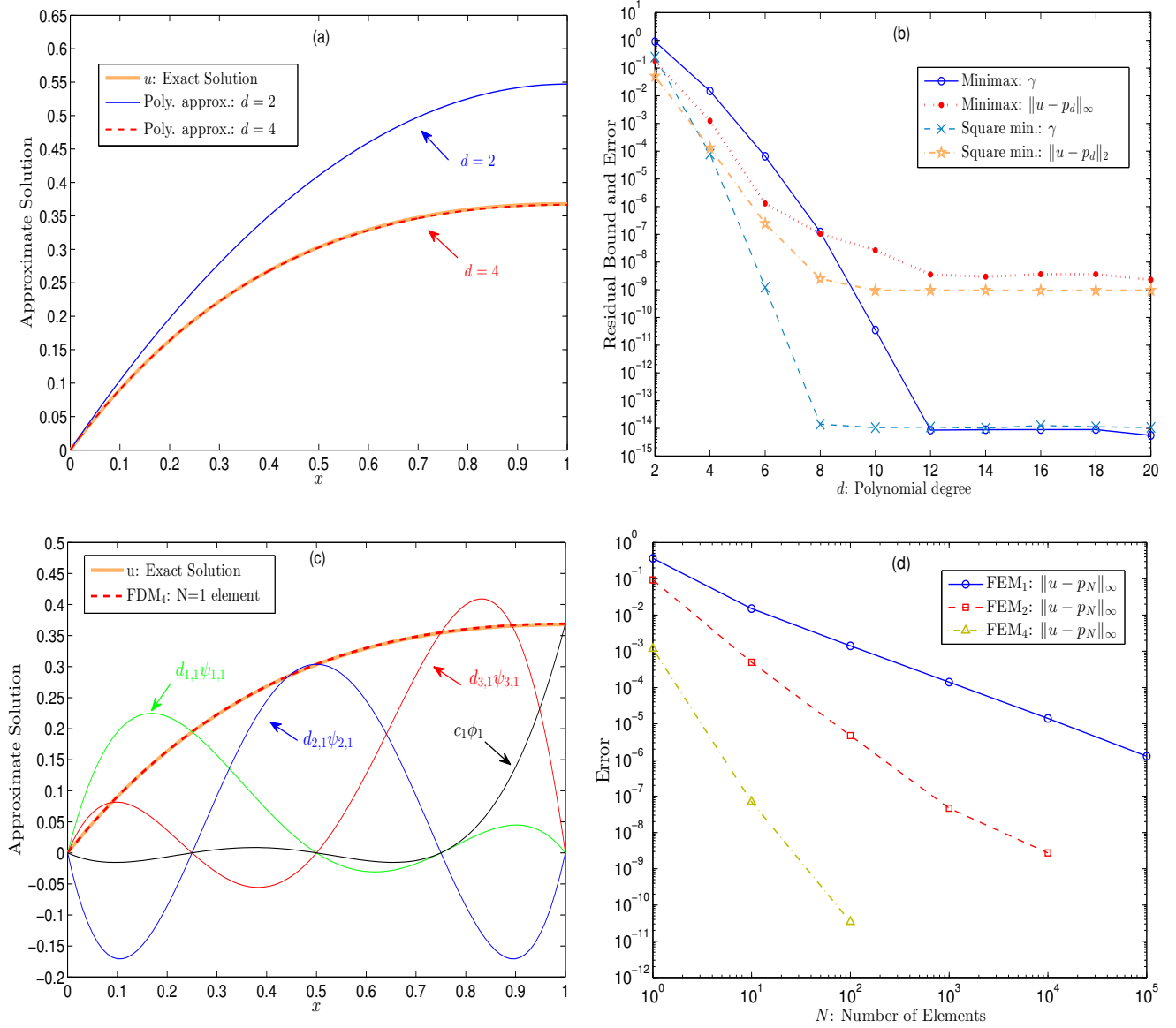


Figure 2.3.3: Polynomial approximation of **Case 2**: Steady Diffusion-Reaction Problem with homogeneous Mixed Dirichlet-Neumann boundary conditions. a) Exact solution $u = u(x)$: solid line, polynomial approximation $p = p(x)$ for polynomial degree $d = 2$: dot line and $d = 6$: dash line. (b) Optimal residual function bound and error. *Minimax* optimization: γ : solid-circle line, $\|u - p\|_\infty$: dot-point line. *Least Squares* optimization: γ : dash-x line, $\|u - p\|_2$: dash-dot-start line. (c) Approximation using the Finite Element Method with piece-wise quartic polynomials (FEM₄). Exact solution $u = u(x)$: solid line, FEM₄ approximation: dash line. (d) L^2 -norm error approximation using FEM₁: solid-circle line, FEM₂: dash-square line and FEM₄: dash-dot-triangle line.

$$\mathbf{P}_{d,N} \left\{ \begin{array}{l} \gamma^* = \inf_{\gamma_j \in \mathbb{R}, c_k^j \in \mathbb{R}} \sum_{j=1}^N \gamma_j \\ \text{subj. to:} \\ \forall j = 1, \dots, N : \left\{ \begin{array}{l} \begin{bmatrix} T_j(x) - s_j^1(x)g_j(x) & \delta_j(x) \\ \delta_j(x) & 1 - s_j^2(x)g_j(x) \end{bmatrix} \in \Sigma_s \\ g_j(x) = (x - x_j)(x_{j+1} - x) \\ s_j^1 \in \Sigma_s, s_j^2 \in \Sigma_s \\ \gamma_j = \int_{x_j}^{x_{j+1}} T_j(x) dx \end{array} \right. \\ \forall j = 1, \dots, N-1 : \left\{ \begin{array}{l} \mathbb{D}_k (\phi_j(x) - \phi_{j+1}(x))|_{x=x_{j+1}} = 0, \forall k = 0, \dots, t \\ \phi_1(0) = 1 \\ \phi_N(0) = 0 \end{array} \right. \end{array} \right. , \quad (2.28)$$

based on Putinar's positivity representation on the compact basic semialgebraic sets (sub-domains): $\Omega_j = \{x \in \mathbb{R} : (x - x_j)(x_{j+1} - x) \geq 0\}$, $\forall j = 1, \dots, N$, where $\mathbb{D} = \frac{d^k}{dx^k}$ is the differential linear operator which enforces continuity ($k = 0$) and a selected degree of smoothness ($k = 1, \dots, t \in \mathbb{N}$).

Figure 2.3.4(a) shows the exact solution and multiple local approximations. Local approximations consist in polynomial basis ϕ_j for each interval $j = 1, \dots, N = 5$, with polynomial degree $d = 4$ and smoothness $t = 1$. For this polynomial degree an approximation error $\|u - p_{d,N}\| \leq 10^{-6}$ is achieved for a number of local intervals $N \geq 16$. As Figure 2.3.4(b) depicts, a monotonic decreasing tendency for the global residual function bound and approximation error is also verified. This optimization ran into numerical problems for polynomial degrees $d \geq 20$ and number of elements (intervals) $N \geq 70$ for an unique polynomial and multiple polynomials, respectively.

Figure 2.3.4(c) shows a local polynomial approximation using a variant of the polytopic decomposition proposed in (2.19), similar to the FEM with quadratic polynomials (FEM₂), where 3 polynomials are used in each sub-domain by means of including:

$$\begin{aligned} \phi_j &= \varphi_{1,j}(x) + \varphi_{2,j}(x) + \varphi_{3,j}(x), \\ \varphi_{1,j}(x_j + h/2) &= \varphi_{1,j}(x_{j+1}) = 0 & \varphi_{k,j} &\in \mathbb{P}_{1,4}[\mathbf{x}] \\ \varphi_{2,j}(x_j) &= \varphi_{2,j}(x_{j+1}) = 0 & , \quad h &= x_{j+1} - x_j \\ \varphi_{3,j}(x_j) &= \varphi_{3,j}(x_j + h/2) = 0 & \forall j &= 1, \dots, N = 5, k = 1, 2, 3. \end{aligned} \quad (2.29)$$

as constraints in the convex formulation (2.19). Figure 2.3.4(d) depicts the $\mathcal{L}^2$ -norm approximation error using the Finite Element Method with piece-wise linear (FEM₁), quadratic (FEM₂) and quartic (FEM₄) polynomials as functions of the number of elements N . In this case the FEM needs around 300 quadratic polynomials (~ 100 Elements, 3 basis per element) to achieve an approximation error of 10^{-6} . This accuracy can be obtained using a polynomial approximation of degree $d = 14$ with an unique polynomial or degree $d = 4$ and $N = 100$ for multiple local polynomials, via the convex optimization approach proposed.

2. FORMULATION OF DIFFERENTIAL EQUATIONS AS CONVEX OPTIMIZATION PROBLEMS

2.3. Computational Examples

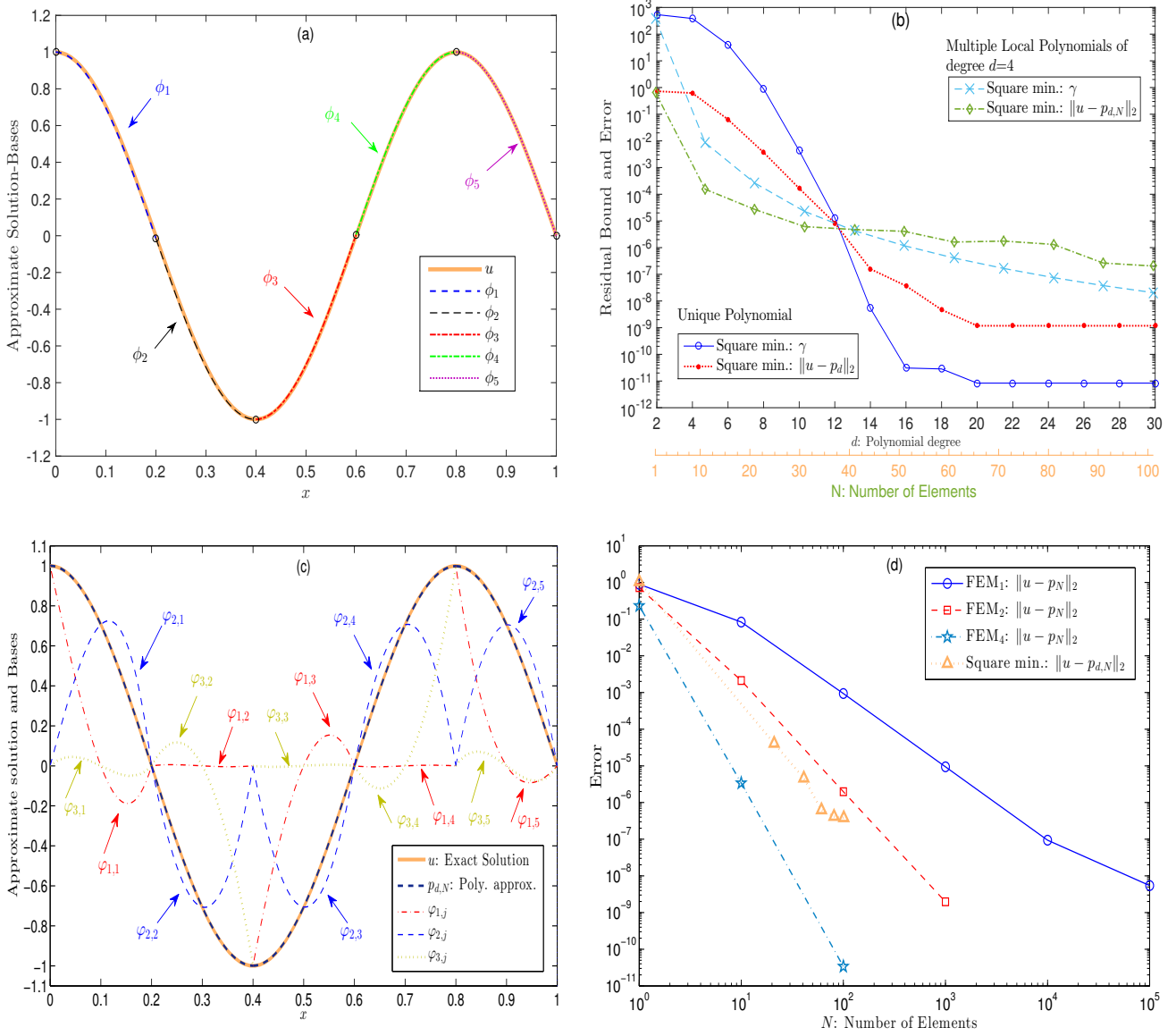


Figure 2.3.4: Polynomial approximation of the Sturm–Liouville Eigenvalue Problem. (a) Exact solution $u = u(x)$: solid line, domain partitioned into $N = 5$ uniform intervals (circles: elements nodes), each interval with a local polynomial approximation ϕ_j of degree $d = 4$. (b) Optimal residual function bound and *Least Squares* optimization error for a unique approximate polynomial and multiple local polynomials. (c) Approximation using constraints (2.29) in (2.19) with local polynomials of degree $d = 4$. Exact solution $u = u(x)$: solid line, multiple local polynomial approximations: dash line. (d) L^2 -norm error approximation using FEM₁: solid-circle line, FEM₂: dash-square line, FEM₄: dash-dot-star line and convex optimization with constraints (2.29) in (2.19): dot-triangle line.

2.3.2 Two-Dimensional PDE-BVP

This section addresses the classical two-dimensional Poisson PDE [176, 143, 169, 180, 181, 182]. This arises as an equilibrium problem in many physical systems such as in the theory of conservative fields (for instance: Electrical, gravitational, magnetic). The Poisson PDE can also be seen as the stationary version of the diffusion problem, slightly more general than the Laplace equation.¹²

2.3.2.1 Poisson 2-Dimensional Equation

This example considers the 2-dimensional Poisson BVP with homogeneous Dirichlet boundary conditions in a square domain, namely

$$\mathbf{P} : \begin{cases} \frac{\partial^2 u}{\partial x^2}(x, y) + \frac{\partial^2 u}{\partial y^2}(x, y) = -1, & \forall (x, y) \in (0, 1) \times (0, 1) \\ u(x, 0) = 0, \quad u(x, 1) = 0, & \forall x \in [0, 1] \\ u(0, y) = 0, \quad u(1, y) = 0, & \forall y \in [0, 1] \end{cases}, \quad (2.30)$$

with an approximate solution in terms of Fourier's series given by [181]:

$$u(x, y) = \frac{16}{\pi^4} \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{\sin((2i+1)\pi x) \sin((2j+1)\pi y)}{(2i+1)(2j+1)((2i+1)^2 + (2j+1)^2)}. \quad (2.31)$$

The problem (2.30) has been approximately solved via both a unique polynomial approximation for the whole domain via the variant *Least Squares* (2.15) in the formulation (2.16), and multiple local polynomials based on a triangular-based polytopic domain decomposition according to Figure 2.1.3(b) and variant *Least Squares* (2.15) in the formulation (2.19) (similar to (2.28)).

2.3.3 Rational Polynomial Functions

2.3.3.1 Reciprocal Function Problem

This section describes an approach to the approximation problem of rational polynomial functions. This is a particular topic of research in approximation theory with a wide range of theoretical and numerical results for specific rational structures [183], also known as “multiplicative inverse functions” or “reciprocal functions”.

Let $1/Q(x)$ be a particular structure of rational real-valued polynomials with $Q > 0$, $Q(x) = \sum_{j=0}^{d_Q} q_j x^j$ for $x \in \Omega \subset \mathbb{R}$ and polynomial degree d_Q . The problem consists in finding the reciprocal polynomial of Q , i.e., a polynomial R such that:

$$\mathbf{P} : \begin{cases} Q(x) \cdot R(x) = 1 \\ \forall x \in \Omega \subset \mathbb{R} \end{cases}. \quad (2.32)$$

¹²For instance, in electrical fields the Poisson equation is: $\Delta u = \text{div}(\nabla u) = -4\pi\rho$, where ρ is density of the charge distribution, which is an inhomogeneous version of the Laplace equation $\Delta u = 0$ obtained for charges located outside of problem's domain ($\rho = 0$). In this particular case its solutions are called “harmonic functions” such as Fourier's series, which can be obtained by the method separation of variables [176, 180].

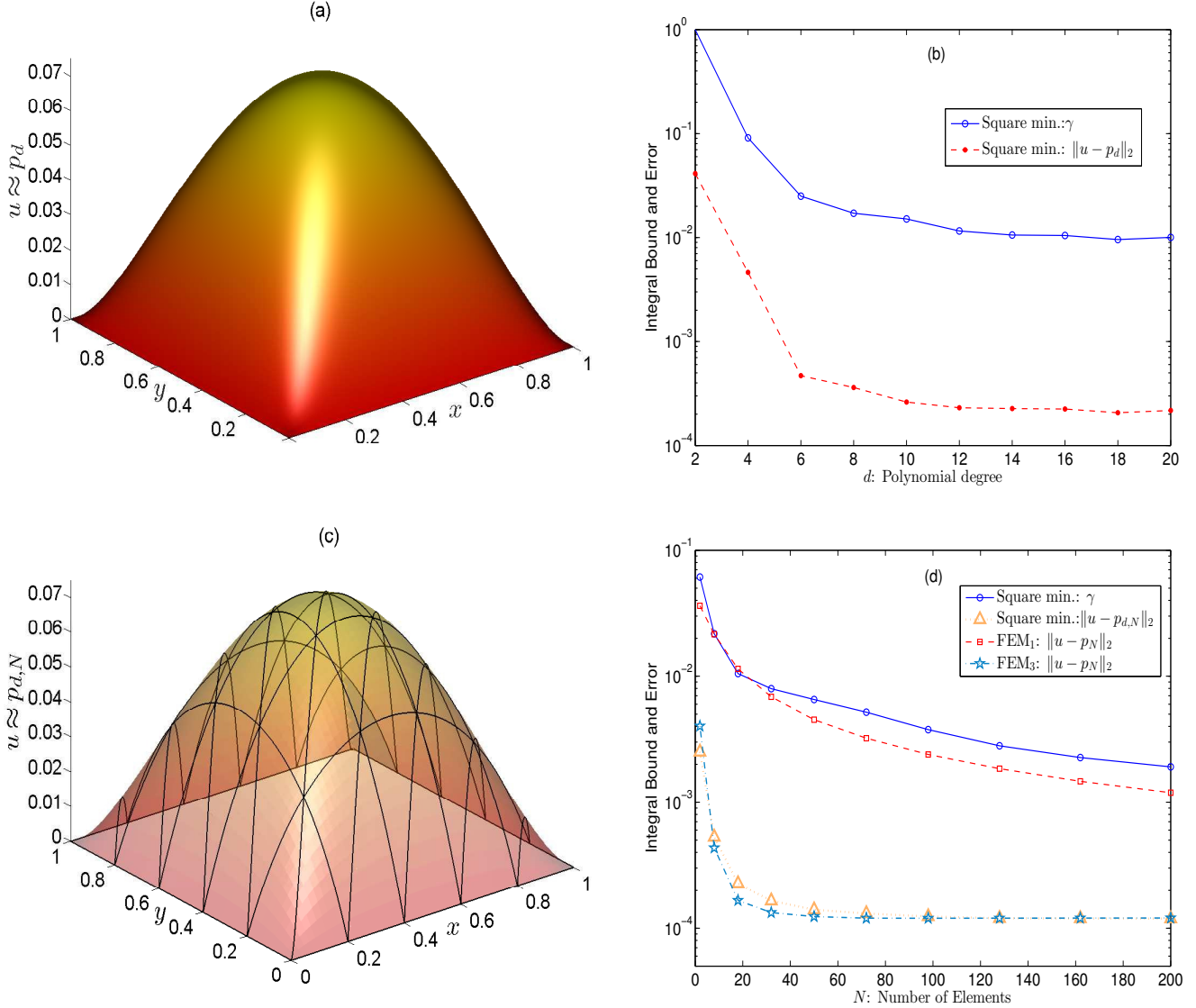


Figure 2.3.5: Polynomial approximation of 2-Dimensional Poisson PDE. (a) Approximate solution for an unique polynomial of degree $d = 14$. (b) Optimal residual function bound and *Least Squares* optimization error for a unique approximate polynomial. (c) Approximation using variant (2.15) in formulation (2.19) with local polynomials of degree $d = 4$. (d) $\mathcal{L}^2$ -norm error approximation using FEM₁: dash-square line, FEM₃: dash-dot-star line and variant (2.15) in formulation (2.19): dot-triangle line. Optimal residual function bound of variant (2.15) in (2.19): solid-circle line.

The approach proposed to solve approximately $\mathbf{P}$ considers a partition of the domain $\bigcup_{j=1}^N \Omega_j = \Omega$, with N subintervals $\Omega_j = [x_j, x_{j+1}]$ (N disjoint finite elements except at their boundaries) so that R is locally approximated in every j -th subinterval: $R \approx R_{d,j} = \sum_{k=0}^d r_k^j x^k$, $\forall x \in \Omega_j$, where d stands for the polynomial degree of R_j with coefficients $r_k^j \in \mathbb{R}$. Thus, the global approximation $R \approx R_d = \sum_{j=1}^N \chi_{\Omega_j}(x) R_{d,j}(x)$ can be computed by a slight modification of the convex optimization problem (2.19), namely

$$\mathbf{P}_{d,N} \left\{ \begin{array}{l} \gamma^* = \inf_{\gamma \in \mathbb{R}, r_k^j \in \mathbb{R}} \gamma \\ \text{subj. to:} \\ \forall j = 1, \dots, N : \left\{ \begin{array}{l} \left(\bar{I}_j - Q(x)R_{d,j}(x) - s_j^1(x)g_j(x) \right) \in \sum_s \\ \left(Q(x)R_{d,j}(x) - \underline{I}_j - s_j^2(x)g_j(x) \right) \in \sum_s \\ g_j(x) = (x - x_j)(x_{j+1} - x) \\ s_j^1 \in \sum_s, s_j^2 \in \sum_s \\ \begin{bmatrix} \gamma & (\underline{I}_j - 1) \\ (\underline{I}_j - 1) & \gamma \end{bmatrix} \succeq 0, \quad \begin{bmatrix} \gamma & (\bar{I}_j - 1) \\ (\bar{I}_j - 1) & \gamma \end{bmatrix} \succeq 0 \end{array} \right. \\ \forall j = 1, \dots, N-1 : \left\{ \begin{array}{l} R_{d,j}(x_{j+1}) = R_{d,j+1}(x_{j+1}) \\ \frac{dR_{d,j}}{dx}(x_{j+1}) = \frac{dR_{d,j+1}}{dx}(x_{j+1}) \\ \gamma \geq 0 \end{array} \right. \end{array} \right. , \quad (2.33)$$

where $\deg(s_j^1) = \deg(s_j^2) = d + \deg(Q) - 2$, $\bar{I}_j \leq Q(x)R_{d,j}(x) \leq \underline{I}_j$, with extreme function values $\bar{I}_j \in \mathbb{R}$ and $\underline{I}_j \in \mathbb{R}$, $\forall j = 1, \dots, N$. For this formulation point-wise continuity of the approximate solution and its gradient at every inter-boundary point x_j has been considered. The reason of including a partition of the domain is to address rational polynomials with strong variations (this normally occurs if the polynomial fluctuates above and below the value of 1) which requires multiplicative inverses of high degree, hard to be achieved by a unique function in the whole domain. Figure 2.3.6(a) shows numerical results of the optimization problem $\mathbf{P}_{d,N}$ for the rational function

$$R(x) = \frac{1}{4x^6 + 5x^5 - 8x^3 + x^2 - x + 1.8}, \quad \forall x \in \Omega = [0, 1],$$

with $d = 10$ and $N = 10$. Circles indicate the boundary nodes x_j selected. The product $Q \cdot R_{d,j}$ is depicted by a dash-dot line, which reflects the hardest multiplicative inversion zone $Q \rightarrow 0$ (dash line). Figure 2.3.6 (b) illustrates the monotone decreasing tendency of the approximation error as the number of elements N increases.

2.3.4 Nonlinear Differential Equations

2.3.4.1 Algebraic Riccati Differential Equation

To illustrate the flexibility of the SOS approach to formulate certain quadratic non-linearities, this section describes an alternative solution to the traditional analytical or numerical methods used to solve the *algebraic (scalar) differential Riccati equation*.

Let

$$\mathbf{S} : \left\{ \begin{array}{l} \dot{x}(t) = ax(t) + bu(t) \\ x(0) = x_0 \end{array} \right. , \quad J(x, u) = \frac{1}{2} \int_0^T (qx^2(t) + ru^2(t)) dt \quad (2.34)$$

be a generic linear dynamical system, with $x \in \mathbb{R}$, where $x_0 \in \mathbb{R}$ is its initial condition and $u \in \mathbb{R}$ is

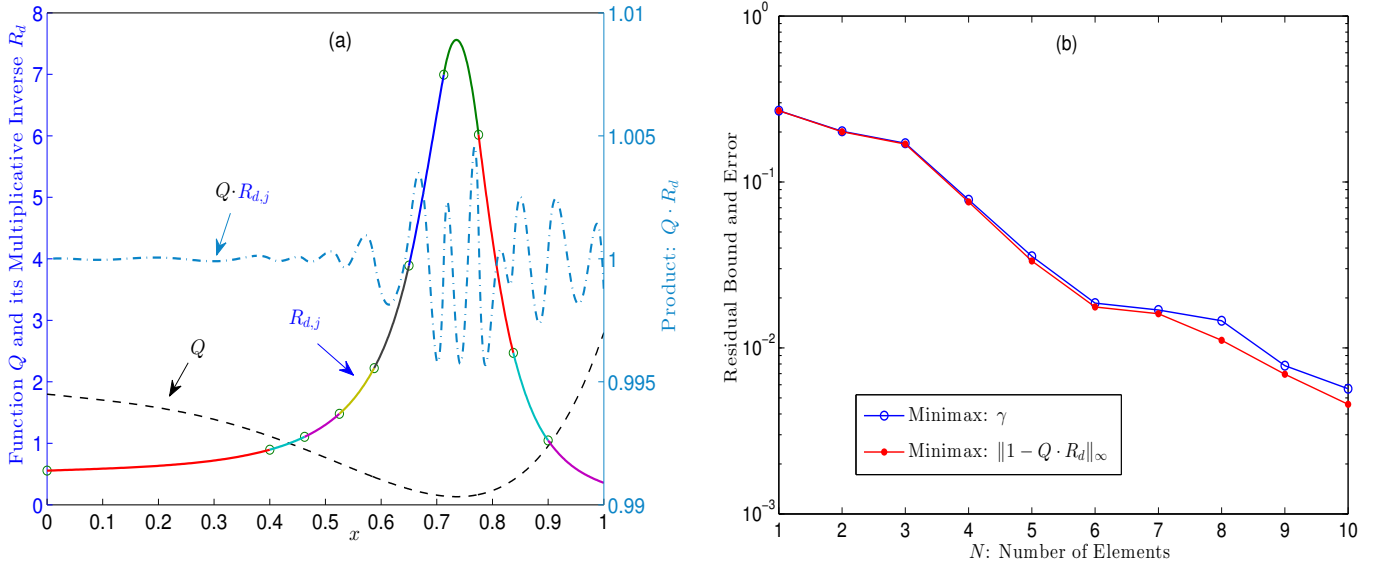


Figure 2.3.6: Reciprocal polynomial approximation. (a) Function Q : dash line and polynomial approximation R_d : solid line (y-axis: left-hand side) (circles: elements nodes), for polynomial degree $d = 10$ and $N = 10$ elements. Product $Q \cdot R_d$: dash-dot line (y-axis: right-hand side). (b) Optimal residual function bound γ : solid-circle line. Norm of the approximation error: $\|1 - Q \cdot R_d\|_\infty$: solid-dot line.

the control action; $a \in \mathbb{R}$ and $b \in \mathbb{R} \setminus \{0\}$ are the system parameters. Let $J = J(x, u)$ be a quadratic performance index with $q \geq 0$ and $r \geq 0$ real scalar weights, and $T > 0$ a finite time of control. Following the classical methodology for solving the Finite Horizon Linear-Quadratic Regulator Problem (LQR), which leads to an optimal feedback $u(t) = -\frac{b}{r}K(t)x(t)$ [184, 185], considering $a = b = 1$, $q = 3$ and $r = 1$, the resulting Riccati equation and its exact solution (K^*) are given by:

$$\mathbf{P} : \begin{cases} \frac{d}{dt}K(t) + 2K(t) - K^2(t) + 3 = 0 \\ K(T) = 0 \end{cases} \Rightarrow K^*(t) = 3 \frac{(1 - e^{-4(T-t)})}{1 + 3e^{-4(T-t)}}. \quad (2.35)$$

To solve the nonlinear differential boundary-value problem $\mathbf{P}$, a polynomial function $K_d(t) = \sum_{j=0}^{z(d)-1} k_j t^j$ is proposed as an approximate solution of K . Due to the particular structure of $\mathbf{P}$ in terms of the quadratic non-linearity in K , this can be formulated as the convex optimization problem:

$$\mathbf{P}_d : \begin{cases} \gamma^* = \inf_{\gamma \in \mathbb{R}^+, k_j \in \mathbb{R}} -\gamma \\ \text{subj. to: } M_d = \begin{bmatrix} \frac{d}{dt}K_d(t) + 2K_d(t) + 3 - s(t)g(t, T) & K_d(t) \\ K_d(t) & 1 \end{bmatrix} \in \Sigma_s \\ \gamma \geq 0, s \in \Sigma_s \\ K_d(T) = 0 \\ \int_0^T K_d(t) dt \geq \gamma \end{cases}, \quad (2.36)$$

for some polynomial $K_d = K_d(t)$ of degree d , SOS polynomial $s = s(t)$ of degree $d - \deg(g) \geq 0$ and

$g = g(t, T)$ time-algebraic description of the finite horizon (interval) of control. In this formulation the main equality is relaxed as $\delta(t, K) = K(t) + 2K(t) + 3 - K^2 \geq 0$ and interpreted as a non-negative “residual function”. In addition, via the Schur’s complement, a sufficient condition of this non-negativity can be formulated by:

$$M = \begin{bmatrix} \frac{d}{dt}K(t) + 2K(t) + 3 & K(t) \\ K(t) & 1 \end{bmatrix} \succeq 0. \quad (2.37)$$

Since the horizon of control $[0, T]$ is equivalent to the set $\Omega_T = \{t \in \mathbb{R}; g(t, T) = t(T - t) \geq 0\} \subset \mathbb{R}$, which is a “compact basic semi-algebraic set” due to the property of compactness of the function g , $\forall t \in \mathbb{R}$, the matrix version of the Putinar’s Positivstellensatz can be applied. Thus, for the approximation K_d of K , the non-negativity of M (2.37) can be formulated via the condition $M_d \in \Sigma_s$ (2.36).

The main idea to solve approximately (2.35) is to exploit the non-negative characteristic of its solution ($K^* > 0$). Since maximizing the integral of K_d implies to increase the upper bounds of the quadratic term K_d^2 , the problem $\mathbf{P}_d$ (2.36) seeks the maximum function K_d^2 so that the SOS condition for M_d is verified. Due to in the Riccati equation (2.35) the term K^2 acts decreasing the residual function $\delta = \delta(t, K)$, to obtain a function K_d^2 with maximum upper bounds is equivalent to reach the lowest bound of the SOS condition $M_d \in \Sigma_s$, leading to $\frac{d}{dt}K_d(t) + 2K_d(t) + 3 - K_d^2(t) = 0$. The results of the optimization problem (2.36) for $T = 1$ are shown in Figure 2.3.7.

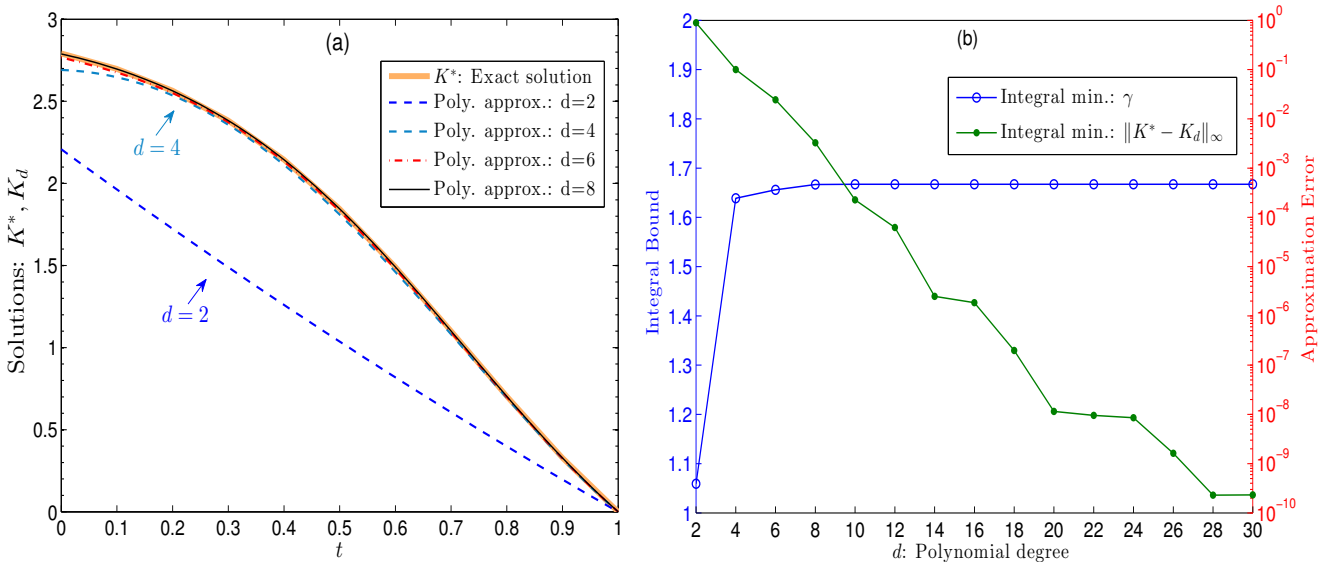


Figure 2.3.7: Polynomial approximation of the Algebraic Riccati Differential Equation (2.35). (a) Exact solution $K^* = K^*(t)$: solid line, polynomial approximation $K_d = K_d(t)$ for polynomial degrees $d = \{2, 4, 6, 8\}$. (b) Lower bound of the integral constraint γ : solid-circle line (y-axis: left-hand side). Norm of the approximation error: $\|K^* - K_d\|_\infty$: dash-dot line (y-axis: right-hand side).

Chapter 3

Convex Optimization Approach for Backstepping PDE Design¹

—David Hilbert (1862-1943)—

In August 1900, city of Paris, D. Hilbert gave one of the main lectures at the second International Congress of Mathematicians. He presented 23 problems on topics covering all branches of mathematics, which in his thought its solution should have significant consequences. The 17th problem concerns the formulation of positive definite rational functions as “sums of quotients of squares”. Its starting point was in 1885 during the doctoral dissertation of Hermann Minkowski, where Hilbert was one of his examiners. H. Minkowski commented that “there exists non-negative real polynomials that cannot be written as finite sums of squares”. The arguments given by H. Minkowski turned D. Hilbert’s mind, and later in 1888, he proved the truth of this statement. However an explicit example of these polynomials was only found in 1967 by Theodore Motzkin, inspired by the Emil Artin’s solution of the 17th Hilbert’s problem (1927). Artin’s proof triggered many further developments, one of these constitutes the foundations of the Sum-of-Squares decomposition and optimization methodology.

In this chapter the Backstepping design for PDEs is recast as a convex optimization problem. Some types of linear one-dimensional parabolic PDEs and a first-order hyperbolic PDE are studied, with particular attention to non-strict feedback structures. The exact matching condition over the resulting Kernel-PDEs is relaxed, providing bounds to guarantee the stability of the target system. Thus, the integral Kernels of the Volterra-type and Fredholm-type transformations involved can be approximated. Due to the compactness of these integral operators, this approximation is carried out by polynomial functions, setting forth differential-BVP problems, linear in term of polynomials coefficients. This characteristic and mainly the SOS-based representation properties of positive polynomials over polytopic domains, enable to formulate an optimization problem and its solution via semidefinite programming. Similarly, inverse Kernels are approximated as the optimal solution of a SOS and Moment problem. Additionally, the uniqueness and invertibility of the Fredholm-type transformation are proved in the

¹This chapter is based on the publications [186, 187].

space of real-analytic functions and continuous functions. Numerical examples illustrate the performance of the method proposed.

3.1 Introduction

Continuous-time Backstepping for PDEs² is a well-established methodology in boundary control/observer design in DPS [21, 50]. It mainly relies on the basis of the *Volterra transformation* [188, 189, 40, 190], to map an “original” system (PDE model of the problem addressed) into a particular “target” system (PDE model with desired features of stability and convergence).

This methodology of design involves two key elements:

- (i) well-posedness/solution of the **Kernel-PDE**,
- (ii) invertibility of the integral transformation,

and has mostly been applied on systems known as “*strict-feedback*” systems. In these cases, the Volterra-type transformation, the invertibility of which is a well-know property [18, 191, 192, 193, 194], is a suitable transformation to manage its causal structure (causal in space [195, 196]). To achieve the “target” system, this transformation leads to the so-called *Kernel-PDE*, which is simple to solve in comparison with the operator Riccati equation derived from the linear quadratic regulator (LQR) approach [12, 43, 44, 45, 22, 48]. Moreover, for some classes of such systems, this problem can be reduced to a *standard* Kernel-PDE form, which allows obtaining a *closed-form solution* [197, 198, 199]. In general (e.g., systems with spatially and temporally variant parameters) a closed-form analytic solution is improbable and simple numeric method cannot be applied directly [49].

To solve the Kernel-PDEs, these are commonly transformed into integral equations to be solved via the *Successive Approximation* method [200]. The mean feature of this methodology is the possibility of finding Kernels with a closed-form solution or recursive formulas for their computation [21, 40, 50, 190, 197, 198, 199, 200, 201]. Since for strict-feedback systems this methodology has provided a useful tool, less research efforts have been devoted to solve the Kernel-PDE in alternative ways.

Recently, Backstepping for PDEs has been applied on systems with “*non-strict*” feedback structure³ via the application of the *Fredholm-type transformation*, for parabolic [202, 205] as well as hyperbolic

²For a complete exposition of this methodology see [21, 50, 51, 18, 49]. A brief summary related to the observer design problem is presented in Appendix B.

³For instance, these systems can be found in dynamics with non-local terms involving the whole spatial domain, in PDE models [202] or finite-dimensional with distributed delays [203]. Similarly, in design-oriented problems such as: control of coupled PDE-ODE (Ordinary Differential Equations) systems by under-actuated schemes [204, 82] (fewer actuators than spatial states [83]; it avoids an additional control action to cancel the non-strict feedback term [198]), or observer design for systems whose output (sensing) comprises the states on the whole domain [84].

PDEs [204, 82]. In these cases, the Volterra-type transformation cannot be used and the Fredholm-type transformation provides a suitable alternative to deal with it. However, its application leads to new and intricate mathematical problems, such as the *operator invertibility* and *Kernel solvability*, where the traditional tools just provide partial answers. For instance:

- the application of the concepts of fixed point theory (Picard sequence of successive approximations [206, 207]) does not lead to a general prove of convergence of an approximate solution of the Kernel-PDE nor invertibility of a Fredholm-type transformation. This is due to the necessary condition of contraction of the resulting operator (Kernels with small spectral radius) which leads to conservative constraints over the system coefficients [204, 82], one of the main drawbacks of this methodology of analysis for addressing general cases [200, 208].
- On the contrary, if a particular Kernel structure is proposed, such as partially separable Kernels, a simplified analysis can be carried out based on the the method of separation of variables [205]. However, the invertibility of the integral transformation and solvability of the resulting Kernel-PDE is limited to a specific class of coefficients of the system.

Since the current methodologies used in Backstepping for PDEs mostly are framed in the context of the Banach's contraction principle ⁴ [211], as it has been pointed out in [82], new methods of analysis and solution of the Kernel equations with integral operators which are not contractions are needed.

The main objective of the proposed approach is to determine Kernels to guarantee stability. Since the target system involves a margin of clearance in its stability, which is characterized as a domain of stability rather than a specific point, this allows relaxing the exact zero matching condition over the Kernel-PDEs. Thus, the resulting Kernel-PDEs can be solved approximately to achieve this stable domain. The essential aims of the method of solution are:

- (a) independence to the spectral characteristics of the resulting operators (thus, it does not impose restrictions over the system),
- (b) flexibility to include a wide range of linear integral transformations,
- (c) boundness and smoothness in the PDE-Kernel solution,

to provide a general approach.

3.1.1 Integral Compact Operators

In general, linear differential equations can be transformed into linear integral equations. Frequently, these equations involve operators which are bounded or compact (completely continuous)

⁴These tools are also typically used to prove existence and uniqueness [190, 201, 208, 209, 210].

[200, 208, 20, 212, 213, 214]. In fact, every linear integral operator $\mathbb{A}_K : \mathcal{X} \rightarrow \mathcal{X}$:

$$\mathbb{A}_K[u(\cdot)](\mathbf{x}) := \int_{\Omega \subset \mathbb{R}^n} K(\mathbf{x}, \mathbf{y}) u(\mathbf{y}) d\mathbf{y} \quad (3.1)$$

with continuous Kernel or weakly singular **Kernel** K , is **compact** on the Banach space of continuous functions.⁵

This is the case of the resulting integral operators derived from the Kernel-PDEs in the Backstepping PDE design, as is indicated in [50] (page 19, footnote 2), where the Kernel is bounded twice continuously differentiable.

In infinite-dimensional spaces, for linear equations of second kind, namely

$$(\mathbb{I} - \mathbb{A}_K)[u(\cdot)](\mathbf{x}) = w(\mathbf{x}), \quad (3.2)$$

two approaches are usually carried out to determine whether there exists a bounded inverse⁶:

- The first method is framed in the context of the “Banach contraction principle” [211, 206, 207], based on the so-called Neumann series, for bounded operator with small spectral radius ($\|\mathbb{A}_K\| < 1$) [208]. This is the **essential tool in the standard Backstepping PDE methodology**, which relies on the inherent *contraction* property of the Volterra operator [217], guaranteeing the uniform convergence of the *Successive Approximation method* [40, 190, 201, 208].
- The second method is a re-statement of the celebrated Fredholm Alternative theorem [208, 20, 213, 218], based on the compactness property of $\mathbb{A}_K$, which is the **core of the proposed approach**. In this case the existence of a unique trivial solution $u = 0$ of the homogeneous equation $u - \mathbb{A}_K u = 0$ implies invertibility and thus the uniqueness of solutions.

In most of traditional Banach spaces and for all Hilbert spaces, *every compact operator is a limit of finite rank operators* [194, 216, 219, 220].⁷ For continuous Kernels, a simple option for establishing this sequence are **polynomial Kernels**, which are a special class of degenerate Kernels⁸[208, 216]. For instance, if $K_N := \sum_{j=0}^N k_j x^{\alpha_j} y^{\beta_j}$ then

$$\mathbb{A}_{K_N}[u(\cdot)](x) = \int_{\Omega} K_N(x, y) u(y) dy = \sum_{j=0}^N k_j x^{\alpha_j} \int_{\Omega} y^{\beta_j} u(y) dy$$

⁵ $\mathbb{A}_K$ is compact in the Banach space of continuous functions: $\mathcal{X} = (\mathcal{C}(\Omega; \mathbb{R}), \|\cdot\|_{\infty})$ and on the Hilbert space of square integrable functions: $\mathcal{X} = (\mathcal{L}^2(\Omega; \mathbb{R}), \langle \cdot, \cdot \rangle_{\mathcal{L}^2})$. Likewise, for square integrable Kernels, $\mathbb{A}_K$ is compact on this Hilbert space [208, 215, 216].

⁶For bounded linear operators, bijectivity is a sufficient and necessary condition for (bounded) invertibility [192, 20].

⁷Compact operators resemble the behaviour of operators in finite-dimensional spaces [20, 19].

⁸Degenerate Kernels has been proposed to simplify the solution of the resulting Kernel-PDEs and for verifying invertibility of the integral transformation involved [205, 21, 50]. However, these approaches are restricted to particular/conservative system properties [205].

is a polynomial in the span of $\{x^{\alpha_j}\}_{j=0}^N$ so that $\mathbb{A}_{K_N}$ is a finite rank operator and hence compact [216]. Moreover, since $K \in \mathcal{C}(\Omega \times \Omega)$ ($\Omega = [0, 1]$), based on the Weierstrass approximation theorem [221], there exists a sequence of polynomials $\{K_N\}_{N=0}^\infty$ such that:

$$\|\mathbb{A}_K - \mathbb{A}_{K_N}\| = \sup_{\|u\| \leq 1} \|\mathbb{A}_{K-K_N} u\| \leq \|K - K_N\|_\infty \xrightarrow{N \rightarrow \infty} 0, \quad (3.3)$$

with $k_j \in \mathbb{R}$, $\alpha_j \in \mathbb{N}$ and $\beta_j \in \mathbb{N}$ for $j = 0, \dots, N$. Equivalent results can be obtained for square integrable Kernels [40, 222]. It is worth nothing that in this case the linear equation (3.2) is reduced to a finite-dimensional problem [190, 201, 221].

3.2 Parabolic PDE and the Volterra-type Operator

3.2.1 Problem Setting

Backstepping for PDEs has mostly been applied on systems with strict-feedback structure. A usual example of such systems is the class of parabolic PDEs [21, 50] described by:

$$\begin{aligned} u_t(x, t) &= \epsilon u_{xx}(x, t) + \lambda(x)u(x, t) \\ u(0, t) &= 0, \quad u(1, t) = U(t), \end{aligned} \quad (3.4)$$

where ϵ is a constant diffusivity coefficient, λ is a spatially varying reactivity parameter and $u(x, 0) = u_0(x) \in \mathcal{C}(\Omega; \mathbb{R})$ is the initial condition.⁹

The objective is to find a control action $U = U(t)$ so that the origin of (3.4) is exponentially stable. For this class of system, the Backstepping PDE methodology proposes a Volterra-type transformation:

$$\begin{aligned} w(x, t) &= u(x, t) - \int_0^x K(x, y)u(y, t)dy \\ &= (\mathbb{I} - \mathbb{V}_K)[u(\cdot, t)](x), \end{aligned} \quad (3.5)$$

where $\mathbb{I}$ is the identity operator and $\mathbb{V}_K: \mathcal{C}(\Omega; \mathbb{R}) \rightarrow \mathcal{C}(\Omega; \mathbb{R})$, to transform the **original system** (3.4) into a **target system**:

$$\begin{aligned} w_t(x, t) &= \epsilon w_{xx}(x, t) - cw(x, t) \\ w(0, t) &= 0, \quad w(1, t) = 0, \end{aligned} \quad (3.6)$$

which is exponentially stable ($\frac{c}{\epsilon} > -\frac{\pi^2}{4}$).

Following the standard Backstepping PDE design procedure detailed in [21], the transformed system

⁹A wide class of reaction-advection-diffusion PDEs can be transform into (3.4) with constant diffusion and no advection terms via the so-called Gauge transformation [21, 50].

(3.4) takes the form:

$$\begin{aligned}
w_t(x, t) - \epsilon w_{xx}(x, t) + cw(x, t) &= \underbrace{\epsilon K(x, 0)}_{\delta_0(x)} u_x(0, t) + \underbrace{\left((\lambda(x) + c) + 2\epsilon \frac{d}{dx} K(x, x) \right)}_{\delta_1(x)} u(x, t) + \\
&\int_0^x \underbrace{(\epsilon K_{xx}(x, y) - \epsilon K_{yy}(x, y) - (\lambda(y) + c)K(x, y))}_{\delta_2(x, y)} u(y, t) dy, \\
w(0, t) &= u(0, t) = 0, \quad w(1, t) = \underbrace{u(1, t)}_{=U(t)} - \int_0^1 K(1, y) u(y, t) dy = 0,
\end{aligned} \tag{3.7}$$

so that the target system (3.6) is achievable if the boundary control action is determined by $U(t) = \int_0^1 K(1, y) u(y, t) dy$ and if the continuous bounded Kernel $K = K(x, y)$ satisfies the so-called **Kernel-PDE**:

$$\delta_2(x, y) = 0, \quad \delta_1(x) = 0, \quad \delta_0(x) = 0, \quad \forall (x, y) \in \Omega_L, \tag{3.8}$$

namely an “exact zero matching” condition over each term δ_i ($i = 1, 2, 3$), which are henceforth considered as **residual functions**. This linear hyperbolic PDE (Klein-Gordon-type) is well-posed and, for a constant reactivity term $\lambda = \lambda_0$, it can be solved in a closed form [21, 50, 197]. Its solution, denoted by K^* :

$$K^*(x, y) = -\bar{\lambda} y \frac{I_1 \left(\sqrt{\bar{\lambda}(x^2 - y^2)} \right)}{\bar{\lambda}(x^2 - y^2)}, \quad \bar{\lambda} = \frac{\lambda_0 + c}{\epsilon},$$

is in terms of the first-order modified Bessel function I_1 .

3.2.2 Kernel-PDE as a Convex Optimization Problem

Proposition 1. *Let*

$$u(x, t) = w(x, t) + \int_0^x L(x, y) w(y, t) dy \tag{3.9}$$

be the inverse transformation of (3.5) [21, 191] and the relation:

$$L(x, y) = \check{L}(x, y) - \sigma, \quad \check{L}(x, y) \geq 0, \quad \sigma \geq 0 \tag{3.10}$$

a positive decomposition of L in the triangular domain Ω_L . Let $m_{0,0}$ be the 0-order moment of $\check{L}$ in accordance with

$$m_{i,j} := \int_0^1 \int_0^x x^i y^j \check{L}(x, y) dy dx. \tag{3.11}$$

Let $c \geq 0$, $\bar{\delta}_1 = \max_{x \in \Omega} |\delta_1(x)|$, $\bar{\delta}_2 = \max_{(x,y) \in \Omega_L} |\delta_2(x,y)|$ and $\delta_0(x) = 0$, the transformed system (3.7) is exponentially stable if the residual functions satisfy:

$$\bar{\delta}_1 + \bar{\delta}_2 \leq \epsilon \left(\frac{4\pi^2}{\pi^2 + 20} \right) \left(\frac{1}{4m_{0,0} + \sigma + 1} \right). \quad (3.12)$$

Proof. Let $V = \frac{1}{2} \int_0^1 w^2(x,t)dx = (1/2)\|w(x)\|^2$ be a Lyapunov functional.¹⁰ Its time-derivative $\dot{V} = \int_0^1 w(x)w_t(x)dx$ along the trajectory (3.7), with $\delta_0(x) = 0$, is given by:

$$\dot{V} = \underbrace{\epsilon \int_0^1 w(x)w_{xx}(x)dx}_{T_1} + \underbrace{\int_0^1 w(x)u(x)\delta_1(x)dx}_{T_2} - c \int_0^1 w^2(x)dx + \underbrace{\int_0^1 w(x) \int_0^x \delta_2(x,y)u(y)dydx}_{T_3}. \quad (3.13)$$

With respect to the term T_1 , using integration by parts, the boundary conditions in (3.6), and splitting the resulting expression by a factor $4/(\pi^2 + 4) \leq \theta < 1$, yields¹¹:

$$T_1 = w(x)w_x(x) \Big|_0^1 - \int_0^1 w_x^2(x)dx \leq \underbrace{-\epsilon\theta\|w_x(x)\|^2}_{T_{1a}} - \underbrace{\epsilon(1-\theta)\|w_x(x)\|^2}_{T_{1b}}. \quad (3.14)$$

Then, applying Wirtinger's inequality on the term T_{1a} and Agmon's and Young's inequalities on the term T_{1b} [21, 50, 223, 224, 225], yields:

$$T_1 \leq -\epsilon\theta\frac{\pi^2}{4}\|w(x)\|^2 + \epsilon(1-\theta)(\|w(x)\|^2 - \bar{w}^2),$$

where $\bar{w}^2 = \max_{x \in \Omega} w^2(x)$, which leads to:

$$\dot{V} \leq -\left(\epsilon\theta\left(\frac{\pi^2}{2} + 2\right) + 2(c-\epsilon)\right)V(t) - \epsilon(1-\theta)\bar{w}^2 + T_2 + T_3. \quad (3.15)$$

As for the term T_2 , substituting $u = u(x,t)$ from (3.9), the inverse Kernel L from (3.10), taking an upper bound by means of the maximum absolute value of some integrand functions and the maximum value of the residual functions, yields:

$$\begin{aligned} T_2 &\leq \int_0^1 w^2(x)|\delta_1(x)|dx + \int_0^1 |w(x)||\delta_1(x)| \int_0^x |L(x,y)||w(y)|dydx \\ &\leq \bar{\delta}_1\|w(x)\|^2 + \bar{\delta}_1\bar{w}^2 \int_0^x \check{L}(x,y)dydx + \underbrace{\sigma\bar{\delta}_1 \int_0^1 |w(x)| \int_0^x |w(y)|dydx}_{T_{2a}}, \end{aligned}$$

¹⁰For a clearer description, the time-dependence in the functions is dropped ($w(x) \equiv w(x,t)$). The norm is the usual in the space of square integrable functions on the domain $\Omega = [0, 1] : \|w(x)\|^2 = \int_\Omega w^2(x)dx$.

¹¹The lower bound of θ corresponds to the condition: $c \geq 0$.

where $\bar{\delta}_1 = \max_{x \in \Omega} |\delta_1(x)|$. Then, using the identity $\int_a^b f(x) \int_a^x f(y) dy dx = (1/2)(\int_a^b f(x) dx)^2$ for f continuous function [226] on the term T_{2a} , and Grüss' Integral inequality [227] on its resulting term, leads to:

$$T_2 \leq (2 + \sigma) \bar{\delta}_1 V(t) + \bar{\delta}_1 \left(\frac{\sigma}{8} + m_{0,0} \right) \bar{w}^2. \quad (3.16)$$

Changing the order of integration in the term T_3 and following a similar procedure as described above for the term T_2 , yields:

$$\begin{aligned} T_3 &= \int_0^1 u(y) \int_y^1 w(x) \delta_2(x, y) dx dy \leq \int_0^1 |w(x)| \int_0^x |w(y)| |\delta_2(x, y)| dy dx + \\ &\quad \int_0^1 \int_0^y \check{L}(y, s) |w(s)| ds \int_y^1 |w(x)| |\delta_2(x, y)| dx dy + \sigma \int_0^1 \int_0^y |w(s)| ds \int_y^1 |w(x)| |\delta_2(x, y)| dx dy \\ &\leq (1 + 2\sigma) \bar{\delta}_2 V(t) + \bar{\delta}_2 \left(\frac{1 + 2\sigma}{8} + m_{0,0} \right) \bar{w}^2, \end{aligned} \quad (3.17)$$

with $\bar{\delta}_2 = \max_{(x,y) \in \Omega_L} |\delta_2(x, y)|$. Finally, using in (3.15) the upper bounds (3.16) and (3.17) for the terms T_2 and T_3 , respectively, grouping terms with respect to $V = \|w(x)\|^2/2$ and $\bar{w}^2$, and choosing $\theta = 20/(\pi^2 + 20)$, the condition $\dot{V} \leq 0$ is satisfied if:

$$-(\bar{\delta}_1 + \bar{\delta}_2) + \frac{\bar{\delta}_1 \sigma + \bar{\delta}_2}{2(\sigma + 1)} + \frac{c + \epsilon \theta_0}{(\sigma + 1)} \geq 0,$$

and

$$-(\bar{\delta}_1 + \bar{\delta}_2) + \frac{\bar{\delta}_1(\sigma + 2) + \bar{\delta}_2}{2(4m_{0,0} + \sigma + 1)} + \frac{\epsilon \theta_0}{(4m_{0,0} + \sigma + 1)} \geq 0,$$

where $\theta_0 = 4\pi^2/(\pi^2 + 20)$ and $c \geq 0$, which leads to the expression (3.12), i.e. a sufficient condition for exponential stability of (3.7). $\square$

Motivated by the result of Proposition 1, which sets forth a “margin of clearance” in the stability of the transformed system (3.7), a *relaxation* of the exact zero matching condition for the residual functions δ_1 and δ_2 can be considered. This allows formulating an approximate solution for the Kernel-PDE (3.8).

Proposition 2. *Let $N(x, y) = \sum_{k=0}^{z(d)-1} n_k x^{\alpha_k} y^{\beta_k}$ be a polynomial approximation of K of arbitrary even degree $d > d_\lambda \in \mathbb{N}$ in accordance with (1.3). Let $Y = Y(x)$, $Z = Z(x, y)$ be polynomials of degree $d_Y = d$ and $d_Z = 2 \left\lceil \frac{d+d_\lambda+2}{2} \right\rceil$, respectively; let $\underline{\rho}_Y$, $\underline{\rho}_Z$, $\bar{\rho}_Y$, and $\bar{\rho}_Z$ be lower and upper bounds of these polynomials in Ω_L ; $\gamma_j \geq 0$, $j = 1, \dots, 4$. For a reactivity term $\lambda = \lambda(x)$ described by a polynomial function of degree d_λ , the Kernel-PDE (3.8) can be formulated as the convex optimization problem:*

$$\begin{aligned} \underset{\gamma_j, N, Y, Z, s_j}{\text{minimize:}} \quad & \gamma_1 + \gamma_2 + \gamma_3 + \gamma_4 \end{aligned} \quad (3.18)$$

$$\text{subject to: } Y(x) - \delta_1(x) = 0, \quad Z(x, y) - \delta_2(x, y) = 0, \quad (3.19)$$

$$\left(Y(x) - \underline{\rho}_Y - s_1(x) \, g_1(x) \right) \in \Sigma_s, \quad (3.20)$$

$$(\bar{\rho}_Y - Y(x) - s_2(x) \, g_1(x)) \in \Sigma_s, \quad (3.21)$$

$$\left(Z(x, y) - \underline{\rho}_Z - [s_3(x, y) \, s_4(x, y)] \, g_L(x, y) \right) \in \Sigma_s, \quad (3.22)$$

$$(\bar{\rho}_Z - Z(x, y) - [s_5(x, y) \, s_6(x, y)] \, g_L(x, y)) \in \Sigma_s, \quad (3.23)$$

$$s_j \in \Sigma_s, \forall j = 1, \dots, 6, \quad (3.24)$$

$$\begin{bmatrix} \gamma_1 & \underline{\rho}_Y \\ \underline{\rho}_Y & \gamma_1 \end{bmatrix} \succeq 0, \quad \begin{bmatrix} \gamma_2 & \bar{\rho}_Y \\ \bar{\rho}_Y & \gamma_2 \end{bmatrix} \succeq 0, \quad \begin{bmatrix} \gamma_3 & \underline{\rho}_Z \\ \underline{\rho}_Z & \gamma_3 \end{bmatrix} \succeq 0, \quad \begin{bmatrix} \gamma_4 & \bar{\rho}_Z \\ \bar{\rho}_Z & \gamma_4 \end{bmatrix} \succeq 0, \quad (3.25)$$

$$\epsilon N_{xx}(x, y) - \epsilon N_{yy}(x, y) - (\lambda(y) + c)N(x, y) = \delta_2(x, y) \quad (3.26)$$

$$(\lambda(x) + c) + 2\epsilon \frac{d}{dx} N(x, x) = \delta_1(x), \quad (3.27)$$

$$N(x, 0) = 0, \quad (3.28)$$

$\forall (x, y) \in \Omega_L$, for some polynomials s_1, s_2 of degree $d_Y - 2$, s_3, s_4, s_5, s_6 of degree $d_Z - 2$ and $g_L = [g_1, g_2]^\top$, with $g_1(x) = x(1 - x)$ and $g_2(x, y) = y(x - y)$. The optimal minimal bounds for the residual functions are: $\bar{\delta}_1 = \max\{\gamma_1, \gamma_2\}$ and $\bar{\delta}_2 = \max\{\gamma_3, \gamma_4\}$.

Proof. Since N is a polynomial approximation of K , as which is indicated above, the residual functions in (3.7) have a polynomial structure determined by (3.26)-(3.28), equivalent to Y and Z by (3.19). In addition, the quadratic module associated to the representation of the domains Ω and Ω_L are Archimedean (see Lemma 1). Based on Putinar's Positivstellensatz [148, 138], via the SOS decomposition (3.20)-(3.24), the unknown extreme values of Y and Z ($\underline{\rho}_Y, \bar{\rho}_Y, \underline{\rho}_Z, \bar{\rho}_Z$) for each residual function can be determined via a polynomial optimization problem, which is convex in terms of polynomial coefficients and solved via semidefinite programming [136]. The absolute value of these upper and lower bounds are given by means of (3.25), so that the linear cost function (3.18) yields $Y \rightarrow 0, Z \rightarrow 0$ in Ω and Ω_L , respectively, which is equivalent, by (3.19), to the minimal optimal value of the residuals functions δ_1 and δ_2 in their respective domains. $\square$

Remark 6. Due to (3.8) is well posed, since $\mathbb{R}_d(\Omega_L)$ is dense in $\mathcal{C}(\Omega_L)$ and $\mathcal{L}^2(\Omega_L)$, if (3.18)-(3.28) establishes a sequence of minimizers (N_d^*) , $d \in \mathbb{N}$, so that $\{\bar{\delta}_1^{d_1}, \bar{\delta}_1^{d_2}, \dots, \rightarrow 0\}$ and $\{\bar{\delta}_2^{d_1}, \bar{\delta}_2^{d_2}, \dots, \rightarrow 0\}$ are monotone decreasing sequences of $\bar{\delta}_1$ and $\bar{\delta}_2$ as $d \rightarrow \infty$, it is expected to achieve $N_d^* \rightarrow K$ as $d \rightarrow \infty$.

Remark 7. In (3.18)-(3.28) two extra polynomials Y and Z equivalent to δ_1 and δ_2 , respectively, are included. These allow managing the necessary even degree condition of SOS [136], which is not always satisfied by (3.26) or (3.27).

3.2.3 Approximate Inverse Transformation

It is well-known that Volterra operators of the second kind (3.5) with continuous Kernels have a unique solution which is globally invertible [18, 191, 208]. According to the standard Backstepping

PDE procedure [21, 50], the inverse Kernel L in (3.9) is determined following the same approach that leads to the Kernel K , and computed via the *Successive Approximation method*. However, this procedure is not suitable for residual functions $\delta_1(x) \neq 0, \delta_2(x, y) \neq 0$.

A tractable alternative to find an approximate smooth solution of the inverse Kernel L in (3.9) is proposed according to the Moment approach, as it has been presented in Chapter 2, Section 2.2 ([137, 161, 162]).

Proposition 3. *Let*

$$\begin{aligned} L(x, y) - K(x, y) - \int_y^x K(x, s)L(s, y)ds &= 0, \\ \mathbb{T}[L(\cdot, \cdot)](x, y) &= 0, \end{aligned} \quad (3.29)$$

$\forall (x, y) \in \Omega_L$ be the relation satisfied by any direct Kernel and its inverse with the Volterra-type transformations (3.5) and (3.9). Let (3.10) be the positive decomposition of L , where σ is its minimum value in Ω_L . Let $N(x, y) = \sum_{k=0}^{z(d)-1} n_k x^{\alpha_k} y^{\beta_k}$ be a known polynomial approximation of K in accordance with (1.3), a sequence of approximate moments $m_{i,j}$ can be determined by the convex optimization problem:

$$\begin{aligned} \underset{\gamma_j, \varrho_j, m_{i,j}, \sigma}{\text{minimize:}} \quad & \gamma_1 + \gamma_2 + \varrho_2 - \varrho_1 \end{aligned} \quad (3.30)$$

$$\text{subject to:} \quad \rho_{s_{ij}} \geq -\gamma_1, \quad \rho_{s_{ij}} \leq \gamma_2, \quad (3.31)$$

$$\sum_{i=0}^{2d_m} \sum_{j=0}^{2d_m} m_{i,j} + \sigma - \varrho_1 \geq 0, \quad \varrho_2 - \sum_{i=0}^{2d_m} \sum_{j=0}^{2d_m} m_{i,j} - \sigma \geq 0, \quad (3.32)$$

$$\begin{aligned} \rho_{s_{ij}} = m_{i,j} - \frac{\sigma}{(j+1)(i+j+2)} - \int_0^1 \int_0^x N(x, y) x^i y^j dy dx - \\ \sum_{k=0}^{z(d)-1} \frac{n_k}{i+\alpha_k+1} \left(m_{\beta_k, j} - m_{i+\alpha_k+\beta_k+1, j} - \frac{\sigma}{j+1} \left(\frac{1}{j+\beta_k+2} + \frac{1}{i+j+\alpha_k+\beta_k+3} \right) \right), \end{aligned} \quad (3.33)$$

$$M_0 \geq 0, \quad M_1 \succeq 0, \quad M_2 \succeq 0, \dots, M_{d_m} \succeq 0, \quad (3.34)$$

$$R_0 \geq 0, \quad R_1 \succeq 0, \quad R_2 \succeq 0, \dots, R_{d_m-1} \succeq 0, \quad (3.35)$$

$$S_0 \geq 0, \quad S_1 \succeq 0, \quad S_2 \succeq 0, \dots, S_{d_m-1} \succeq 0, \quad (3.36)$$

$$T_0 \geq 0, \quad T_1 \succeq 0, \quad T_2 \succeq 0, \dots, T_{d_m-2} \succeq 0, \quad (3.37)$$

$$M_0 = m(I_0) = [m_{0,0}], \quad M_1 = m(I_1) = \begin{bmatrix} M_0 & m_{1,0} & m_{0,1} \\ m_{1,0} & m_{2,0} & m_{1,1} \\ m_{0,1} & m_{1,1} & m_{0,2} \end{bmatrix}, \quad M_2 = m(I_2), \dots,$$

$$M_{d_m} = m(I_{d_m}), \quad (3.38)$$

$$R_r = m(I_r + g_1(1)) - m(I_r + g_1(2)), \quad (3.39)$$

$$S_r = m(I_r + g_2(1)) - m(I_r + g_2(2)), \quad (3.40)$$

$$T_r = -m(I_r + g_3(1)) + m(I_r + g_3(2)) + m(I_r + g_3(3)) - m(I_r + g_3(4)), \quad (3.41)$$

$$I_r = \begin{bmatrix} \begin{bmatrix} [0,0] & 1,0 & 0,1 \\ 1,0 & 2,0 & 1,1 \\ 0,1 & 1,1 & 0,2 \end{bmatrix} & \begin{bmatrix} 2,0 & 1,1 & 0,2 \\ 3,0 & 2,1 & 1,2 \\ 2,1 & 1,2 & 0,3 \end{bmatrix} & \begin{bmatrix} 3,0 & \dots \\ 4,0 & \dots \\ 3,1 & \dots \end{bmatrix} \\ \begin{bmatrix} 2,0 & 3,0 & 2,1 \\ 1,1 & 2,1 & 1,2 \\ 0,2 & 1,2 & 0,3 \end{bmatrix} & \begin{bmatrix} 4,0 & 3,1 & 2,2 \\ 3,1 & 2,2 & 1,3 \\ 2,2 & 1,3 & 0,4 \end{bmatrix} & \begin{bmatrix} 5,0 & \dots \\ 4,1 & \dots \\ 3,2 & \dots \end{bmatrix} \\ \begin{bmatrix} 3,0 & 4,0 & 3,1 \\ \vdots & \vdots & \vdots \end{bmatrix} & \begin{bmatrix} 5,0 & 4,1 & 3,2 \\ \vdots & \vdots & \vdots \end{bmatrix} & \begin{bmatrix} 6,0 & \dots \\ \vdots & \ddots \end{bmatrix} \end{bmatrix}, \quad (3.42)$$

$$g_1 = [(1,0), (2,0)], \quad g_2 = [(1,1), (0,2)], \quad (3.43)$$

$$g_3 = [(1,2), (2,1), (2,2), (3,1)], \quad (3.44)$$

where M_r are moment matrices of dimension $z(r) = \frac{(r+1)(r+2)}{2}$, R_r, S_r, T_r are localizing matrices (their sequence is limited by $r \leq d_m - \lceil \max(\deg(g_i))/2 \rceil$) associated to the compact basic semi-algebraic set description of Ω_L [137], the entries of which are indexed by I_r according to the order of the powers in the canonical polynomial basis Φ_r (1.3); $s_{ij} = i(2d_m+1)+j$, $\forall i = 0, \dots, 2d_m-d-1$, $j = 0, \dots, 2d_m$, for an arbitrary moment order $d_m \geq \lceil \frac{d+1}{2} \rceil$ and $\gamma_1 \geq 0$, $\gamma_2 \geq 0$, $\varrho_1 \geq 0$ and $\varrho_2 \geq 0$.

Proof. Since the zero function is the only function orthogonal to every element in an inner product space: $\langle \mathbb{T}[L], v \rangle_{\mathcal{L}^2} = 0$, $\forall v \in \mathcal{L}^2(\Omega_L) \Leftrightarrow \mathbb{T}[L] = 0$ (weak formulation) [176], and due to the canonical polynomial basis $\Phi_r(x, y) = \{\phi_i(x)\phi_j(y) = x^i y^j; i+j \leq r, \forall i, j, r \in \mathbb{N}\}$ generates a dense subset of the (separable) Hilbert space $\mathcal{L}^2(\Omega)$, an approximate solution of (3.29) can be found by means of the set of linear equations $\langle \mathbb{T}[L], \phi_i(x)\phi_j(y) \rangle_{\mathcal{L}^2(\Omega_L)} = 0$ (finite dimensional problem):

$$\begin{aligned} & \int_0^1 \phi_i(x) \int_0^x L(x, y) \phi_j(y) dy dx - \int_0^1 \phi_i(x) \int_0^x N(x, y) \phi_j(y) dy dx \\ & - \int_0^1 \phi_i(x) \int_0^x N(x, y) \int_0^y L(y, s) \phi_j(s) ds dy dx = 0, \end{aligned} \quad (3.45)$$

for N a known polynomial approximation of K . Interchanging the order of integration in the last term of (3.45), substituting L by (3.10) and plugging in the polynomial series N and test functions $\phi_i(x) = x^i, \phi_j(y) = y^j$ yields:

$$\begin{aligned} & \int_0^1 \int_0^x \check{L}(x, y) x^i y^j dy dx - \frac{\sigma}{(j+1)(i+j+2)} - \int_0^1 \int_0^x N(x, y) x^i y^j dy dx \\ & - \int_0^1 \left(\int_0^y (\check{L}(y, s) - \sigma) s^j ds \right) \sum_{k=0}^{z(d)-1} n_k y^{\beta_k} \left(\frac{1-y^{i+\alpha_k+1}}{i+\alpha_k+1} \right) dy = 0, \end{aligned} \quad (3.46)$$

$\forall i \in \mathbb{N}, j \in \mathbb{N}, (i+j) \leq r$. Thus, expanding the last term in (3.46), applying definition (3.11) and keeping the integral of third term for numerical computation (N is known), the expression (3.33) is

obtained. This equation is written as a residual function $\rho_{s_{ij}}$, the upper and lower bounds of which are set forth by (3.31) and included in the optimization index (3.30), so that $\rho_{s_{ij}} \rightarrow 0$ as required. Since L is not necessarily positive, it is decomposed by (3.10), where $\check{L} \geq 0$ is considered as a density function for a Borel measure μ : $d\mu = \check{L}dx$, $m_{i,j} = \int_{\Omega_L} x^i y^j d\mu$. Based on the Schmüdgen's representation¹² of non-negative polynomials in a compact basic semialgebraic set [148, 161] (which takes into account all combinations of the functions g_i , $\{1, g_1, g_2, g_3 = g_1 g_2\}$, described by (3.43)-(3.44) in terms of polynomial powers), $\mathbf{y}_r = (m_{i,j})_{(i+j) \leq r}$ is a sequence of moments if and only if the Moment matrices M_r and Localizing matrices R_r, S_r, T_r are semi-definite positive $\forall r \in \mathbb{N}$. However, in practice, a truncated Moment problem can be solved ($r = 0, \dots, d_m$), which is formulated by conditions (3.34)-(3.37). Finally, via the cost function (3.30) with respect to ϱ_1 and ϱ_2 , and the conditions of minimum and maximum (3.32), a bounded optimization problem is enforced, the solution of which is a “valid”¹³ sequence of approximate moments minimizing (3.33). $\square$

Remark 8. *The truncated Moment problem can be solved imposing rank constraints on the Moment and Localizing matrices [137]. This condition sets up a non-convex problem.*

Remark 9. *Alternatively, an optimization problem similar to Proposition 2 can be applied directly to (3.29) to find an approximation of L . This is formulated in Section 3.3.4 for the Fredholm-type operator. Instead of using (3.29), the standard Backstepping PDE approach computes the inverse Kernel following the same approach that leads to the direct Kernel K [21].*

Proposition 4. *Let $\check{L}_r = \sum_{k=0}^{z(r)-1} l_k x^{\alpha_k} y^{\beta_k}$ be a polynomial approximation of $\check{L}$ (3.10) according to definition (1.3). Given a “valid” sequence of approximate moments $\mathbf{y}_r = (m_{i,j})_{(i+j) \leq r}$, $\check{L}$ can be approximated as $\check{L}_r = \theta^\top \Phi_r(x, y)$ by the solution of the convex optimization problem:*

$$\begin{aligned} \underset{\gamma_j \geq 0, \theta, s_j}{\text{minimize:}} \quad & \gamma_1 + \gamma_2 \end{aligned} \quad (3.47)$$

$$\text{subject to:} \quad (\mathbf{M}_r \theta - \mathbf{y}_r) + \gamma_1 \geq 0, \quad \gamma_2 - (\mathbf{M}_r \theta - \mathbf{y}_r) \geq 0, \quad (3.48)$$

$$\theta^\top \Phi_r(x, y) - s_0(x, y) - [s_1(x, y) \ s_2(x, y)] g(x, y) = 0, \quad (3.49)$$

$$s_0, s_1, s_2 \in \Sigma_s, \quad (3.50)$$

where $\gamma_1 \geq 0, \gamma_2 \geq 0, g = [g_1, g_2]^\top$ with $g_1(x) = x(1-x) \geq 0$ and $g_2(x, y) = y(x-y) \geq 0$; polynomials s_0 of degree r and s_1, s_2 of degree $r-2$; $\theta = [l_0, l_1, \dots, l_{z(r)-1}]^\top$ and $\mathbf{M}_r = \int_0^1 \int_0^x \Phi_r(x, y) \Phi_r^\top(x, y) dy dx \in \mathbb{S}_+^{z(r)}$, the elements of which are determined by:

$$\mathbf{M}_r(i, j) = \frac{1}{(\alpha_i + \alpha_j + \beta_i + \beta_j + 2)(\beta_i + \beta_j + 1)}, \quad (3.51)$$

¹²The Putinar's positivity representation leads to a similar formulation and numerical results.

¹³Proposition 3 states conditions to obtain a sequence of real numbers which is a “valid” sequence of moments, i.e., it corresponds to a moments of a non-negative function [161].

where the coefficients α_i, β_i correspond to the powers of the canonical polynomial basis Φ_r (1.3). In addition, the sequence of minimizers $\check{L}_r^*, \forall r \in \mathbb{N}$, is such that $\|\check{L} - \check{L}_r^*\|_{\mathcal{L}^2(\Omega_L)} \rightarrow 0$ as $r \rightarrow \infty$, and therefore $L(x, y) \approx \check{L}_r^*(x, y) - \hat{\sigma}$.

Proof. Let $\phi_k(x, y) = x^{\alpha_k} y^{\beta_k}$ be the k -th element of the vector of monomial basis Φ_r (1.3). The (i, j) -th element (row and column, respectively) of the Moment matrix is given by

$$\mathbf{M}_r(i, j) = \int_0^1 \int_0^x \left(x^{\alpha_i} y^{\beta_i} \right) \left(x^{\alpha_j} y^{\beta_j} \right) dy dx,$$

where (3.51) is its direct computation. Since $\check{L}_r^2 = (\sum_{k=0}^{z(r)-1} l_k \phi_k) (\sum_{t=0}^{z(r)-1} l_t \phi_t) = \theta^\top \Phi_r \Phi_r^\top \theta$, where $\theta = [l_0, l_1, \dots, l_{z(r)-1}]^\top$, following [162], the mean square error of the approximation of $\check{L}$ can be upper bounded by¹⁴ $E(\check{L}_r) = \|\check{L} - \check{L}_r\|_{\mathcal{L}^2(\Omega_L)}^2 \leq \theta^\top \mathbf{M}_r \theta - 2\theta^\top \mathbf{y}_r := 2J(\theta)$, where the “known” valid sequence of approximate moments $\mathbf{y}_r = \left(\int_{\Omega_L} \phi_k(x, y) d\mu \right)_{\alpha_k + \beta_k \leq r}$ is ordered according to

$$\mathbf{y}_r = [m_{0,0}, \dots, m_{z(r),0}, m_{0,1}, \dots, m_{z(r),1}, \dots, m_{0,z(r)}, \dots, m_{z(r),z(r)}]^\top.$$

Since J is quadratic in terms of the unknown vector θ , its minimizer satisfies $\frac{d}{d\theta} \left(\frac{1}{2} \theta^\top \mathbf{M}_r \theta - \theta^\top \mathbf{y}_r \right) = \theta(\mathbf{M}_r + \mathbf{M}_r^\top)/2 - \mathbf{y}_r^\top = 0$ ($\mathbf{M}_r \theta = \mathbf{y}_r$), which has a global minimizer since $\frac{d^2 J(\theta)}{d\theta^2} = \mathbf{M}_r \in \mathbb{S}_+^{z(r)}$ [228]. To avoid the matrix inversion in $\theta = \mathbf{M}_r^{-1} \mathbf{y}_r$ and instead of a SDP formulation of J via the Schur complement [162], the approach proposed defines a lower bound γ_1 and an upper bound γ_2 via (3.48), which by means of (3.47) imposes $\mathbf{M}_r \theta \rightarrow \mathbf{y}_r$.

Thus, since the quadratic module associated to the representation of Ω_L is Archimedean (see Lemma 1), $\check{L}_r = \theta^\top \Phi_r(x, y) \geq 0, \forall (x, y) \in \Omega_L$, if this satisfies the Putinar’s Positivstellensatz representation given by (3.49)-(3.50) for some polynomials s_0, s_1 and s_2 with SOS decomposition. A proof of $\|\check{L} - \check{L}_r^*\|_{\mathcal{L}^2(\Omega_L)} \rightarrow 0$ as $r \rightarrow \infty$ is found in [162, Proposition 4]. $\square$

3.3 Hyperbolic PIDE and the Fredholm-type Operator

3.3.1 Problem Setting

Recently, Backstepping for PDEs has been applied to systems with “non-causal” structure. For instance, consider the class of first-order hyperbolic PIDEs (Partial Integral Differential Equations)

¹⁴According to Chapter 2, Section 2.2:

$$\begin{aligned} E(\check{L}_r) &= \|\check{L} - \check{L}_r\|_{\mathcal{L}^2(\Omega_L)}^2 = \int_{\Omega_L} \left(\check{L} - \check{L}_r \right)^2 dx = \int_{\Omega_L} \check{L}_r^2 dx - 2 \int_{\Omega_L} \check{L}_r \underbrace{\check{L} dx}_{d\mu} + \underbrace{\int_{\Omega_L} \check{L}^2 dx}_{\geq 0} \\ &\leq \theta^\top \left(\int_{\Omega_L} \Phi_r \Phi_r^\top dy dx \right) \theta - 2 \sum_{k=0}^{z(r)-1} l_k \int_{\Omega_L} \phi_k(x, y) d\mu \\ &= \theta^\top \mathbf{M}_r \theta - 2\theta^\top \mathbf{y}_r := 2J(\theta), \end{aligned}$$

(see details in [204, 82]):

$$\begin{aligned} u_t(x, t) &= \epsilon u_x(x, t) + f(x)u(0, t) + \int_0^x h_1(x, y)u(y, t)dy + \int_x^1 h_2(x, y)u(y, t)dy, \\ u(1, t) &= U(t), \end{aligned} \quad (3.52)$$

where $u(x, t) = u_0(x) \in \mathcal{L}^2(\Omega)$ is the initial condition and f, h_1, h_2 are real-valued continuous functions. The objective is to find a control action $U = U(t)$ so that the origin of (3.52) is exponentially stable. For this class of system [204, 82] (see also [205] for parabolic systems) a Fredholm-type transformation has been proposed, namely

$$\begin{aligned} w(x, t) &= u(x, t) - \int_0^x P(x, y)u(y, t)dy - \int_x^1 Q(x, y)u(y, t)dy, \\ &= (\mathbb{I} - \mathbb{F})[u(\cdot, t)](x), \end{aligned} \quad (3.53)$$

where $\mathbb{F}: \mathcal{L}^2(\Omega; \mathbb{R}) \rightarrow \mathcal{L}^2(\Omega; \mathbb{R})$ is a linear operator in terms of Kernels P and Q in the lower Ω_L and upper Ω_U triangular domain, respectively, to transform the **original system** (3.52) into the **target system**:

$$\begin{aligned} w_t(x, t) &= w_x(x, t), \\ w(1, t) &= 0, \end{aligned} \quad (3.54)$$

which is exponentially stable. Following the standard Backstepping PDE design procedure (detailed in [82]), the transformed system (3.52) takes the form:

$$\begin{aligned} w_t(x, t) - w_x(x, t) &= \underbrace{\left(f(x) + P(x, 0) - \int_0^x P(x, y)f(y)dy - \int_x^1 Q(x, y)f(y)dy \right)}_{\delta_0(x)} u(0, t) - \underbrace{Q(x, 1)}_{\delta_3(x)} u(1, t) + \\ &\quad \int_0^x u(y, t) \underbrace{\left(h_1(x, y) + P_x(x, y) + P_y(x, y) - \int_0^y P(x, s)h_2(s, y)ds - \int_y^x P(x, s)h_1(s, y)ds - \int_x^1 Q(x, s)h_1(s, y)ds \right)}_{\delta_1(x, y)} dy + \\ &\quad \int_x^1 u(y, t) \underbrace{\left(h_2(x, y) + Q_x(x, y) + Q_y(x, y) - \int_0^x P(x, s)h_2(s, y)ds - \int_x^y Q(x, s)h_2(s, y)ds - \int_y^1 Q(x, s)h_1(s, y)ds \right)}_{\delta_2(x, y)} dy, \\ w(1, t) &= \underbrace{u(1, t)}_{=U(t)} - \int_0^1 P(1, y)u(y, t)dy, \end{aligned} \quad (3.55)$$

so that the target system (3.54) is achievable if the boundary feedback control is determined by $U(t) = \int_0^1 P(1, y)u(y, t)dy$ and if the continuous Kernels P and Q satisfy the **Kernel-PIDE**:

$$\begin{aligned}\delta_1(x, y) &= 0, & \forall (x, y) \in \Omega_L, \\ \delta_2(x, y) &= 0, & \forall (x, y) \in \Omega_U, \\ \delta_0(x) &= 0, \delta_3(x) &= 0, & \forall x \in \Omega.\end{aligned}\tag{3.56}$$

For these coupled hyperbolic PIDEs, a method of analysis, computation and an equivalent sufficient (conservative) condition for a unique solution have been given in [82].

3.3.2 Existence, Uniqueness and Invertibility

In contrast to the Volterra Operator (3.5), existence, uniqueness and invertibility of the Fredholm-type transformation (3.53) have been proved mostly relying on the Banach contraction mapping principle (see [209] and references therein). In this context, [204, 82] proposes a contraction mapping in terms of a system of integral equations equivalent to (3.56), which is used to calculate the Kernels by Picard's iterative method, the convergent of which is guaranteed only for specific conditions on the system under analysis. This kind of *conservative* conditions can be circumvented if the analysis is restricted to the space of real-analytic functions and Kernels with polynomials structure. Moreover, this analysis also remains valid on the space of continuous functions.

Lemma 2. *If $u = u(\cdot, t)$ is a real-analytic function $\forall t \geq 0$, solution of the integral equation (3.53), and the Kernels P and Q are (bounded) polynomials, $P \neq Q$, then the homogeneous equation $(\mathbb{I} - \mathbb{F})[u(\cdot, t)](x) = 0$ has only the trivial solution $u = 0$.*

Proof. Let $P(x, y) = \sum_{j=0}^{z(d_P)-1} p_j x^{\alpha_j} y^{\beta_j}$ and $Q(x, y) = \sum_{j=0}^{z(d_Q)-1} q_j x^{\alpha_j} y^{\beta_j}$ be polynomials of degree $d_P, d_Q \in \mathbb{N}$, respectively, in accordance with definition 1.3; $d = \max\{d_P, d_Q\}, j = 0, \dots, d, p_j = 0, \forall (\alpha_j + \beta_j) > d_P$ or $q_j = 0, \forall (\alpha_j + \beta_j) > d_Q$ as required. For each $t \geq 0$, since an analytic function has a unique series representation $u(x) = \sum_{k=0}^{\infty} a_k x^k$ [229] (without loss of generality, the analysis considers this function real-analytic at $x = 0$), the homogeneous integral equation (3.53) (with $w = 0$) can be formulated as:

$$\sum_{k=0}^{\infty} a_k \left(x^k + \sum_{j=0}^{s(d)-1} \frac{q_j - p_j}{\beta_j + k + 1} x^{\alpha_j + \beta_j + k + 1} - \sum_{j=0}^{s(d)-1} \frac{q_j}{\beta_j + k + 1} x^{\alpha_j} \right) = 0. \tag{3.57}$$

This expression is equivalent to a linear (independent) combination of monomials of the polynomial basis $\Psi_x = [1, x, \dots, x^d, x^{d+1}, x^{d+2}, \dots, x^{d+N+1}, \dots]$; namely

$$\sum_{k=0}^{\infty} x^k \left(\bar{\nu}_k [a_0, a_1, \dots, a_N, \dots]^{\top} \right) = 0 \Leftrightarrow \Psi_x V \mathbf{a} = 0, \tag{3.58}$$

where $\mathbf{a} = [a_0, a_1, \dots, a_N, \dots]^{\top}$ and $\bar{\nu}_k = [\nu_{k,0}, \dots, \nu_{k,N}, \dots]$ is the k -th row of the matrix:

$$V = \begin{bmatrix} \nu_{0,0} & \nu_{0,1} & \cdots & \nu_{0,N} & \cdots \\ \vdots & \vdots & \cdots & \vdots & \cdots \\ \nu_{d,0} & \nu_{d,1} & \cdots & \nu_{d,N} & \cdots \\ \hline \sum_{j=j_0}^m \frac{q_j - p_j}{\beta_j + 1} & \nu_{d+1,1} & \cdots & \nu_{d+1,N} & \cdots \\ 0 & \sum_{j=j_0}^m \frac{q_j - p_j}{\beta_j + 2} & \cdots & \nu_{d+2,N} & \cdots \\ 0 & 0 & \ddots & \vdots & \cdots \\ \vdots & \vdots & \cdots & \sum_{j=j_0}^m \frac{q_j - p_j}{\beta_j + N + 1} & \cdots \\ 0 & 0 & \cdots & 0 & \ddots \end{bmatrix}, \quad (3.59)$$

$\forall k \in \mathbb{N}$, $m = z(d) - 1$, $j_0 = m - d$, the elements of which are given by:

$$\nu_{k,i} = \chi_{[k]}(i) + \chi_{[0,d+k+1]}(i) \sum_{j=0}^{i-k-1} \frac{q_{s(i,j,k)} - p_{s(i,j,k)}}{j+k+1} - \chi_{[0,d]}(i) \sum_{j=i}^d \frac{q_{r(i,j)}}{j+k+1-i}, \quad (3.60)$$

with $\chi_{[a]}(i) = \{1, i = a; 0, i \neq a\}$ and $\chi_{[a,b]}(i) = \{1, i \in [a, b]; 0, i \notin [a, b]\}$ indicator functions, $s(i, j, k) = j + (i - k)(i - k - 1)/2$, $\forall i \geq k + 1$ and $r(i, j) = j(j + 3)/2 - i$, $\forall i \leq d$ sub-indexes of polynomial coefficients.

Thus, considering $u(x) = \sum_{k=0}^N a_k x^k$ as a finite series, due to the particular upper triangular structure of V (3.59), is clear that $\text{rank}(V) = N + 1 = \dim(\mathbf{a})$ (below the horizontal line of V , for $p_j \neq q_j, \forall j = j_0, \dots, m$, only one diagonal element could be zero, i.e, there is at least N independent rows. The extra row can be taken from above this line). In addition, it can be verified that $\nu_{d+k+1,d+k+1} = 1$, $\forall k = 0, \dots, N$, so that $\lim_{N \rightarrow \infty} \text{rank}(V) \rightarrow \dim(\mathbf{a})$. Therefore, the unique solution of (3.58) is the trivial one $\mathbf{a} = [0, 0, \dots]^\top$ [230], equivalent to $u(\cdot, t) = 0$, $\forall t \geq 0$. $\square$

Theorem 1. Let $\mathbb{T} = \mathbb{I} - \mathbb{F} : \mathcal{X} \rightarrow \mathcal{X}$ be the linear operator in (3.53), where $\mathcal{X} = \mathcal{A}_{\mathcal{B}} = (\mathcal{A} \subset \mathcal{C}(\Omega; \mathbb{R}), \|\cdot\|_{\infty})$ is the Banach space of real-analytic functions or $\mathcal{X} = \mathcal{A}_{\mathcal{H}} = (\mathcal{A} \subset \mathcal{L}_2(\Omega; \mathbb{R}), \langle \cdot, \cdot \rangle_{\mathcal{L}^2})$ is the Hilbert space of square integrable real-analytic functions. If $u = u(\cdot, t) \in \mathcal{X}$, real-analytic function $\forall t \geq 0$, and the Kernels P and Q are bounded polynomials, $P \neq Q$, then the integral equation $\mathbb{T}[u(\cdot, t)](x) = w(x, t)$ (3.53) has unique solution and the operator $\mathbb{T}$ is boundedly invertible in $\mathcal{X}$.

Proof. Let $\mathbb{F}$ be the linear operator with P and Q bivariate (bounded) polynomials of finite basis as in Lemma 2. Reordering terms, these polynomials can be alternatively expressed by

$$P(x, y) = \sum_{n=0}^{d_P} x^n \varphi_n(y), \quad \varphi_n(y) = \sum_{k=n}^{d_P} p_{s(k,n)} y^{k-n}, \quad (3.61)$$

$$Q(x, y) = \sum_{n=0}^{d_Q} x^n \psi_n(y), \quad \psi_n(y) = \sum_{k=n}^{d_Q} q_{s(k,n)} y^{k-n}, \quad (3.62)$$

where $s(k, n) = k(k + 3)/2 - n$. Taking $d = \max\{d_P, d_Q\}$ and assigning $p_{s(k, n)} = 0, \forall s(k, n) > z(d_P) - 1$ or $q_{s(k, n)} = 0, \forall s(k, n) > z(d_Q) - 1$ as required, the linear operator can be written as:

$$\begin{aligned}\mathbb{F}[u(\cdot, t)](x) &= \int_0^1 \sum_{n=0}^d \sum_{k=n}^d (\chi_{[0, x]}(y) p_{s(k, n)} + \chi_{[x, 1]}(y) q_{s(k, n)}) x^n y^{k-n} u(y, t) dy \\ &= \int_0^1 F(x, y) u(y, t) dy = \sum_{n=0}^d x^n \int_0^1 \bar{F}(n, y) u(y, t) dy,\end{aligned}\tag{3.63}$$

where $\chi_{[a, b]}(y) = \{1, y \in [a, b]; 0, y \notin [a, b]\}$ is the indicator function ($[a, b] \subset \mathbb{R}$) and $\bar{F} = \sum_{k=n}^d y^{k-n} (\chi_{[0, x]}(y) p_{s(k, n)} + \chi_{[x, 1]}(y) q_{s(k, n)})$. Since F is a bounded polynomial Kernel (a special class of degenerate Kernels [208]), it is immediate that $\mathbb{F}$ is a finite rank operator, as (3.63) shows.¹⁵ Its compactness can be proved in some Banach and Hilbert spaces. In the case of the Banach space of continuous functions, since F is bounded in $S = [0, 1]^2$ and continuous except possibly along the curve $x = y$ (also known as mildly discontinuous Kernel [232]), $\mathbb{F}$ is a compact operator [217, 231] (also denominated as completely continuous operator [222]). In the Hilbert space of square integrable functions, since $\mathbb{F}$ is a finite rank operator [192], it is a Hilbert-Schmidt operator ($\mathbb{F}[u](\cdot, t) = \sum_{j=0}^{d-1} \psi_j(y) \langle \phi_j(x), u(y, t) \rangle_{\mathcal{L}^2(\Omega)}$, for some $\{\phi_j\}, \{\psi_j\}$ finite orthonormal systems in $\mathcal{L}^2(\Omega)$) and therefore compact [216, 222, 231, 233]. Thus, based on Lemma 2 and on the property of compactness of $\mathbb{F}$, according to a particular feature of the Fredholm Alternative Theorem ([208, Corollary 3.5], [20, Corollary 7.27]),¹⁶ the solution of (3.53) is unique and the operator $\mathbb{T}$ is boundedly invertible in $\mathcal{A}_{\mathcal{B}}$ and $\mathcal{A}_{\mathcal{H}}$. $\square$

Remark 10. *The non-zero solution of the first-order hyperbolic PDE (3.54) is not real-analytic for any function $f_0 \in \mathcal{A}(\Omega)$ as initial condition, $\forall t \geq 0$ ($w(x, t) = \{f_0(x + t), t \leq 1 - x; 0, t > 1 - x\}$). However, Theorem 1 is still valid if the domain of analysis is restricted to $(x, t) \in \{(v, \tau) \in [0, 1] \times [0, 1] | \tau \geq v\}$.*

Remark 11. *It is worth studying how the Fredholm-type operator $\mathbb{F}$ (3.53) splits the standard polynomial basis in two particular linearly independent sets (LI) and, based on this fact, analyse an extension of Lemma 2 to the Banach space of continuous functions $\mathcal{X} = (\mathcal{C}(\Omega; \mathbb{R}), \|\cdot\|_{\infty})$.*

Lemma 3. *If $f \in \mathcal{B}(\Omega \setminus \{0\})$, $\mathcal{B} = (\mathcal{C} - \mathcal{A})$, is a continuous real non-analytic function then $F(x) = x^i \int_0^x y^j f(y) dy$ is a continuous real non-analytic function, $\forall x \in (\Omega \setminus \{0\})$, with $i \in \mathbb{N}$ and $j \in \mathbb{N}$ finite powers.*

Proof. By contradiction (contrapositive), assuming $F(x) = x^i \int_0^x y^j f(y) dy \in \mathcal{A}(\Omega \setminus \{0\})$ as a continuous real analytic function, since $\frac{dF}{dx}(x) \in \mathcal{A}(\Omega \setminus \{0\})$ (derivative of a continuous real analytic function is continuous real analytic [229]), by the second fundamental theorem of calculus [226]:

$$\mathcal{A}(\Omega \setminus \{0\}) \ni \frac{dF}{dx}(x) = ix^{i-1} \int_0^x y^j f(y) dy + x^{i+j} f(x),$$

¹⁵ $\mathbb{F}$ maps into a finite-dimensional space [208, 20, 231].

¹⁶The uniqueness of $\mathbb{F}u = u$ (trivial solution) implies the existence of the solution of $(\mathbb{I} - \mathbb{F})[u(\cdot, t)](x) = w(x, t)$.

$$\mathcal{A}(\Omega \setminus \{0\}) \ni \left(\underbrace{x^{-(i+j)}}_{A_1} \underbrace{\frac{dF}{dx}(x)}_{A_2} \right) = i \underbrace{\left(x^{-(i+j-1)} \right)}_{B_1} \underbrace{\left(x^i \int_0^x y^j f(y) dy \right)}_{B_2} + f(x). \quad (3.64)$$

Since the terms A_1 and B_1 are continuous real analytic functions in $\Omega \setminus \{0\}$ (functions with convergent Taylor's series and radius of convergence: $0 < x < 2x_0, \forall x_0 \in \Omega$), $A_2 \in \mathcal{A}(\Omega \setminus \{0\})$, $B_2 \in \mathcal{A}(\Omega \setminus \{0\})$ by hypothesis, and base on the property that the product of analytic functions is also an analytic function [229], it is inferred that $f \in \mathcal{A}(\Omega \setminus \{0\})$. $\square$

Lemma 4. *If $u = u(\cdot, t) \in \mathcal{C}(\Omega)$, $\forall t \geq 0$ in the integral equation (3.53), and the Kernels P and Q be bounded bivariate polynomials in $\Omega \times \Omega$, $P \neq Q$, then the homogeneous equation $(\mathbb{I} - \mathbb{F})[u(\cdot, t)](x) = 0$ has only the trivial solution $u = 0$.*

Proof. Consider $u = u(\cdot, t) \in \mathcal{B}(\Omega)$, $\forall t \geq 0$ with $\mathcal{B} = (\mathcal{C} - \mathcal{A})$ (the case $u \in \mathcal{A}(\Omega)$ has been proved in Lemma 2), with $\mathbb{F} : \mathcal{C}(\Omega; \mathbb{R}) \rightarrow \mathcal{C}(\Omega; \mathbb{R})$ as in (3.53). For polynomial Kernels $P \neq Q$ formulated according the standard bivariate basis (1.3), with polynomial degrees: $\deg(P) = d_P$, $\deg(Q) = d_Q$, $d = \max(d_P, d_Q)$, it is straightforward to verify that the homogeneous equation $(\mathbb{I} - \mathbb{F})[u(\cdot, t)](x) = 0$ can be written as¹⁷:

$$u(x) - \underbrace{\sum_{n=0}^d \psi_n(x) \int_0^x y^n u(y) dy}_{T_1} - \underbrace{\sum_{i=0}^{d_Q} \theta_i x^i}_{T_2} = 0, \quad (3.65)$$

with

$$\theta_i = \sum_{k=i}^{d_Q} q_{s(k,i)} \left(\int_0^1 y^{k-i} u(y) dy \right), \quad \psi_n(x) = \sum_{k=n}^d \underbrace{\left(q_{r(k,n)} - p_{r(k,n)} \right)}_{c_{r(k,n)}} x^{k-n}, \quad (3.66)$$

$$s(k, i) = k(k+3)/2 - i, \quad r(k, n) = n + k(k+1)/2.$$

As for the term T_1 , consider the following set of functions:

$$\begin{aligned} \Theta &= \left\{ \psi_0(x), \psi_1(x)y, \psi_2(x)y^2, \dots, \psi_n(x)y^n, \dots, \psi_{d-1}(x)y^{d-1}, \psi_d(x)y^d \right\}, \\ &= \left\{ \left(c_0 + c_1x + c_3x^2 + \dots + c_{r(d,0)}x^d \right), \left(c_2 + c_4x + c_7x^2 + \dots + c_{r(d,1)}x^{d-1} \right) y, \right. \\ &\quad \left(c_5 + c_8x + c_{12}x^2 + \dots + c_{r(d,2)}x^{d-2} \right) y^2, \dots, \left(\sum_{k=n}^d c_{r(k,n)}x^{k-n} \right) y^n, \dots, \\ &\quad \left. \left(c_{r(d-1,d-1)} + c_{r(d,d-1)}x \right) y^{d-1}, c_{r(d,d)}y^d \right\}. \end{aligned} \quad (3.67)$$

¹⁷For notational simplicity the time-dependence in the functions is dropped.

Since Θ is a linearly independent (LI) set of functions (bivariate polynomials of different degrees),

$$\bar{\Theta} = \Theta \circ u(y) = \left\{ \psi_0(x), \dots, \psi_n(x)y^n, \dots, \psi_d(x)y^d \right\} \circ u(y), \quad (3.68)$$

is also a LI set of functions $\forall u \in \mathcal{C}(\Omega)$, $u(x, t) \neq 0$, $\forall x \in \Omega$, $t \geq 0$, where $\circ$ denotes the Hadamard product (element-wise).

Let $\mathbb{V} : \mathcal{C}(\Omega; \mathbb{R}) \rightarrow \mathcal{C}(\Omega; \mathbb{R})$ be the standard Volterra operator given by: $\mathbb{V}[f(\cdot, t)](x) = \int_0^x f(y, t)dy$, $f \in \mathcal{C}(\Omega)$. Due to the Volterra operator $\mathbb{V}$ is linear and injective [235], this maps (3.68) to

$$\Gamma_1 = \mathbb{V}(\bar{\Theta}) = \left\{ \left(c_0 + c_1x + \dots + c_{r(d,0)}x^d \right) \int_0^x u(y)dy, \dots, \left(\sum_{k=n}^d c_{r(k,n)}x^{k-n} \right) \int_0^x y^n u(y)dy, \dots, c_{r(d,d)} \int_0^x y^d u(y)dy \right\}, \quad (3.69)$$

which also is a set of LI functions (a linear injective operator preserves linear independence)[236]. In addition, based on Lemma 3, restricting the domain Ω to $\bar{\Omega} = \Omega \setminus \{0\}$, Γ_1 is a set of continuous real non-analytic functions.

Thus, since the term T_2 in (3.66) is a linear combination of elements of the standard LI univariate polynomial basis $\Gamma_2 = \{1, x, x^2, \dots, x^{d_Q}\}$, and this basis cannot span to the space $\mathcal{B}(\bar{\Omega})$, it is inferred that $\Gamma_1 \cup \Gamma_2$ is a set of LI functions, $\forall x \in \bar{\Omega}$.

Moreover, it is immediate to verify that if $u \in \mathcal{B}(\bar{\Omega})$, u is not a linear combination of the elements of Γ_2 . Similarly, regarding Γ_1 , for $\alpha_1, \alpha_2 \in \mathbb{R}$, $\alpha_1 u(x) + \alpha_2 x^j \int_0^x y^n u(y)dy = 0 \Leftrightarrow \alpha_1 = \alpha_2 = 0$ or $u = 0$, $\forall x \in \bar{\Omega}$, since it is a homogeneous Volterra integral equation of the second kind (this can also be obtained including u_y to Θ and then applying $\mathbb{V}$. However it requires $u \in \mathcal{C}^1$). Therefore, $\Gamma = \{u(x)\} \cup \Gamma_1 \cup \Gamma_2$ is a set of LI functions in $\bar{\Omega}$.

Finally, since $u = 0$ is solution of (3.65) for $x = 0$ ($\theta_0 = 0$), and (3.65) is a linear combination of LI functions in Γ , $\forall x \in \bar{\Omega}$, it is concluded that $(\mathbb{I} - \mathbb{F})[u(\cdot, t)](x) = 0$ has only the trivial solution $u = u(\cdot, t) = 0$, $\forall x \in \Omega$, $\forall t \geq 0$. $\square$

Remark 12. *It is worth noting that the mutually linear independence between elements of Γ_1 and Γ_2 . Taking the Wronskian of arbitrary elements of these sets yields*

$$W = \begin{vmatrix} x^i & x^j \int_0^x y^n u(y)dy \\ ix^{i-1} & jx^{j-1} \int_0^x y^n u(y)dy + x^{j+n}u(x) \end{vmatrix} = x^{i+j-1} \left((j-i) \int_0^x y^n u(y)dy + x^{n+1}u(x) \right) \quad (3.70)$$

$$= x^{i+j+n} \left((j-i)(1/x) \int_{\tau(x,n)}^x u(y)dy + u(x) \right) \quad (3.71)$$

$$= 0 \Leftrightarrow u(x) = 0, \quad \forall x \in \bar{\Omega}, \quad (3.72)$$

with $0 < \tau(x, n) < x \leq 1$, $\forall i = 0, \dots, d_Q$, $\forall j = 0, \dots, d$ and $n \in [0, d]$, where the equivalence from (3.70) to (3.71) relies on the Bonnet's mean value theorem based on the fact that y^n is a nonnegative increasing function for $n \geq 1$ [237], namely

$$\exists \text{ at least one } \tau \in (0, x) \text{ such that: } \int_0^x y^n v(y) dy = x^n \int_\tau^x v(y) dy,$$

and (3.72) due to (3.71) involves an homogeneous Volterra integral equation of the second kind [190].

However “linearly independent functions may have an identically zero Wronskian”. That is not the case for analytic functions (see Theorem 1 in [238]).

Theorem 2. Let $\mathbb{T} = \mathbb{I} - \mathbb{F} : \mathcal{X} \rightarrow \mathcal{X}$ be the integral linear operator in (3.53), where $\mathcal{X} = (\mathcal{C}(\Omega; \mathbb{R}), \|\cdot\|_\infty)$ is the Banach space of continuous functions. If the Kernels P and Q are bounded polynomials, $P \neq Q$, then the integral equation $w(x, t) = \mathbb{T}[u(\cdot, t)](x)$ (3.53) has a unique solution and the operator $\mathbb{T}$ is boundedly invertible in $\mathcal{X}$.

Proof. [The proof follows the same arguments of the one of Theorem 1]. Let $\mathbb{F}$ be as in (3.53), this can be written as:

$$\mathbb{F}[u(\cdot, t)](x) = \int_0^1 \underbrace{\left(\sum_{n=0}^d x^n \sum_{k=n}^d (\chi_{[0,x]}(y) p_{s(k,n)} + \chi_{[x,1]}(y) q_{s(k,n)}) y^{k-n} \right)}_F u(y, t) dy. \quad (3.73)$$

In the Banach space of continuous functions, since F is bounded in $\Omega \times \Omega$ and continuous except possibly along the curve $x = y$, $\mathbb{F}$ is a compact operator [217, 231]. Thus, based on Lemma 2 for continuous real analytic functions and Lemma 4 for continuous non-analytic real functions, according to a particular feature of the Fredholm Alternative Theorem [208, Corollary 3.5], the solution of $w(x, t) = \mathbb{T}[u(\cdot, t)](x)$ is unique and the operator $\mathbb{T}$ is boundedly invertible in $\mathcal{X}$. $\square$

3.3.3 Kernel-PIDE as a Convex Optimization Problem

Proposition 5. Let

$$u(x, t) = w(x, t) + \int_0^x R(x, y) w(y, t) dy + \int_x^1 S(x, y) w(y, t) dy \quad (3.74)$$

be the inverse transformation of (3.53) in terms of the Kernels R and S (as it is proposed in [205, 204, 82] under specific conditions on the system (3.52)). Let $\Delta_1 = \int_0^1 \int_0^x \delta_1^2(x, y) dy dx$ and $\Delta_2 = \int_0^1 \int_x^1 \delta_2^2(x, y) dy dx$ be the mean square of the residual functions detailed in (3.55). Considering $\delta_0(x) = 0$ and $\delta_3(x) = 0$ in (3.56), the transformed system (3.55) is exponentially stable if the residual functions satisfy:

$$\sqrt{\Delta_1} + \sqrt{\Delta_2} \leq \frac{e^{-1}}{(1 + \sqrt{\sigma_1} + \sqrt{\sigma_2})}, \quad (3.75)$$

where $\sigma_1 = \int_0^1 \int_0^x R^2(x, y) dy dx$ and $\sigma_2 = \int_0^1 \int_x^1 S^2(x, y) dy dx$.

Proof. Let $V = \frac{1}{2} \int_0^1 e^{\alpha x} w^2(x, t) dx$ be a Lyapunov functional for some $\alpha > 0$. Its time-derivative $\dot{V} = \int_0^1 e^{\alpha x} w(x) w_t(x) dx$ along the trajectory (3.55), with $\delta_0(x) = 0$ and $\delta_3(x) = 0$, is given by:

$$\dot{V} \leq \underbrace{\int_0^1 e^{\alpha x} w(x) w_x(x) dx}_{T_1} + \underbrace{\int_0^1 e^{\alpha x} w(x) \int_0^x u(y) \delta_1(x, y) dy dx}_{T_2} + \underbrace{\int_0^1 e^{\alpha x} w(x) \int_x^1 u(y) \delta_2(x, y) dy dx}_{T_3}. \quad (3.76)$$

Using integration by parts and the boundary condition in (3.54) on the term T_1 yields:

$$T_1 \leq e^{\alpha x} w^2(x) \Big|_{x=0}^{x=1} - \alpha/2 \int_0^1 e^{\alpha x} w^2(x) dx \leq -\alpha V(t). \quad (3.77)$$

Regarding the term T_2 , changing the order of integration and plugging in $u = u(y, t)$ the expression given by the inverse transformation (3.74), this can be written as:

$$\begin{aligned} T_2 = & \underbrace{\int_0^1 w(y) \int_y^1 e^{\alpha x} w(x) \delta_1(x, y) dx dy}_{T_{2a}} + \underbrace{\int_0^1 \left(\int_0^y R(y, s) w(s) ds \right) \left(\int_y^1 e^{\alpha x} w(x) \delta_1(x, y) dx \right) dy}_{T_{2b}} \\ & + \underbrace{\int_0^1 \left(\int_y^1 S(y, s) w(s) ds \right) \left(\int_y^1 e^{\alpha x} w(x) \delta_1(x, y) dx \right) dy}_{T_{2c}}. \end{aligned} \quad (3.78)$$

Using the Cauchy-Schwarz integral inequality [239], yields:

$$\begin{aligned} T_{2a}^2 & \leq \left(\int_0^1 e^{2\alpha x} w^2(x) dx \right) \left(\int_0^1 \left(\int_0^x w(y) \delta_1(x, y) dy \right)^2 dx \right) \\ & \leq \max_{x \in \Omega} \{e^{\alpha x}\} \max_{y \in \Omega} \{e^{-\alpha y}\} 4V^2(t) \cdot \left(\int_0^1 \left(\int_0^x e^{\alpha y} w^2(y) dy \right) \left(\int_0^x \delta_1^2(x, y) dy \right) dx \right) \Rightarrow \\ T_{2a} & \leq 2e^{\alpha/2} V(t) \sqrt{\Delta_1}, \end{aligned} \quad (3.79)$$

$$\begin{aligned} T_{2b}^2 & \leq \left(\int_0^1 \left(\int_0^y R(y, s) w(s) ds \right)^2 dy \right) \left(\int_0^1 \left(\int_y^1 e^{\alpha x} w(x) \delta_1(x, y) dx \right)^2 dy \right) \\ & \leq \max_{x \in \Omega} \{e^{\alpha x}\} \max_{s \in \Omega} \{e^{-\alpha s}\} \left(\int_0^1 \left(\int_0^y R^2(y, s) ds \right) \cdot \left(\int_0^y e^{\alpha s} w^2(s) ds \right) dy \right) \cdot \\ & \quad \left(\int_0^1 \left(\int_y^1 e^{\alpha x} w^2(x) dx \right) \cdot \left(\int_y^1 \delta_1^2(x, y) dx \right) dy \right) \Rightarrow \\ T_{2b} & \leq 2e^{\alpha/2} V(t) \sqrt{\Delta_1} \left(\int_0^1 \int_0^y R^2(y, s) ds dy \right)^{\frac{1}{2}}, \end{aligned} \quad (3.80)$$

and

$$T_{2c}^2 \leq \left(\int_0^1 \left(\int_y^1 S(y, s) w(s) ds \right)^2 dy \right) \left(\int_0^1 \left(\int_y^1 e^{\alpha x} w(x) \delta_1(x, y) dx \right)^2 dy \right)$$

$$\begin{aligned} &\leq 4e^{\alpha}V^2(t) \left(\int_0^1 \left(\int_y^1 S^2(y,s)ds \right) dy \right) \left(\int_0^1 \left(\int_y^1 \delta_1^2(x,y)dx \right) dy \right) \Rightarrow \\ T_{2c} &\leq 2e^{\alpha/2}V(t)\sqrt{\Delta_1} \left(\int_0^1 \int_y^1 S^2(y,s)dsdy \right)^{\frac{1}{2}}, \end{aligned} \quad (3.81)$$

where $\Delta_1 = \int_0^1 \int_0^x \delta_1^2(x,y)dydx$. With respect to the term T_3 , this can be written as:

$$\begin{aligned} T_3 &= \underbrace{\int_0^1 w(y) \int_0^y e^{\alpha x} w(x) \delta_2(x,y) dx dy}_{T_{3a}} + \underbrace{\int_0^1 \left(\int_0^y R(y,s) w(s) ds \right) \left(\int_0^y e^{\alpha x} w(x) \delta_2(x,y) dx \right) dy}_{T_{3b}} + \\ &\quad \underbrace{\int_0^1 \left(\int_y^1 S(y,s) w(s) ds \right) \left(\int_0^y e^{\alpha x} w(x) \delta_2(x,y) dx \right) dy}_{T_{3c}}, \end{aligned} \quad (3.82)$$

where upper bounds for every term, in relation with $\Delta_2 = \int_0^1 \int_x^1 \delta_2^2(x,y)dydx$, can be found following the same procedure described above in (3.79)-(3.81). Finally, using in (3.76) the upper bounds for the terms T_1 , T_2 and T_3 , the condition $\dot{V} \leq 0$ is satisfied provided:

$$\alpha - 2e^{\alpha/2} \left(\sqrt{\Delta_1} + \sqrt{\Delta_2} \right) (1 + \sqrt{\sigma_1} + \sqrt{\sigma_2}) \geq 0,$$

with $\sigma_1 = \int_0^1 \int_0^x R^2(x,y)dydx$ and $\sigma_2 = \int_0^1 \int_x^1 S^2(x,y)dydx$. Since the factor $(1/2)\alpha e^{-\alpha/2}$ reaches the maximum value of e^{-1} at $\alpha = 2$, the expression (3.75) is obtained, i.e. a sufficient condition for exponential stability of (3.55). $\square$

Based on this result, similar to Proposition 2, a relaxation on the zero matching condition for the residual functions δ_1 and δ_2 can be considered and the Kernel-PIDE (3.56) can be solved approximately in terms of polynomial Kernels.

Proposition 6. *Let $N(x,y) = \sum_{k=0}^{z(d)-1} n_k x^{\alpha_k} y^{\beta_k}$ and $M(x,y) = \sum_{k=0}^{z(d)-1} m_k x^{\alpha_k} y^{\beta_k}$ be polynomial approximations of P and Q , respectively, of arbitrary even degree $d \in \mathbb{N}$, with coefficients n_k and m_k and powers in accordance with (1.3). Let $Y = Y(x,y)$, $Z = Z(x,y)$ be polynomials of degree $d_\delta = 2 \left\lceil \frac{\max\{d+d_{h_1}, d+d_{h_2}\}+1}{2} \right\rceil$ and $\gamma_1 \geq 0$, $\gamma_2 \geq 0$. For any functions f, h_1, h_2 described for polynomials of degree d_f, d_{h_1}, d_{h_2} , respectively, the Kernel-PIDE (3.56) can be formulated as the convex optimization problem:*

$$\underset{\gamma_j, N, M, T_1, T_2, Y, Z, s_j}{\text{minimize:}} \quad \gamma_1 + \gamma_2 \quad (3.83)$$

$$\text{subject to:} \quad Y(x,y) - \delta_1(x,y) = 0, \quad Z(x,y) - \delta_2(x,y) = 0, \quad (3.84)$$

$$\begin{bmatrix} 2\gamma_1 - T_1(x,y) - s_1(x,y)g_1(x) & Y(x,y) \\ Y(x,y) & \gamma_1 - s_2(x,y)g_2(x,y) \end{bmatrix} \in \Sigma_s^{2 \times 2}, \quad (3.85)$$

$$\begin{bmatrix} 2\gamma_2 - T_2(x, y) - s_3(x, y)g_1(x) & Z(x, y) \\ Z(x, y) & \gamma_2 - s_4(x, y)g_3(x, y) \end{bmatrix} \in \Sigma_s^{2 \times 2}, \quad (3.86)$$

$$s_1, s_2, s_3, s_4 \in \Sigma_s, \quad (3.87)$$

$$\int_0^1 \int_0^x T_1(x, y) dy dx = 0, \quad \int_0^1 \int_x^1 T_2(x, y) dy dx = 0 \quad (3.88)$$

$$\left. \begin{aligned} \delta_1 &= \delta_1(x, y) \Big|_{Q \approx M}^{P \approx N}, \\ \delta_2 &= \delta_2(x, y) \Big|_{Q \approx M}^{P \approx N}, \\ \delta_0 &= \delta_0(x, y) \Big|_{Q \approx M}^{P \approx N} = 0, \\ \delta_3 &= \delta_3(x) \Big|_{Q \approx M} = M(x, 1) = 0, \end{aligned} \right\} \text{ as in (3.55),} \quad (3.89)$$

for some polynomials $s_j, j = 1, \dots, 4$ of degree $d_\delta - 2$, T_1 and T_2 of degree $2d_\delta$, $g_1(x) = x(1 - x)$, $g_2(x, y) = y(x - y)$ and $g_3(x, y) = (1 - y)(y - x)$. The optimal root mean square bounds of the residual functions are: $\sqrt{\Delta_1} \leq \gamma_1$ and $\sqrt{\Delta_2} \leq \gamma_2$.

Proof. The convex optimization problem formulation follows similar arguments as the ones given in the proof of Proposition 2. Regarding the optimal mean square bounds for δ_1 and δ_2 , let

$$A_1 = \begin{bmatrix} 2\gamma_1 - T_1(x, y) & Y(x, y) \\ Y(x, y) & \gamma_1 \end{bmatrix} \succ 0, \forall (x, y) \in \Omega_L \quad (3.90)$$

be a symmetric real polynomial positive definite matrix on Ω_L (pointwise condition). Taking the Schur's complement of A_1 and its integration on the domain Ω_L yields

$$\begin{aligned} A_1 \succ 0 &\Leftrightarrow 2\gamma_1^2 - \gamma_1 T_1(x, y) - Y^2(x, y) > 0, \quad \forall (x, y) \in \Omega_L, \\ \int_0^1 \int_0^x Y^2(x, y) dy dx &< \gamma_1^2 - \gamma_1 \int_0^1 \int_0^x T_1(x, y) dy dx. \end{aligned} \quad (3.91)$$

If there exists a polynomial function T_1 satisfying (3.88), it is clear that γ_1 is an upper bound of the root mean square value of Y in Ω_L . The positivity condition (3.90) is made computationally tractable via the matrix-polynomial version of Putinar's Positivstellensatz¹⁸[138, 157], based on the representation of Ω_L as in Proposition 1, namely, for some $\rho > 0$, $A(x, y) \succ \rho I \succ 0$, $\forall (x, y) \in \Omega_L$, then:

$$\begin{aligned} \left(A(x, y) - g_1(x)S_1(x, y) - g_2(x, y)S_2(x, y) \right) &\in \Sigma_s^{2 \times 2} \\ S_1, S_2 &\in \Sigma_s^{2 \times 2}. \end{aligned} \quad (3.92)$$

Thus, conditions (3.85) and (3.86) are obtained if the following particular forms for S_1 and S_2 are

¹⁸This condition can also be "scalarized", i.e., expressed in terms of scalar polynomials [136, 138, 157].

considered:

$$\begin{aligned} S_1 &= \begin{bmatrix} s_1(x, y) & 0 \\ 0 & 0 \end{bmatrix} \in \Sigma_s^{2 \times 2}, \quad s_1 \in \Sigma_s, \\ S_2 &= \begin{bmatrix} 0 & 0 \\ 0 & s_2(x, y) \end{bmatrix} \in \Sigma_s^{2 \times 2}, \quad s_2 \in \Sigma_s, \end{aligned} \quad (3.93)$$

the SOS matrix condition of which is immediately verified since $S_1 = BB^T$, $S_2 = CC^T$ with:

$$\begin{aligned} B &= \begin{bmatrix} b_1(x, y) & \cdots & b_{m_1}(x, y) & 0 \\ 0 & \cdots & 0 & 0 \end{bmatrix} \in \mathbb{R}[\mathbf{x}]^{2 \times (m_1+1)}, \\ C &= \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & c_1(x, y) & \cdots & c_{m_2}(x, y) \end{bmatrix} \in \mathbb{R}[\mathbf{x}]^{2 \times (m_2+1)}, \end{aligned} \quad (3.94)$$

where $s_1 = \sum_{j=1}^{m_1} b_j^2(x, y) \in \Sigma_s$ and $s_2 = \sum_{j=1}^{m_2} c_j^2(x, y) \in \Sigma_s$; $b_j \in \mathbb{R}[\mathbf{x}]$, $\forall j = 1, \dots, m_1$, $c_j \in \mathbb{R}[\mathbf{x}]$, $\forall j = 1, \dots, m_2$, for some finite $m_1 \in \mathbb{N}$ and $m_2 \in \mathbb{N}$. Following the same arguments, conditions for Z on Ω_U can be deduced. Therefore, according to the equality (3.84) and the optimization objective (3.83), the root mean square error of δ_1 and δ_2 are minimized. $\square$

3.3.4 Approximate Inverse Transformation

For known Kernels P and Q , the inverse transformation of (3.53) can be found by means of the direct substitution of (3.74) in (3.53), which yields

$$\int_0^x w(y, t) \delta_1(x, y) dy + \int_x^1 w(y, t) \delta_2(x, y) dy = 0, \quad (3.95)$$

equality satisfied if the residual functions

$$\begin{aligned} \delta_1(x, y) &= R(x, y) - P(x, y) - \int_0^y P(x, s) S(s, y) ds - \int_y^x P(x, s) R(s, y) ds - \\ &\quad \int_x^1 Q(x, s) R(s, y) dy, \quad \forall (x, y) \in \Omega_L, \end{aligned} \quad (3.96)$$

$$\begin{aligned} \delta_2(x, y) &= S(x, y) - Q(x, y) - \int_0^x P(x, s) S(s, y) ds - \int_x^y Q(x, s) S(s, y) ds - \\ &\quad \int_y^1 Q(x, s) R(s, y) ds, \quad \forall (x, y) \in \Omega_U. \end{aligned} \quad (3.97)$$

are identically zero in their respective triangular domains. Since (3.95) does not depend on any original and target systems (expression equivalent to (3.29) for the Volterra operator), it can be used to find an approximation of the inverse Kernels R and S , given the approximate direct ones $P \approx N$, $Q \approx M$.

Proposition 7. Let $A(x, y) = \sum_{k=0}^{z(d)-1} a_k x^{\alpha_k} y^{\beta_k}$ and $B(x, y) = \sum_{k=0}^{z(d)-1} b_k x^{\alpha_k} y^{\beta_k}$ be polynomial approximations of R and S , respectively, of arbitrary even degree $d \in \mathbb{N}$, with coefficients a_k and b_k and powers in accordance with (1.3). Let Y and Z be polynomials of degree $d_\delta = 2 \left\lceil \frac{\max\{d_N, d_M\} + d + 1}{2} \right\rceil$ and $\gamma_j \geq 0, j = 1, \dots, 4$. For given direct approximate Kernels N and M of P and Q with degrees d_N and d_M , respectively (solution of (3.83)-(3.89)), the integral equation (3.95)-(3.97) can be formulated as the convex optimization problem:

$$\begin{array}{ll} \text{minimize:} & \gamma_1 + \gamma_2 + \gamma_3 + \gamma_4 \\ \gamma_j, A, B, Y, Z, s_j & \end{array} \quad (3.98)$$

$$\text{subject to:} \quad Y(x, y) - \delta_1(x, y) = 0, \quad Z(x, y) - \delta_2(x, y) = 0, \quad (3.99)$$

$$\left(Y(x, y) - \underline{\rho}_Y - [s_1(x, y) \ s_2(x, y)] \ g_L(x, y) \right) \in \Sigma_s, \quad (3.100)$$

$$(\bar{\rho}_Y - Y(x, y) - [s_3(x, y) \ s_4(x, y)] \ g_L(x, y)) \in \Sigma_s, \quad (3.101)$$

$$\left(Z(x, y) - \underline{\rho}_Z - [s_5(x, y) \ s_6(x, y)] \ g_U(x, y) \right) \in \Sigma_s, \quad (3.102)$$

$$(\bar{\rho}_Z - Z(x, y) - [s_7(x, y) \ s_8(x, y)] \ g_U(x, y)) \in \Sigma_s, \quad (3.103)$$

$$s_1, s_2, s_3, s_4, s_5, s_6, s_7, s_8 \in \Sigma_s, \quad (3.104)$$

$$\begin{bmatrix} \gamma_1 & \underline{\rho}_Y \\ \underline{\rho}_Y & \gamma_1 \end{bmatrix} \succeq 0, \quad \begin{bmatrix} \gamma_2 & \bar{\rho}_Y \\ \bar{\rho}_Y & \gamma_2 \end{bmatrix} \succeq 0, \quad \begin{bmatrix} \gamma_3 & \underline{\rho}_Z \\ \underline{\rho}_Z & \gamma_3 \end{bmatrix} \succeq 0, \quad \begin{bmatrix} \gamma_4 & \bar{\rho}_Z \\ \bar{\rho}_Z & \gamma_4 \end{bmatrix} \succeq 0, \quad (3.105)$$

$$\delta_1 = \delta_1(x, y) \Big|_{\substack{P \approx N, Q \approx M \\ R \approx A, S \approx B}} \text{ as in (3.96),} \quad (3.106)$$

$$\delta_2 = \delta_2(x, y) \Big|_{\substack{P \approx N, Q \approx M \\ R \approx A, S \approx B}} \text{ as in (3.97),} \quad (3.107)$$

for some polynomials $s_j, j = 1, \dots, 8$ of degree $d_\delta - 2$, $g_L = [g_1, g_2]$, $g_U = [g_1, g_3]$, with $g_1(x) = x(1 - x)$, $g_2(x, y) = y(x - y)$ and $g_3(x, y) = (1 - y)(y - x)$. The optimal minimal bounds for the residual functions are: $\bar{\delta}_1 = \max\{\gamma_1, \gamma_2\}$ and $\bar{\delta}_2 = \max\{\gamma_3, \gamma_4\}$.

Proof. The proof follows the same arguments of the one of Proposition 3, hence omitted. $\square$

3.4 Numerical Results¹⁹

3.4.1 Parabolic PDE with constant reactivity term

This example considers $\lambda(x) = 20$ and $\epsilon = 1$ in the system (3.4) and $c = 0$ in target system (3.6). The direct Kernel K in (3.5) is approximated solving the convex optimization problem (3.18)-(3.28). The bounds of the residual functions δ_1 and δ_2 are indicated in Figure 3.4.1. The approximate Kernel N for $d = 12$ is depicted in Figure 3.4.2(a).

Regarding the inverse transformation (3.9), using the previous direct approximate Kernel, a moment

¹⁹The numerical solution of the convex optimization problems has been obtained via the Yalmip toolbox for Matlab [163] and SOSTOOLS [164], using the semi-definite programming solver SeDuMi [165] and the SDP package part of the Mosek solver [166].

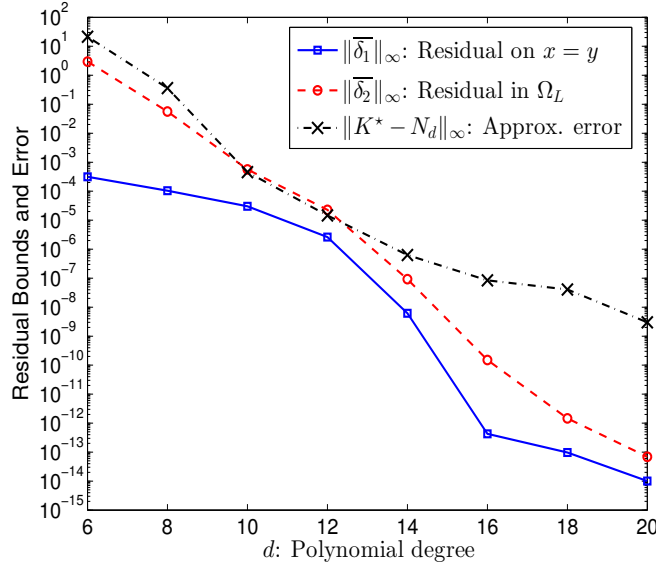


Figure 3.4.1: Bounds for the residual functions in (3.7), solution of (3.18)-(3.28), as a function of the polynomial degree d .

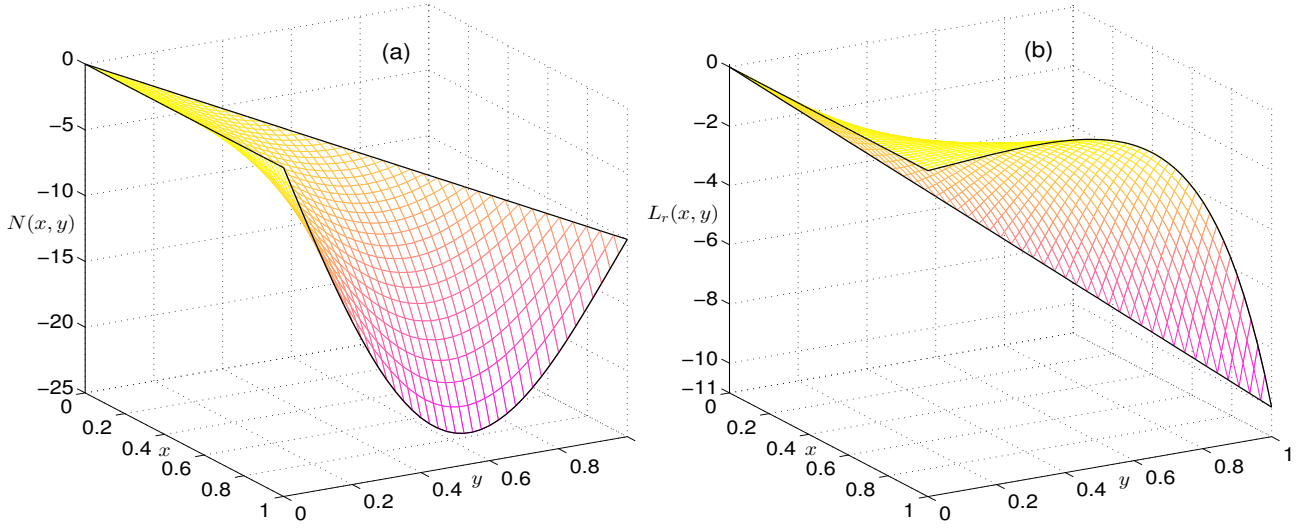


Figure 3.4.2: (a) Direct approximate Kernel N for $d = 12$ ($K \approx N$ in (3.5)). (b) Inverse approximate Kernel L_r for $r = 13$ ($L \approx L_r$ in (3.9)).

$m_{i,j}$	$m_{0,0}$	$m_{0,1}$	$m_{0,2}$	$m_{0,3}$	$m_{0,4}$	$m_{1,0}$	$m_{1,1}$	$m_{1,2}$	$m_{1,3}$	$m_{1,4}$	σ
$\check{L}^*$	4.1903	1.2286	0.5517	0.3011	0.1843	2.7834	0.9177	0.4394	0.2497	0.1572	10.0000
$\check{L}$	4.1940	1.2298	0.5523	0.3015	0.1846	2.7859	0.9187	0.4399	0.2500	0.1574	10.0076

Table 3.1: Extract of the moment sequence (3.11), optimal solution of (3.30)-(3.44), according to $r = 0, \dots, 13$.

sequence (3.11) has been calculated solving the convex optimization problem (3.30)-(3.44) for $d_m = 13$. Table 3.1 shows a small part of this sequence, which is compared with the inverse Kernel L^* obtained via the Successive Approximation method, the positive component of which is given by $\check{L}^*(x, y) = -\bar{\lambda}y \frac{J_1(\sqrt{\bar{\lambda}(x^2-y^2)})}{\sqrt{\bar{\lambda}(x^2-y^2)}} + 10$, where J_1 is the first order Bessel function [50]. The whole moment sequence has two digit of precision with respect to $\check{L}^*$. Finally, based on the whole previous moment sequence, an approximate inverse Kernel $L_r = \check{L}_r - \sigma$ has been determined solving (3.47)-(3.50) for $r = 13$, the result of which is shown in Figure 3.4.2(b).

3.4.2 Parabolic PDE with spatially varying reactivity term

This example considers two cases of spatially varying reactivity in the system (3.4):

$$\text{Case 1: } \lambda(x) = 10 + \frac{8}{\cosh^2(\sigma(x - 0.5))}, \quad (3.108)$$

$$\begin{aligned} \text{Case 2: } a(x) &= 1 + 0.3 \sin(3\pi x), \\ \lambda(x) &= 15 + \frac{1}{4} \frac{d^2 a}{dx^2}(x) - \frac{3}{16} \frac{da}{dx}(x). \end{aligned} \quad (3.109)$$

The first case is a particular example of the “one-peak” functions described in [50], where its associated Kernel-PDE has a closed-form solution. On the contrary, the second case requires a numerical solution of its Kernel-PDE [21]. For both cases the direct Kernel K in (3.5) has been approximated solving the convex optimization problem (3.18)-(3.28). In particular, for the second case, the variant described in Chapter 2, Section 2.1.4 has been applied. The reactivity terms have been approximated by a polynomial $\lambda(x) \approx \sum_{k=0}^r c_k x^k$.

B.1) Case 1: High Order Polynomials

For the first case a degree of $r = 10$ for the reactivity term has been selected, with a maximum approximation error of 0.11%. Its results are indicated in Table 3.2, which shows the high Kernel polynomial degrees required to achieve small bounds for the residual functions $\bar{\delta} = \max\{\bar{\delta}_1, \bar{\delta}_2\}$.

d	18	20	22	24
$\bar{\delta}$	$1.1 \cdot 10^{-3}$	$4.63 \cdot 10^{-4}$	$1.06 \cdot 10^{-4}$	$5.52 \cdot 10^{-5}$

Table 3.2: Maximum bounds of residual functions for spatially varying reactivity term (3.108).

B.2) Case 2: Multiple Kernel Approach

For the second case a polynomial approximation for the reactivity term, with parameter bounds $|c_k| < 10^6$, requires a degree $r > 20$. These conditions, (i) large number of Kernel parameters ($z(r) > \binom{2+r+d}{r+d}$ with d degree of K) and (ii) large parameter magnitudes, make unreasonable and unsuccessful the application of the approach proposed with a single Kernel. However, the

property of representation of non-negative polynomials in compact basic semi-algebraic sets leads to a *Multiple Kernel Approach* via a “domain decomposition” of the whole domain in small tractable elements, as it has been proposed in Section 2.1.4.

Consider the lower 2-dimensional triangular domain Ω_L (1-dimensional case is straightforward). Let $\bar{\Omega}_x = [x_1, \dots, x_j, \dots, x_{m+1}]$ and $\bar{\Omega}_y = [y_1, \dots, y_i, \dots, y_{m+1}]$ be partitions in m intervals for Ω in relation with the variable x and y , respectively. For $N_e = m(m+1)/2$ elements $\Omega_{i,j} = [x_j, x_{j+1}] \times [y_i, y_{i+1}] \subset \Omega_L$, $\Omega_L = \bigcup_{i,j} \Omega_{i,j}$, $j = 1, \dots, m$, $i = 1, \dots, j$, the residual function $\delta = \delta(x, y)$ can be minimized by the convex optimization problem:

$$\begin{aligned}
& \underset{\gamma_{i,j} \geq 0, \rho_{i,j} \geq 0, \varrho_{i,j} \geq 0}{\text{minimize:}} && \sum_{j=1}^m \sum_{i=1}^j \gamma_{i,j} + \rho_{i,j} + \varrho_{i,j} \\
& \text{subject to:} && (\delta_{i,j}(x, y) - \bar{s}_{i,j}^1(x, y) E_{i,j}(x, y)) \in \Sigma_s, \\
& && (\gamma_{i,j} - \delta_{i,j}(x, y) - \bar{s}_{i,j}^2(x, y) E_{i,j}(x, y)) \in \Sigma_s, \\
& && (F_{i,j}^k(x) - \tau_{i,j}^1(x) E_{i,j}^1(x)) \in \Sigma_s, \\
& && (\rho_{i,j} - F_{i,j}^k(x) - \tau_{i,j}^2(x) E_{i,j}^1(x)) \in \Sigma_s, \\
& && (G_{i,j}^k(y) - v_{i,j}^1(y) E_{i,j}^2(y)) \in \Sigma_s, \\
& && (\varrho_{i,j} - G_{i,j}^k(y) - v_{i,j}^2(y) E_{i,j}^2(y)) \in \Sigma_s,
\end{aligned} \tag{3.110}$$

with

$$\begin{aligned}
F_{i,j}^k(x) &= \mathbb{Y}_k(N_{i,j}(x, y) - N_{i-1,j}(x, y))|_{y=y_i}, \\
G_{i,j}^k(y) &= \mathbb{X}_k(N_{i,j}(x, y) - N_{i,j-1}(x, y))|_{x=x_j}, \\
\mathbb{Y}_k &= \frac{\partial^k}{\partial y^k}, \quad \mathbb{X}_k = \frac{\partial^k}{\partial x^k}, \quad k = 0, 1, \dots, r_s, \\
\bar{s}_{i,j}^n &= [s_{i,j}^{k,1}(x, y), s_{i,j}^{k,2}(x, y)], \quad s_{i,j}^{n,1} \in \Sigma_s, \quad s_{i,j}^{n,2} \in \Sigma_s, \quad \forall n = 1, 2, \\
\tau_{i,j}^m &\in \Sigma_s, \quad v_{i,j}^m \in \Sigma_s, \quad \forall m = 1, 2, \\
E_{i,j}(x, y) &= [E_{i,j}^1(x), E_{i,j}^2(y)]^\top, \\
E_{i,j}^1(x) &= (x - x_j)(x_{j+1} - x), \quad E_{i,j}^2(y) = (y - y_i)(y_{i+1} - y),
\end{aligned} \tag{3.111}$$

$\forall x \in [x_j, x_{j+1}]$, $\forall y \in [y_i, y_{i+1}]$, where $\delta(x, y) = \delta_{i,j}(x, y)$, $\forall (x, y) \in \Omega_{i,j}$ and the operators $\mathbb{X}$ and $\mathbb{Y}$ enforce the continuity ($k = 0$) and a selected smoothness ($k = r_s \in \mathbb{N}$) between the approximate Kernel solutions on every element.²⁰

A partition of the domain allows decreasing the polynomial approximation degree to $r = 6$ or lower, for a maximum approximation error less than 3% in the reactivity term. Results for the Kernel solution, including (3.110)-(3.111) in the convex optimization problem (3.18)-(3.28), are indicated in the first row of Table 3.3, where $J(\delta) = (1/N_e)(\sum_{j=1}^m \sum_{i=j}^{N_e} \gamma_{i,j}^2)^{(1/2)}$. For a partition of $N_e = 15$

²⁰For simplicity, the optimization scheme (3.110)-(3.111) only shows square elements $\Omega_{i,j}$. A slight modification of this formulation allows including triangular elements on the diagonal $x = y$.

elements ($m = 5$), Figure 3.4.3 depicts the approximate Kernel solution.

N_e	1	3	6	10	15
Case 2: $J(\delta)$	— — —	$2.43 \cdot 10^{-1}$	$1.81 \cdot 10^{-2}$	$9.7 \cdot 10^{-3}$	$3.1 \cdot 10^{-3}$
Example $\lambda = \text{cte}$: $J(\delta)$	1.55	$2.86 \cdot 10^{-2}$	$3.4 \cdot 10^{-3}$	$1.5 \cdot 10^{-3}$	$5.6 \cdot 10^{-4}$

Table 3.3: Root mean square of residual functions. Domain partition in N_e elements, Kernel with degree $d = 6$. First row, example from Section 3.4.2, Case 2. Second row, example with constant reactivity term from Section 3.4.1.

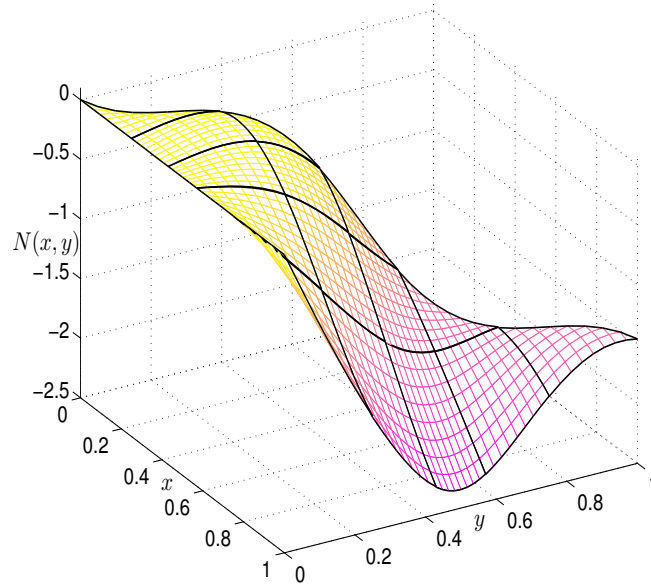


Figure 3.4.3: Approximate Kernel: $K \approx N$ in (3.5). Polynomial degree $d = 6$ and Ω_L partitioned in $N_e = 15$ elements ($m = 5$).

3.4.3 Hyperbolic PIDE

This example considers the problem presented in [204][82]:

$$f(x) = a + \frac{b\sigma}{\sqrt{c} \cosh(\sqrt{c})} \sinh(\sqrt{c}(1-x)) \quad (3.112)$$

$$h_2(x) = -\frac{b\sigma}{\cosh(\sqrt{c})} \cosh(\sqrt{c}x) \cosh(\sqrt{c}(1-y)) \quad (3.113)$$

$$h_1(x) = h_2(x) + b\sigma \cosh(\sqrt{c}(x-y)) \quad (3.114)$$

with $a = 1.25$, $b = 0.1$, $c = 0.01$, $\sigma = 7.5$ (the application of the proposed approach is not limited to this case, which has been selected for comparison purposes). This problem (equations (71)-(73)

of [82]) corresponds to a first-order PDE coupled with a second order ODE, equivalent to the 1-dimensional hyperbolic PDE (3.52). The direct Kernels P and Q in (3.53) have been approximated solving the convex optimization problem (3.83)-(3.89). To implement this approach, the functions f , h_1 and h_2 in (3.112)-(3.114) have been approximated by a combination of univariate polynomials of degree 4, with maximum approximation error $< 10^{-9}\%$. The approximate Kernels M and N for $d = 10$ are shown in Figure 3.4.4(a), with root mean square bounds for the residual functions: $\gamma_1 = 5.28 \cdot 10^{-11}$ and $\gamma_2 = 7.20 \cdot 10^{-11}$. Using this result, the inverse Kernels R and S have been approximated solving (3.98)-(3.107) for $d = 10$. The result is shown in Figure 3.4.4(b) with bounds of the residual functions: $\bar{\gamma} = \max\{\gamma_1, \gamma_2, \gamma_3, \gamma_4\} \leq 3.33 \cdot 10^{-11}$.

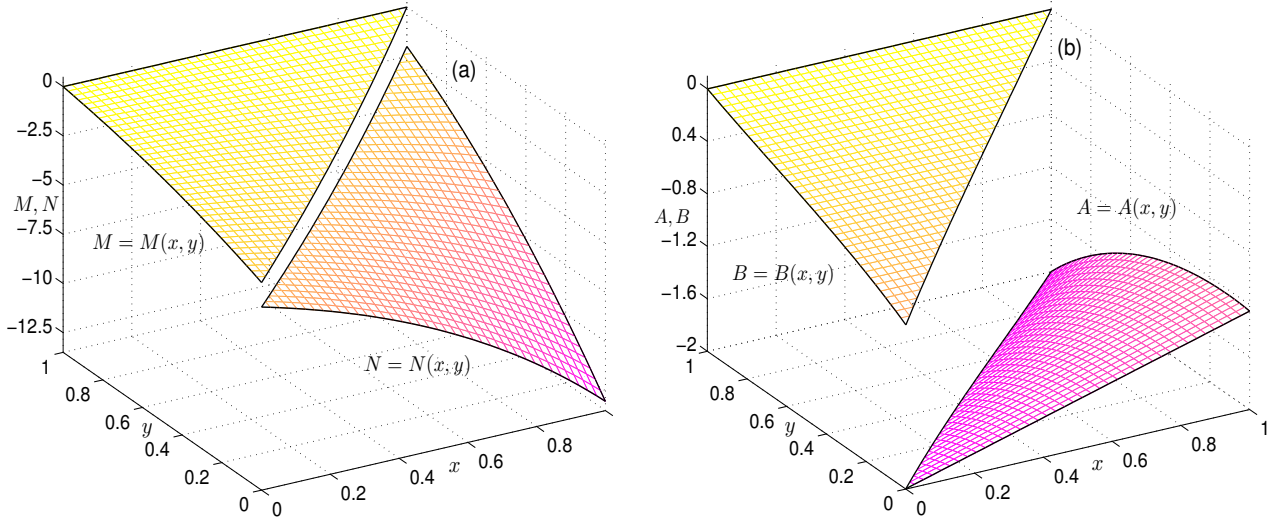


Figure 3.4.4: (a) Direct approximate Kernels M and N for $d = 10$ ($P \approx N$, $Q \approx M$ in (3.53)). (b) Inverse approximate Kernels A and B for $d = 10$ ($R \approx A$, $S \approx B$ in (3.74)).

Chapter 4

Adaptive Observer Design for a Class of Parabolic PDEs¹

—Vito Volterra (1860-1940)—

V. Volterra became interested in mathematics and science in his preadolescent years, passion which prevailed despite the dismay of his family. His interest in pure mathematics was inspired by his geometry instructor Cesare Arzelà, at the technical high school in Florence, and his calculus teacher Ulisse Dini, at the Scuola Normale Superiore in Pisa.

V. Volterra made important contributions in several fields. A real gem is considered his discovering of the “pointwise discontinuous functions”, two decades before René Baire’s category theorem, when he still was a student at Pisa. However, he is mostly known for his work in linear integral equations applying a principle which he called “the passage from the discrete to the continuous”, the limiting case of which was used by Ivar Fredholm and David Hilbert. He introduced the modern concept of “Functional” in his 1887 paper “On functions that depend on other functions”, name given later by Jacques Hadamard. He is considered one of the founders of the branch of mathematics that came to be called “Functional Analysis”.

In this chapter the observer design problem, for simultaneous state and parameter estimation, for a class of one-dimensional linear parabolic PDEs is addressed. The type of PDEs considered involves an uncertain “boundary” control gain parameter or an uncertain “in-domain” reactivity term. The design is based on the Backstepping PDE methodology, including a modified integral transformation to compensate for the parameter uncertainty. The solution of the resulting coupled Kernel-PDE/ODEs is obtained using the methodology proposed in Chapter 3, namely via its formulation as a convex optimization problem based on Sum-of-Squares decomposition and its computation via polynomial optimization techniques. This methodology allows computing – in a fast and direct way – the state and parameter observer gains at every fixed sampling time. Similar to the observer based on the Volterra-type transformation, an observer design method with two boundary measures via the Fredholm-type transformation is presented. Numerical simulations illustrate the performance of the approach proposed.

¹This chapter is based on the publications [240, 187].

4.1 Introduction

In the field of continuous-time boundary control/observer design for linear PDEs, Backstepping for PDEs stands out as a simple and systematic design methodology [21, 50, 49, 51, 18]. As dual to control design [21], Backstepping PDE observer design² mimics the feedback structure of the *Luenberger-type state observer* for finite dimensional systems [50]. In this methodology the observer error equation is mapped into a target PDE with given stability and convergence properties. This transformation is achieved for specific output injection observer gains which are *functions of the integral Kernels* governed by the resulting Kernel-PDE [86].

For certain class of PDE models, mainly due to the contraction properties of the Volterra integral transformation, a closed-form solution of the Kernel-PDE can be obtained [197], allowing a direct computation of the observer gains, scheme which has been commonly implemented (see for instance [241, 132, 11]). However, this PDE-Kernel computation approach, framed in the context of the Banach's contraction principle [211] and carried out via the Successive Approximation method [208], is not suitable for general cases such as causal systems with spatially and temporally variant parameters [49, 242], where also direct numerical methods cannot be applied [243]. On the other hand, the application of this type of recursive methods remains valid only for particular or conservative conditions in non-strict feedback systems (non-causal in space), which require a Fredholm-type integral transformation [205, 82].

In general, adaptive methods in Backstepping PDE design need state measurement on the full space domain [50, 122, 123, 124]. A feasible case of adaptive observer design based on boundary measurements requires a particular system structure denominated “canonical observer form”³ [125], which includes “in-domain” terms constituted of multiplications of unknown parameters and output measurements of the system [126]. Despite this structure can be obtained by an integral transformation on the original system, enabling to design adaptive output-feedback control schemes [126, 244], this is not useful to estimate the uncertain parameters of the original system, since it turns out to be an inverse under-determined problem.

A common approach proposed to overcome this problem consist of approximating the original PDE model as a finite dimensional input-output model, suitable for applying parametric identification techniques. However, in this case the resulting over-parameterized structure does not lead to determine a unique solution (see for instance [11]). As has been commented in [11], currently “there is no clear way” to address systems with general “in-domain” uncertain terms and the adaptive observer design remain as a *significant fundamental challenge* in IDS.

The main objectives of the approach proposed in this chapter are:

- (a) avoid any pre-transformation and/or finite-dimensional formulation, enabling a direct para-

²A brief summary on Backstepping PDE-based observer design is presented in Appendix B.

³It is worth noting that this type of structure can be seen as part of the “*Positive Real*” IDS [6].

meter identification from the original PDE model,

- (b) performs a simultaneous state and parameter estimation considering only measurements at the boundary, considering:
 - one point of measurement via the Volterra-type integral transformation,
 - two points of measurement via the Fredholm-type integral transformation (non-strict feedback error system).

4.2 Observer Design for one Uncertain Boundary Parameter

4.2.1 Problem Setting

Consider the one-dimensional linear parabolic PDE with an uncertain parameter in the mixed boundary conditions, namely

$$\begin{aligned}
u_t(x, t) &= \epsilon u_{xx}(x, t) + \lambda u(x, t), \\
u(0, t) &= 0, \quad u_x(1, t) = u(1, t) + \theta(t)U(t), \\
\theta_t(t) &= \sigma_1 \alpha(t) + \eta_1(\tau_i)\omega(\tau_i), \quad \forall t \in [t_a, t_b], \\
\alpha_t(t) &= \sigma_2 \alpha(t) + \eta_2(\tau_i)\omega(\tau_i), \quad \forall t \in [t_a, t_b], \\
y(t) &= u(1, t).
\end{aligned} \tag{4.1}$$

where $u(x, 0) = u_0(x) \in \mathcal{C}(\Omega)$ is the state initial condition, $\epsilon > 0$ is a constant diffusivity term, λ is a constant reactivity term, U is the control action, gain of which (θ) is considered as an uncertain time-varying parameter modelled as a double integrator; $\sigma_1 \geq 0$ and $\sigma_2 \leq 0$ with $\theta(t_a) = \theta_0$ and $\alpha(t_a) = \alpha_0$ unknown initial conditions of the parameters.

With respect to the parametric model (double integrator), this can be seen as a particular case of the model proposed in [245, 246]. This type of model considers the variation of θ as the output of a linear system excited by random jumps (random time τ_i) of unknown magnitude (η_1 and η_2), where ω is a completely unknown sequence of isolated impulses.⁴ Thus, this model can exhibit various waveform patterns according to the number of integrators and combination of their states (see details in [247, 245]). In particular, the structure in (4.1), for $\sigma_1 \geq 0$ and $\sigma_2 \leq 0$, provides a combination of steps and ramps, where the parameter α vanishes under no excitation.⁵ In comparison with the standard models which consider θ as piecewise constant function of time [127], the proposed model leads to an observer with lower feedback gains and more accurate estimates of the transient dynamic [245, 246].

⁴As is pointed out in [247], the model includes the signal ω and magnitudes η_1 and η_2 only for the symbolic purpose of providing a mathematical idea of the model. These variables are not used in the observer design.

⁵This model allows taking into account dynamics such as piece-wise constant functions of time with smooth transitions between their values.

4.2.2 Adaptive Observer

The proposed adaptive observer has a Luenberger-type collocated structure based on [86, 21, 50], namely

$$\begin{aligned}\hat{u}_t(x, t) &= \epsilon \hat{u}_{xx}(x, t) + \lambda \hat{u}(x, t) + h(x) \tilde{u}(1, t), \\ \hat{u}(0, t) &= 0, \quad \hat{u}_x(1, t) = \hat{u}(1, t) + p_b \tilde{u}(1) + \hat{\theta}(t) U(t), \\ \hat{\theta}_t(t) &= \sigma_1 \hat{\alpha}(t) + \rho_1 S_1(q(1)) \tilde{u}(1, t), \quad \forall t \in [t_a, t_b], \\ \hat{\alpha}_t(t) &= \sigma_2 \hat{\alpha}(t) + \rho_2 S_2(p(1)) \tilde{u}(1, t), \quad \forall t \in [t_a, t_b], \\ \hat{y}(t) &= \hat{u}(1, t),\end{aligned}\tag{4.2}$$

where $\hat{u}$, $\hat{\theta}$ and $\hat{\alpha}$ are the estimate state and control gain parameters, respectively, initial conditions of which are $\hat{u}(x, 0) = \hat{u}_0(x) \in \mathcal{C}(\Omega)$, $\hat{\theta}(t_a) = \hat{\theta}_0$ and $\hat{\alpha}(t_a) = \hat{\alpha}_0$; $\tilde{u}(1, t) = y(t) - \hat{y}(t)$ is the (measurable) output estimation error; $\rho_1 < 0$, $\rho_2 < 0$; S_1 and S_2 are real valued functions, $q \in \mathcal{C}^2(\Omega)$ and $p \in \mathcal{C}^2(\Omega)$. The function $h = h(x)$ and the constant p_b are the observer gains for the state and the boundary condition, respectively.

4.2.3 Design via the Volterra-type Transformation

The observer error equation between (4.1) and (4.2) can be written as:

$$\begin{aligned}\tilde{u}_t(x, t) &= \epsilon \tilde{u}_{xx}(x, t) + \lambda \tilde{u}(x, t) - h(x) \tilde{u}(1, t), \\ \tilde{u}(0, t) &= 0, \quad \tilde{u}_x(1, t) = (1 - p_b) \tilde{u}(1, t) + \tilde{\theta}(t) U(t), \\ \tilde{\theta}_t(t) &= \sigma_1 \tilde{\alpha}(t) - \rho_1 S_1(q(1)) \tilde{u}(1, t), \quad \forall t \in [t_a, t_b] \\ \tilde{\alpha}_t(t) &= \sigma_2 \tilde{\alpha}(t) - \rho_2 S_2(p(1)) \tilde{u}(1, t), \quad \forall t \in [t_a, t_b], \\ \tilde{u}(1, t) &= y(t) - \hat{y}(t),\end{aligned}\tag{4.3}$$

where $\tilde{u}(x, t) = u(x, t) - \hat{u}(x, t)$, $\tilde{\theta}(t) = \theta(t) - \hat{\theta}(t)$ and $\tilde{\alpha}(t) = \alpha(t) - \hat{\alpha}(t)$ are the state and parameters estimation errors, respectively. The design of the Backstepping PDE-based adaptive observer (4.2) consists basically in transforming the “**original**” estimation error system (4.3) into a “**target**” system with prescribed features of stability. In this case, the following exponentially stable system:

$$\begin{aligned}\tilde{w}_t(x, t) &= \epsilon \tilde{w}_{xx}(x, t) - c \tilde{w}(x, t) \\ \tilde{w}(0, t) &= 0, \quad \tilde{w}_x(1, t) = 0, \\ \frac{d}{dt} \begin{bmatrix} \tilde{\theta}(t) \\ \tilde{\alpha}(t) \end{bmatrix} &= \Lambda \begin{bmatrix} \tilde{\theta}(t) \\ \tilde{\alpha}(t) \end{bmatrix} + \Gamma \tilde{w}(1, t), \quad \forall t \in [t_a, t_b],\end{aligned}\tag{4.4}$$

it is proposed, where $c > 0$, $\Lambda \in \mathbb{R}^{2 \times 2}$ (with bounded elements), the eigenvalues of which have non-positive real parts, and $\Gamma \in \mathbb{R}^{2 \times 1}$ (with bounded elements).

Remark 13. The PDE (4.4) with respect to the variable $\tilde{w}$ is a standard parabolic PDE, with homogeneous boundary conditions, which is exponentially stable for $c > -\epsilon \frac{\pi^2}{4}$. This can be verified as follows. Let $V = \frac{1}{2} \int_0^1 \tilde{w}^2(x, t) dx$ be a Lyapunov functional.⁶ Its time-derivative $\dot{V} = \int_0^1 \tilde{w}(x) \tilde{w}_t(x) dx$ along the trajectory (4.4) is given by:

$$\begin{aligned} \dot{V} &= \int_0^1 \tilde{w}(x) (\epsilon \tilde{w}_{xx}(x, t) - c \tilde{w}(x, t)) dx = \cancel{\epsilon \tilde{w}(x) \tilde{w}_x(x) \Big|_{x=0}^{x=1}}^0 - \underbrace{\epsilon \int_0^1 \tilde{w}_x^2(x) dx}_{T_1} - c \int_0^1 \tilde{w}^2(x) dx, \\ &\leq - \left(\epsilon \frac{\pi^2}{4} + c \right) \int_0^1 \tilde{w}^2(x) dx \leq 0, \end{aligned} \quad (4.5)$$

where a variation of Wirtinger's inequality has been applied on the term T_1 [50]. Thus, if matrix Λ is Hurwitz, the coupled PDE-ODE system (4.4) is exponentially stable for any Γ bounded gain.

To achieve this state mapping the *modified* Volterra-type transformation

$$\tilde{u}(x, t) = q(1) \tilde{w}(x, t) - \int_x^1 K(x, y) \tilde{w}(y, t) dy - q(x) \tilde{\theta}(t) - p(x) \tilde{\alpha}(t) \quad (4.6)$$

with polynomial kernel $K \in \mathcal{C}^2(\Omega \times \Omega)$ is proposed, where the polynomial functions q and p are included to compensate for the parameters uncertainty. Following the standard procedure of Backstepping PDE observer design [21, 50], the transformation (4.6) leads to:

$$\begin{aligned} \tilde{u}_t(x, t) &= q(1) \tilde{w}_t(x, t) - \int_x^1 K(x, y) \tilde{w}_t(y, t) dy - q(x) \tilde{\theta}_t(t) - p(x) \tilde{\alpha}_t(t) \\ &= q(1) (\epsilon \tilde{w}_{xx}(x, t) - c \tilde{w}(x, t)) - \underbrace{\int_x^1 \epsilon K(x, y) \tilde{w}_{yy}(y, t) dy}_{T_0} + c \int_x^1 K(x, y) \tilde{w}(y, t) dy - \\ &\quad q(x) \tilde{\theta}_t(t) - p(x) \tilde{\alpha}_t(t), \\ &= q(1) (\epsilon \tilde{w}_{xx}(x, t) - c \tilde{w}(x, t)) + \int_x^1 (cK(x, y) - \epsilon K_{yy}(x, y)) \tilde{w}(y, t) dy + \\ &\quad \epsilon \frac{\partial}{\partial y} K(x, 1) \tilde{w}(1, t) - \epsilon \frac{\partial}{\partial y} K(x, x) \tilde{w}(x, t) - \cancel{\epsilon K(x, 1) \tilde{w}_x(1)}^0 + \epsilon K(x, x) \tilde{w}_x(x, t) - \\ &\quad q(x) \left(\sigma_1 \tilde{\alpha}(t) - \rho_1 S_1(q(1)) \left(q(1) [\tilde{w}(1, t) - \tilde{\theta}(t)] - p(1) \tilde{\alpha}(t) \right) \right) - \\ &\quad p(x) \left(\sigma_2 \tilde{\alpha}(t) - \rho_2 S_2(p(1)) \left(q(1) [\tilde{w}(1, t) - \tilde{\theta}(t)] - p(1) \tilde{\alpha}(t) \right) \right), \end{aligned} \quad (4.7)$$

where $\tilde{w}$ has been substituted by its “target” dynamic (4.4) and the time derivatives of the parameters estimation errors by the expressions stated in (4.3); the transformation (4.6) has been evaluated at the boundary $x = 1$, namely

$$\tilde{u}(1, t) = q(1) [\tilde{w}(1, t) - \tilde{\theta}(t)] - p(1) \tilde{\alpha}(t), \quad (4.8)$$

and integration by parts has been applied on the term T_0 . Substituting the second spatial derivative of (4.6), namely:

⁶For a clearer description, the time-dependence in the functions is dropped ($\tilde{w}(x) \equiv \tilde{w}(x, t)$).

$$\begin{aligned}\tilde{u}_{xx}(x, t) = & q(1)\tilde{w}_{xx}(x, t) - \int_x^1 K_{xx}(x, y)\tilde{w}(y, t)dy + 2\frac{\partial}{\partial x}K(x, x)\tilde{w}(x, t) + \frac{\partial}{\partial y}K(x, x)\tilde{w}(x, t) + \\ & K(x, x)\tilde{w}_x(x, t) - q_{xx}(x)\tilde{\theta}(t) - p_{xx}(x)\tilde{\alpha}(t),\end{aligned}\quad (4.9)$$

the expressions (4.7), (4.8) and the modified Volterra-type transformation (4.6) into the “original” estimation error system (4.3) leads to:

$$\begin{aligned}\tilde{u}_t(x, t) - \epsilon\tilde{u}_{xx}(x, t) - \lambda\tilde{u}(x, t) + h(x)\tilde{u}(1, t) = & \underbrace{\left(\epsilon q_{xx}(x) + q(x)[\lambda - \rho_1 S_1(q(1))q(1)] - p(x)\rho_2 S_2(p(1))q(1) - h(x)q(1)\right)}_{T_1} \tilde{\theta}(t) + \\ & \underbrace{\left(\epsilon p_{xx}(x) - \sigma_1 q(x) + p(x)[\lambda - \sigma_2 - \rho_2 S_2(p(1))p(1)] - q(x)\rho_1 S_1(q(1))p(1) - h(x)p(1)\right)}_{T_2} \tilde{\alpha}(t) - \\ & \underbrace{\left(\epsilon K_y(x, 1) + q(x)\rho_1 S_1(q(1))q(1) + p(x)\rho_2 S_2(p(1))q(1) + h(x)q(1)\right)}_{T_3} \tilde{w}(1, t) - \\ & \underbrace{\left(2\epsilon\left(\frac{\partial}{\partial y}K(x, x) + \frac{\partial}{\partial x}K(x, x)\right) + q(1)[\lambda + c]\right)}_{T_4} \tilde{w}(x, t) + \\ & \int_x^1 \left(\epsilon K_{xx}(x, y) - \epsilon K_{yy}(x, y) + [\lambda + c]K(x, y)\right) \tilde{w}(y, t)dy = 0.\end{aligned}\quad (4.10)$$

Thus, comparing the terms T_1 and T_2 with respect to T_3 and using the identity $\frac{d}{dx}K(x, x) = \frac{\partial}{\partial x}K(x, x) + \frac{\partial}{\partial y}K(x, x)$ in T_4 [21] yields

$$\begin{aligned}\tilde{u}_t(x, t) - \epsilon\tilde{u}_{xx}(x, t) - \lambda\tilde{u}(x, t) + h(x)\tilde{u}(1, t) = & \delta_0(x)\tilde{\theta}(t) - \delta_1(x)\tilde{w}(x, t) + \\ & \int_x^1 \delta_2(x, y)\tilde{w}(y, t)dy + \delta_3(x)\tilde{\alpha}(t) = 0,\end{aligned}\quad (4.11)$$

with boundary conditions:

$$\begin{aligned}\tilde{u}(0, t) = & q(1)\tilde{w}(0, t) - \int_0^1 K(0, y)\tilde{w}(y, t)dy - q(0)\tilde{\theta}(t) - p(0)\tilde{\alpha}(t) = 0 \\ \tilde{u}_x(1, t) = & q(1)\tilde{w}_x(1, t) + K(1, 1)\tilde{w}(1, t) - q_x(1)\tilde{\theta}(t) - p_x(1)\tilde{\alpha}(t) \\ = & (1 - p_b)\underbrace{\tilde{u}(1, t)}_{(q(1)\tilde{w}(1, t) - q(1)\tilde{\theta}(t) - p(1)\tilde{\alpha}(t))} + \tilde{\theta}(t)U(t),\end{aligned}\quad (4.12)$$

equivalent, at each time t , to the set of differential boundary value problems, here denominated **Kernel-PDE/ODE**, composed by the coupled PDE-ODE BVP:

$$\delta_0(x) = \epsilon q_{xx}(x) + \lambda q(x) + \epsilon K_y(x, 1) \quad (4.13)$$

$$\delta_1(x) = 2\epsilon \frac{d}{dx}K(x, x) + q(1)(\lambda + c), \quad (4.14)$$

$$\delta_2(x, y) = \epsilon K_{xx}(x, y) - \epsilon K_{yy}(x, y) + (c + \lambda)K(x, y), \quad \forall (x, y) \in \Omega_U, \quad (4.15)$$

$$K(0, y) = 0, \quad \forall y \in \Omega, \quad (4.16)$$

$$q(0) = 0, \quad q_x(1, t) = K(1, 1) - U(t), \quad (4.17)$$

and the ODE-BVP:

$$\delta_3(x) = \epsilon p_{xx}(x) + (\lambda - \sigma_2)p(x) - \sigma_1 q(x) + p(1) \frac{\epsilon K_y(x, 1)}{q(1)} \quad (4.18)$$

$$p(0) = 0, \quad p_x(1) = p(1) \frac{K(1, 1)}{q(1)}, \quad (4.19)$$

$\forall q(1) \neq 0$, where the observer gains are given by:

$$p_b = 1 - \frac{K(1, 1)}{q(1)}, \quad (4.20)$$

$$h(x) = -\frac{1}{q(1)} (\epsilon K_y(x, 1) + \rho_1 S_1(q(1))q(1)q(x) + \rho_2 S_2(p(1))q(1)p(x)) \quad (4.21)$$

for a zero matching condition on the residual functions δ_j , $\forall j = 0, \dots, 3$.

Remark 14. *The boundary condition (4.17) implies that the function q is not only dependent on the spatial variable but also on time. This continuous time-dependence implies that the validity of posing (4.14)-(4.19) is restricted to signals U with particular features, such as for instance “piece-wise constant” dynamics. Thus, despite the scheme of solving (4.14)-(4.19) at “each fixed time” can achieve good results, this does not necessarily lead to a stable and convergent estimation. Therefore, for general cases, it is necessary to pose a full coupled PDE/ODE (initial value problem) with time-varying boundary conditions.*

Regarding the “target” parameter error dynamic stated in (4.3), substituting (4.8) into the parameter estimation error of the “original” system (4.2) yields

$$\frac{d}{dt} \begin{bmatrix} \tilde{\theta}(t) \\ \tilde{\alpha}(t) \end{bmatrix} = \underbrace{\begin{bmatrix} \rho_1 S_1(q(1))q(1) & \sigma_1 + \rho_1 S_1(q(1))p(1) \\ \rho_2 S_2(p(1))q(1) & \sigma_2 + \rho_2 S_2(p(1))p(1) \end{bmatrix}}_{\Lambda} \begin{bmatrix} \tilde{\theta}(t) \\ \tilde{\alpha}(t) \end{bmatrix} + \underbrace{\begin{bmatrix} -\rho_1 S_1(q(1))q(1) \\ -\rho_2 S_2(p(1))q(1) \end{bmatrix}}_{\Gamma} \tilde{w}(1, t),$$

$\forall t \in [t_a, t_b]$. Thus, a sufficient condition to set forth a matrix Λ with negative eigenvalues is $S_1(q(1))q(1) > 0$ and $S_2(p(1))p(1) > 0$ ($\rho_1 < 0$, $\rho_2 < 0$, $\sigma_1 \geq 0$, $\sigma_2 \leq 0$), from a suitable selection of the functions S_1 and S_2 , for instance $S_i(f) = f$ or $S_i(f) = \text{sign}(f)$. As for the invertibility of (4.6), for bounded continuous functions $q = q(x)$ and $p = p(x)$, it is provided by the invertibility property of the Volterra integral operator on the Banach space of continuous functions [191, 208], which is a similar case of the transformation presented in [248].

Remark 15. *In general, since the function q and therefore p are dependent on time, the matrix Λ is function of time. In this case, the non-positivity condition on its eigenvalues is not sufficient to guarantee the stability of (4.4). A sufficient condition can be obtained if the matrix Λ is also differentiable or with Lipschitz continuity, with sufficiently small Lipschitz constants [127].*

4.2.3.1 Kernel-PDE/ODE as a Convex Optimization Problem

Similar to Chapter 3, Section 3.2.2, a relaxation of the exact zero matching condition for the residual functions δ_i can be considering to solve approximately the Kernel-PDE/ODE (4.13)-(4.19).

Proposition 8. *Let $N(x, y) = \sum_{k=0}^{z(d)-1} n_k x^{\alpha_k} y^{\beta_k}$ be a polynomial approximation of K in (4.6) in accordance with (1.3) and $f(x) = \sum_{k=0}^d f_k x^k$ be polynomial approximation of q in (4.6) of arbitrary even degree $d \in \mathbb{N}$. Let η_j and σ_j be lower and upper bounds of δ_j in (4.13)-(4.14), respectively, $\forall x \in \Omega$, $\forall j = 0, 1$. Let η_3 and σ_3 be lower and upper bounds of δ_2 in (4.15), respectively, $\forall (x, y) \in \Omega_U$; $\gamma_j \geq 0$, $\xi_j \geq 0$, $\forall j = 0, \dots, 2$. The coupled PDE-ODE BVP (4.13)-(4.17) can be formulated as the convex optimization problem:*

$$\begin{aligned} \underset{\gamma_j, \xi_j, N, f, s_{ij}}{\text{minimize:}} \quad & \sum_{j=0}^3 \gamma_j + \xi_j \end{aligned} \quad (4.22)$$

$$\text{subject to:} \quad (\delta_j(x) - \eta_j - s_{0j}(x) g_1(x)) \in \Sigma_s, \quad \forall j = 0, 1, \quad (4.23)$$

$$(\sigma_j - \delta_j(x) - s_{1j}(x) g_1(x)) \in \Sigma_s, \quad \forall j = 0, 1, \quad (4.24)$$

$$(\delta_2(x, y) - \eta_2 - [s_{20}(x, y) \ s_{21}(x, y)] g_U(x, y)) \in \Sigma_s, \quad (4.25)$$

$$(\sigma_2 - \delta_2(x, y) - [s_{22}(x, y) \ s_{23}(x, y)] g_U(x, y)) \in \Sigma_s, \quad (4.26)$$

$$s_{0j}, s_{1j}, s_{2i} \in \Sigma_s, \quad \forall j = 0, 1, \ i = 0, \dots, 3, \quad (4.27)$$

$$\begin{bmatrix} \gamma_j & \eta_j \\ \eta_j & \gamma_j \end{bmatrix} \succeq 0, \quad \begin{bmatrix} \xi_j & \sigma_j \\ \sigma_j & \xi_j \end{bmatrix} \succeq 0, \quad \forall j = 0, \dots, 3, \quad (4.28)$$

$$\delta_0 = \delta_0(x) \Big|_{q \approx f}^{K \approx N} \quad \text{as in (4.13)}, \quad (4.29)$$

$$\delta_1 = \delta_1(x) \Big|_{q \approx f}^{K \approx N} \quad \text{as in (4.14)}, \quad (4.30)$$

$$\delta_2 = \delta_2(x, y) \Big|_{K \approx N} \quad \text{as in (4.15)}, \quad (4.31)$$

$$N(0, y) = 0, \quad \forall y \in \Omega \quad (4.32)$$

$$f(0) = 0, \quad f_x(1, t) = N(1, 1) - U(t), \quad (4.33)$$

$\forall (x, y) \in \Omega_U$, $\forall t \geq 0$, for some polynomials s_{0j} , s_{1j} and s_{2i} of degree $d-2$, $\forall j=0, 1$, $\forall i=0, \dots, 3$; $g_U = [g_1, g_3]$ with $g_1(x) = x(1-x)$ and $g_3(x, y) = (y-x)(1-y)$ (see (2.8)). The optimal minimal bounds for the residual functions are: $\bar{\delta}_j = \max\{\gamma_j, \xi_j\}$, $\forall j = 0, \dots, 2$.

Proof. [The proof follows the same arguments exposed in Proposition 2.] Let N be the polynomial approximation of K (4.6), as it is defined above. Then the residual functions (4.13)-(4.16) have polynomial structure determined by (4.29)-(4.33). Since the quadratic module associated to the compact basic semi-algebraic representations of the domains Ω and Ω_U are Archimedean (see Lemma 1), based on Putinar's Positivstellensatz [148, 138] and on the SOS decomposition of (4.23)-(4.27), the unknown extreme values of each δ_j (η_j, ξ_j) can be determined as a polynomial optimization problem, which is convex in terms of their polynomial coefficients [136]. The absolute value of the lower and upper bounds of each δ_j are given by means of (4.28), so that the linear cost function (4.22) yields $\delta_j \rightarrow 0$ in Ω_U . $\square$

Regarding the ODE-BVP (4.18)-(4.19), this can also be approximately solved via polynomial optimization based on known polynomial approximations of the solutions K and q of the coupled PDE-ODE BVP (4.13)-(4.17).

Proposition 9. *Let N and f be known polynomial approximations of degree d of the solutions P and q of (4.13)-(4.17), respectively. Let $l(x) = \sum_{k=0}^{d_l} l_k x^k$ be a polynomial approximation of p in (4.6) of even degree $d_l \geq d$. Let η and σ be lower and upper bounds of δ_3 in (4.18), respectively, $\forall x \in \Omega$; $\gamma \geq 0$, $\xi \geq 0$. The coupled ODE-BVP (4.18)-(4.19) can be formulated as the convex optimization problem:*

$$\begin{aligned}
& \underset{\gamma, \xi, l, s_0, s_1}{\text{minimize:}} && \gamma + \xi \\
& \text{subject to:} && (\delta_3(x) - \eta - s_0(x) g_1(x)) \in \Sigma_s, \\
& && (\sigma - \delta_3(x) - s_1(x) g_1(x)) \in \Sigma_s, \\
& && s_0, s_1 \in \Sigma_s, \begin{bmatrix} \gamma & \eta \\ \eta & \gamma \end{bmatrix} \succeq 0, \quad \begin{bmatrix} \xi & \sigma \\ \sigma & \xi \end{bmatrix} \succeq 0, \\
& && l(0) = 0, \quad l_x(1) = l(1)N(1,1)/f(1),
\end{aligned} \tag{4.34}$$

$\forall x \in \Omega$, for some polynomials s_0 and s_1 of degree $d_l - 2$; $g_1(x) = x(1-x)$ (see (2.6)). The optimal minimal bound for the residual function is: $\overline{\delta_3} = \max\{\gamma, \xi\}$.

Proof. The proof follows the same arguments exposed in Proposition 2 and (8), hence it is omitted. $\square$

Remark 16. *It is worth noting that the ODE-BVP (4.18)-(4.19) has a boundary value ($p(1)$) as part of the “in-domain” equation. Proposition (9) allows solving this particular structure where other approaches, such as the “shooting method”, can not be directly applied.*

4.3 Design for an Uncertain Reactivity Parameter

4.3.1 Problem Setting

The objective is to design an adaptive observer for the one-dimensional linear parabolic PDE with strict-feedback structure and Neumann boundary conditions described by

$$\begin{aligned}
u_t(x, t) &= \epsilon u_{xx}(x, t) + \lambda(t)u(x, t), \\
u_x(0, t) &= u(0), \quad u_x(1, t) = U(t), \\
\frac{d\lambda(t)}{dt} &= 0, \quad \forall t \in [t_a, t_b], \quad \lambda(t_a) = \lambda_0, \\
y(t) &= u(1, t) \quad \text{or} \quad y_0(t) = u(0, t), \quad y_1(t) = u(1, t),
\end{aligned} \tag{4.35}$$

to simultaneously estimate the state u and the uncertain reactivity parameter λ . This parameter is considered as a piece-wise constant function of time (as it is described in Section 4.2.1, this

observer can easily be modified to include smooth time dynamics via a double integrator). For this system $u(x, 0) = u_0(x) \in \mathcal{C}(\Omega; \mathbb{R})$ is its initial condition and $U = U(t)$ denotes the control action. As it is indicated in (4.35) (last equation), the observer design considers two cases: one point of measurement y and two boundary points: y_0 and y_1 .

4.3.2 Target System

Proposition 10. *Consider the following “target” coupled PDE-ODE:*

$$\begin{aligned}\tilde{w}_t(x, t) &= \epsilon \tilde{w}_{xx}(x, t) - (c - \tilde{\lambda}(t)) \tilde{w}(x, t) - \phi(x) \tilde{\lambda}^2(t), \\ \tilde{w}_x(0, t) &= 0, \\ \tilde{w}_x(1, t) &= 0, \\ \tilde{\lambda}_t(t) &= -\rho \tilde{\lambda}(t) - a_0 \tilde{w}(0, t) - a_1 \tilde{w}(1, t),\end{aligned}\tag{4.36}$$

with initial conditions $\tilde{w}(x, 0) = \tilde{w}_0$ and $\tilde{\lambda}(0) = \tilde{\lambda}_0$; $\rho > 0$, $a_0 \in \mathbb{R}$, $a_1 \in \mathbb{R}$ ⁷ and $\phi \in \mathcal{L}^2(\Omega)$. Assuming $|\lambda| \leq M$ a known bound, if the conditions

$$c \geq 2M + 2\alpha M^2 + \max\{|a_0|\gamma_0, |a_1|\gamma_1\}(1 + \sigma_0),\tag{4.37}$$

$$\rho \geq \frac{1}{2} \left(\frac{|a_0|}{\gamma_0} + \frac{|a_1|}{\gamma_1} \right) + \frac{1}{2\alpha} \int_0^1 \phi^2(x) dx,\tag{4.38}$$

are satisfied for some selected constants $\alpha > 0$, $\gamma_0 > 0$, and $\gamma_1 > 0$, and for $\sigma_0 \geq \frac{\max\{|a_0|\gamma_0, |a_1|\gamma_1\}}{\epsilon}$, then the PDE-ODE system (4.36) is stable.

Proof. Let $V = \frac{1}{2} \int_0^1 \tilde{w}^2(x, t) dx + \frac{1}{2} \tilde{\lambda}(t)^2$ be a Lyapunov functional.⁸ Its time-derivative $\dot{V} = \int_0^1 w(x) w_t(x) dx + \tilde{\lambda} \tilde{\lambda}_t$ along the trajectory (4.36) is given by:

$$\begin{aligned}\dot{V} &= \int_0^1 w(x) (\epsilon \tilde{w}_{xx}(x) - c \tilde{w}(x)) dx + \int_0^1 \tilde{\lambda} \tilde{w}^2(x) dx - \int_0^1 \tilde{w}(x) \phi(x) \tilde{\lambda}^2 dx + \tilde{\lambda} \tilde{\lambda}_t, \\ &\leq \cancel{\epsilon \tilde{w}(x) \tilde{w}_x(x) \Big|_{x=0}^{x=1}} - \epsilon \int_0^1 w_x^2(x) dx - c \int_0^1 \tilde{w}^2(x) dx + |\tilde{\lambda}| \int_0^1 \tilde{w}^2(x) dx + \tilde{\lambda}^2 \underbrace{\left| \int_0^1 \tilde{w}(x) \phi(x) dx \right|}_{T_1} + \tilde{\lambda} \tilde{\lambda}_t,\end{aligned}\tag{4.39}$$

which via the Cauchy-Schwarz integral inequality [239] on the term T_1 and Young's inequality (special case)[21] on its resulting relations for some $\alpha > 0$, and considering $|\tilde{\lambda}| \leq 2M$ yields

$$\begin{aligned}\dot{V} &\leq -\epsilon \int_0^1 w_x^2(x) dx - (c - |\tilde{\lambda}|) \int_0^1 \tilde{w}^2(x) dx + \frac{\tilde{\lambda}^2}{2} \left(\alpha \int_0^1 \tilde{w}^2(x) dx + \frac{1}{\alpha} \int_0^1 \phi^2(x) dx \right) + \tilde{\lambda} \tilde{\lambda}_t, \\ &\leq -\epsilon \int_0^1 \tilde{w}_x^2(x) dx - (c - 2M - 2\alpha M^2) \int_0^1 \tilde{w}^2(x) dx + \frac{\tilde{\lambda}^2}{2\alpha} \int_0^1 \phi^2(x) dx + \tilde{\lambda} \tilde{\lambda}_t.\end{aligned}\tag{4.40}$$

⁷The coefficient a_0 or a_1 can take zero value to exclude the boundary measurements $u(0, t)$ or $u(1, t)$, respectively.

⁸For a clearer description, the time-dependence in the functions is dropped ($\tilde{w}(x) \equiv w(x, t)$, $\tilde{\lambda} \equiv \tilde{\lambda}(t)$).

In last equation of the target system (4.36) it is verified:

$$\begin{aligned}
\tilde{\lambda}\tilde{\lambda}_t &= \tilde{\lambda} \left(-\rho\tilde{\lambda}(t) - a_0\tilde{w}(0) - a_1\tilde{w}(1) \right) \\
&\leq -\rho\tilde{\lambda}^2 + \underbrace{|a_0||\tilde{\lambda}||\tilde{w}(0)|}_{T_2} + \underbrace{|a_1||\tilde{\lambda}||\tilde{w}(1)|}_{T_3}, \\
&\leq -\underbrace{\tilde{\lambda}^2 \left(\rho - \frac{1}{2} \left(\frac{|a_0|}{\gamma_0} + \frac{|a_1|}{\gamma_1} \right) \right)}_{\sigma} + \frac{|a_0|\gamma_0}{2}\tilde{w}^2(0) + \frac{|a_1|\gamma_1}{2}\tilde{w}^2(1),
\end{aligned} \tag{4.41}$$

which is derived from using Young's inequality on terms T_2 and T_3 for some $\gamma_0 > 0$, $\gamma_1 > 0$. Then, applying over the differential expressions $\frac{d}{dx}((1-x)\tilde{w}^2(x))$ and $\frac{d}{dx}(x\tilde{w}^2(x))$ an integration along $x \in [0, 1]$, in relation with the terms $\tilde{w}(0)$ and $\tilde{w}(1)$, respectively, yields

$$\begin{aligned}
\tilde{\lambda}\tilde{\lambda}_t &\leq -\sigma\tilde{\lambda}^2 + \frac{|a_0|\gamma_0}{2} \left(\int_0^1 \tilde{w}^2(x)dx + 2 \int_0^1 (x-1)\tilde{w}(x)\tilde{w}_x(x)dx \right) + \\
&\quad \frac{|a_1|\gamma_1}{2} \left(\int_0^1 \tilde{w}^2(x)dx + 2 \int_0^1 x\tilde{w}(x)\tilde{w}_x(x)dx \right), \\
&\leq -\sigma\tilde{\lambda}^2 + \bar{\beta} \int_0^1 \tilde{w}^2(x)dx + \bar{\beta} \left(\left| \int_0^1 (x-1)\tilde{w}(x)\tilde{w}_x^2(x)dx \right| + \left| \int_0^1 x\tilde{w}(x)\tilde{w}_x^2(x)dx \right| \right),
\end{aligned} \tag{4.42}$$

where $\bar{\beta} = \max\{|a_0|\gamma_0, |a_1|\gamma_1\} > 0$. Thus, applying once more time the Cauchy-Schwarz integral inequality on the last term in (4.42) and Young's inequality on its resulting expressions, leads to:

$$\begin{aligned}
\tilde{\lambda}\tilde{\lambda}_t &\leq -\sigma\tilde{\lambda}^2 + \bar{\beta} \int_0^1 \tilde{w}^2(x)dx + \\
&\quad \bar{\beta} \left(\int_0^1 \tilde{w}_x^2(x)dx \right)^{1/2} \left(\left(\max_{x \in [0,1]} (x-1)^2 \right)^{1/2} + \left(\max_{x \in [0,1]} x^2 \right)^{1/2} \right) \left(\int_0^1 \tilde{w}^2(x)dx \right)^{1/2}, \\
&\leq -\sigma\tilde{\lambda}^2 + \bar{\beta}(1 + \sigma_0) \int_0^1 \tilde{w}^2(x)dx + \frac{\bar{\beta}}{\sigma_0} \int_0^1 \tilde{w}_x^2(x)dx,
\end{aligned} \tag{4.43}$$

for some $\sigma_0 > 0$. Finally substituting (4.43) into (4.40) is obtained:

$$\begin{aligned}
\dot{V} &\leq -\left(\epsilon - \frac{\bar{\beta}}{\sigma_0} \right) \int_0^1 \tilde{w}_x^2(x)dx - (c - 2M - 2\alpha M^2 - \bar{\beta}(1 + \sigma_0)) \int_0^1 \tilde{w}^2(x)dx - \\
&\quad \tilde{\lambda}^2 \left(\sigma - \frac{1}{2\alpha} \int_0^1 \phi^2(x)dx \right) \leq 0,
\end{aligned} \tag{4.44}$$

if conditions (4.37) and (4.38) are satisfied for $\sigma_0 \geq \frac{\bar{\beta}}{\epsilon}$, which are sufficient conditions for the stability of (4.36). $\square$

4.3.3 Design for one boundary measurement: Volterra-type Transformation

Consider the system (4.35) with a unique point of measurement $y = u(1, t)$. For a simultaneous estimation of u and λ , a Luenberger-type observer with collocated setup structure based on [86] is proposed, namely

$$\begin{aligned}
\hat{u}_t(x, t) &= \epsilon \hat{u}_{xx}(x, t) + \hat{\lambda}(t) \hat{u}(x, t) + \frac{p(x)}{q(1)} \tilde{u}(1, t), \\
\hat{u}_x(0, t) &= \hat{u}(0, t), \\
\hat{u}_x(1, t) &= U(t) + p_b \tilde{u}(1, t) \\
\frac{d\hat{\lambda}}{dt}(t) &= \frac{\rho_1}{q(1)} S_1 \tilde{u}(1, t), \quad \forall t \in [t_a, t_b], \quad \hat{\lambda}(t_a) = \hat{\lambda}_0, \\
\hat{y}(t) &= \hat{u}(1, t),
\end{aligned} \tag{4.45}$$

where $\hat{u}$ and $\hat{\lambda}$ are the estimate state and reactivity parameter, respectively, initial conditions of which are $\hat{u}(x, 0) = \hat{u}_0(x) \in \mathcal{C}(\Omega)$ and $\hat{\lambda}(t_a) = \hat{\lambda}_0$; $y(t) - \hat{y}(t) = \tilde{u}(1, t) = u(1, t) - \hat{u}(1, t)$ is the measurable output estimation error; $\rho_1 < 0$; $p = p(x)$ and p_b are the observer gains for the state and boundary condition, respectively, S_1 is a real valued function and $q \in \mathcal{C}^2(\Omega)$ with $q(1) \neq 0$.

The observer error equation can be written as:

$$\begin{aligned}
\tilde{u}_t(x, t) &= \epsilon \tilde{u}_{xx}(x, t) - \frac{p(x)}{q(1)} \tilde{u}(1, t) + \underbrace{\left[\hat{u}(x, t) \tilde{\lambda}(t) + \tilde{u}(x, t) \hat{\lambda}(t) + \tilde{u}(x, t) \tilde{\lambda}(t) \right]}_{\Delta(x, t)} \\
\tilde{u}_x(0, t) &= \tilde{u}(0, t), \\
\tilde{u}_x(1, t) &= -p_b \tilde{u}(1, t), \\
\frac{d\tilde{\lambda}}{dt}(t) &= -\frac{\rho_1}{q(1)} S_1 \tilde{u}(1, t), \quad \forall t \in [t_a, t_b], \quad \tilde{\lambda}(t_a) = \tilde{\lambda}_0, \\
\tilde{u}(1, t) &= y(t) - \hat{y}(t),
\end{aligned} \tag{4.46}$$

where $\tilde{u}(x, t) = u(x, t) - \hat{u}(x, t)$ and $\tilde{\lambda}(t) = \lambda(t) - \hat{\lambda}(t)$ are the state and parameter estimation error, respectively. The term $\Delta(x, t) = u(x, t) \lambda(t) - \hat{u}(x, t) \hat{\lambda}(t)$ can be considered as a particular case obtained from the Taylor's series approximation of the bivariate function $g = u(x, t) \lambda(t)$ around the point $(\hat{u}(x, t), \hat{\lambda}(t))$.

The design of the adaptive observer (4.45) consists in transforming the “original” estimation error system (4.46) into the “target” system (4.36), which is assumed to be stable for a suitable selection of the parameter “ c ”, for $\mu = 1$ and $a_0 = 0$. To achieve this mapping a modification of the Volterra-type transformation (3.5), with integral Kernel $P \in \mathcal{C}^2(\Omega_U)$,

$$\tilde{u}(x, t) = q(1) \tilde{w}(x, t) - \int_x^1 P(x, y) \tilde{w}(y, t) dy - q(x) \tilde{\lambda}(t) \tag{4.47}$$

is proposed, where the function q allows rejecting the λ parameter uncertainty. Similarly, as it has been detailed in Section 4.2.3, following the standard procedure of Backstepping PDE observer design [21, 50], the transformation leads to:

$$\tilde{u}_t(x, t) = q(1) \tilde{w}_t(x, t) - \int_x^1 P(x, y) \tilde{w}_t(y, t) dy - q(x) \tilde{\lambda}_t(t)$$

$$\begin{aligned}
 &= q(1) \left(\epsilon \tilde{w}_{xx}(x, t) - [c + \tilde{\lambda}(t)] \tilde{w}(x, t) - \phi(x) \tilde{\lambda}^2(t) \right) - \underbrace{\int_x^1 \epsilon P(x, y) \tilde{w}_{yy}(y, t) dy}_{T_0} + \\
 &\quad [c - \tilde{\lambda}(t)] \int_x^1 P(x, y) \tilde{w}(y, t) dy + \tilde{\lambda}^2(t) \int_x^1 P(x, y) \phi(y) dy - q(x) \tilde{\lambda}_t(t), \tag{4.48} \\
 &= q(1) \left(\epsilon \tilde{w}_{xx}(x, t) - [c + \tilde{\lambda}(t)] \tilde{w}(x, t) \right) + \int_x^1 \left([c - \tilde{\lambda}(t)] P(x, y) - \epsilon P_{yy}(x, y) \right) \tilde{w}(y, t) dy + \\
 &\quad \epsilon \frac{\partial}{\partial y} P(x, 1) \tilde{w}(1, t) - \epsilon \frac{\partial}{\partial y} P(x, x) \tilde{w}(x, t) - \cancel{\epsilon P(x, 1) \tilde{w}_x(1)} \overset{0}{+} \epsilon P(x, x) \tilde{w}_x(x, t) - \\
 &\quad \left(q(1) \phi(x) - \int_x^1 P(x, y) \phi(y) dy \right) \tilde{\lambda}^2(t) - q(x) \left(-\frac{\rho_1}{q(1)} S_1 \left(q(1) [\tilde{w}(1, t) - \tilde{\lambda}(t)] \right) \right),
 \end{aligned}$$

where $\tilde{w}$ has been substituted by its “target” dynamic (4.36) and the time derivative of the parameter estimation error by its dynamic stated in (4.46); the transformation (4.47) has been evaluated at the boundary $x = 1$, namely

$$\tilde{u}(1, t) = q(1) [\tilde{w}(1, t) - \tilde{\lambda}(t)], \tag{4.49}$$

and integration by parts has been applied on the term T_0 . Substituting the modified Volterra-type transformation (4.47), its second spatial derivative and the expressions (4.48) and (4.49) into the “original” estimation error system (4.46) leads to

$$\begin{aligned}
 &\tilde{u}_t(x, t) - \epsilon \tilde{u}_{xx}(x, t) + \frac{p(x)}{q(1)} \tilde{u}(1, t) - \Delta(x, t) = \\
 &\quad \underbrace{\left(\epsilon q_{xx}(x) + q(x) [\hat{\lambda}(t) - \rho_1 S_1] - p(x) - \hat{u}(x, t) \right)}_{T_1} \tilde{\lambda}(t) + \underbrace{\left(\epsilon P_y(x, 1) + q(x) \rho_1 S_1 + p(x) \right)}_{T_2} \tilde{w}(1, t) - \\
 &\quad \underbrace{\left(2\epsilon \left(\frac{\partial}{\partial y} P(x, x) + \frac{\partial}{\partial x} P(x, x) \right) + q(1) [c + \hat{\lambda}(t)] \right)}_{T_3} \tilde{w}(x, t) + \tag{4.50} \\
 &\quad \int_x^1 \left(\epsilon P_{xx}(x, y) - \epsilon P_{yy}(x, y) + [c + \hat{\lambda}(t)] P(x, y) \right) \tilde{w}(y, t) dy + \\
 &\quad \left(q(x) - q(1) \phi(x) + \int_x^1 P(x, y) \phi(y) dy \right) \tilde{\lambda}^2(t).
 \end{aligned}$$

Thus, comparing T_1 with respect to T_2 and using the identity $\frac{d}{dx} K(x, x) = \frac{\partial}{\partial y} K(x, x) + \frac{\partial}{\partial x} K(x, x)$ in T_3 [21] yields

$$\begin{aligned}
 &\tilde{u}_t(x, t) - \epsilon \tilde{u}_{xx}(x, t) + \frac{p(x)}{q(1)} \tilde{u}(1, t) + \Delta(x, t) = \left(q(x) - q(1) \phi(x) + \int_x^1 P(x, y) \phi(y) dy \right) \tilde{\lambda}^2(t) + \\
 &\quad \delta_0(x, t) \tilde{\lambda}(t) - \delta_1(x, t) \tilde{w}(x, t) + \int_x^1 \delta_2(x, y, t) \tilde{w}(y, t) dy = 0, \tag{4.51}
 \end{aligned}$$

with constraints obtained from the boundary conditions in (4.46), namely

$$\begin{aligned}
\tilde{u}_x(0, t) &= q(1)\tilde{u}_x(0, t) - \int_0^1 P_x(0, y)\tilde{w}(y, t)dy + P(0, 0)\tilde{w}(0, t) - q_x(0)\tilde{\lambda}(t) \\
&= \tilde{u}(0, t) = q(1)\tilde{w}(0, t) - \int_0^1 P(0, y)\tilde{w}(y, t)dy - q(0)\tilde{\lambda}(t), \\
\tilde{u}_x(1, t) &= q(1)\tilde{u}_x(1, t) + P(1, 1)\tilde{w}(1, t) - q_x(1)\tilde{\lambda}(t) \\
&= -p_b\tilde{u}(1, t) = -p_b \left(q(1) \left(\tilde{w}(1, t) - \tilde{\lambda}(t) \right) \right),
\end{aligned} \tag{4.52}$$

which are equivalent, at each time t , to the **Kernel-PDE/ODE**:

$$\delta_0(x, t) = \epsilon q_{xx}(x) + \hat{\lambda}(t)q(x) - \hat{u}(x, t) + \epsilon P_y(x, 1), \tag{4.53}$$

$$\delta_1(x, t) = 2\epsilon \frac{d}{dx} P(x, x) + q(1)(c + \hat{\lambda}(t)) \tag{4.54}$$

$$\delta_2(x, y, t) = \epsilon P_{xx}(x, y) - \epsilon P_{yy}(x, y) + (c + \hat{\lambda}(t))P(x, y), \quad \forall (x, y) \in \Omega_U \tag{4.55}$$

$$q(1) = P(0, 0), \quad P(0, y) = P_x(0, y), \tag{4.56}$$

$$q(0) = q_x(0), \quad q_x(1) = P(1, 1), \tag{4.57}$$

and function $\phi(x)$ satisfying the Volterra's integral equation of second kind:

$$\frac{q(x)}{q(1)} = \phi(x) - \int_x^1 \frac{P(x, y)}{q(1)} \phi(y) dy, \tag{4.58}$$

$\forall q(1) \neq 0$, for a zero matching condition on the residual functions δ_j , $\forall j = 0, \dots, 2$. In this case the observer gains are given by:

$$p_b = \frac{P(1, 1)}{q(1)}, \tag{4.59}$$

$$p(x) = -(\epsilon P_y(x, 1) + \rho_1 S_1 q(x)). \tag{4.60}$$

Remark 17. Due to (4.53) includes $\hat{u}$ and $\hat{\lambda}$, and (4.55) considers $\hat{\lambda}$, the Kernel P and function q are not only dependent on the spatial variable but also on time. As it has been commented in Remark 14, this continuous time-dependence implies that the validity of posing (4.53)-(4.57) is restricted to signals $\hat{u}$ and $\hat{\lambda}$ with particular features. Therefore, despite a practical scheme of solving (4.14)-(4.19) at “each fixed time” can achieve good results, this formulation does not necessarily lead to a stable and convergent estimation.

Regarding the “target” parameter error dynamic stated in (4.46), substituting (4.49) into the parameter estimation error of the “original” system (4.45) yields

$$\frac{d\tilde{\lambda}}{dt}(t) = -\frac{\rho_1}{q(1)} S_1 \left(q(1)[\tilde{w}(1, t) - \tilde{\lambda}(t)] \right) = (\rho_1 S_1) \tilde{\lambda}(t) - (\rho_1 S_1) \tilde{w}(1, t),$$

$\forall t \in [t_a, t_b]$. Thus, considering $\tilde{w}$ as a bounded signal and $\rho_1 < 0$, a sufficient condition to obtain a stable dynamic for $\tilde{\lambda}$ in (4.46) is the selection of a positive function $S_1 > 0$, $\forall t \geq 0$.

With respect to the integral equation (4.58), given a bounded function q and a Kernel P , the

existence and uniqueness of a bounded continuous function $\phi = \phi(x)$ are guaranteed by the well-known invertibility property of the Volterra integral operator on some Banach and Hilbert spaces [18, 208, 191]. As for equations (4.53)-(4.57), an approximate solution can be determined, at each fixed time, via the methodology developed in Chapter 2, for instance setting up an equivalent algorithm to the convex optimization problem described in Proposition 2.

4.3.3.1 Kernel-PDE/ODE as a Convex Optimization Problem

Similar to Section 4.2.3.1, the Kernel-PDE/ODE (4.53)-(4.57) can be approximately solved via an optimization problem.

Proposition 11. *Let $N(x, y) = \sum_{k=0}^{z(d)-1} n_k x^{\alpha_k} y^{\beta_k}$ be a polynomial approximation of P and $f(x) = \sum_{k=0}^d f_k x^k$ be polynomial approximation of q in (11), of arbitrary even degree $d > d_{\hat{u}} \in \mathbb{N}$, in accordance with (1.3). Let η_j and σ_j be lower and upper bounds of δ_j , respectively, in Ω , $\forall j = 0, 1$. Let η_2 and σ_2 be the lower and upper bound of δ_2 , respectively, in Ω_U ; $\gamma_j \geq 0$, $\xi_j \geq 0$, $\forall j = 0, 1, 2$. For $\hat{\lambda}$ and $\hat{u}$ estimated by the observer (4.45) and the former approximated by a polynomial function of degree $d_{\hat{u}}$, the Kernel-PDE (4.53)-(4.57) can be formulated as the convex optimization problem:*

$$\begin{aligned}
& \underset{\gamma_j, \xi_j, N, f, h, s_{ij}}{\text{minimize:}} && \sum_{j=0}^2 \gamma_j + \xi_j \\
& \text{subject to:} && (\delta_j(x, t) - \eta_j - s_{0j}(x) \ g_1(x)) \in \Sigma_s, \quad \forall j = 0, 1, \\
& && (\sigma_j - \delta_j(x, t) - s_{1j}(x) \ g_1(x)) \in \Sigma_s, \quad \forall j = 0, 1, \\
& && (\delta_2(x, y, t) - \eta_2 - [s_{20}(x, y) \ s_{21}(x, y)] \ g_U(x, y)) \in \Sigma_s, \\
& && (\sigma_2 - \delta_2(x, y, t) - [s_{22}(x, y) \ s_{13}(x, y)] \ g_U(x, y)) \in \Sigma_s, \\
& && s_{0j}, s_{1j}, s_{2i} \in \Sigma_s, \quad \forall j = 0, 1, \ i = 0, \dots, 3, \\
& && \begin{bmatrix} \gamma_j & \eta_j \\ \eta_j & \gamma_j \end{bmatrix} \succeq 0, \quad \begin{bmatrix} \xi_j & \sigma_j \\ \sigma_j & \xi_j \end{bmatrix} \succeq 0, \quad \forall j = 0, \dots, 2, \\
& && \delta_0 = \delta_0(x, t) \Big|_{q \approx f}^{P \approx N} \quad \text{as in (4.53),} \\
& && \delta_1 = \delta_1(x, t) \Big|_{q \approx f}^{P \approx N} \quad \text{as in (4.54),} \\
& && \delta_2 = \delta_2(x, y, t) \Big|_{P \approx N} \quad \text{as in (4.55),} \\
& && f(1) = N(0, 0), \quad N(0, y) = N_x(0, y), \\
& && f(0) = f_x(0), \quad f_x(1) = N(1, 1),
\end{aligned} \tag{4.61}$$

$\forall (x, y) \in \Omega_U$, $\forall t \geq 0$, for some polynomials s_{0j}, s_{1j} of degree $d - 2$, $\forall j = 0, \dots, 2$; s_{i3} of degree $d - 2$, $\forall i = 0, \dots, 3$ and $g_U = [g_1, g_3]$ with $g_1(x) = x(1 - x)$ and $g_3(x, y) = (y - x)(1 - y)$ (see (2.8)). The optimal minimal bounds for the residual functions are: $\bar{\delta}_j = \max\{\gamma_j, \xi_j\}$, $\forall j = 0, \dots, 2$.

Proof. The proof follows the same arguments exposed in Proposition 2, equivalent to Proposition 8, hence it is omitted. $\square$

4.3.4 Design for two boundary measurements: Fredholm-type Transformation

Consider the adaptive observer design problem for the system (4.1) with two points of measurement at the boundary: $y_0(t) = u(0, t)$ and $y_1(t) = u(1, t)$. In this case, similarly to (4.45), a Luenberger-type structure is proposed, namely

$$\begin{aligned}\hat{u}_t(x, t) &= \epsilon \hat{u}_{xx}(x, t) + \hat{\lambda}(t) \hat{u}(x, t) + \frac{h_0(x)}{\beta} \tilde{u}(0, t) + \frac{h_1(x)}{\beta} \tilde{u}(1, t), \\ \hat{u}_x(0, t) &= p_0 \tilde{u}(0), \\ \hat{u}_x(1, t) &= U(t) + p_1 \tilde{u}(1), \\ \frac{d}{dt} \hat{\lambda}(t) &= \frac{\rho_0}{\beta} S_0 \tilde{u}(0, t) + \frac{\rho_1}{\beta} S_1 \tilde{u}(1, t), \quad \forall t \in [t_a, t_b], \quad \hat{\lambda}(0) = \hat{\lambda}_0, \\ \hat{y}_0(t) &= \hat{u}(0, t), \quad \hat{y}_1(t) = \hat{u}(1, t),\end{aligned}\tag{4.62}$$

where $\hat{u}$ and $\hat{\lambda}$ are the estimate state and reactivity parameter, respectively, for a combination of the measurable output estimation errors $\tilde{u}(0, t) = u(0, t) - \hat{u}(0, t)$ and $\tilde{u}(1, t) = u(1, t) - \hat{u}(1, t)$ via the states observer gains $h_1 = h_1(x)$ and $h_2 = h_2(x)$, $\forall x \in \Omega$. $\hat{u}(x, 0) = \hat{u}_0(x) \in \mathcal{C}(\Omega)$ and $\hat{\lambda}(t_a) = \hat{\lambda}_0$ are the initial conditions for the states and parameter, respectively; $\rho_0 < 0$, $\rho_1 < 0$ and $\beta \neq 0$; S_0 and S_1 are real valued functions; p_0 and p_1 are the observer gains for the boundary conditions. The observer error equation can be written as:

$$\begin{aligned}\tilde{u}_t(x, t) &= \epsilon \tilde{u}_{xx}(x, t) - \underbrace{\left[\frac{h_0(x)}{\beta} \tilde{u}(0, t) + \frac{h_1(x)}{\beta} \tilde{u}(1, t) \right]}_{\Psi(x, t)} + \underbrace{\left(\hat{u}(x, t) \tilde{\lambda}(t) + \tilde{u}(x, t) \hat{\lambda}(t) + \tilde{u}(x, t) \tilde{\lambda}(t) \right)}_{\Delta(x, t)}, \\ \tilde{u}_x(0, t) &= -p_0 \tilde{u}(0, t), \\ \tilde{u}_x(1, t) &= -p_1 \tilde{u}(1, t), \\ \frac{d}{dt} \tilde{\lambda}(t) &= -\frac{\rho_0}{\beta} S_0 \tilde{u}(0, t) - \frac{\rho_1}{\beta} S_1 \tilde{u}(1, t), \quad \forall t \in [t_a, t_b], \quad \tilde{\lambda}(t_a) = \tilde{\lambda}_0,\end{aligned}\tag{4.63}$$

which is a non strict-feedback system (non-causal in space [195]) due to the inclusion of two measurement points, as stated by the term Ψ in the first equation of (4.63); $\Delta = u(x, t) \tilde{\lambda}(t) - \hat{u}(x, t) \hat{\lambda}(t)$. In this case, to achieve the “target” system (4.36), which is assumed to be stable for a suitable selection of the parameter “c”, with $a_0 \neq 0$ and $a_1 \neq 0$, a modification of the Fredholm-type integral transformation (3.53), with integral Kernels $P \in \mathcal{C}^2(\Omega_L)$ and $Q \in \mathcal{C}^2(\Omega_U)$,

$$\begin{aligned}\tilde{u}(x, t) &= \beta \tilde{w}(x, t) - \int_0^x P(x, y) \tilde{w}(y, t) dy - \int_x^1 Q(x, y) \tilde{w}(y, t) dy - q(x) \tilde{\lambda}(t) \\ &= (\beta \mathbb{I} - \mathbb{F})[\tilde{w}(\cdot, t)](x) - q(x) \tilde{\lambda}(t)\end{aligned}\tag{4.64}$$

is proposed, where the function $q \in \mathcal{C}^2(\Omega)$ is included to compensate for the λ parameter uncertainty (the non-modified version of this Fredholm-type transformation has been presented in [205, 204, 82]); $\mathbb{I}$ and $\mathbb{F}$ denote the identity and integral Fredholm-type operator, respectively. In accordance

with the standard methodology in Backstepping observer design for PDEs [21, 50], equivalent to the procedure followed in Sections 4.2.3 and 4.3.3, this transformation leads to:

$$\begin{aligned}
& \tilde{u}_t(x, t) - \epsilon \tilde{u}_{xx}(x, t) + \Psi(x, t) - \Delta(x, t) = \\
& \left(h_0(x) + q(x)\rho_0 S_0 - \epsilon P_y(x, 0) \right) \tilde{w}(0, t) + \left(h_1(x) + q(x)\rho_1 S_1 + \epsilon Q_y(x, 1) \right) \tilde{w}(1, t) + \\
& \left(2\epsilon \frac{d}{dx} P(x, x) - 2\epsilon \frac{d}{dx} Q(x, x) - \beta(c + \hat{\lambda}(t)) \right) \tilde{w}(x, t) + \\
& \left(\epsilon q_{xx}(x) + q(x)\hat{\lambda}(t) - \hat{u}(x, t) - \frac{q(0)}{\beta} (h_0(x) + q(x)\rho_0 S_0) - \frac{q(1)}{\beta} (h_1(x) + q(x)\rho_1 S_1) \right) \tilde{\lambda}(t) + \\
& \cancel{\tilde{w}_x(0, t)}^0 (\epsilon P(x, 0)) - \cancel{\tilde{w}_x(1, t)}^0 (\epsilon Q(x, 1)) + \\
& \int_0^x \left(\epsilon P_{xx}(x, y) - \epsilon P_{yy}(x, y) + P(x, y)(c + \hat{\lambda}(t)) \right) \tilde{w}(y, t) dy + \\
& \int_x^1 \left(\epsilon Q_{xx}(x, y) - \epsilon Q_{yy}(x, y) + Q(x, y)(c + \hat{\lambda}(t)) \right) \tilde{w}(y, t) dy + \\
& \left(\int_0^x P(x, y)\phi(y) dy + \int_x^1 Q(x, y)\phi(y) dy + q(x) - \beta\phi(x) \right) \tilde{\lambda}^2(t) - \\
& \frac{1}{\beta} \left((h_0(x) + \rho_0 S_0 q(x)) \int_0^1 Q(0, y) \tilde{w}(y, t) dy + (h_1(x) + \rho_1 S_1 q(x)) \int_0^1 P(1, y) \tilde{w}(y, t) dy \right) = 0,
\end{aligned} \tag{4.65}$$

and constraints from the boundary conditions in (4.63) given by:

$$\begin{aligned}
\tilde{u}_x(0, t) &= \beta \cancel{\tilde{w}_x(0, t)}^0 + [Q(0, 0) - P(0, 0)] \tilde{w}(0, t) - \int_0^1 Q_x(0, y) \tilde{w}(y, t) dy - q_x(0) \tilde{\lambda}(t) \\
&= -p_0 \tilde{u}(0, t) = -p_0 \beta \tilde{w}(0, t) + p_0 \int_0^1 Q(0, y) \tilde{w}(y, t) dy + p_0 q(0) \tilde{\lambda}(t), \\
\tilde{u}_x(1, t) &= \beta \cancel{\tilde{w}_x(1, t)}^0 + [Q(1, 1) - P(1, 1)] \tilde{w}(1, t) - \int_0^1 P_x(1, y) \tilde{w}(y, t) dy - q_x(1) \tilde{\lambda}(t) \\
&= -p_1 \tilde{u}(1, t) = -p_1 \beta \tilde{w}(1, t) + p_1 \int_0^1 P(1, y) \tilde{w}(y, t) dy + p_1 q(1) \tilde{\lambda}(t).
\end{aligned} \tag{4.66}$$

Due to the complexity of the coupled PDE-ODE (4.65), in particular its last term which involves the interaction of the observer function gains: h_0 , h_1 , q and Kernels: P , Q in the whole domain Ω , the symmetry of its relations allows considering the special case of: $h_0 = h_1 = h$ and $\rho_0 = \rho_1 = \bar{\rho}$, which simplifies (4.65)-(4.66) leading to a more tractable form described by:

$$\begin{aligned}
& \tilde{u}_t(x, t) - \epsilon \tilde{u}_{xx}(x, t) + \Psi(x, t) - \Delta(x, t) = \\
& \left(q(x) - \beta\phi(x) + \int_0^x P(x, y)\phi(y) dy + \int_x^1 Q(x, y)\phi(y) dy \right) \tilde{\lambda}^2(t) + \\
& \delta_0(x) \tilde{\lambda}(t) + \delta_1(x) \tilde{w}(0, t) + \delta_2(x) \tilde{w}(1, t) + \delta_3(x) \tilde{w}(x, t) + \\
& \int_0^x \delta_7(x, y) \tilde{w}(y, t) dy + \int_x^1 \delta_8(x, y) \tilde{w}(y, t) dy = 0,
\end{aligned} \tag{4.67}$$

with constraints:

$$\begin{aligned} P(1, y) + Q(0, y) &= 0, \\ Q_x(0, y) + p_0 Q(0, y) &= 0, \quad P_x(1, y) + p_1 P(1, y) = 0, \\ Q(0, 0) - P(0, 0) &= -p_0 \beta, \quad Q(1, 1) - P(1, 1) = -p_1 \beta, \\ q_x(0) &= p_0 q(0), \quad q_x(1) = -p_1 q(1), \end{aligned} \tag{4.68}$$

$\forall y \in \Omega$, which are equivalent for $S_0 = S_1 = S$, $q(0) = q(1) = \beta\mu/2$ with $\mu \in \mathbb{R}$, at each time t , to the **Kernel-PDE/ODE**:

$$\delta_0(x) = \epsilon q_{xx}(x) + q(x) \left(\hat{\lambda}(t) - \bar{\rho} \mu S \right) - \hat{u}(x, t) - \mu h(x), \tag{4.69}$$

$$\delta_1(x) = h(x) + q(x) \bar{\rho} S - \epsilon P_y(x, 0), \tag{4.70}$$

$$\delta_2(x) = h(x) + q(x) \bar{\rho} S + \epsilon Q_y(x, 1), \tag{4.71}$$

$$\delta_3(x) = 2\epsilon \frac{d}{dx} P(x, x) - \epsilon \frac{d}{dx} Q(x, x) - \beta(c + \hat{\lambda}(t)), \tag{4.72}$$

$$\delta_4(x) = Q_x(0, y) + p_0 Q(0, y) |_{y=x}, \tag{4.73}$$

$$\delta_5(x) = p_1 Q_x(0, y) + p_0 P_x(1, y) |_{y=x}, \tag{4.74}$$

$$\delta_6(x) = P(1, y) + Q(0, y), \tag{4.75}$$

$$\delta_7(x, y) = \epsilon P_{xx}(x, y) - \epsilon P_{yy}(x, y) + (c + \hat{\lambda}(t)) P(x, y), \quad \forall (x, y) \in \Omega_L \tag{4.76}$$

$$\delta_8(x, y) = \epsilon Q_{xx}(x, y) - \epsilon Q_{yy}(x, y) + (c + \hat{\lambda}(t)) Q(x, y), \quad \forall (x, y) \in \Omega_U \tag{4.77}$$

$$q_x(0) = (\mu/2)(Q(0, 0) - P(0, 0)), \quad (\mu/2)\beta = q(0), \tag{4.78}$$

$$q_x(1) = (\mu/2)(Q(1, 1) - P(1, 1)), \quad (\mu/2)\beta = q(1), \tag{4.79}$$

and function $\phi(x)$ satisfying satisfying the Fredholm's-type integral equation of second kind:

$$\frac{q(x)}{\beta} = \phi(x) - \int_0^x \frac{P(x, y)}{\beta} \phi(y) dy - \int_x^1 \frac{Q(x, y)}{\beta} \phi(y) dy, \tag{4.80}$$

$\forall y \in \Omega, \forall x \in \Omega, \beta \neq 0$, for an exact zero matching condition on the residual functions δ_j , $\forall j=0, \dots, 8$. In this case the boundary observer gains are given by $p_0 = -\frac{q_x(0)}{q(0)}$ and $p_1 = -\frac{q_x(1)}{q(1)}$. Regarding the “target” parameter error dynamic stated in (4.36), substituting the Fredholm-type transformation (4.64) evaluated at both boundaries into the “original” parameter error dynamic given in (4.63) for $\rho_0 = \rho_1 = \bar{\rho}$ yields

$$\begin{aligned} \frac{d\hat{\lambda}}{dt}(t) &= -\frac{\bar{\rho}}{\beta} S (\tilde{u}(0, t) + \tilde{u}(1, t)) \\ &= -\frac{\bar{\rho}}{\beta} \left(\left[\beta \tilde{w}(0, t) - \int_0^1 Q(0, y) \tilde{w}(y, t) dy - q(0) \tilde{\lambda}(t) \right] + \left[\beta \tilde{w}(1, t) - \int_0^1 P(1, y) \tilde{w}(y, t) dy - q(1) \tilde{\lambda}(t) \right] \right) \\ &= \frac{\bar{\rho}}{\beta} S (q(0) + q(1)) \tilde{\lambda}(t) - \bar{\rho} S (\tilde{w}(0, t) + \tilde{w}(1, t)) - \frac{\bar{\rho}}{\beta} S \left(\int_0^1 (Q(0, y) + P(1, y)) \tilde{w}(y, t) dy \right) \end{aligned}$$

$$= [\bar{\rho}\mu S] \tilde{\lambda}(t) - \bar{\rho}S\tilde{w}(0, t) - \bar{\rho}S\tilde{w}(1, t),$$

with $\mu = q(0)/(2\beta)$, $q(0) = q(1)$, $\forall t \in [t_a, t_b]$. Thus, considering $\tilde{w}$ as bounded signal and $\bar{\rho} < 0$, a sufficient condition of stability for $\tilde{\lambda}$ in (4.36) is the selection of a function S such that $\mu S > 0$, $\forall t \geq 0$.

With respect to the integral equation (4.80), given a continuous bounded function q and continuous bounded Kernels P and Q , the existence and uniqueness of a bounded continuous function $\phi = \phi(x)$ relies on the invertibility of the Fredholm-type operator $\mathbb{I} - \mathbb{F}$. This has been proved for the Banach and Hilbert spaces of real-analytic functions $\mathcal{A}(\Omega)$ in Theorem 1 and Banach spaces of continuous functions $\mathbb{C}(\Omega)$ in Theorem 2.

4.3.4.1 Kernel-PDE/ODE as a Convex Optimization Problem

Similar to Sections 4.2.3.1 and 4.3.3.1, the Kernel-PDE/ODE (4.69)-(4.79) can be approximately solved based on the approach presented in Chapter 2. For instance, setting up a similar formulation to the convex optimization problem described in Chapter 3, Proposition 6, and computing its solution at “each fixed time”.

Proposition 12. *Let $N(x, y) = \sum_{k=0}^{z(d)-1} n_k x^{\alpha_k} y^{\beta_k}$ and $M(x, y) = \sum_{k=0}^{z(d)-1} m_k x^{\alpha_k} y^{\beta_k}$ be polynomial approximations of P and Q in (4.64), respectively, of arbitrary even degree $d > d_{\hat{u}} \in \mathbb{N}$ in accordance with (1.3). Let $f(x) = \sum_{k=0}^d f_k x^k$ and $l(x) = \sum_{k=0}^d l_k x^k$ be polynomial approximations of q in (4.64) and $h = h_1 = h_2$ in (4.62), respectively. Let $T_j = T_j(x)$, $\forall j = 0, \dots, 6$ and $T_j = T_j(x, y)$, $\forall j = 7, 8$, be polynomials of degree $2d_{\delta_j}$; $\gamma_j \geq 0, \forall j = 0, 8$. Let $\hat{u}$ be estimated by the observer (4.62) and approximated by a polynomial function of degree $d_{\hat{u}}$. The Kernel-PDE (4.69)-(4.79) can be formulated as the “bilinear” convex optimization problem:*

$$\underset{\gamma_j, N, M, T_j, f, l, s_{ij}}{\text{minimize:}} \quad \sum_{j=0}^8 \gamma_j \quad (4.81)$$

$$\text{subject to:} \quad \begin{bmatrix} \gamma_j - T_j(x) - s_{0j}(x, y)g_1(x) & \delta_j(x) \\ \delta_j(x) & \gamma_j \end{bmatrix} \in \Sigma_s, \quad \forall j = 0, \dots, 6 \quad (4.82)$$

$$\begin{bmatrix} 2\gamma_7 - T_7(x, y) - s_{07}(x, y)g_1(x) & \delta_7(x, y) \\ \delta_7(x, y) & \gamma_7 - s_{17}(x, y)g_2(x, y) \end{bmatrix} \in \Sigma_s, \quad (4.83)$$

$$\begin{bmatrix} 2\gamma_8 - T_8(x, y) - s_{08}(x, y)g_1(x) & \delta_8(x, y) \\ \delta_8(x, y) & \gamma_8 - s_{18}(x, y)g_3(x, y) \end{bmatrix} \in \Sigma_s \quad (4.84)$$

$$s_{0j}, s_{17}, s_{18} \in \Sigma_s, \quad \forall j = 0, \dots, 8, \quad (4.85)$$

$$\int_0^1 T_j(x) dx = 0, \quad \forall j = 0, \dots, 6, \quad (4.86)$$

$$\int_0^1 \int_0^x T_7(x, y) dy dx = 0, \int_0^1 \int_x^1 T_8(x, y) dy dx = 0, \quad (4.87)$$

$$\delta_j = \delta_j(x) \Big|_{\substack{q \approx f, h \approx l \\ P \approx N, Q \approx M}}, \quad \forall j = 0, \dots, 6, \quad \text{as in (4.69)-(4.75)}, \quad (4.88)$$

$$\delta_j = \delta_j(x, y) \Big|_{\substack{P \approx N \\ Q \approx M}}, \quad \forall j = 7, 8, \quad \text{as in (4.76)-(4.77)}, \quad (4.89)$$

$$f_x(0) = (\mu/2)(M(0, 0) - N(0, 0)), \quad f(0) = (\mu/2)\theta, \quad (4.90)$$

$$f_x(1) = (\mu/2)(M(1, 1) - N(1, 1)), \quad f(1) = (\mu/2)\theta, \quad (4.91)$$

for some polynomials s_{0j}, s_{17}, s_{18} of degree $d_{\delta_j} - 2, \forall j = 0, \dots, 8$; g_1, g_2 and g_3 according to (2.7)-(2.8). The optimal mean square bounds for the residual functions are: $\int_0^1 \delta_j^2(x) dx \leq \gamma_j^2, \forall j = 0, \dots, 6$; $\int_0^1 \int_0^x \delta_7^2(x, y) dy dx \leq \gamma_7^2$ and $\int_0^1 \int_x^1 \delta_8^2(x, y) dy dx \leq \gamma_8^2$.

Proof. Following similar arguments as the ones given in the proof of Proposition 6, the problem (4.81)-(4.91) can be formulated as a “bilinear” convex optimization problem.⁹ In particular, regarding the optimal mean square bounds for $\gamma_i, \forall i = 1, \dots, 8$, the formulation of the matrix inequalities and the existence of T_i are justified by (3.91) and (3.92)-(3.94) $\forall i = 1, \dots, 8$, in the respective domains Ω_L or Ω_U . The bilinearity appears in the equations (4.73)-(4.74) due to the products of the unknowns p_0 and p_1 with the Kernels P and Q . Therefore, according to the optimization objective (4.81), the root mean square error of each δ_j is minimized. $\square$

Remark 18. The problem (4.81)-(4.91) can be solved by means of the software PENOPT (with the package to solve bilinear matrix inequalities: PENBMI) [249, 250]. To speed up the computation of the solution, alternatively, this optimization problem can be implemented iteratively via gradient methods, in particular, for instance, using the result proposed by Barzilai and Borwein [251, 252].

4.4 Numerical Results¹⁰

4.4.1 System with an Uncertain Boundary Parameter

To implement the observer (4.2), the Kernel PDE-ODE BVP (4.13)-(4.19) has been solved approximately, at each fixed time, in two steps. Firstly, solving the problem (4.13)-(4.17) via the implementation convex optimization problem formulated in Proposition 8. Secondly, based on the polynomial solutions K and q obtained, solving the ODE-BVP (4.18)-(4.19) via Proposition 9. The performance of this observer is depicted in Figure 4.4.1. A polynomial degree $d = 12$ for the kernel K and functions q and p has been selected. The scheme denominated “Obs. 1st order” corresponds to $\sigma_1 = \sigma_2 = 0$, with $\rho_1 = -300$, and $S_1(\cdot) = \text{sign}(\cdot)$. For the scheme denominated “Obs. 2nd order”, $\sigma_1 = 50, \sigma_2 = -1$, with $\rho_1 = -300$ and $\rho_2 = -300$; $S_2(\cdot) = \text{sign}(\cdot)$. For both cases $\epsilon = 1, \lambda = -1.5$ and $c = 15$ have been considered.

⁹Based on the SOS decomposition, the resulting convex optimization problem is formulated via matrix inequalities, which are considered bilinear in terms of the unknowns.

¹⁰The numerical solution of the convex optimization problems has been obtained via the Yalmip toolbox for Matlab [163] and SOSTOOLS [164], using the semi-definite programming solver SeDuMi [165] and the SDP package part of the Mosek solver [166].

The adaptive observer proposed has been compared with a Lyapunov-based adaptive observer [127, 253]. It has been designed based via model reduction of the *Transcendental Transfer Function* [11, 254] of the parabolic PDE (4.3), namely

$$H(s) = \frac{\mathcal{L}(u(1, t))}{\mathcal{L}(\theta(t)U(t))} = \frac{\sinh(\sqrt{(s - \lambda)/\epsilon})}{((s - \lambda)/\epsilon) \cosh(\sqrt{(s - \lambda)/\epsilon}) - \sinh(\sqrt{(s - \lambda)/\epsilon})}, \quad (4.92)$$

where $\mathcal{L}$ denotes the Laplace transformation, considering a 1st order model for the parameter dynamic, i.e., $\sigma_1 = \sigma_2 = 0$. This adaptive observer has the Luenberger structure:

$$\begin{aligned} \dot{\hat{v}}(t) &= A\hat{v}(t) + \hat{\theta}(t)BU(t) + L\tilde{y}(t), \\ \hat{y}(t) &= C\hat{v}(t), \\ \dot{\hat{\theta}}(t) &= \rho U(t)F\tilde{y}(t), \quad \rho > 0, \end{aligned} \quad (4.93)$$

where (A, B, C) is a space state representation of the second order *Padé approximation* of (4.92) around $s = \epsilon + \lambda$ (see Figure 4.4.1(d)). Using the Lyapunov function $V = \tilde{v}(t)^\top P \tilde{v}(t) + (1/\rho)\tilde{\theta}^2(t)$, the observer gains are determined via the solution of the Linear Matrix Inequality/Equality problem:

$$\begin{aligned} A^\top P + PA - C^\top Z^\top - ZC &\leq -\sigma P, \quad P \succeq 0, \sigma > 0, \\ B^\top P &= FC, \quad L = P^{-1}Z. \end{aligned} \quad (4.94)$$

For $\sigma = 10$ and $\rho = 300/\|F\|$, Figures 4.4.1(a)-(c) show the effect of this model reduction on the parameter estimation, considering a control action $U(t) = \sin(30\pi t)$.

Clearly, the 2nd order observer gives better results than the 1st order observer and the Lyapunov-based adaptive observer, in particular, for the transient dynamic of the parameter θ .

4.4.2 System with an Uncertain Reactivity Parameter

The adaptive observer for (4.35) has been implemented via two schemes:

- “Observer 1”: Adaptive observer (4.45), with a unique measured point, based on the Volterra-type operator. The observer gains (4.59)-(4.60) have been determined via the solution of the convex optimization problem formulated in Proposition 11.
- “Observer 2”: Adaptive observer (4.62), with two measured points, based on the Fredholm-type operator. The observer gains have been determined by means of the solution of the Kernel-PDE (4.69)-(4.79) via Proposition 12.

At every sampling time the estimated state $\hat{u} = \hat{u}(x, t)$ has been approximated as a polynomial function of degree 6. For $t = 0.1$, Figure 4.4.2(a)-(b) illustrates the approximate integral Kernels for both adaptive observers, considering polynomial Kernels with degree $d = 16$, $\hat{u}(x) = 3.667 +$

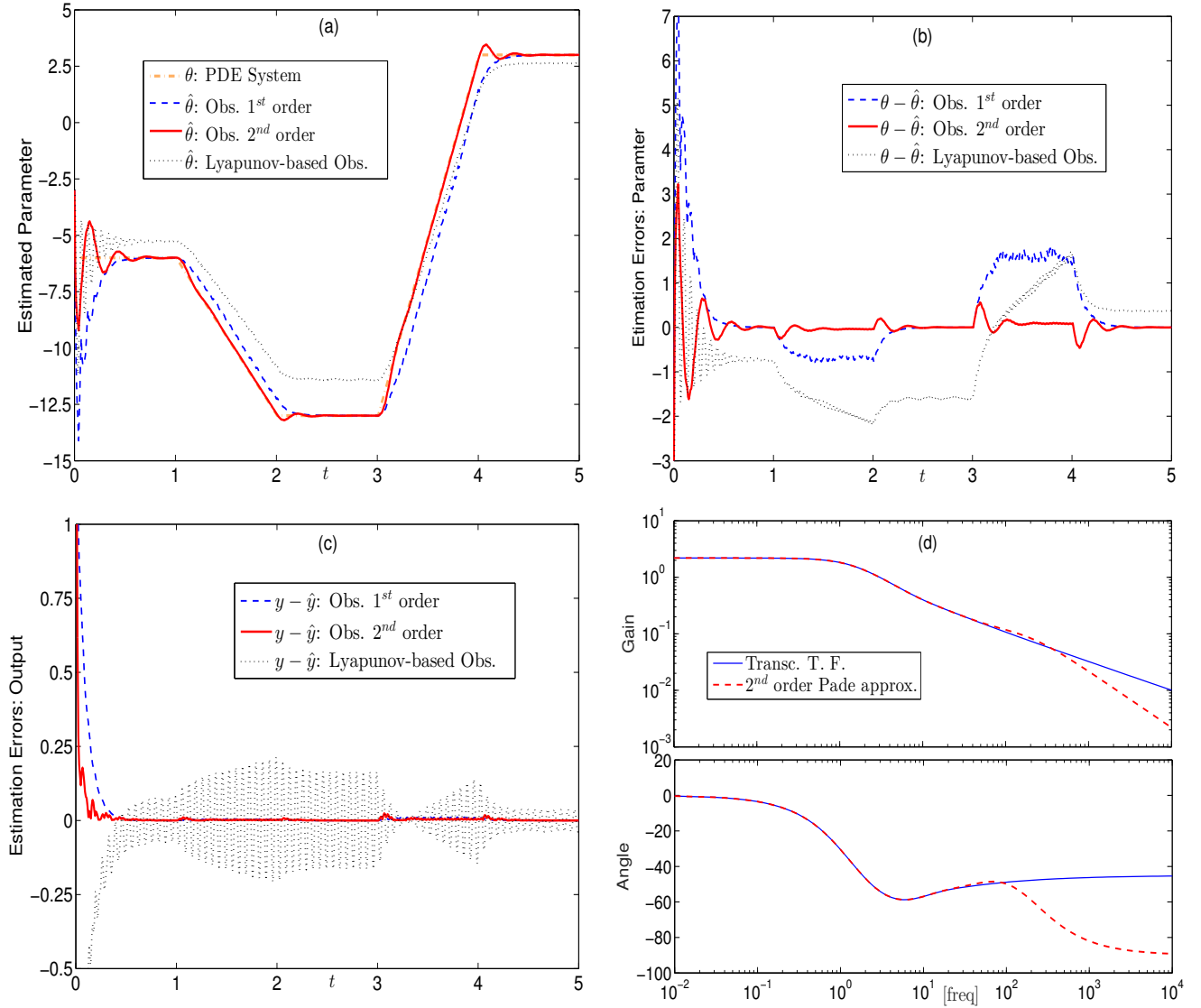


Figure 4.4.1: Adaptive Observer Design for PDE with Boundary Uncertain Parameter. (a), (b) and (c): results from observer 1st order model: dash line, observer 2nd order model: solid line and Lyapunov-based observer design: dot line. True parameter: dash-dot line. (a) Estimated parameters. (b) Parameter estimation errors. (c) Output estimation errors. (d) Bode plot. Second order Padé approximation: dash-line. Transcendental transfer function (4.92): solid line.

$2.42x + 0.15x^2 + 0.82x^3 + 4.87x^4 - 12.75x^5 + 5.41x^6$, $\hat{\lambda} = 1.5$, $c = 17$, $\rho = -0.14 \int_0^1 \hat{u}(x, t) dx$ and $\mu = 2$. For the period of time $t = \overline{0, 3}$, the performance of the simultaneous state and parameter estimation is shown in Figure 4.4.2 (c)-(d), with respect to the mean square state estimation error $E_2(t) = (1/N) \|u(\cdot, t) - \hat{u}(\cdot, t)\|_2^{1/2}$, with $N = 40$ spatial points in Ω , and the estimation of the reactivity parameter λ . For the Observer 1, the parameters $\rho_1 = -0.32$, $S_1 = \int_0^1 \hat{u}(x, t) dx > 0$ and $c = 17$ have been selected. For the Observer 2, lower gains $\bar{\rho} = -0.14$, $S = \text{sing}(\mu) \int_0^1 \hat{u}(x, t) dx$, $c = 17$ and $\mu = 2$ are considered.

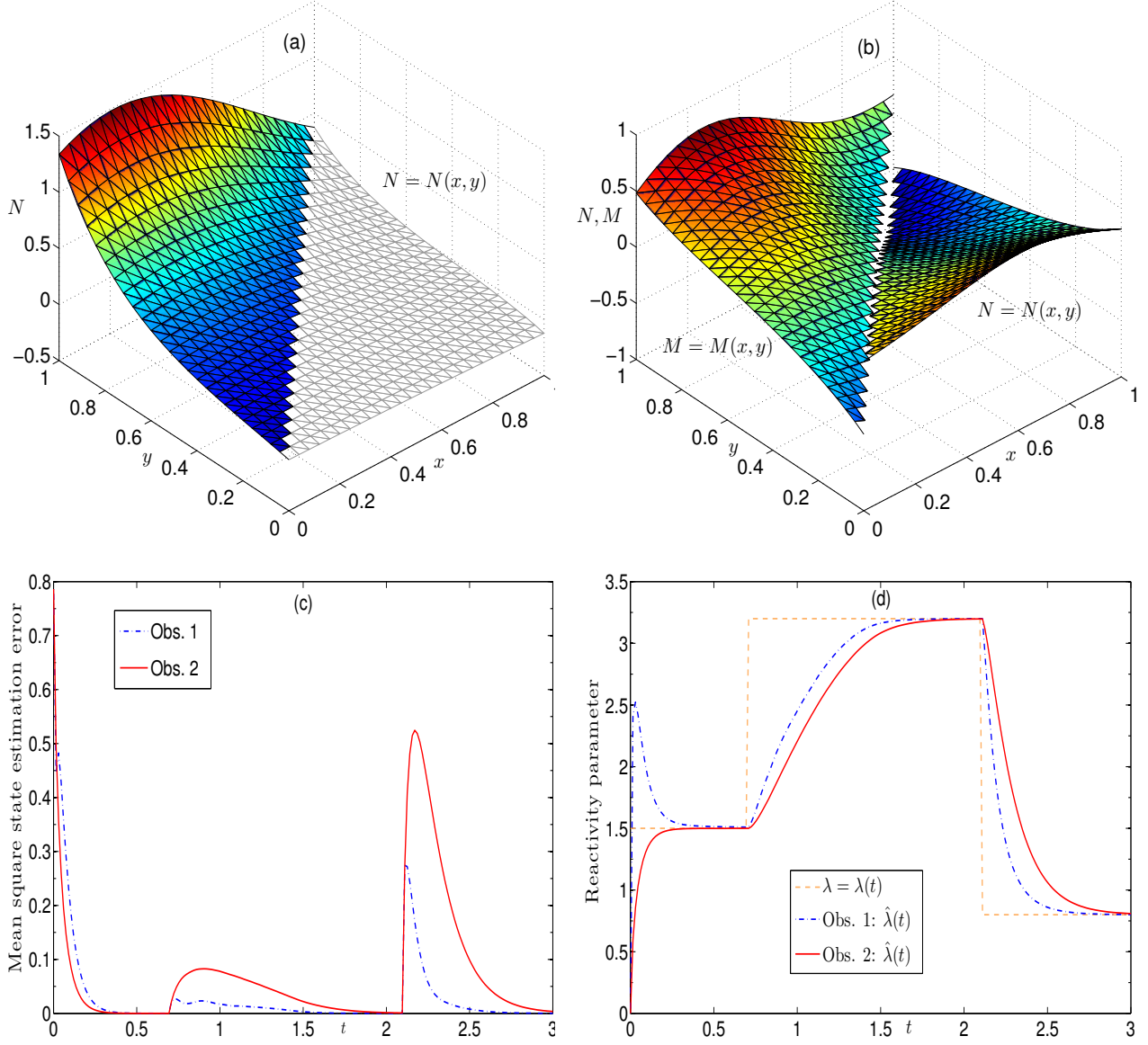


Figure 4.4.2: Adaptive Observer Design for PDE with Reactivity Uncertain Parameter. (a) Approximate Kernel N for the adaptive "Observer 1" (one boundary measurement). Integral Kernel out of the upper triangular domain in grey color. (b) Approximate kernels N and M for the adaptive "Observer 2" (two boundary measurement points). (c)-(d) Simulation results for a step-type change in the reactivity parameter. Adaptive "Observer 1": dash-dot line and "Observer 2": solid line. (c) Mean square state estimation error E_2 . (d) Estimation of the reactivity parameter. True parameter λ (dash line) and its estimate $\hat{\lambda}$ (via "Observer 1" and "2").

Chapter 5

Adaptive Observer for a Model of Lithium-Ion Batteries¹

—Michael Faraday (1791-1867)—

M. Faraday, a man with virtually no formal education, is well known for his contributions to the principles of the “electromagnetic induction” and “electrochemistry”. His is considered as one of the most influential scientists of all time with a legacy which goes further beyond his outstanding experimental investigations, registered in a manuscript which comprises about 3500 pages, from science to technology. For instance, his concepts about electric and magnetic fields were mathematically developed a generation later by James Maxwell, leading to the conception of the special theory of relativity by Albert Einstein. M. Faraday coined many familiar words such as “electrode”, “anode”, “cathode” and “ions”. His intense work can be seen in his advice to a younger scientist to “work, finish, publish”. Despite this, he still found time for visiting and helping poor and sick people as elder of his local church in London. M. Faraday did not commercialize his discoveries of the principles of the electric motor, the dynamo, the transformer, etc., which have led to enormous changes in our daily lives.

In this Chapter the observer design problem for the simultaneous estimation of the solid Lithium concentration and of the diffusivity parameter, for a model of Lithium-Ion batteries, is addressed. The battery model is based on electrochemical principles, formulated via PDEs and nonlinear functions, and considers a “single solid particle” as representative element of its electrodes. Following the design proposed in Chapter 4, the adaptive observer design is based on the Backstepping PDE methodology, including a modified Volterra-type transformation to compensate for the diffusivity uncertainty. The resulting coupled Kernel-PDE/ODE is solved using the convex optimization methodology proposed in Chapter 2 and formulated in Chapter 3. In addition, a simple variation of the Volterra-type transformation is studied, which leads to an uncoupled Kernel-PDE and ODE-BVP formulation. This allows computing the state and parameter observer gains, at each fixed time, by means of the well-known closed-form solution of the Backstepping Kernel-PDE and an efficient numerical solution for the ODE-BVP involved. On the other hand, a novel scheme of inversion of the nonlinear output

¹This chapter is based on the publication [255].

mapping, which relates the measured voltage and internal variables of the model, is presented. The effectiveness of this approach is illustrated by numerical simulations.

5.1 Introduction

Physics-based electrochemical models provide an accurate prediction of the internal dynamics of Lithium-ion batteries. The most established model of this type, the so-called pseudo two-dimensional (P2D) model, also known as Dolez-Fuller-Newman (DFN) model [256], includes both macro-scale dynamics in the electrolyte and micro-scale diffusion of Lithium in spherical particles in the porous active material, which are formulated by a set of nonlinear Partial Differential Algebraic Equations (PDAEs). Due to its complexity [257] and the significant computational burden [258], more tractable models have been used for control, identification and state estimation [10, 259]. For instance:

- **Equivalent-Circuit Models:** Common electronic elements are used to define a circuit that closely matches the measured behaviour of the battery cell. In general these models cannot predict long-term behaviour of a battery [10].
- **Single Particle Models:** Phenomenological electrochemical-based model which considers a single particle of active material to describe the main dynamics in the electrodes (see [260] and the references cited therein).

Despite the simple formulation of the *Single Particle Model* (SPM) and consequently its limited operating range to low current rates, this model captures the *main fundamental challenges* of the full DFN model [260, 261]. One of these consists in the estimation of the electrochemical internal state variables (which are not directly measurable) from measurements such as current, voltage and temperature, which ultimately leads to the estimation of the State-of-Charge (SOC) of the battery.

In this context two problems are addressed in this chapter:

- *Estimation of the Diffusivity Coefficient:* Uncertain parameters inherently characterize battery models [262]. In particular, the diffusivity coefficient [11], which plays one of the main roles to describe the dynamic of insertion/extraction of Lithium ions in/from solid particles. This parameter can be time/physically dependent on the internal states of the battery [263, 264].
- *Inversion of the Nonlinear Output Mapping:* To estimate the Lithium-ion concentrations, PDE-based linear observers mostly use particle surface values as feedback signals (boundary values of the PDE). These variables are not physically accessible and are related to the output measured variables (voltage, current and temperature) via nonlinear functions, which describe the exchange of flux of ions in terms of the overpotentials along the battery.

As far as the state estimation is concerned, Backstepping PDE observer design [21, 86, 50] provides a simple and “fully” infinite-dimensional approach for addressing the spatial and temporal estimation of Lithium-ion concentrations [262, 11, 264, 265, 266]. The observer mimics the structure of the classical Luenberger-type state observer, mapping the error equation into a stable target PDE system via output injection gains, which, in turn, are computed as functions of the integral Kernels governed by the resulting Kernel-PDE [86, 50].

Under parameter uncertainty, as it has been commented in Chapter 4, adaptive methods in Backstepping PDE design mostly need measurements of the states on the full spatial domain [50, 122, 123, 124] to provide a convergent parameter estimation. On the contrary, its input-output observer design approach is limited to a particular system structure denominated “canonical observer form” [125, 126, 82]. This structure includes additive terms where the unknown parameters and boundary measurements are factors, which allows formulating a direct parameter estimation. However, this structure is not present in PDE-based battery models, as in many general cases.

On the other hand, Backstepping PDE observer design needs a linear output feedback from the boundary measurements [21, 50]. For PDE-based battery models this involves the *inversion* of the nonlinear output mapping so that the Lithium-ion surface concentration is formulated as a function of the measured variables (voltage and current). This inversion has commonly been carried out by linear approximations (via Taylor’s series representation) [11], and therefore its performance is limited by the gradient computation of the equilibrium potentials. For some batteries chemistries these potentials exhibit a monotonic behavior so that it is feasible to compute the gradient, even if this has a small magnitude. However, for some electrode materials such as graphite, common in Lithium-ion batteries, these potentials have a nearly zero value ($< 10^{-6}$) in wide zones of its operational range [261, 267].

In this chapter above-described problems are studied via a systematic application of the results obtained through Chapters 2, 3 and 4, namely exploiting the properties of polynomial optimization based on SOS decomposition and Moment Theory.

The main objectives addressed in this chapter are:

- (a) perform a simultaneous estimation of the Lithium-ion concentration and diffusivity uncertain parameter considering only measurements of voltage and current,
- (b) develop a nonlinear mapping inversion scheme, avoiding gradient-based computations, providing an accurate estimation of the Lithium-ion surface concentration to be used by the observer feedback.

To describe the battery cell behaviour electrochemical-based models commonly consider **three physical domains**:

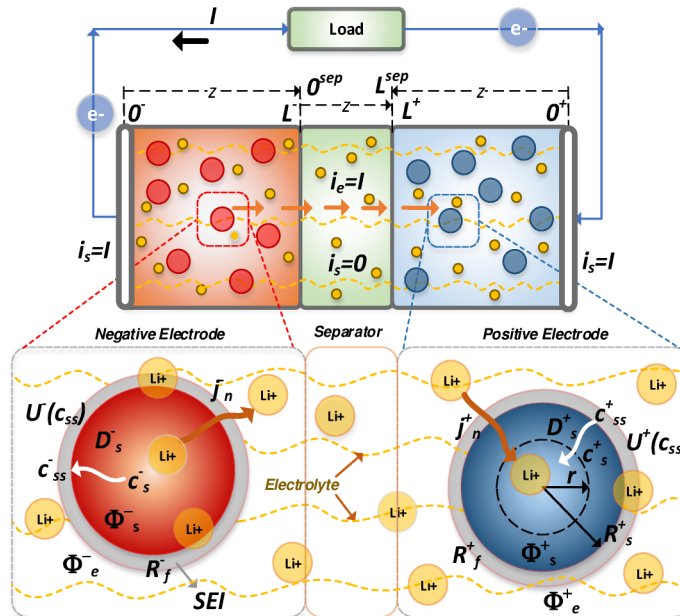


Figure 5.2.1: Schematic of an Intercalation Cell and its Single Particle Model.

- *Negative Electrode*: This usually is a pure metal or an alloy. Most of Lithium-ion batteries currently has *Graphite* as principal element. However, to expand the capacity of the cell, elements such as *Silicon* has been included as an additional intercalation material.
- *Separator*: This layer is a thin porous medium, which is an ionic conductor but electronically separates both electrodes, i.e., this only allows ionic charge transfers preventing internal short circuits and rapidly self-discharge.
- *Positive Electrode*: This include compounds of various chemistries, often metallic oxides of its blend, such as Li_xCoO_2 , $\text{Li}_x\text{Mn}_2\text{O}_4$ and Li_xFePO_4 . The Lithium ions are the main components taking part of the oxidation and reduction chemical reactions and of the mass transport between electrodes.

This model also considers **two disjoint phases (states)** for Lithium ions (porous electrode theory [268]):

- *Solid*: The agglomeration of lattice sites can be modeled as spherical *solid* particles capable of storing Lithium. In the *Electrodes*, the transfer of Lithium ions is modeled as intercalation process, i.e., an insertion or extraction from interstitial places of these *porous solid* particles, which are immerse in the *Electrolyte*.

- *Dissolved*: The *Electrodes* and the *Separator* are all immerse in the *Electrolyte* filling all pores of the solid material. When a Lithium ion leaves a solid particle, this is considered dissolved in the electrolyte phase. In the *Separator*, Lithium ions only exist dissolved in the *Electrolyte*.
 - *Electrolyte*: This is a concentrated solution, solid or liquid, which is the medium for a transfer of ionic charge between electrodes. This can also act as an electronic insulator. Its potential gradient is not negligible for currents above 0.5C.²

The DFN model [256] describes the dynamics of a Lithium-ion battery via a set of PDAEs (see Appendix C for a summary of these equations, description of variables and parameters associated). This mainly involves 5 coupled equations to describe the movement of Lithium ions and potentials in the solid and electrolyte states, and the rate of reaction at the solid-electrolyte boundary. Since this model considers that Lithium can exist at every point along the battery, either in the solid phase or in the dissolved state, the intercalation dynamics has to be solved at every point of the spacial domain with the consequent high computational burden and complexity of the model.

A simplification of this model is the so-called *Single Particle Model* (SPM). This considers that the diffusion in the solid particles is the slowest process so that its dynamic dominates over the transport in the electrolyte. Thus, each electrode can be modeled via a *single* representative spherical particle of active material where the intercalation occurs (see Figure 5.2.1) (see [11, 260] and the references cited therein for an in-depth description). This simplification involves the following assumptions:

- (i) the electrolyte dynamic is neglected considering the Lithium-ion concentrations in the electrolyte phase constant and known ($c_e(z, \tau) = c_e^0$),
- (ii) there is no spatial variation of the Lithium-ion concentrations in the solid phase ($c_s^\pm(z, r, \tau) = c_s^\pm(r, \tau)$, $j_n^\pm(z, \tau) = j_n^\pm(\tau)$),
- (iii) the moles of Lithium ions are preserved in the electrolyte ($\frac{d\eta_{Li,e}}{d\tau} = 0$) and in the solid phase ($\frac{d\eta_{Li,s}}{d\tau} = 0$),

where the variable z is the spatial position, r is the radial spherical coordinate and τ denotes time; c_s is the Lithium-ion concentration in the solid particle, c_e is its concentration in the electrolyte and j_n is pore-wall molar ion flux; the upper-index “ $\pm$ ” relates a variable to its spatial domain, i.e., “ $+$ ”: positive electrode and “ $-$ ”: negative electrode; $\eta_{Li,e}$ and $\eta_{Li,s}$ stand for the total moles of Lithium ions in the electrolyte and solid particle state, respectively (refer to Appendix C, Tables C.1 and C.2, for a list of variables and parameters, respectively).

These assumptions lead to a *proportionally* between the pore-wall molar ion flux ($j_n^\pm$) and the current applied to the battery: I , namely

²C-rate: ratio of the current (A) to the nominal battery capacity (A-h). For instance if a battery has a nominal capacity of 2 (A-h), 0.5C correspond to a current magnitude of 1(A). If this battery supplies 1C of current per 1 hour, after that it will be fully discharge.

$$\frac{\partial c_e}{\partial z} = 0 = \frac{\partial c_e}{\partial \tau} \Rightarrow i_e(\tau) = i_e, \quad c_e(\tau) = c_e^0; \quad \frac{\partial i_e^l}{\partial z}(z, \tau) = a^l F j_n^l(z, \tau),$$

- Positive Electrode:

$$i_e(0^+, \tau) = 0, \quad i_e(L^+, \tau) = -I(\tau),$$

$$\int_{0^+}^{L^+} \frac{\partial i_e(z, \tau)}{\partial z} dz = -I(\tau) = \int_{0^+}^{L^+} a^+ j_n^+(\tau) dz \Rightarrow I(\tau) = -a^+ L^+ F j_n^+(\tau),$$

- Negative Electrode:

$$i_e(L^-, \tau) = I(\tau), \quad i_e(0^-, \tau) = 0,$$

$$\int_{0^-}^{L^-} \frac{\partial i_e(z, \tau)}{\partial z} dz = I(\tau) = \int_{0^-}^{L^-} a^- j_n^-(\tau) dz \Rightarrow I(\tau) = a^- L^- F j_n^-(\tau).$$

derived from the property of *conservation of charge* (see Appendix C, equation (C.6)), where a is the particle specific interfacial area, L is the layer thickness and F denotes the Faraday's constant.

Thus, the resulting PDE model for the intercalation process of Lithium ions in the solid particle is given by the diffusion dynamic (see Appendix C, equation (C.4)) :

$$\begin{aligned} \frac{\partial c_s^l}{\partial \tau}(r, \tau) &= \frac{1}{r^2} \frac{\partial}{\partial r} \left(D_s^l(\tau) r^2 \frac{\partial c_s^l}{\partial r}(r, \tau) \right), \quad r \in (0, R_s^l) \\ \frac{\partial c_s^l}{\partial r}(0, \tau) &= 0, \\ \frac{\partial c_s^l}{\partial r}(R_s^l, \tau) &= l \frac{I(\tau)}{D_s^l(\tau) a^l L^l F}, \end{aligned} \tag{5.1}$$

where the upper index l associates the Lithium-ion concentration in the solid phase c_s to an electrode ($l = \{+, -\}$, positive electrode and negative electrode, respectively), R_s^l is the radius of the particles and D_s^l stands for the diffusivity parameter considered as a *piece-wise constant function of time*.

Similarly, the assumptions (i)-(iii) lead to a model output which corresponds to a nonlinear mapping between the voltage V (measurable) and the Lithium-ion concentration on the particle surface $c_{ss}^l(\tau) = c_s^l(R_s^l, \tau)$ (non-measurable) described by:

$$V(\tau) = \eta_s^+(\tau) - \eta_s^-(\tau) + U^+(c_{ss}^+(\tau)) - U^-(c_{ss}^-(\tau)) + F R_f^+ j_n^+(\tau) - F R_f^- j_n^-(\tau), \tag{5.2}$$

$$\eta_s^l(\tau) = \frac{2RT}{F} \sinh^{-1} \left(-l \frac{I(\tau)}{2a^l L^l i_o^l(\tau)} \right), \tag{5.3}$$

$$i_o^l(\tau) = k^l \left[c_e^0 c_{ss}^l(\tau) \left(c_{s, \max}^l - c_{ss}^l(\tau) \right) \right]^{1/2}, \tag{5.4}$$

$$j_n^l(\tau) = -l \frac{I(\tau)}{a^l L^l F}, \tag{5.5}$$

where η_s^l corresponds to the solid-phase intercalation reaction overpotential, i_o^l is the exchange current density and U^l stands for the Open-Circuit equilibrium Potential (OCP).

5.3 Adaptive Observer Design

The objective of the observer design problem is to carry out a simultaneous estimation of the uncertain parameter D_s^l and state variable c_s^l in (5.1) from measurements of V and I according to (5.2)-(5.5). The design hereafter considered is based on a slight modification of the Backstepping PDE design methodology presented in Chapter 4.

5.3.1 SPM Formulation for Observer Design

To simplify the observer design the scale and state transformation [11, 265]:

$$x = \frac{r}{R_s^l}, \quad t = \frac{\overline{D}_s^l \tau}{(R_s^l)^2}, \quad c_n^l(x, t) = x \frac{c_s^l(x, t)}{c_{s,\max}^l}, \quad (5.6)$$

are carried out,³ where $\overline{D}_s^l$ denotes a (known) *nominal value of the diffusion parameter*, so that the transformed (normalized) states satisfy $c_n^l \in [0, 1]$ and the radial spatial variable satisfies $x \in [0, 1]$. The transformed system from (5.1) via (5.6) is a parabolic PDE model with mixed boundary conditions, namely

$$\begin{aligned} \frac{\partial c_n^l}{\partial t}(x, t) &= \epsilon^l(t) \frac{\partial^2 c_n^l}{\partial x^2}(x, t), \\ c_n^l(0, t) &= 0, \\ \frac{\partial c_n^l}{\partial x}(1, t) &= c_n^l(1, t) + \overbrace{\left(\frac{1}{\epsilon^l(t)} \right)}^{\theta^l(t)} \overbrace{\left(l \frac{R_s^l}{\overline{D}_s^l a^l L^l F c_{s,\max}^l} \right)}^{g^l} I(t), \\ \frac{d}{dt} \epsilon^l(t) &= 0, \quad \forall t \in [t_a, t_b], \quad \epsilon^l(0) = \epsilon_0^l, \\ y(t) &= c_n^l(1, t) = \varphi(V(t), I(t)), \end{aligned} \quad (5.7)$$

where $\epsilon^l(t) = (D_s^l(t)/\overline{D}_s^l)$ represents the *uncertain diffusivity ratio* parameter (piece-wise constant function of time), the reciprocal of which $\theta^l(t)$ appears as part of the gain of the current density I at one boundary condition. The model output $y(t) = c_{n,ss}^l(t) = \varphi(V(t), I(t))$ corresponds to the inverse of the nonlinear mapping (5.2).

5.3.2 Target System

Proposition 13. *Consider the following “target” coupled PDE-ODE:*

$$\begin{aligned} \tilde{w}_t(x, t) &= \epsilon(t) \tilde{w}_{xx}(x, t) - c\epsilon(t) \tilde{w}(x, t) - \psi(x) \tilde{\epsilon}(t) \tilde{w}(1, t) - \phi(x) \tilde{\epsilon}^2(t), \\ \tilde{w}(0, t) &= 0, \quad \tilde{w}_x(1, t) = 0, \\ \tilde{\epsilon}_t(t) &= -\rho \tilde{\epsilon}(t) - a_1 w(1, t), \end{aligned} \quad (5.8)$$

³The new coordinates x and t depend on the electrode parameters. It is worth noting that, between the negative and positive electrode diffusion dynamics, these new coordinates could have a considerable different range. For notational clarity this dependence is omitted.

with initial conditions $\tilde{w}(x, 0) = \tilde{w}_0$ and $\tilde{\epsilon}(0) = \tilde{\epsilon}_0$; $\rho > 0$, $a_1 \in \mathbb{R}$, $\psi(x) \in \mathcal{L}^2(\Omega)$ and $\phi \in \mathcal{L}^2(\Omega)$. Assuming ϵ such that: $0 < \epsilon_0 \leq \epsilon \leq M$, with ϵ_0 and M a known bounds, if the conditions

$$c > \frac{M^2}{2\epsilon_0} \left(\sigma_3 + \frac{1}{\sigma_1} \int_0^1 \psi^2(x) dx \right) + \frac{M\alpha}{\epsilon_0 \sigma_2} \left(1 + \frac{1}{\sigma_2} \right) - (1 - \alpha) \frac{\pi^2}{4}, \quad (5.9)$$

$$\rho \geq \frac{a_1^2}{2\sigma_0} + \frac{1}{2\sigma_3} \int_0^1 \phi^2(x) dx, \quad (5.10)$$

are satisfied for some selected constants $0 < \alpha \leq 1$, $\sigma_0 > 0$, $\sigma_1 > 0$, $\sigma_2 \geq 2 \left(\frac{\alpha M}{\sigma_0 + \sigma_1} \right)$ and $\sigma_3 > 0$, then target system (5.8) is stable.

Proof. Let $V = \frac{1}{2} \int_0^1 \tilde{w}^2(x, t) dx + \frac{1}{2} \tilde{\epsilon}(t)^2$ be a Lyapunov functional.⁴ Its time-derivative $V_t = \int_0^1 \tilde{w}(x) \tilde{w}_t(x) dx + \tilde{\epsilon} \tilde{\epsilon}_t$ along the trajectory (5.8) is given by:

$$\begin{aligned} V_t &= \int_0^1 \tilde{w}(x) \left(\epsilon \tilde{w}_{xx}(x) - c \epsilon \tilde{w}(x) \right) dx - \tilde{\epsilon} \tilde{w}(1) \int_0^1 \tilde{w}(x) \psi(x) dx - \tilde{\epsilon}^2 \int_0^1 \tilde{w}(x) \phi(x) dx + \tilde{\epsilon} \tilde{\epsilon}_t, \\ &\leq \underbrace{\epsilon \tilde{w}(x) \tilde{w}_x(x) \Big|_{x=0}^{x=1}}_{B_0} \xrightarrow{0} \underbrace{\epsilon \int_0^1 \tilde{w}_x^2(x) dx - c \epsilon \int_0^1 \tilde{w}^2(x) dx}_{T_0} + \tilde{\epsilon} (-\rho \tilde{\epsilon} - a_1 \tilde{w}(1)) + \\ &\quad \tilde{\epsilon} \tilde{w}(1) \int_0^1 \tilde{w}(x) \psi(x) dx + \tilde{\epsilon}^2 \int_0^1 \tilde{w}(x) \phi(x) dx, \\ &\leq -\epsilon \alpha \int_0^1 \tilde{w}_x^2(x) dx - \epsilon \left(c + (1 - \alpha) \frac{\pi^2}{4} \right) \int_0^1 \tilde{w}^2(x) dx - \rho \tilde{\epsilon}^2 + \\ &\quad \underbrace{|\tilde{w}(1)| |a_1 \tilde{\epsilon}|}_{T_1} + \underbrace{|\tilde{w}(1)| \left| \tilde{\epsilon} \int_0^1 \tilde{w}(x) \psi(x) dx \right|}_{T_2} + \underbrace{\left| \int_0^1 \tilde{w}(x) \phi(x) dx \right| \tilde{\epsilon}^2}_{T_3}, \end{aligned} \quad (5.11)$$

where the term T_0 has been split by some factor $0 < \alpha < 1$ and Wirtinger's inequality [50] has been applied on one of its resulting expressions. Thus, applying the Cauchy-Schwarz integral inequality [239] on each term: T_1 , T_2 and T_3 , Young's inequality (special case) [21] on its resulting relations for some $\sigma_0 > 0$, $\sigma_1 > 0$ and $\sigma_3 > 0$, and considering $|\tilde{\epsilon}| \leq M$ yields

$$\begin{aligned} V_t &\leq -\epsilon \alpha \int_0^1 \tilde{w}_x^2(x) dx - \epsilon \left(c + (1 - \alpha) \frac{\pi^2}{4} \right) \int_0^1 \tilde{w}^2(x) dx - \rho \tilde{\epsilon}^2 + \frac{1}{2} \left(\sigma_0 \tilde{w}^2(1) + \frac{1}{\sigma_0} a_1^2 \tilde{\epsilon}^2 \right) + \\ &\quad \frac{1}{2} \left(\sigma_1 \tilde{w}^2(1) + \frac{1}{\sigma_1} M^2 \int_0^1 \psi^2(x) dx \int_0^1 \tilde{w}^2(x) dx \right) + \frac{1}{2} \left(\sigma_3 M^2 \int_0^1 \tilde{w}^2(x) dx + \frac{1}{\sigma_3} \tilde{\epsilon}^2 \int_0^1 \phi^2(x) dx \right). \end{aligned} \quad (5.12)$$

With respect to the boundary term $\tilde{w}(1, t)$ in (5.12), via integration by parts of $\frac{d}{dx} (x \tilde{w}^2(x))$ along $x \in [0, 1]$, and using the Cauchy-Schwarz integral inequality in conjunction with Young's inequality leads to:

⁴For a clearer description, the time-dependence in the functions is dropped ($\tilde{w}(x) \equiv w(x, t)$, $\epsilon \equiv \epsilon(t)$ and $\tilde{\epsilon} \equiv \tilde{\epsilon}(t)$).

$$\begin{aligned}
 \tilde{w}^2(1) &= \int_0^1 \tilde{w}^2(x) dx + 2 \int_0^1 x \tilde{w}(x) \tilde{w}_x(x) dx, \\
 &\leq \int_0^1 \tilde{w}^2(x) dx + 2 \left(\left(\max_{x \in [0,1]} x^2 \right)^{1/2} \right) \left(\int_0^1 \tilde{w}^2(x) dx \right)^{1/2} \left(\int_0^1 \tilde{w}_x^2(x) dx \right)^{1/2}, \\
 &\leq \left(1 + \frac{1}{\sigma_2} \right) \int_0^1 \tilde{w}^2(x) dx + \sigma_2 \int_0^1 \tilde{w}_x^2(x) dx,
 \end{aligned} \tag{5.13}$$

for some $\sigma_2 > 0$. Thus, substituting (5.13) into (5.12) and re-grouping terms is verified:

$$\begin{aligned}
 V_t &\leq - \left[\alpha \epsilon - \left(\frac{\sigma_0 + \sigma_1}{2} \right) \sigma_2 \right] \int_0^1 \tilde{w}_x^2(x) dx - \left[\epsilon \left(c + (1 - \alpha) \frac{\pi^2}{4} \right) - \frac{M^2}{2} \left(\sigma_3 + \frac{1}{\sigma_1} \int_0^1 \psi^2(x) dx \right) - \right. \\
 &\quad \left. \frac{\sigma_0 + \sigma_1}{2} \left(1 + \frac{1}{\sigma_2} \right) \right] \int_0^1 \tilde{w}^2(x) dx - \tilde{\epsilon}^2 \left(\rho - \frac{a_1^2}{2\sigma_0} - \frac{1}{2\sigma_3} \int_0^1 \phi^2(x) dx \right) \leq 0,
 \end{aligned} \tag{5.14}$$

if conditions (5.9) and (5.10) are satisfied for $\frac{\sigma_0 + \sigma_1}{2} \leq \frac{\epsilon \alpha}{\sigma_2} \leq \frac{\alpha M}{\sigma_2}$, which are sufficient conditions for the stability of (5.8). $\square$

5.3.3 Design via the Volterra-type Transformation

The proposed adaptive observer for the system (5.7) with respect to the negative electrode diffusion dynamic⁵ has a Luenberger-type collocated structure [86, 21], namely

$$\begin{aligned}
 \frac{\partial \hat{c}_n}{\partial t}(x, t) &= \hat{\epsilon}(t) \frac{\partial^2 \hat{c}_n}{\partial x^2}(x, t) + h(x) \tilde{c}_n(1, t), \\
 \frac{\partial \hat{c}_n}{\partial x}(0, t) &= 0, \\
 \frac{\partial \hat{c}_n}{\partial x}(1, t) &= \hat{c}_n(1, t) + \hat{\theta}(t) \varrho I(t) + p_b \tilde{c}_n(1, t), \\
 \frac{d\hat{\epsilon}}{dt}(t) &= \rho_1 S(q(1)) \tilde{c}_n(1, t), \quad \hat{\epsilon}(0) = \hat{\epsilon}_0, \\
 \hat{y}(t) &= \hat{c}_n(1, t),
 \end{aligned} \tag{5.15}$$

where $\hat{c}_n$ and $\hat{\epsilon}$ are the estimated Lithium-ion solid concentration and diffusivity ratio parameter in the negative electrode, respectively; $\tilde{c}_n(1, t) = y(t) - \hat{c}_n(1, t)$ denotes the output estimation error considering as unique indirect measurable point $y(t) = c_n(1, t)$ from the inversion of (5.2); $\rho_1 < 0$; S is a real valued function and $q \in \mathcal{C}^2(\Omega)$. The function $h = h(x)$ and the constant p_b are the observer feedback gains for the state and the boundary condition, respectively. The observer error equation can be written as:

$$\begin{aligned}
 \frac{\partial \tilde{c}_n}{\partial t}(x, t) &= \underbrace{\left[\epsilon(t) \frac{\partial^2 \tilde{c}_n}{\partial x^2}(x, t) + \tilde{\epsilon}(t) \frac{\partial^2 \hat{c}_n}{\partial x^2}(x, t) \right]}_{\Delta(x, t)} - h(x) \tilde{c}_n(1, t) \\
 \tilde{c}_n(0, t) &= 0,
 \end{aligned} \tag{5.16}$$

⁵For the sake of notational simplicity the negative electrode upper index in the variables is dropped.

$$\begin{aligned}\frac{\partial \tilde{c}_n}{\partial x}(1, t) &= (1 - p_b)\tilde{c}_n(1, t) + \tilde{\theta}(t)qI(t), \\ \frac{d\tilde{\epsilon}}{dt}(t) &= -\rho_1 S(q(1))\tilde{c}_n(1, t), \quad \forall t \in [t_a, t_b], \quad \tilde{\epsilon}(0) = \tilde{\epsilon}_0,\end{aligned}$$

where $\tilde{c}_n(x, t) = c_n(x, t) - \hat{c}_n(x, t)$ is the concentration estimation error, $\tilde{\epsilon}(t) = \epsilon(t) - \hat{\epsilon}(t)$ is the diffusivity ratio estimation error and $\tilde{\theta}(t) = (1/\epsilon(t) - 1/\hat{\epsilon}(t))$. The term⁶ $\Delta(x, t) = \epsilon(t)\frac{\partial^2 c_n}{\partial x^2}(x, t) - \hat{\epsilon}(t)\frac{\partial^2 \hat{c}_n}{\partial x^2}(x, t)$.

The backstepping PDE-based adaptive observer design for the system (5.7) consists in transforming the “original” estimation error system (5.16) into the “target” system (5.8), which is assumed to be stable. To achieve this state mapping, in accordance with the integral transformation (4.47) proposed in Chapter 4, the modified Volterra-type transformation

$$\tilde{c}_n(x, t) = q(1)\tilde{w}(x, t) - \int_x^1 P(x, y)\tilde{w}(y, t)dy - q(x)\tilde{\epsilon}(t) \quad (5.17)$$

with Kernel $P \in \mathcal{C}^2(\Omega \times \Omega)$ is proposed. The function q is included to reject the time-varying uncertainty on the diffusivity ratio parameter. In addition, to simplify the boundary condition in (5.16) $\tilde{\theta}(t)$ is approximated by a first order Taylor approximation, i.e., $\tilde{\theta}(t) \approx -\tilde{\epsilon}(t)/\hat{\epsilon}^2(t)$.

Remark 19. A higher order Taylor’s approximation of $\tilde{\theta}$ in (5.16) can also be considered. The approximation error of this term implies that, in Proposition (13), Lyapunov time-derivative (5.11), one boundary condition in the term B_0 does not vanish. However, this residual term can included in the main expressions via (5.13) to be compensated by a suitable selection of the parameter c .

Following the standard procedure of Backstepping PDE observer design [21, 50] applied in Chapter 4, Sections 4.2.3 and 4.3.3, the transformation leads to:

$$\begin{aligned}\frac{\partial \tilde{c}_n}{\partial t}(x, t) &= q(1)\tilde{w}_t(x, t) - \int_x^1 P(x, y)\tilde{w}_t(y, t)dy - q(x)\tilde{\epsilon}_t(t), \\ &= q(1)\left(\epsilon(t)\tilde{w}_{xx}(x, t) - c\epsilon(t)\tilde{w}(x, t) - \psi(x)\tilde{\epsilon}(t)\tilde{w}(1, t) - \phi(x)\tilde{\epsilon}^2(t)\right) - \underbrace{\int_x^1 \epsilon(t)P(x, y)\tilde{w}_{yy}(y, t)dy}_{T_0} - \\ &\quad \epsilon(t)\int_x^1 cP(x, y)\tilde{w}(y, t)dy + \tilde{\epsilon}(t)\tilde{w}(1, t)\int_x^1 P(x, y)\psi(y)dy + \tilde{\epsilon}^2(t)\int_x^1 P(x, y)\phi(y)dy - q(x)\tilde{\epsilon}_t(t), \\ &= q(1)\left(\epsilon(t)\tilde{w}_{xx}(x, t) - c\epsilon(t)\tilde{w}(x, t)\right) + \epsilon(t)\int_x^1 (cP(x, y) - P_{yy}(x, y))\tilde{w}(y, t)dy + \quad (5.18) \\ &\quad \epsilon(t)\frac{\partial}{\partial y}P(x, 1)\tilde{w}(1, t) - \epsilon(t)\frac{\partial}{\partial y}P(x, x)\tilde{w}(x, t) - \cancel{\epsilon(t)P(x, 1)\tilde{w}_x(1)} \xrightarrow{0} \epsilon(t)P(x, x)\tilde{w}_x(x, t) - \\ &\quad \left(q(1)\psi(x) - \int_x^1 P(x, y)\psi(y)dy\right)\tilde{w}(1, t)\tilde{\epsilon}(t) - \left(q(1)\phi(x) - \int_x^1 P(x, y)\phi(y)dy\right)\tilde{\epsilon}^2(t) - \\ &\quad q(x)\left(-\rho_1 S(q(1))(q(1)[\tilde{w}(1, t) - \tilde{\epsilon}(t)])\right),\end{aligned}$$

⁶As it has been commented in Chapter 4, the term $\Delta(x, t)$ in (5.16) corresponds to a particular (exact) case of the Taylor’s series approximation of the bivariate function $F = \epsilon(t)\frac{\partial^2 c_n}{\partial x^2}(x, t)$ around the point $\left(\frac{\partial^2 \hat{c}_n}{\partial x^2}(x, t), \hat{\epsilon}(t)\right)$.

where $\tilde{w}$ has been substituted by its “target” dynamic (5.8) and $\tilde{\epsilon}_t$ by its dynamic stated in (5.16); the transformation (4.47) has been evaluated at the boundary $x = 1$, namely

$$\tilde{c}_n(1, t) = q(1)[\tilde{w}(1, t) - \tilde{\epsilon}(t)], \quad (5.19)$$

and on the term T_0 has been applied integration by parts. Substituting the modified Volterra-type transformation (5.17), its second spatial derivative and the expressions (5.18) and (5.19) into the “original” estimation error system (5.16) yields

$$\begin{aligned} \frac{\partial \tilde{c}_n}{\partial t}(x, t) - \Delta(x, t) + h(x)\tilde{c}_n(1, t) = & \underbrace{\left(\epsilon(t)q_{xx}(x) - [q(x)\rho_1 S(q(1))q(1) + q(1)h(x)] - \frac{\partial^2 \tilde{c}_n}{\partial x^2}(x, t) \right)}_{T_1} \tilde{\epsilon}(t) + \\ & \underbrace{\left(\epsilon(t)P_y(x, 1) + [q(x)\rho_1 S(q(1))q(1) + q(1)h(x)] - q(1)\psi(x)\tilde{\epsilon}(t) \right)}_{T_2} \tilde{w}(1, t) - \\ & \epsilon(t) \underbrace{\left(2 \left(\frac{\partial}{\partial y} P(x, x) + \frac{\partial}{\partial x} P(x, x) \right) + q(1)c \right)}_{T_3} \tilde{w}(x, t) + \\ & \epsilon(t) \left(\int_x^1 (P_{xx}(x, y) - P_{yy}(x, y) + cP(x, y)) \tilde{w}(y, t) dy \right) + \\ & \left(\int_x^1 P(x, y)\psi(y) dy \right) \tilde{w}(1, t)\tilde{\epsilon}(t) - \left(q(1)\phi(x) - \int_x^1 P(x, y)\phi(y) dy \right) \tilde{\epsilon}^2(t). \end{aligned} \quad (5.20)$$

Thus, comparing the terms in brackets of T_1 and T_2 , using the identity $\frac{d}{dx}K(x, x) = \frac{\partial}{\partial x}K(x, x) + \frac{\partial}{\partial y}K(x, x)$ in T_3 [21] and decomposing $\epsilon(t)q_{xx}(x) = \hat{\epsilon}(t)q_{xx}(x) + \tilde{\epsilon}(t)q_{xx}(x)$ and $\epsilon(t)P_y(x, 1) = \hat{\epsilon}(t)P_y(x, 1) + \tilde{\epsilon}(t)P_y(x, 1)$, leads to

$$\begin{aligned} \frac{\partial \tilde{c}_n}{\partial t}(x, t) - \Delta(x, t) + h(x)\tilde{c}_n(1, t) = & \left(P_y(x, 1) - q(1)\psi(x) + \int_x^1 P(x, y)\psi(y) dy \right) \tilde{w}(1, t)\tilde{\epsilon}(t) + \\ & \left(q_{xx}(x) - q(1)\phi(x) + \int_x^1 P(x, y)\phi(y) dy \right) \tilde{\epsilon}^2(t) + \\ & \delta_0(x, t)\tilde{\epsilon}(t) - \delta_1(x, t)\epsilon(t)\tilde{w}(x, t) + \epsilon(t) \int_x^1 \delta_2(x, y, t)\tilde{w}(y, t) dy = 0, \end{aligned} \quad (5.21)$$

with constraints obtained from the boundary conditions in (5.16), namely

$$\begin{aligned} \tilde{c}_n(0, t) &= q(1)\tilde{w}(0, t) - \int_0^1 P(0, y)\tilde{w}(y, t) dy - q(0)\tilde{\epsilon}(t) = 0 \\ \frac{\partial \tilde{c}_n}{\partial x}(1, t) &= q(1)\tilde{w}_x(1, t) + P(1, 1)\tilde{w}(1, t) - q_x(1)\tilde{\epsilon}(t) \\ &= -p_b\tilde{u}(1, t) = (1 - p_b) (q(1) (\tilde{w}(1, t) - \tilde{\epsilon}(t))) + \overbrace{\left[-\frac{\tilde{\epsilon}(t)}{\tilde{\epsilon}^2(t)} \right]}^{\approx \hat{\theta}(t)} \varrho I(t), \end{aligned} \quad (5.22)$$

which are equivalent, at each time t , to the coupled **Kernel-PDE/ODE**:

$$\delta_0(x, t) = \hat{\epsilon}(t)q_{xx}(x) - \frac{\partial^2 \hat{c}_n}{\partial x^2}(x, t) + \hat{\epsilon}(t)P_y(x, 1), \quad (5.23)$$

$$\delta_1(x) = 2\frac{dP}{dx}(x, x) + q(1)c, \quad \forall x \in \Omega, \quad (5.24)$$

$$\delta_2(x, y) = P_{xx}(x, y) - P_{yy}(x, y) + cP(x, y), \quad \forall (x, y) \in \Omega_U \quad (5.25)$$

$$q(0) = 0, \quad P(0, y) = 0, \quad \forall y \in \Omega \quad (5.26)$$

$$q_x(x, t)|_{x=1} = P(1, 1) + \frac{\varrho I(t)}{\hat{\epsilon}^2(t)}, \quad (5.27)$$

and functions $\psi(x)$ and $\phi(x)$ satisfying the Volterra's integral equations of second kind:

$$\frac{P_y(x, 1)}{q(1)} = \psi(x) - \int_x^1 \frac{P(x, y)}{q(1)} \psi(y) dy, \quad (5.28)$$

$$\frac{q_{xx}(x)}{q(1)} = \phi(x) - \int_x^1 \frac{P(x, y)}{q(1)} \phi(y) dy, \quad (5.29)$$

$\forall q(1) \neq 0$, for a zero matching condition on the residual functions δ_j , $\forall j = 0, \dots, 2$. In this case the observer gains are given by:

$$p_b = 1 - \frac{P(1, 1)}{q(1)} \quad (5.30)$$

$$h(x) = - \left(\frac{1}{q(1)} \hat{\epsilon}(t) P_y(x, 1) + \rho_1 q(x) S(q(1)) \right). \quad (5.31)$$

Regarding the “target” parameter error dynamic (5.16), substituting (5.19) into the parameter estimation error of the “original” error system (5.15) yields

$$\frac{d\tilde{\epsilon}}{dt}(t) = -\rho_1 S(q(1)) (q(1)[\tilde{w}(1, t) - \tilde{\epsilon}(t)]) = [\rho_1 S(q(1))q(1)]\tilde{\epsilon}(t) - [\rho_1 S(q(1))q(1)]\tilde{w}(1, t),$$

$\forall t \in [t_a, t_b]$. Thus, considering $\tilde{w}$ as a bounded signal and $\rho_1 < 0$, a sufficient condition to obtain a stable dynamic for $\tilde{\epsilon}$ in (5.16) is the selection of a function S such that $S(q(1))q(1) > 0$, $\forall t \geq 0$.

With respect to the integral equations (5.28)-(5.29), given a continuous bounded function q , a continuous bounded P and P_y , the existence and uniqueness of bounded continuous functions $\psi = \psi(x)$ and $\phi = \phi(x)$ are guaranteed by the invertibility property of the Volterra integral operator on some Banach and Hilbert spaces [18, 208, 191].

5.3.4 Coupled PDE-ODE as a Convex Optimization Problem

Similar to the methodology described in Section 4.3.3.1, the coupled Kernel-PDE/ODE (5.23)-(5.27) can be approximately solved considering a relaxation of the exact zero matching condition on its residual functions δ_i , via its SOS decomposition and its formulation as an optimization problem.

Proposition 14. *Let $N(x, y) = \sum_{k=0}^{z(d)-1} n_k x^{\alpha_k} y^{\beta_k}$ be a polynomial approximation of P in (5.17) of arbitrary even degree $d > d_{\hat{c}_{xx}} \in \mathbb{N}$, in accordance with (1.3). Let $\mu(x) = \sum_{k=0}^d \mu_k x^k$ be a polynomial approximation of q in (5.17) with domain constrained to (1.4). Let η_j, σ_j be lower and upper bounds of δ_j , respectively, in Ω , $\forall j = \overline{0, 1}$. Let η_2, σ_2 be lower and upper bound of δ_2 , respectively, in Ω_U . Let $\frac{\partial^2 \hat{c}_n}{\partial x^2}(x, t)$ be estimated by the observer (5.15) and approximated at each fixed time by a polynomial function of degree $d_{\hat{c}_{xx}}$. The coupled Kernel-PDE/ODE (5.23)-(5.27) can be recast as the convex optimization problem:*

$$\begin{aligned}
& \underset{\eta_j, \sigma_j, N, \mu, s_{ij}}{\text{minimize:}} && \sum_{j=0}^2 \eta_j + \sigma_j, \\
& \text{subject to:} && (\delta_j(x) - \eta_j - s_{0j}(x) g_1(x)) \in \Sigma_s, \quad \forall j = 0, 1, \\
& && (\sigma_j - \delta_j(x) - s_{1j}(x) g_1(x)) \in \Sigma_s, \quad \forall j = 0, 1, \\
& && (\delta_2(x, y) - \eta_2 - [s_{20}(x, y) s_{21}(x, y)] g_U(x, y)) \in \Sigma_s, \\
& && (\sigma_2 - \delta_2(x, y) - [s_{22}(x, y) s_{23}(x, y)] g_U(x, y)) \in \Sigma_s, \\
& && s_{0j}, s_{1j}, s_{2i} \in \Sigma_s, \quad \forall j = 0, 1, 2, \quad i = 0, \dots, 3 \\
& && \eta_j \geq 0, \quad \sigma_j \geq 0, \quad \forall j = 0, 1, 2, \\
& && \delta_0 = \delta_0(x) \big|_{q \approx \mu}^{P \approx N} \quad \text{as in (5.23),} \\
& && \delta_1 = \delta_1(x) \big|_{q \approx \mu}^{P \approx N} \quad \text{as in (5.24),} \\
& && \delta_2 = \delta_2(x, y) \big|_{P \approx N} \quad \text{as in (5.25),} \\
& && \mu(0) = 0, \quad N(0, y) = 0, \\
& && \mu_x(1) = N(1, 1) + \frac{\varrho I(t)}{\hat{c}^2(t)},
\end{aligned} \tag{5.32}$$

for some polynomials s_{0j}, s_{1j}, s_{2i} of degree $d - 2$, $\forall j = 0, 1, \forall i = 0, \dots, 3$ and $g_U = [g_1, g_2]$ according to (2.8), with $\|\delta_j(\mathbf{x})\|_\infty = \max\{\eta_j, \sigma_j\}$, $\forall j = 0, 1, 2$.

Proof. The proof follows the same arguments exposed in Propositions 2 and 8, hence it is omitted. $\square$

5.3.5 Uncoupled Kernel-PDE/ODE via Convex Optimization

The Kernel-PDE/ODE (5.23)-(5.27) can be uncoupled considering a small modification of the integral transformation (5.17), namely

$$\tilde{c}_n(x, t) = w(x, t) - \int_x^1 P(x, y) w(y, t) dy - q(x) \tilde{\epsilon}(t), \tag{5.33}$$

(the factor $q(1)$ has been dropped from the first term in the right-hand side of (5.17)). In this case, following the same procedure described in Section 5.3.3, the modified Volterra-type transformation, at each fixed time t , leads to the set of equations:

$$\delta_0(x, t) = \hat{\epsilon}(t)q_{xx}(x) - \frac{\partial^2 \hat{c}_n}{\partial x^2}(x, t) + q(1)\hat{\epsilon}(t)P_y(x, 1), \quad (5.34)$$

$$q(0) = 0, \quad q_x(1) = P(1, 1)q(1) + \frac{\rho I(t)}{\hat{\epsilon}^2(t)}, \quad (5.35)$$

$$\delta_1(x) = 2 \frac{d^2 P}{dx^2}(x, x) + c, \quad \forall x \in \Omega, \quad (5.36)$$

$$\delta_2(x, y) = P_{xx}(x, y) - P_{yy}(x, y) + cP(x, y), \quad \forall (x, y) \in \Omega_U \quad (5.37)$$

$$P(0, y) = 0, \quad \forall y \in \Omega, \quad (5.38)$$

and functions $\psi(x)$ and $\phi(x)$ satisfying the Volterra's integral equations of second kind:

$$\frac{\partial P}{\partial y}(x, 1) = \psi(x) - \int_x^1 P(x, y)\psi(y)dy, \quad (5.39)$$

$$\frac{d^2 q}{dx^2}(x) = \phi(x) - \int_x^1 P(x, y)\phi(y)dy, \quad (5.40)$$

for a zero matching condition on the residual functions $\delta_j, \forall j = 0, \dots, 2$. In this case the observer gains are given by:

$$p_b = 1 - P(1, 1) \quad (5.41)$$

$$h(x) = - \left(\hat{\epsilon}(t) \frac{\partial P}{\partial y}(x, 1) + \rho q(x)S(q(1)) \right). \quad (5.42)$$

Thus, at each fixed time t , the ODE BVP (5.34)-(5.35) can be solved for a known Kernel P , which is the solution of the well-known Kernel-PDE (5.36)-(5.38) and arises from the Backstepping PDE observer design for a collocated setup. Its solution is unique and can be obtained in closed-form via the Successive Approximation Method [197, 21, 50]:

$$\Theta(x, y) = \sqrt{c(y^2 - x^2)}, \quad P(x, y) = -cx \frac{I_1(\Theta(x, y))}{\Theta(x, y)} \quad (5.43)$$

$$\frac{\partial P}{\partial y}(x, y) = \frac{-c^2 xy}{\Theta^2(x, y)} \left(\frac{(I_0 + I_2)(\Theta(x, y))}{2} - \frac{I_1(\Theta(x, y))}{\Theta(x, y)} \right), \quad (5.44)$$

where I_j stands for the j -th order modified Bessel function.

It is worth noting that the ODE BVP (5.34)-(5.35) includes the boundary term $q(1)$ in the “in-domain” ODE expression. This problem can be effectively solved via SOS decomposition and polynomial optimization as follows.

Proposition 15. *Let $\mu(x) = \sum_{k=0}^d \mu_k x^k$ be a polynomial approximation of q in (5.33) of arbitrary even degree $d > d_{P_y} \in \mathbb{N}$. Let η, σ be lower and upper bounds of δ_0 in (5.34) in Ω , respectively. Considering a polynomial approximation of $P_y(x, 1)$ from (5.43)-(5.44) of degree d_{P_y} , the ODE BVP (5.34)-(5.35) can be recast, at each fixed time t , as the convex optimization problem:*

$$\begin{aligned}
 & \underset{\eta, \sigma, \mu, s_j}{\text{minimize :}} \quad \eta + \sigma, \\
 & \text{subject to :} \quad (\delta_0(x) - \eta - s_0(x) \ g_1(x)) \in \Sigma_s, \\
 & \quad (\sigma - \delta_0(x) - s_1(x) \ g_1(x)) \in \Sigma_s, \\
 & \quad s_0, s_1 \in \Sigma_s, \quad \eta \geq 0, \quad \sigma \geq 0, \\
 & \quad \mu(0) = 0, \quad \left. \frac{d\mu}{dx}(x) \right|_{x=1} = \frac{-c}{2} + \frac{\varrho I(t)}{\hat{\epsilon}^2(t)},
 \end{aligned} \tag{5.45}$$

for some SOS polynomials s_0, s_1 of degree $d-2$ and g_1 according to (2.6), with $\|\delta_0(x)\|_\infty = \max\{\eta, \sigma\}$.

Proof. It follows similar arguments of the proof of Propositions 2 and 8, considering univariate polynomials in Ω , hence it is omitted. $\square$

5.4 Output Mapping Inversion

The linear feedback problem in the observer (5.15) involves the indirect measurement of $y(t) = c_{n,ss}^l(t) = c_s^l(R_s^l, t)/c_{s,\max}^l = \varphi(V(t), I(t))$ from the inversion of (5.2). This output mapping inversion can be addressed assuming the positive electrode diffusion dynamic as *instantaneous* in comparison with the negative electrode dynamics [265], and on the basis of a simplification of the Lithium conservation property [11], which leads to a linear relation between negative/positive electrode surface concentrations; i.e.,

$$c_{n,ss}^+(t) = - \left(\frac{\varepsilon^- L^- c_{s,\max}^-}{\varepsilon^+ L^+ c_{s,\max}^+} \right) c_{n,ss}^-(t) + \frac{\eta_{\text{Li},s}}{\varepsilon^+ L^+ c_{s,\max}^+}. \tag{5.46}$$

Thus, the nonlinear output mapping (5.2) is a function of only one unknown variable: $c_{n,ss}^-(t)$, namely

$$V(t) = G(c_{n,ss}^-(t), I(t)) \tag{5.47}$$

$$\begin{aligned}
 & = \Delta\eta_s(i_o^+(t), i_o^-(t), I(t)) + U^+(c_{n,ss}^+(t)) - U^-(c_{n,ss}^-(t)) - \left(R_f^+/(a^+ L^+) + R_f^-/(a^- L^-) \right) I(t) \\
 & i_o^l(t) = k^l c_{s,\max}^l \sqrt{c_e^o} \left[c_{n,ss}^l(t) (1 - c_{n,ss}^l(t)) \right]^{1/2},
 \end{aligned} \tag{5.48}$$

with $\Delta\eta_s = \eta_s^+(t) - \eta_s^-(t)$, for η_s^l according to (5.3) in terms of (5.48) and $c_{n,ss}^+$ given by (5.46).

The method of inversion of (5.47) is crucial due to the high sensitivity of the observer performance in relation to the accuracy of its computation [262]. To solve the nonlinear equation $E(t) = V(t) - G(c_{n,ss}^-(t), I(t)) = 0$, standard gradient-based approaches [11] are severely affected due to the nearly zero gradient of G with respect to the Lithium-ion surface concentration in some intervals of its domain and for some type of electrode's materials [269, 267].

5.4.1 Inversion via Moment Approach

To solve (5.47), in contrast to gradient-based schemes and their inherent limitations [127], the proposed approach is based on an inverse Moment problem formulation [137, 162]. Let $V_t = V(t)$ and $I_t = I(t)$ be measured values and $v(x) = c_{n,ss}^-(x, t) \in [v_{a_k}, v_{b_k}] = \Omega_k \subset \Omega$ at each fixed time. Considering all parameters in (5.47)-(5.48) and the U^l functions known, the quadratic error

$$J(v) = (G(v, I_t) - V_t)^2, \quad v \in \Omega_k \subset \Omega, \quad (5.49)$$

is interpreted as a (non-negative) density function $J \in \mathcal{L}^2(\Omega_k, \lambda)$ with respect to the Lebesgue measure $\lambda(v) = v \in \Omega_k$. This allows computing finitely many *moments* (sequence) $\mathbf{Y}_d = \{y_i\}$, $\forall i = 0, \dots, d$ of the Borel measure $d\mu(v) = J(v)d\lambda(v)$ (J : Radon-Nikodym derivative of μ with respect to λ [137, 138]) as:

$$y_i = \int_{v_{a_k}}^{v_{b_k}} \Psi_i(v) J(v) dv, \quad (5.50)$$

where Ψ_i are elements of a polynomial basis in Ω_k . The objective is to recover from $\mathbf{Y}_d$ a non-negative approximate polynomial J_d of J in Ω_k , to compute a polynomial optimization: $\min_{v \in \Omega_k} \{J_d\} \rightarrow c_{n,ss}^* = \varphi(V_t, I_t)$ at each fixed time t .

Proposition 16. *Let $J_d = \sum_{i=0}^d \vartheta_i \Psi_i(v) = \Theta^\top \mathbf{B}(v)$ be a polynomial approximation of J in (5.49) of degree $d \in \mathbb{N}$, with $\Theta \in \mathbb{R}^{d+1}$ and $\mathbf{B}(v) = [\Psi_0, \dots, \Psi_d]^\top$ a selected polynomial basis of monomials Ψ_i in Ω_k . Given a sequence of moments $\mathbf{Y}_d = \{y_i\}$, $\forall i = 0, \dots, d$, J can be approximated by the solution of the convex optimization problem:*

$$\begin{aligned} \underset{\gamma_j \geq 0, \Theta, s_j}{\text{minimize:}} \quad & \gamma_1 + \gamma_2 \\ \text{subject to:} \quad & (\mathbf{M}_d \Theta - \mathbf{Y}_d) + \gamma_1 \geq 0, \quad \gamma_2 - (\mathbf{M}_d \Theta - \mathbf{Y}_d) \geq 0, \\ & \Theta^\top \mathbf{B}(v) - s_0(v) - s_1(v)g_1(v) = 0, \\ & s_0, s_1 \in \Sigma_s, \end{aligned} \quad (5.51)$$

where $\gamma_1 \geq 0$, $\gamma_2 \geq 0$, $\Theta = [\vartheta_0, \dots, \vartheta_d]^\top$, the moment matrix

$$\mathbf{M}_d = \int_{v_{a_k}}^{v_{b_k}} \Phi(v) \Phi^\top(v) dv \in \mathbb{S}_+^{d+1}, \quad (5.52)$$

$g_1(v) = (v - v_{a_k})(v_{b_k} - v) \geq 0$, and polynomials s_0 of degree d and s_1 of degree $d - 2$. In addition, the sequence of minimizers J_r^* , $\forall r \in \mathbb{N}$, is such that $\|J - J_r^*\|_{\mathcal{L}^2(\Omega_k, \lambda)} \rightarrow 0$ as $r \rightarrow \infty$.

Proof. The convex optimization problem described above is an alternative formulation of the problem presented in [162], using linear constraints in (5.51), hence its proof is omitted. $\square$

If $\mathbf{B}$ is selected as the canonical polynomial basis, the elements of the moment matrix correspond to:

$$\mathbf{M}_d(i, j) = \left(v_{b_k}^{i+j+1} - v_{a_k}^{i+j+1} \right) / (i + j + 1), \quad (5.53)$$

$\forall i = 0, \dots, d$ and $j = 0, \dots, d$. The problem (5.51) could lead to an ill-conditioned matrix inversion. However, using the shifted Legendre monomials as basis in Ω_k (orthogonal basis), a diagonal moment matrix can be obtained, namely

$$\mathbf{M}_d = [m_{0,0}, m_{1,1}, \dots, m_{1+d,1+d}] I \in \mathbb{S}_+^{d+1}, \quad (5.54)$$

$$m_{i,i} = \int_{v_{a_k}}^{v_{b_k}} \Psi_i^2(v) dv = \frac{v_{b_k} - v_{a_k}}{1 + 2i}, \quad (5.55)$$

$$\Psi_i(v) = \frac{1}{2^i} \sum_{j=0}^{\lfloor i/2 \rfloor} (-1)^j \binom{i}{j} \binom{2i-2j}{i} (c_1 v + c_0)^{i-2j}, \quad (5.56)$$

with $c_1 = 2/(v_{b_k} - v_{a_k})$, $c_0 = 1 - c_1 v_{b_k}$ and $I \in \mathbb{R}^{d+1 \times d+1}$ the identity matrix. It is worth noting that for $\Theta = \mathbf{M}_d^{-1} \mathbf{Y}_d$ the non-negativity of J_d is not guaranteed. To provide this condition, this has to be included as in Proposition 16 via its SOS representation.

To compute $\min_{v \in \Omega} \{J_d\}$, non-gradient methods can be used such as the Golden Section search [270]. For some kind of OCPs with nearly zero gradient in some intervals [267], (5.49) could present more than one global minimum, which can be computed by the Moment-SOS approach [137, 138]. The scheme of inversion, at each fixed time, consists in a systematic reduction of intervals: $\Omega \supset \Omega_1 \supset \dots \supset \Omega_k \dots \supset \Omega_N$, with $\Omega_k = [v_{k-1}^* - \delta_k, v_{k-1}^* + \delta_k]$, $\delta_{k-1} > \delta_k > 0$, centered at the local optimal point v_{k-1}^* . Based on an accurate approximation $J_d = J_d(v), \forall v \in \Omega_k$, recovered from its respective local moments, the sequence $\{v_k^*\} \rightarrow \min_{v \in \Omega} \{J\}$.

5.5 Numerical Results⁷

The adaptive observer (5.15) has been implemented by means of two schemes:

- “Observer 1”: Based on the Volterra-type transformation (5.17), the resulting coupled Kernel-PDE/ODE (5.23)-(5.27) has been solved via Proposition 14. The Observer gains have been computed by (5.30)-(5.31).
- “Observer 2”: Based on the Volterra-type transformation (5.33), the resulting uncoupled Kernel-PDE (5.36)-(5.38) and ODE-BVP (5.34)-(5.35) have been solved via (5.43)-(5.44) and Proposition 15, respectively. The Observer gains have been computed by (5.41)-(5.42).

The SPM electrochemical parameters mostly correspond to the DFN model DUALFOIL [271, 242]. To illustrate the scheme proposed for the nonlinear mapping inversion under nearly zero gradient, the OCPs functions have been taken from [267], which corresponds to a Graphite/LiFePO₄ Cell.

⁷The numerical solution of the convex optimization problems has been obtained via the Yalmip toolbox for Matlab [163] and SOSTOOLS [164], using the semi-definite programming solver SeDuMi [165] and the SDP package part of the Mosek solver [166].

Regarding the inversion of the nonlinear output mapping (5.2), the performance of the scheme proposed in Section 5.4 is shown in Figure 5.5.1, for a 1C constant current discharge (C-rate: ratio of the current (A) to the nominal battery capacity (A-h)). At the first time instant ($t = 0$), for initial condition $c_{n,ss}^-(0) = 0.712100857944318$, the scheme yields the value $\varphi(V_0, I_0) = 0.712100857944327$ after $k = 6$ steps of reduction of the intervals Ω_k (Figure 5.5.1 (a)). As it is depicted in Figure 5.5.1 (b), for the whole range of $c_{n,ss}$, the scheme of inversion achieves a precision of 13 decimals.

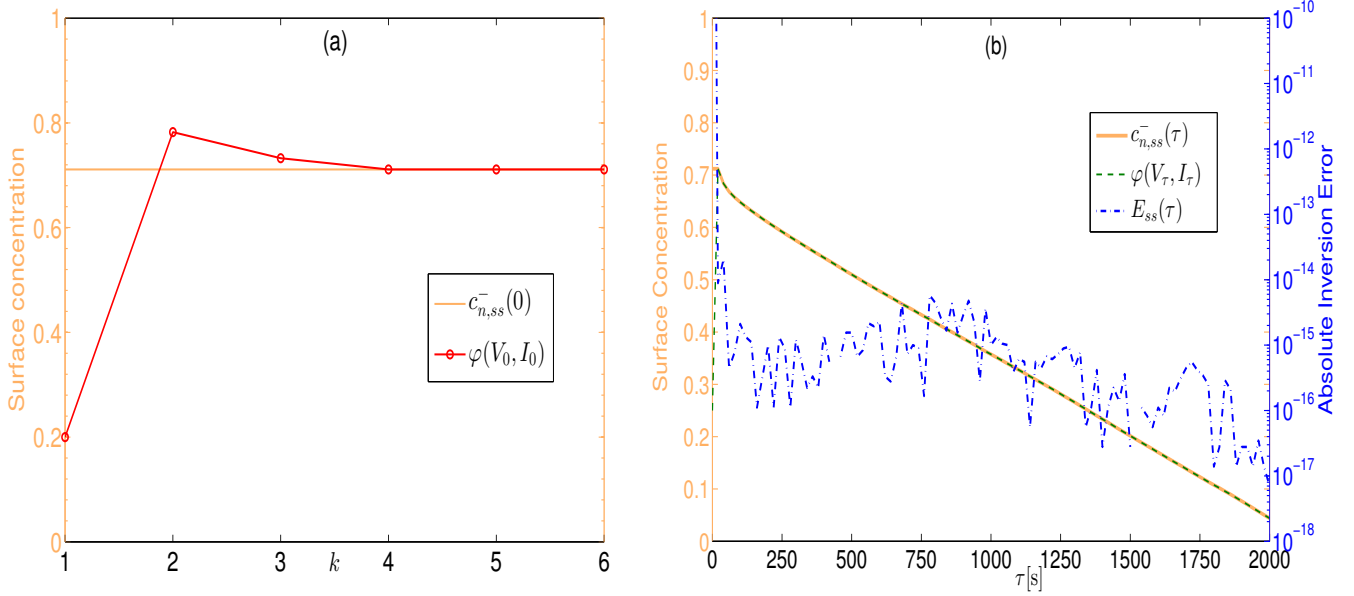


Figure 5.5.1: Inversion of the Nonlinear Output Mapping in Lithium-Ion Batteries. Case of constant discharge current of 1C. SPM (actual) surface concentration (left y-axis) $c_{n,ss}^-$: solid line. (a) Six first optimization steps at the instant $t = 0$: dash-circle line. (b) Surface concentration obtained by the inversion scheme (y-axis: left-hand side): $\varphi(V_\tau, I_\tau)$: dash line. Absolute inversion error (y-axis: right-hand side) $E_{ss} = |c_{n,ss}^-(\tau) - \varphi(V_\tau, I_\tau)|$: dash-dot line.

As for the observer (5.15), a degree $d = 12$ for the polynomial Kernels has been selected. At each fixed time, the Hessian of the estimated Lithium-ion concentration $\frac{\partial^2 \hat{c}_n}{\partial x^2}(x, t)$ has been approximated as a polynomial of degree $d_{\hat{c}_{xx}} = 3$. For the “Observers 1” and “2”, the performance of the simultaneous state and diffusivity ratio parameter estimation is shown in Figure 5.5.2, where a discharge/charge current generated via an Urban Dynamometer Driving Schedules (UDDS) has been applied [11, 242]. In Figure 5.5.2(a), the mean absolute estimation error: $E(t) = (1/N) \sum_{x \in \Omega} |c_n^-(\cdot, t) - \hat{c}_n^-(\cdot, t)|$, with $N = 100$ spatial points, is depicted as well as the estimated surface concentration. In Figure 5.5.2(b), the estimated diffusivity ratio parameter and its actual value ($\epsilon(t) = 1, \forall t \geq 0$) are shown. For “Observer 1”, the parameters $c = 30$ and $\rho_1 = -1.2 \cdot 10^3$ have been selected. For “Observer 2”, higher gains $c = 30$ and $\rho_1 = -5.65 \cdot 10^3$ have been considered. For both observers, the function S is selected as $S(q(1)) = q(1)$. Figure 5.5.3 depicts the estimated diffusivity ratio parameter for a step-type variation from $\epsilon(t) = 1$ to $\epsilon(t) = 0.95$ at $t = 300$ [s]. In this case, a multi-frequency-type charge/discharge current has been applied and, for the Observer 2, the parameters $c = 30$ and $\rho_1 = -4.2 \cdot 10^3$ have been selected.

Based on the numerical implementation of the observers designed and on the results obtained, Observer 2 provides a better performance than Observer 1.

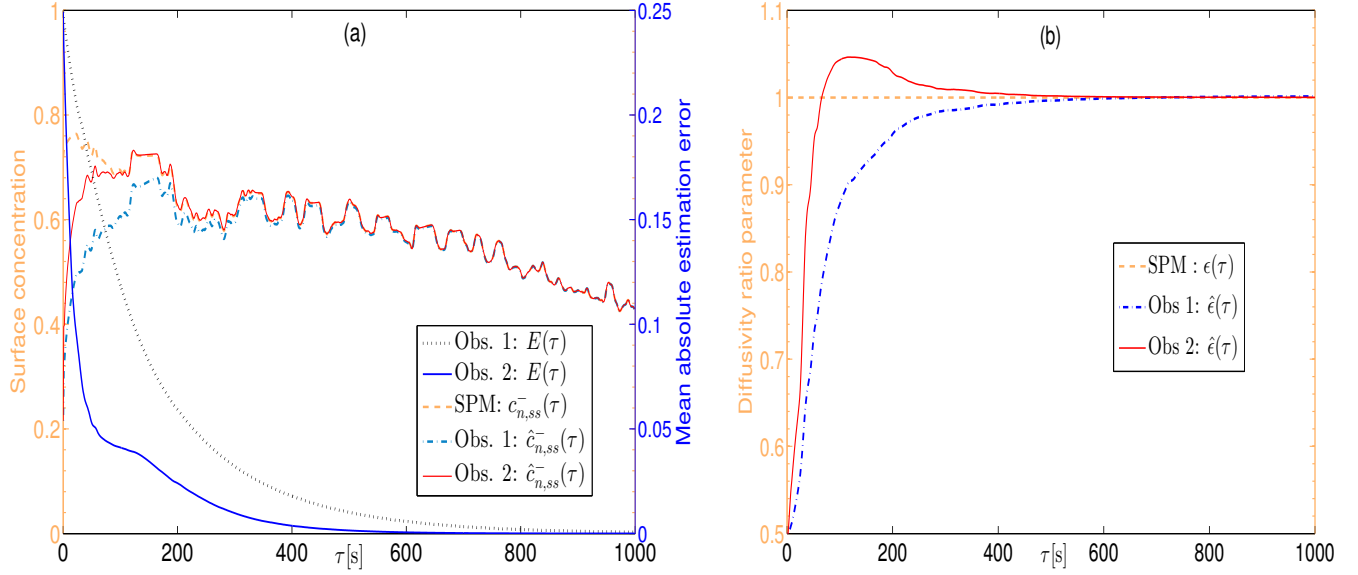


Figure 5.5.2: Adaptive Observer Design for Lithium-Ion Batteries. Observer results for UDDS charge/discharge. Case of constant diffusivity parameter. (a) Mean absolute estimation error. Observer 1: dot line and Observer 2: solid line. Estimated surface concentration (y-axis: right-hand side). SPM (actual) value: dash line, Observer 1: dash-dot line and Observer 2: solid line (y-axis: left-hand side). (b) Estimated diffusivity ratio parameter. SPM value: dash line, Observer 1: dash-dot line and Observer 2: solid line.

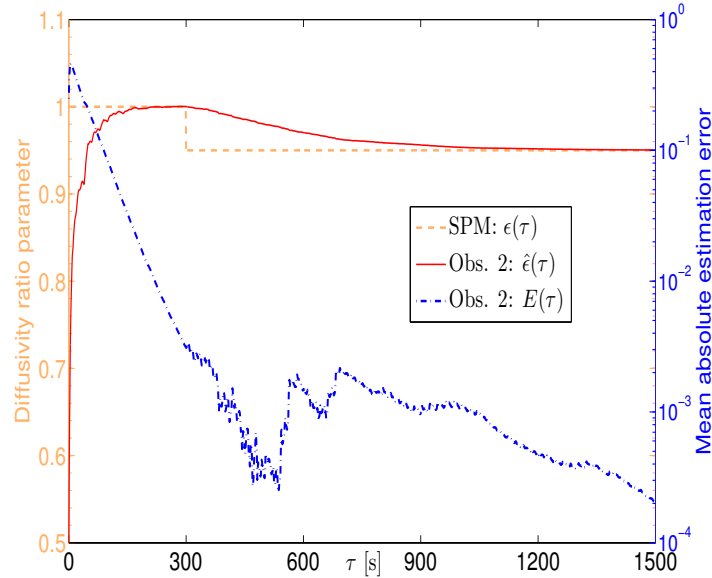


Figure 5.5.3: Adaptive Observer Design for Lithium-Ion Batteries. Observer 2: results for a step-type change of diffusivity parameter. SPM (actual) value: dash line. Estimated diffusivity ratio parameter: solid line (y-axis: left-hand side). Mean absolute estimation error: dash-dot line (y-axis: right-hand side).

Chapter 6

Conclusions

—George Stokes (1819–1903)—

G. Stokes mainly focused his carrier on natural philosophy but with various excursions into pure mathematics, to develop methods or validate mathematical techniques to solve particular physical problems. The so-called Stokes's theorem was first appeared in print in 1854, included by himself as part an exam at Cambridge University, where Clerk Maxwell was a student. Stokes's and the Divergence theorem were used through the Maxwell's Treatise on Electricity and Magnetism, but only its vector form appeared in the Josiah W. Gibbs's work on vector analysis. These forms were not obviously related each other and to the George Green's theorem, but Vito Volterra in 1889 was able to unite them and, along with Henri Poincaré, generalize to higher dimensions. Although Elie Cartan realized that these three theorems can be easily stated using differential forms, it was Edouard Goursat in 1917 who first noted that these could be written in the simple form: $\int_S w = \int_T dw$, today called the "generalized Stokes theorem".

Backstepping methodology for PDEs basically bases its analysis and design on the Volterra-type transformation and its synthesis on the solution of the so-called Kernel-PDE. In this thesis, the notable simplicity of its central idea has been used to propose an intuitive modification of the Volterra-type and Fredholm-type transformations to compensate for the parameter uncertainties, for a class of one-dimensional linear parabolic PDEs. In addition, to solve the resulting coupled Kernel-PDE/ODE new notable results in polynomial optimization from SOS and Moment theory have been exploited. The methodology proposed allows a “simultaneous state and parameter estimation”, which does not require any pre-transformation to some canonical form and/or finite dimensional formulation, enabling a direct parameter identification from the original PDE model, considering only boundary measurements.

6.1 Differential Equations as Convex Optimization Problems

In Chapter 2, a methodology to recast differential equations-boundary value problems with polynomial terms as convex optimization problems has been proposed. This formulation is intuitive in its conception and can be carried out in a straightforward manner. The variety of numerical examples

analyzed illustrates the high accuracy of the method proposed. In particular, for an approximate solution with a unique polynomial, lower polynomial degrees are needed to achieve similar accuracies to the standard Finite Element Method using hundreds of elements with local basis such as linear, quadratic or quartics. The flexibility of the approach proposed also allows defining domain decompositions and solve local optimal problems with similar results to the FEM.

One advantage of this approach based on SOS is its simple numerical programming, which does not need to formulate and build up particular discrete algebraic structures verifying its conditioning and managing its boundary conditions. On the other hand, it is also possible to extend the approach proposed to solve PDE-initial value problems, for instance combining it with the Crank-Nicolson's method or other time-marching schemes.

The main limitation of the method proposed is the current state of SOS tools, regarding the type of monomials used in the decompositions, and the convex optimization tools, in relation with managing a large number of parameters and parameters with big magnitudes. For problems with dimensions greater than one, SOS polynomials are a small subset of positive polynomials (see Chapter 2, Remark 5), so that is theoretically harder to find an optimal solution with accurate results as the dimension increase.

6.2 Convex Formulation of Backstepping Design for PDEs

In Chapter 3, based on polynomial functions and SOS decomposition, a novel methodology has been proposed to recast the so-called Kernel-PDEs, core of the synthesis in Backstepping for PDEs, as a convex optimization problem. In contrast to the standard Successive Approximation Method used to solve Kernel-PDEs and its inherited restrictions, the method proposed allows managing systems with strict and non-strict feedback structure, involving the Volterra and Fredholm-type transformations, in a unified manner not subject to convergence conditions.

For polynomial Kernels and analytic functions or continuous functions, uniqueness and invertibility of the Fredholm-type transformation have been proved, which allows applying the methodology proposed to a wide class of problems without restriction on the spectral characteristic of some integral operators.

In practice, the method proposed does not impose conservative constraints on the parameters of the system to accomplish contraction properties on the operators involved so that achieving convergence (Banach Contraction Principle). On the contrary, it is flexible to manage linear problems with operators of different structure and objectives. The resulting polynomial approximation problem can be formulated as a convex optimization problem and solved directly resorting semidefinite programming, obtaining accurate approximate solutions with a theoretically arbitrary precision.

6.3 Adaptive Observer Design

In Chapter 4, a simple modification of the Volterra-type transformation has been proposed to compensate for the time-varying gain uncertainty of the boundary control action. The resulting coupled Kernel-PDE/ODE is formulated as an convex optimization problem, the numerical solution of which is fast enough to be solved at every sampling time. A numerical example illustrates the capability of the design proposed to perform an accurate “simultaneous state and parameter estimation” only from one boundary measurement.

The above adaptive observer proposed has been extended to the challenge problem of “on-line” estimation of “in-domain” uncertain parameters such as *reactivity* and *diffusivity*. Similar to the case of boundary unknown parameter, the observer design relies on a modification of the Volterra/Fredholm-type transformation, where additive extra terms have been included to account for the parameter uncertainties. Based on this modified transformation, a systematic methodology of observer design has been formulated. Its intuitive mathematical steps seek to reject the time-variant uncertainties, the associate components of which are dynamically cancelled out by means of solving differential boundary value problems.

The uniqueness and invertibility of the Fredholm-type operator in the space of analytical functions and continuous functions proved in Chapter 3, allows designing adaptive observers, using polynomial Kernels, via two boundary points of measurement. Despite the resulting coupled Kernel-PDE/ODE problem is more complex than for one measurable point, its approximate numerical solution based on the method proposed in Chapter 2 does not involve higher complexities. On the other hand, although the resulting differential problem not necessarily has a unique global solution, the method of solution allows finding the optimal local dependence between the observer gains and integral Kernels, which can be computed in a simple and fast way at every fixed sampling time.

In contrast with the case of a boundary unknown parameter, the modified integral transformation for “in-domain” parameter uncertainties leads to “target systems” which are not necessarily stable. This stability depends on the selection of one crucial parameter “ c ” which is also the main factor in the coupled Kernel-PDE/ODE. Despite a bound for this parameter can be found to guarantee stability and convergence of the estimation error, this depends on some polynomial Kernels which are also indirect functions of it. This recursive dependence is the main drawback of the proposed observer design methodology, which involves analysing bounds of integral equations.

Numerical examples illustrate the ability of the design proposed to perform an accurate simultaneous state and “boundary” or “in-domain” parameter estimation only from boundary measurements. The parameter adaptation laws are similar in structure to the resulting gains derived from stability analysis using Lyapunov-like arguments and Barbălat’s lemma in LPS (for instance, gradient methods). Similarly, the identification problem is also subject to standard condition of persistent excitation. The main parameter of design in the coupled Kernel-PDE/ODE (“ c ”) has

a simple direct proportionality with the observer gain for the estates of the system, allowing a faster convergence as its value increases. However, the selection of its magnitude, beside stability conditions, is a delicate balance of convergence between states and parameters, which is not clear to set forth in advance for different systems.

The observer design developed proposes time-invariant polynomial Kernels in the modified Volterra-type or Fredholm-type transformation to compensate for uncertainties. This is a simplified approach since a dynamical rejection of theses uncertainties is needed and therefore, the scheme has to carry out a computation of these Kernels at every time instant, which is not physically realizable. In practice, for slow and smooth systems such as some types of parabolic PDEs, the method proposed achieves satisfactory results computing theses Kernels at every fixed sampling time, as it has been illustrated in different numerical examples. However, despite assuming that these Kernels behave as piece-wise constant function of time, the steps of design do not lead necessarily to a stable and convergent estimation.

6.3.1 Adaptive Observer for Lithium-Ion Batteries

In Chapter 5, an adaptive observer has been designed to address the relevant problem of estimating the Lithium-ion concentrations in solid particles under an uncertain diffusivity coefficient. The Volterra-type transformation and the steps of design formulated in Chapter 4 has been extended to this case where both “in-domain” and “boundary” uncertainties has been considered, showing the generality of the method proposed. Numerical simulations confirm the capability of this design to achieve a convergent and accurate simultaneous parameter and estate estimation.

The observer design proposes a Luenberger-type structure. In contrast to the standard methodologies based on gradient schemes, based on polynomial optimization and Moment theory, a novel methodology of inversion of the nonlinear map between voltage and solid particle concentration has been proposed. Thus, the problem derived from the flat shape of Open Circuit Potential functions has been circumvented and the observer linear feedback term can be dynamically determined with high precision for the whole domain. In addition, this method allows finding multiple global minimums for the inversion problem.

In contrast to the observer detailed in Chapter 4, for a diffusivity uncertainty the observer design involves an ODE which includes the second derivative of the estimated state variable with respect to the spatial coordinate. Since, for the battery model considered, this state can behave similarly to a linear function, the second derivative provides signals (information) with small magnitudes. Therefore, an increase of the polynomial observer gain is required to capture small variations in the parameters and achieve a faster convergence. However, this kind of gains makes the scheme highly sensitive to noisy measurements.

6.4 Future Research Directions

The adaptive observer design proposed in this thesis is not complete and a series of rigorous mathematical analyses is still pending to provide consistency with results obtained from numerical simulations. The insights achieved, from the limitations and potentials of the methodology developed in this thesis, contribute to inspire potential new lines of further research that are worth exploring.

- Backstepping design for PDEs proposes in advance a parameterized “target” system. However its methodology to achieve this target does not consider degrees of freedom to optimize the parameters involved, at least keeping the simplicity of the design. The formulation of this method as a direct optimal problem, for instance to design state estimators as optimal filters, has not been explored. In this problem, the features of the convex formulation via SOS could contribute with new ideas to set forth a tractable optimal design.
- Backstepping design for PDEs considers only one “target” system. Naturally one question arises: would it be possible to consider multiple “targets”? The exact matching condition set by the current method excludes this possibility. However, as it has shown in this thesis, that condition can be relaxed to transform into stable systems. This is an unexplored subject with relevant implications, in particular with respect to Lyapunov stability analysis. For instance, a set of *target PDEs* could constitute a domain of stability which could be parameterizable via SOS. Thus, integral transformations could lead to matrix inequalities instead of an exact target, tractable of solving and optimizing.
- All above mentioned points are connected to provide robustness to the control/observer problem, which in DPS is limited to each particular case under study. The features of the convex formulation via SOS could contribute to advance to formulate more general or unified tools. For instance, looking for formulations in DPS which allows an equivalent Lyapunov stability of polytopic structures in LPS.
- The observer design proposed leads to some type of “target” systems which involve certain combinations of states, parameters, signal errors and Kernels. The analysis of stability of these particular structures, via the Lyapunov’s methodology, could be useful to study the recursive dependence in the Kernels of integral transformations involved and look for its natural derivation and ways for its compensation.
- The methodology of Backstepping design for PDEs has remarkable similarities to a method of solving the identifiability problem framed in the context of the Sturm-Liouville problem, via the Gel’fand-Levitan’s theory and “deformation formulas”. This connection has not been explored and it could provide new insights to the adaptive observer design problem.

Chapter 7

Bibliography

- [1] F. Duarte and A. Silva. “An introduction to inverse problems with applications”. *Springer*, 2013.
- [2] G. C. Goodwin and K. S. Sin. “Adaptive Filtering Prediction and Control”. *Dover Publications*, 1984.
- [3] H. T. Banks, R. C Smith and Y. Wang. “Smart material structures: modeling, estimation and control”. *Wiley*, 1996.
- [4] P. Christofides. “Nonlinear and Robust Control of PDE Systems: Methods and Applications to Transport-Reaction Processes”. *Birkhäuser*, 2001.
- [5] H. T. Banks and K. Kunisch. “Estimation techniques for distributed parameter systems”. *Birkhäuser*, 1989.
- [6] M. A. Demetriou. “Adaptive observers for non-square positive real infinite dimensional systems”. *American Control Conference (ACC)*. 2016, pp. 3441-3448.
- [7] M. A. Demetriou. “Adaptive Parameter Estimation of Abstract Parabolic and Hyperbolic Distributed Parameter Systems”. *Ph.D. thesis, Departments of Electrical Engineering-Systems and Mathematics, University of Southern California, Los Angeles*. 1993.
- [8] M. A. Demetriou and I.G. Rosen. “Adaptive identification of second-order distributed parameter systems”. *Inverse Problems*. 1994, vol. 10, no. 2, pp. 261-294.
- [9] Jung-Ki Park. “Principles and Applications of Lithium Secondary Batteries”. *Wiley VCH*, 2012.
- [10] Gregory L. Plett. “Battery Management Systems, Volume I: Battery Modeling; Volume II: Equivalent-Circuit Methods”. *Artech House*, 2015.

- [11] S. Moura, N. Chaturvedi and M. Krstic. “Adaptive PDE observer for battery SOC/SOH estimation via an electrochemical model”. *ASME Journal of Dynamic Systems, Measurement, and Control*. 2014, vol. 136, no.1, paper 011015.
- [12] R. F. Curtain and H. J. Zwart. “An Introduction to Infinite Dimensional Linear Systems Theory”. *Springer-Verlag*, 1995.
- [13] Han-Xiong Li and Chenkun Qi. “Modeling of distributed parameter systems for applications-A synthesized review from time-space separation”. *Journal of Process Control*. 2010, vol.20, pp. 891-901
- [14] Han-Xiong Li and Chenkun Qi. “Spatio-Temporal Modeling of Nonlinear Distributed Parameter Systems”. *Springer*, 2011.
- [15] V. Kecman. “State-Space Models of Lumped and Distributed Systems”. *Springer-Verlag*, 1988.
- [16] W. Harmon Ray. “Some recent applications of distributed parameter systems theory-A survey”. *Proceedings of the Second IFAC Symposium, Coventry, Great Britain*. 1977, pp. 21-32.
- [17] M. Athans. “Toward a practical theory for distributed parameter systems”. *IEEE Transactions on Automatic Control*. 1970, vol. 15, no. 2, pp. 245-247.
- [18] W. J. Liu. “Elementary Feedback Stabilization of the Linear Reaction-Convection-Diffusion Equation and the Wave Equation”. *Springer. Series of Mathématiques & Applications*, 2009.
- [19] Dung Minh Ha. “Functional Analysis. Volume I: A Gentle Introduction”. *Matrix Editions*, 2006.
- [20] B. Rynne and M. Youngson. “Linear Functional Analysis”. *Springer Undergraduate Mathematics Series*. Second Edition, 2008.
- [21] M. Krstic and A. Smyshlyaev. “Boundary Control of PDEs: A Course on Backstepping Designs”. *Society for Industrial and Applied Mathematics (SIAM)*, 2008.
- [22] I. Lasiecka and R. Triggiani. “Control Theory for Partial Differential Equations: Continuous and Approximation Theories. I: Abstract Parabolic Systems. II: Abstract Hyperbolic-like Systems over a Finite Time Horizon”. *Cambridge University Press. Encyclopedia of Mathematics and its Applications*, 2000.
- [23] I. Lasiecka. “Control of systems governed by partial differential equations: a historical perspective”. *Proceedings of IEEE Conference on Decision and Control*. 1995, pp. 2792-2796.
- [24] W. Harmon Ray. “Advanced Process Control”. *McGraw-Hill Inc*, 1981.

- [25] Tilman Utz and Knut Graichen. “Two-degrees-of-freedom Optimization-based Control and Estimation of a Parabolic Equation System”. *Proceedings of the American Control Conference (ACC)*, 2013, pp. 384-389.
- [26] J. Castillo and G. Miranda. “Mimetic Discretization Methods”. *Chapman and Hall/CRC*, 2013.
- [27] J. Hyman and J. Scovel. “Deriving Mimetic Difference Approximations to Differential Operators Using Algebraic Topology”. *Los Alamos National Laboratory*. 1988, unpublished report.
- [28] P. Bochev and J. Hyman. “Principles of Mimetic Discretizations of Differential Operators”. *IMA Volumes in Mathematics and its Applications*. 2006, v.142, pp. 89-119.
- [29] S.H. Christiansen, H. Z. Munthe-Kaas and B. Owren. “Topics in structure-preserving discretization”. *Acta Numerica*. 2011, vol. 20, pp 1-119.
- [30] J. B. Perot. “The Relationship Between Discrete Calculus Methods and Other Mimetic Approaches”. *Woudschoten Conference, Zeist, Netherlands*. 2011.
- [31] Leo J. Grady and J. Polimeni. “Discrete Calculus: Applied Analysis on Graphs for Computational Science”. *Springer*, 2010.
- [32] Thomas Meurer. “On the Extended Luenberger-Type Observer for Semilinear Distributed Parameter Systems”. *IEEE Transaction on Automatic Control*. 2013, vol. 58, no. 7, pp. 1732-1742.
- [33] D.J. Cooper, W.F. Ramirez and D.E Clough. “Comparison Of Linear Distributed-Parameter Filters To Lumped Approximants”. *Aiche Journal*. 1986, vol.32, no. 2, pp. 186-194.
- [34] A. Bensoussan, G. Da Prato, M.C. Delfour and S.K Mitter. “Representation and Control of Infinite Dimensional Systems”. *Birkhäuser. Systems & Control: Foundations & Applications*, Second Edition, 2007.
- [35] M. Tucsnak and G. Weiss. “Observation and Control for Operator Semigroups”. *Birkhäuser*, 2009.
- [36] A. Pazy. “Semigroups of Linear Operators and Applications to Partial Differential Equations”. *Springer, Applied Mathematical Sciences*, 1983.
- [37] K.-J. Engel and R. Nagel. “One-Parameter Semigroups for Linear Evolution Equations”. *Springer, Graduate Text in Mathematics*, 2000.
- [38] H.T. Banks. “A Functional Analysis Framework for Modeling, Estimation and Control in Science and Engineering”. *Chapman and Hall-CRC Press*, 2012.

- [39] R. Datko. “Extending a theorem of A. M. Liapunov to Hilbert spaces”. *Journal Math. Anal. Appl.* 1970, vol. 32, pp. 610-616.
- [40] F. Riesz and B. Sz.-Nagy. “Functional Analysis”. *Dover Publications*, 1990.
- [41] Z. Emirsjlow and S. Townley. “From PDEs with Boundary Control to the Abstract State Equation with an Unbounded Input Operator: A Tutorial”. *European Journal of Control*, 2000, vol. 6, Issue 1, pp. 27-49.
- [42] C. Harkort and J. Deutscher. “An Approach for the Construction of the Control Operator Associated to Boundary Control Systems”. *Proceedings of the 8th International Workshop on Multidimensional Systems (nDS)*. 2013, pp. 1-6.
- [43] J. L. Lions. “Optimal Control of Systems Governed by Partial Differential Equations”. *Springer-Verlag*, 1971.
- [44] A. Bensoussan, M. C. Delfour, and S. K. Mitter. “The linear quadratic optimal control problem for infinite dimensional systems over an infinite horizon; survey and examples”. *Proceedings of the IEEE Conference on Decision and Control*, Dec. 1976, pp. 746-751.
- [45] X. Li and J. Yong. “Optimal Control Theory for Infinite Dimensional Systems”. *Birkhäuser*, 1995.
- [46] K. Hulsing. “Methods for Computing Functional Gains for LQR Control of Partial Differential Equations”. *PhD thesis, Virginia Polytechnic Institute and State University*, 1999.
- [47] J. Saak. “Efficient Numerical Solution of Large Scale Algebraic Matrix Equations in PDE Control and Model Order Reduction”. *PhD thesis, TU Chemnitz*, 2009.
- [48] K. Morris and C. Navasca. “Approximation of low rank solutions for linear quadratic control of partial differential equations”. *Computational Optimization and Applications*. 2010, vol. 46, no. 1, pp. 93-111.
- [49] T. Meurer. “Control of Higher-Dimensional PDEs: Flatness and Backstepping Designs”. *Springer. Series Communications and Control Engineering*, 2013.
- [50] A. Smyshlyaev and M. Krstic. “Adaptive Control of Parabolic PDEs”. *Princeton University Press*, 2010.
- [51] R. Vazquez and M. Krstic. “Control of Turbulent and Magnetohydrodynamic Channel Flows. Boundary Stabilization and State Estimation”. *Birkhäuser*, 2007.
- [52] Norbert Wiener. “The Extrapolation, Interpolation and Smoothing of Stationary Time Series”. *John Wiley and Sons*, 1949.

- [53] Rudolph Emil Kalman. "A New Approach to Linear Filtering and Prediction Problems". *Transactions of the ASME, Journal of Basic Engineering*. 1960, vol. 82, series D, pp. 35-45.
- [54] Thomas Kailath, Ali H. Saye and Babak Hassibi. "Linear Estimation". *Prentice Hall*, 2000.
- [55] Sigeru Omatu and John H. Seinfeld. "Distributed Parameter Systems: Theory and Applications". *Oxford Science Publications*, 1989.
- [56] Ruth Curtain and Anthony Pritchard. "Infinite dimensional linear systems theory". *Springer-Verlag*, 1978.
- [57] William J. Terrell. "Stability and Stabilization: An Introduction". Princeton University Press, 2009.
- [58] Eduardo D. Sontag. "Mathematical Control Theory: Deterministic Finite Dimensional Systems". Springer-Verlag. 1998, second edition.
- [59] David G. Luenberger. "Observing the State of a Linear System". *IEEE Transactions on Military Electronics*. 1964, vol. 8, no. 2, pp. 74-80.
- [60] Z. Hidayat, R. Babuska and B. De Schutter. "Observers for Linear Distributed-Parameter Systems: A survey". *IEEE International Symposium on Robot and Sensors Environments (ROSE)*. 2011
- [61] M. Vidyasagar, R. Goodson and R. Klein. "Comments on A definition and some results for distributed system observability". *IEEE Transactions on Automatic Control*. 1971, vol. 16, no. 1, pp. 106.
- [62] P. K. C. Wang, "Control of distributed parameter systems". *Advances in Control Systems*, vol. I, C.T. Leondes, Ed. New York: Academic Press. 1964, 75-172.
- [63] P. K. C. Wang and F. Tung, "Optimum control of distributed parameter system". *Trans. ASME J. Basic Eng.*, 1964, pp. 67-79.
- [64] R. Goodson and R. Klein. "A definition and some results for distributed system observability". *IEEE Transactions on Automatic Control*. 1970, vol. 15, no. 2, pp. 165-170.
- [65] Susanne C. Brenner and Ridgway Scott. "The Mathematical Theory of Finite Element Methods". *Springer*. 2008, third edition.
- [66] M. Balas. "Do all linear flexible structures have convergent second-order observers?". *Journal Of Guidance Control And Dynamics*. 1999, vol. 22, no. 6, pp. 905-908.
- [67] L. A. Gould. "Chemical Process Control: Theory and Applications". *Addison-Wesley*, 1969.

- [68] Stig Larsson. “Partial Differential Equations With Numerical Methods”. *Springer*, 2008, second edition.
- [69] M. A. Demetriou. “Fault diagnosis of hyperbolic distributed parameter systems”. *Proceedings of IEEE International Symposium on Intelligent Control*. 1996, pp. 194-196.
- [70] H. T. Banks and D. J. Inman. “On Damping Mechanisms In Beams”. *Journal Of Applied Mechanics-Transactions of rhe Asme*. 1991, vol. 58, no. 3, pp. 716-723.
- [71] M. Balas. “Active Control of Flexible Systems”. *Journal of Optimization Theory and Applications*. 1978, vol. 25, no. 3, pp. 415-436.
- [72] L. Meirovitch and H. Baruh. “Effect of damping on observation spillover instability”. *Journal of Optimization Theory and Applications*. 1981, vol. 35, no. 1, pp. 31-44.
- [73] Jean-Michel Coron. “Control and Nonlinearity”. *American Mathematical Society*, 2007.
- [74] Georges Bastin and Jean-Michel Coron. “Stability and Boundary Stabilization of 1-D Hyperbolic Systems”. *Birkhäuser*, 2016.
- [75] Zheng-Hua Luo, Bao-Zhu Guo and Orner Morgul. “Stability and Stabilization of Infinite Dimensional Systems with Applications”. *Springer*, 1999.
- [76] Christopher D. Rahn. “Mechatronic Control of Distributed Noise and Vibration: A Lyapunov Approach”. *Springer*, 2001.
- [77] Emilia Fridman and Yury Orlov. “An LMI approach to H_1 boundary control of semilinear parabolic and hyperbolic systems”. *Automatica*. 2009, vol. 45, pp. 2060-2066.
- [78] Emilia Fridman. “Observers and initial state recovering for a class of hyperbolic systems via Lyapunov method”. *Automatica*. 2013, vol. 49, pp. 2250-2260.
- [79] Aditya Gahlawat. “Analysis and Control of Parabolic Partial Differential Equations with Application to Tokamaks using Sum-of-Squares Polynomials”. *PhD Thesis, Arizona State University*, 2016.
- [80] A. Gahlawat and M. M. Peet. “A Convex Approach to Output Feedback Control of Parabolic PDEs Using Sum-of-Squares”. To Appear in *IEEE Transactions on Automatic Control*.
- [81] G. Valmorbida, M. Ahmadi and A. Papachristodoulou. “Stability Analysis for a Class of Partial Differential Equations via Semidefinite Programming”. *IEEE Transactions on Automatic Control*. 2016, vol. 61, no. 6, pp. 1649-1654.
- [82] F. Bribiesca-Argomedo and M. Krstic. “Backstepping-Forwarding Control and Observation for Hyperbolic PDEs with Fredholm Integrals”. *IEEE Transactions on Automatic Control*, 2015, vol. 60, no. 8, pp. 2145-2160.

- [83] R. Vazquez. “Boundary Control Laws and Observer Design for Convective, Turbulent and Magnetohydrodynamic Flows”. *PhD Thesis, University of California, San Diego*. 2006.
- [84] D. Tsubakino and S. Harab. “Backstepping observer design for parabolic PDEs with measurement of weighted spatial averages”. *Automatica*, 2015, vol. 53, pp. 179-187.
- [85] R. Khosroushahi, and H. Marquez. “Reducing domain structural complexity in PDE backstepping boundary observer design using conformal mapping”. *European Control Conference (ECC)*. 2013, pp. 3113-3118.
- [86] A. Smyshlyaev and M. Krstic. “Backstepping Observers for a Class of Parabolic PDEs”. *Systems and Control Letters*, 2005, vol. 54, no. 7, pp.613-625.
- [87] A. Kirsch. “An introduction to the mathematical theory of inverse problems”. *Springer*, 2011.
- [88] A. Tarantola. “Inverse Problem Theory and Methods for Model Parameter Estimation”. *Society for Industrial and Applied Mathematics (Siam)*. 2005.
- [89] H.T. Banks and H.T. Tran. “Mathematical and Experimental Modeling of Physical and Biological Processes”. CRC Press, 2009.
- [90] H. T. Banks, Shuhua Hu, and W. Clayton Thompson “Modeling and Inverse Problems in the Presence of Uncertainty”. CRC Press, 2014.
- [91] M. P. Polis and R. E. Goodson. “Parameter Identification in Distributed Systems: A Synthesizing Overview”. *Proceedings of the IEEE*. 1976, vol. 64, no.1, pp. 45-61.
- [92] H. T. Banks (Editor). “Control and Estimation in Distributed Parameter Systems”. *Society for Industrial and Applied Mathematics (SIAM)*, 1992.
- [93] G . A . Phillipson. “Identification of Distributed Systems”. *American Elsevier*, 1971.
- [94] M. Courdresses, M. Polis and M. Amouroux. “On identifiability of parameters in a class of parabolic distributed systems”. *IEEE Transactions on Automatic Control*. 1981, vol. 26, no. 2, pp. 474-477.
- [95] S. Kitamura and S. Nakagiri. “Identifiability of Spatially-Varying and Constant Parameters in Distributed Systems of Parabolic Type”. *SIAM Journal Control and Optimization*. 1997, vol. 15, no. 5, pp. 785-802.
- [96] G. Freiling and V. Yurko. “Inverse Sturm-Liouville Problems and Their Applications”. *Nova Biomedical*, 2001.
- [97] A. Pierce. “Unique identification of eigenvalues and coefficients in a parabolic problem”. *SIAM J. Control and Optimization*. 1979, vol. 17, no. 4, pp. 494-499.

- [98] S. Nakagiri. "Review of Japanese work of the last 10 years on identifiability in distributed". *Inverse Problems*. 1993, vol. 9, pp. 143-191.
- [99] Costas Kravaris and John H. Seinfeld. "Identifiability of Spatially-Varying Conductivity from Point Observation as an Inverse Sturm-Liouville Problem". *SIAM Journal on Control and Optimization*. 1986, vol. 24, no. 3, pp. 522-542.
- [100] Semion Gutman and Junhong Ha. "Identifiability of piecewise constant conductivity and its stability". *Proceedings of the 47th IEEE Conference on Decision and Control*. 2008, pp. 2728-2733.
- [101] Semion Gutman and Junhong Ha. "Parameter identifiability for heat conduction with a boundary input". *Mathematics and Computers in Simulation*. 2009, vol. 79, pp. 2192-2210.
- [102] T. Suzuki. "Gel'fand-Levitan's theory deformation formulas and inverse problems". *J.Fac. Sci. Univ. Tokyo, Sect. IA, Math*. 1985, vol. 32, pp. 223-271.
- [103] I. M. Gelfand and B. M. Levitan. "On the determination of a differential equation from its spectral function". *Izvestiya Akad. Nauk SSSR*. 1951, 15, pp. 309-360; English translation: *Amer. Math. Soc. Transl.*. 1955, 2-1, pp. 253-304.
- [104] C. Farrar, S. Doebling and D. Nix. "Vibration-based structural damage identification". *Philosophical Transactions Of The Royal Society Of London Series A-Mathematical Physical And Engineering Sciences*. 2001, vol. 359, no. 1778, pp. 131-149.
- [105] H. T. Banks, D. J. Inman, D. J. Leo and Y. Wang. "An experimentally validated damage detection theory in smart structures". *Journal of Sound and Vibration*. 1996, vol. 191, no. 5, pp. 859-880.
- [106] H. T. Banks and A. Criner. "Thermal based methods for damage detection and characterization in porous materials". *Inverse Problems*. 2012, vol. 28, no. 6.
- [107] Y. Zou, L. Tong and G. Steven. "Vibration-Based Model-Dependent Damage (Delamination) Identification And Health Monitoring For Composite Structures -A Review". *Journal Of Sound And Vibration*. 2000, vol. 230, no. 2, pp. 357-378.
- [108] M. Friswell. "Damage Identification Using Inverse Methods". *Philosophical Transactions Of The Royal Society A-Mathematical Physical And Engineering Sciences*. 2007, vol. 365, no. 1851, pp. 393-410.
- [109] S. A. Billings. "Nonlinear System Identification: NARMAX Methods in the Time, Frequency, and Spatio-Temporal Domains". *Wiley*, 2013.

- [110] L. Z. Guo and S. A. Billings. "Detection of fatigue cracks in a beam using a spatio-temporal dynamical system identification method". *Journal of Sound and Vibration*. 2007, vol. 299, no. 1-2, pp. 22-35.
- [111] Y. Guo, L. Z. Guo, S. A. Billings, D. Coca and Z. Q. Lang. "Characterizing Nonlinear Spatio-Temporal Systems In The Frequency Domain". *International Journal Of Bifurcation And Chaos*. 2012, vol. 22, no. 2, 1230009.
- [112] J. Baumeister and W. Scondo. "Asymptotic embedding methods for parameter estimation". *Proceedings of the 26th IEEE Conference on Decision and Control*. 1987, pp. 170-174.
- [113] W. Scondo. "Ein Modellabgleichsverfahren zur adaptiven Parameteridentifikation in Evolutionsgleichungen". *Ph.D. thesis, Johann Wolfgang Goethe-Universitat zu Frankfurt am Main, Frankfurt am Main, Germany*. 1987.
- [114] M. A. Demetriou. "Adaptive consensus filters for collocated infinite dimensional systems". *Proceedings of the 50th IEEE Conference on Decision and Control and European Control Conference*. 2011, pp. 597-602.
- [115] M. A. Demetriou, Kazufumi Ito and Ralph C. Smith. "Adaptive techniques for the MRAC, adaptive parameter identification, and on-line fault monitoring and accommodation for a class of positive real infinite dimensional systems". *International Journal of Adaptive Control and Signal Processing*. 2009, vol. 23, no. 2, pp.193-215.
- [116] M. A. Demetriou and K. Ito. "Optimal online parameter estimation for a class of infinite dimensional systems using Kalman filters". *American Control Conference*. 2003, vol. 3, pp. 2708-2713.
- [117] M. A. Demetriou and I.G. Rosen. "On-line robust parameter identification for parabolic systems". *International Journal of Adaptive Control and Signal Processing*. 2001, vol. 15, no. 6, pp. 615-631.
- [118] R. F. Curtain, M. A. Demetriou and K. Ito. "Adaptive observers for slowly time varying infinite dimensional systems". *Proceedings of the 37th IEEE Conference on Decision and Control*. 1998, vol. 4, pp. 4022-4027.
- [119] M. A. Demetriou and I.G. Rosen. "Robust adaptive estimation schemes for parabolic distributed parameter systems". *Proceedings of the 36th IEEE Conference on Decision and Control*. 1997, vol. 4, pp.3448-3453.
- [120] M. A. Demetriou. "Adaptive Parameter Estimation for a Class of Distributed Parameter Systems with Persistence of Excitation". *Proceedings of the 31st IEEE Conference on Decision and Control*. 1992, pp. 1742-1743.

- [121] J. Baumeister, W. Scondo, M. A. Demetriou and I. G. Rosen. “On-Line Parameter Estimation for Infinite-Dimensional Dynamical Systems”. *SIAM J. Control Optim.* 1997, vol. 35, no. 2, pp. 678-713.
- [122] M. Krstic. “Systematization of approaches to adaptive boundary control of PDEs”. *International Journal of Robust and Nonlinear Control*. 2006, vol. 16, pp. 801-818.
- [123] M. Krstic and A. Smyshlyaev, “Adaptive boundary control for unstable parabolic PDEs - Part I: Lyapunov design”. *IEEE Transactions on Automatic Control*, 2008, vol. 53, no.7, pp. 1575-1591.
- [124] A. Smyshlyaev and M. Krstic. “Adaptive boundary control for unstable parabolic PDEs - Part II: Estimation-based designs”. *Automatica*. 2007, vol. 43, pp. 1543-1556.
- [125] A. Smyshlyaev and M. Krstic. “Output-feedback adaptive control for parabolic PDEs with spatially varying coefficients”. *IEEE Conference on Decision and Control*. 2006.
- [126] A. Smyshlyaev and M. Krstic. “Adaptive boundary control for unstable parabolic PDEs - Part III: Output-feedback examples with swapping identifiers”. *Automatica*. 2007, vol. 43, pp. 1557-1564.
- [127] P. Ioannou and J. Sun. “Robust Adaptive Control”. *Prentice-Hall, NJ*, 1996.
- [128] D. Wexler. “On frequency domain stability for evolution equations in hilbert space via the algebraic riccati equation”. *SIAM J. Control and Optimization*. 1980, vol. 11, pp. 969-983.
- [129] R. F. Curtain. “Old and new perspectives on the positive-real lemma in systems and control theory”. *ZAMM Journal of Applied Mathematics and Mechanics*. 1999. vol. 79, no. 9, pp. 579-590.
- [130] Agus Hasan. “Adaptive boundary control and observer of linear hyperbolic systems with application to managed pressure drilling”. *Proceedings of the ASME 2014 Dynamic Systems and Control Conference*. 2014
- [131] Agus Hasan. “Adaptive boundary observer for nonlinear hyperbolic systems: Design and field testing in managed pressure drilling”. *American Control Conference*. 2015, pp. 2606-2612.
- [132] Agus Hasan, Ole Morten Aamo and Miroslav Krstic. “Boundary observer design for hyperbolic PDE-ODE cascade systems”. *Automatica*. 2016, vol. 68, pp. 75-86.
- [133] T. Ahmed-Al, F. Giri, M. Krstic, M., L. Burlion and F. Lamnabhi-Lagarigue. “Adaptive observer for parabolic PDEs with uncertain parameter in the boundary condition”. *European Control Conference*. 2015, pp. 1343-1348.

- [134] T. Ahmed-Al, F. Giri, M. Krstic, M., F. Lamnabhi-Lagarrigue and L. Burlion. “Adaptive observer for a class of parabolic PDEs”. *IEEE Transactions on Automatic Control*. 2016, vol. 61, no. 10, pp. 3083-3090.
- [135] T. Ahmed-Al, F. Giri, M. Krstic, M., L. Burlion and F. Lamnabhi-Lagarrigue. “Adaptive boundary observer for parabolic PDEs subject to domain and boundary parameter uncertainties”. *Automatica*. 2016, vol.72, pp. 115-122.
- [136] G. Blekherman, P. Parrilo and R. Thomas. “Semidefinite Optimization and Convex Algebraic Geometry”. *Society for Industrial and Applied Mathematics*, 2013.
- [137] J. B. Lasserre. “Moments, Positive Polynomials and Their Applications”. *Imperial College Press Optimization Series*, 2010.
- [138] J. B. Lasserre. “An Introduction to Polynomial and Semi-Algebraic Optimization”. *Cambridge University Press, Cambridge Texts in Applied Mathematics*, 2015.
- [139] Friedrich Sauvigny. “Partial Differential Equations: Vol. 1 Foundations and Integral Representations”. *Springer*, 2006.
- [140] J. W. Thomas. “Numerical Partial Differential Equations: Finite Difference Methods”, *Springer*, 1995.
- [141] Endre Süli and David F. Mayers. “An Introduction to Numerical Analysis”. *Cambridge University Press*. 2003.
- [142] O.C. Zienkiewicz, R.L. Taylor and J.Z. Zhu. “The Finite Element Method: Its Basis and Fundamentals”. *Elsevier Butterworth-Heinemann*, 2005, sixth edition.
- [143] Mark S. Gockenbach. “Partial Differential Equations: Analytical and Numerical Methods”. *Society for Industrial and Applied Mathematics*, 2010, second edition.
- [144] Mark S. Gockenbach. “Understanding and Implementing the Finite Element Method”. *Society for Industrial and Applied Mathematics*, 2006.
- [145] K.W. Morton and David F. Mayers. “Numerical Solution of Partial Differential Equations: An Introduction”. *Cambridge University Press*, 2005, second edition.
- [146] Alfio Quarteroni. “Numerical Approximation of Partial Differential Equations”. *Springer*, 2008.
- [147] Mark H. Holmes. “Introduction to Numerical Methods in Differential Equations”. *Springer*, 2007.

- [148] M. Laurent. “Sums of squares, moment matrices and optimization over polynomials”. *Springer, in Emerging Applications of Algebraic Geometry, IMA Volumes in Mathematics and its Applications*, 2009, vol. 149, pp. 157-270. [Online]. Available: <http://homepages.cwi.nl/~monique/>
- [149] M. Marshall. “Positive Polynomials and Sums of Squares”. *Mathematical Surveys and Monographs, American Mathematical Society*, 2008, vol. 146.
- [150] W. A. Sutherland. “Introduction to Metric and Topological Spaces”. *Oxford University Press*, 1981, second edition.
- [151] Nick Trefethen. “Approximation Theory and Approximation Practice”. *Society for Industrial and Applied Mathematics*, 2012.
- [152] Michael J. D. Powell. “Approximation Theory and Methods”. *Cambridge University Press*, 1981.
- [153] Tom M. Apostol. “Mathematical Analysis”. *Pearson*, 1974, second edition.
- [154] John H. Hubbard and Barbara Burke Hubbard. “Vector Calculus, Linear Algebra, and Differential Forms: A Unified Approach”. *Matrix Editions*, 2009, fourth edition.
- [155] P. Olver. “Applied Mathematics Lecture Notes”. 2014. [Online]. Available: <http://www.math.umn.edu/~olver/appl.html>
- [156] Alan Jeffrey. “Applied partial differential equations. An introduction”. *Academic Press*, 2002.
- [157] C.W. Scherer and C.W.J. Hol. “Matrix Sum-of-Squares Relaxations for Robust Semi-Definite Programs”. *Mathematical Programming*. 2006, vol. 107, issue 1-2, pp.189-211.
- [158] Jacob Fish and Ted Belytschko. “A First Course in Finite Elements”. *John Wiley & Sons*, 2007.
- [159] H. L. Royden and P.M Fitzpatrick. “Real Analysis”. *Macmillan*, fourth edition, 2010.
- [160] C. Caramanis. “Solving Linear Partial Differential Equations via Semidefinite Optimization”. *MSc. thesis, Massachusetts Institute of Technology*, 2001.
- [161] D. Bertsimas and C. Caramanis. “Bounds on linear PDEs via Semidefinite Optimization”. *Springer-Verlag. Journal Mathematical Programming. Series A*, 2006, vol. 108, issue 1, pp. 135-158.
- [162] D. Henrion, J. Lasserre and M. Mevissen. “Mean Squared Error Minimization for Inverse Moment Problems”. *Springer. Applied Mathematics & Optimization*, 2014, vol. 70, n.1, pp. 83-110.

- [163] J. Löfberg. “YALMIP : A Toolbox for Modeling and Optimization in MATLAB,” in *Proc. CACSD Conference*, 2004. [Online]. Available: <http://users.isy.liu.se/johanl/yalmip>
- [164] A. Papachristodoulou, J. Anderson, G. Valmorbida, S. Prajna, P. Seiler and P. A. Parrilo, *SOSTOOLS: Sum of squares optimization toolbox for MATLAB, version 3.01*, July 2016. [Online]. Available: <http://www.eng.ox.ac.uk/control/sostools>
- [165] J. F. Sturm, “Using SeDUMi 1.02, a MATLAB toolbox for optimization over symmetric cones,” *Optimization Methods and Software*, vol. 11 & 12, pp. 625-653, August 1999. [Online]. Available: <http://sedumi.ie.lehigh.edu/>
- [166] *The MOSEK optimization toolbox for MATLAB, version 8*. Denmark: MOSEK ApS. [Online]. Available: <https://www.mosek.com/>
- [167] J. David Logan. “An Introduction to Nonlinear Partial Differential Equations”. *Wiley-Interscience*, 2008, second edition.
- [168] Richard L. Burden and J. Douglas Faires. “Numerical Analysis”. *Cengage Learning*, 2011, ninth edition.
- [169] Arieh Iserles. “A First Course in the Numerical Analysis of Differential Equations”. *Cambridge University Press*, 2008, second edition.
- [170] G. M. Phillips and Peter J. Taylor. “Theory and Applications of Numerical Analysis”. *Academic Press*, 1996, second edition.
- [171] L. Fox and D. F. Mayers. “Numerical Solution of Ordinary Differential Equations for Scientists and Engineers”. *Springer*, 2013.
- [172] Kendall Atkinson, Weimin Han and David E. Stewart. “Numerical Solution of Ordinary Differential Equations”. *Wiley-Blackwell*, 2009.
- [173] E. Ward Cheney and David R. Kincaid. “Numerical Mathematics and Computing”. *Cengage Learning*, 2012, seventh edition.
- [174] K. Eriksson, D. Estep P. Hansbo and C. Johnson. “Computational Differential Equations”. *Cambridge University Press*. 1996, second edition.
- [175] J. Nathan Kutz. “Data-Driven Modeling and Scientific Computation: Methods for Complex Systems and Big Data”. *OUP Oxford*, 2013.
- [176] P. Olver. “Introduction to Partial Differential Equations”. *Springer. Series of Undergraduate Texts in Mathematics*, 2014.
- [177] M.A. Al-Gwaiz. “Sturm-Liouville Theory and its Applications”. *Springer*, 2008.

- [178] Alfio Quarteroni. “Numerical Models for Differential Problems”. *Springer*, 2014, second edition.
- [179] Joseph E. Flaherty. “Course Notes - Finite Element Analysis”. *Rensselaer Polytechnic Institute*. [Online]. Available: <http://www.cs.rpi.edu/flaherje/>
- [180] Sandro Salsa. “Partial Differential Equations in Action From Modelling to Theory”. *Springer*, 2016, second edition.
- [181] Nakhle H. Asmar. “Partial Differential Equations with Fourier Series and Boundary Value Problems”. *Dover Books on Mathematics*, 2016, third edition.
- [182] Richard Haberman. “Applied Partial Differential Equations with Fourier Series and Boundary Value Problems”. *Prentice-Hall Inc.*, 1998, third edition.
- [183] M. Dehghana and M.R. Eslahchib. “Best uniform polynomial approximation of some rational functions”. *Computers and Mathematics with Applications*, 2010, vol. 59, issue 1, pp. 382-390
- [184] B. Anderson and J. B. Moore. “Optimal Control: Linear Quadratic Methods”. *Dover Publications*, 2007.
- [185] D.S. Bernstein and P. Tsiotras. “A Course in Classical Optimal Control”. *Pre-print Georgia Institute of Technology*, 2009.
- [186] P. Ascencio, A. Astolfi and T. Parisini, “Backstepping PDE Design, Volterra and Fredholm Operators: a Convex Optimization Approach,” *IEEE IEEE 55th Conference on Decision and Control*, 2015, pp. 7048-7053.
- [187] P. Ascencio, A. Astolfi and T. Parisini, “Backstepping PDE Design: A Convex Optimization Approach,” Submitted to *IEEE Transactions on Automatic Control*, 2016.
- [188] V. Volterra. “Theory of Functionals, and of Integral and Integro-Differential Equations”. *Blackie and Son, London*, 1930.
- [189] K. Yosida. “Lectures on Differential and Integral Equations”. *Dover Publications*, 1991.
- [190] S. M. Zemlyan. “The Classical Theory of Integral Equations: A Concise Treatment”. *Birkhäuser Mathematics*, 2012.
- [191] W. J. Liu. “Boundary feedback stabilization of an unstable heat equation”. *SIAM Journal on Control and Optimization*. 2003, vol. 42, no. 3, pp. 1033-1043.
- [192] I. Gohberg, S. Goldbergand and M. Kaashoek. “Basic Classes of Linear Operators”. *Birkhäuser*, 2003.
- [193] C. Corduneanu. “Integral Equations and Applications”. *Cambridge University Press*, 1991

- [194] J. B. Conway. “A Course in Functional Analysis”. *Springer*, 1997.
- [195] C. Corduneanu. “Functional Equations with Causal Operators”. *CRC Press, Taylor & Francis*, 2002.
- [196] V. Lakshmikantham, S. Leela, Z. Drici and F. McRae. “Theory of causal differential equations”. *Atlantis Press*, 2010.
- [197] A. Smyshlyaev and M. Krstic. “Closed-Form Boundary State Feedbacks for a Class of 1-D Partial Integro-Differential Equations”. *IEEE Transactions on Automatic Control*. 2004, vol. 49, no. 12, pp. 2185-2202.
- [198] M. Krstic and A. Smyshlyaev. “Backstepping boundary control for first-order hyperbolic PDEs and application to systems with actuator and sensor delays”. *Systems & Control Letters*. 2008, vol. 57, issue 9, pp. 750-758.
- [199] A. Smyshlyaev, E. Cerpa and M. Krstic, “Boundary stabilization of a 1-D wave equation with in-domain antidamping”. *SIAM Journal of Control and Optimization*. 2010, vol. 48, pp. 4014-4031.
- [200] Ram P. Kanwal. “Linear Integral Equations. Theory and Practice”. *Birkhäuser*, 1996, second edition.
- [201] A. Jerri. “Introduction to Integral Equations with Applications”. *Wiley*, 1999, second edition.
- [202] C. Guo, C. Xie and C. Zhou. “Stabilization of a spatially non-causal reaction-diffusion equation by boundary control”. *International Journal of Robust and Nonlinear Control*. 2012, vol. 24, issue 1, pp. 1-17.
- [203] N. Bekiaris-Liberis and M. Krstic. “Lyapunov Stability of Linear Predictor Feedback for Distributed Input Delays”. *IEEE Transactions on Automatic Control*. 2011, vol. 56, no. 3, pp. 655-660.
- [204] F. Bribiesca-Argomedeo and M. Krstic. “Backstepping-Forwarding Boundary Control Design for First-Order Hyperbolic Systems With Fredholm Integrals”. *American Control Conference*. 2014, pp. 5428-5433.
- [205] D. Tsubakino, F. Bribiesca Argomedeo and M. Krstic. “Backstepping-Forwarding Control of Parabolic PDEs with Partially Separable Kernels”. *IEEE Conference on Decision and Control*. 2014, pp. 5236-5241.
- [206] S. Almezal, Q. H. Ansari and M. A. Khamsi (Editors). “Topics in Fixed Point Theory”. *Springer*, 2014.
- [207] V. Berinde. “Iterative Approximation of Fixed Points”. *Springer*, 2007, second edition.

- [208] R. Kress. "Linear Integral Equations". *Springer. Series Applied Mathematical Sciences*, 2014, third edition.
- [209] F. Cali, E. Marchetti and V. Muresan. "On some Volterra-Fredholm Integral Equation". *International Journal of Pure and Applied Mathematics*. 2006, vol. 31, no. 2, pp. 173-184.
- [210] S. Moradi, M. Mohammadi Anjedani and E. Analoei. "On Existence and Uniqueness of Solutions of a Nonlinear Volterra-Fredholm Integral Equation". *International Journal of Nonlinear Analysis and Applications*. 2015, vol. 6, no. 1, pp. 62-68.
- [211] M. A. Khamsi and W. A. Kirk. "An Introduction to Metric Spaces and Fixed Point Theory". *John Wiley & Sons. Pure and Applied Mathematics*, 2001.
- [212] H. Hochstadt. "Integral Equations". *Wiley*, 1989.
- [213] I. Stakgold and M. J. Holst. "Green's Functions and Boundary Value Problems". *Wiley*, 2011, third edition.
- [214] J. R. Ringrose. "Compact Linear Operators of Volterra Type". *Mathematical Proceedings of the Cambridge Philosophical Society*. 1955, vol. 51, issue 01, pp 44-55.
- [215] P. R. Halmos and V. S. Sunder. "Bounded integral operators on L_2 spaces". *Springer-Verlag*, 1978.
- [216] A. Bowers, N.J. Kalton. "An Introductory Course in Functional Analysis". *Springer*, 2014.
- [217] A.N. Kolmogorov and S.V. Fomin. "Elements of the Theory of Functions and Functional Analysis. Volume 1: Metric and Normed Spaces". *Graylock Press*, 1957.
- [218] E. Kreyszig. "Introductory Functional Analysis with Applications". *Wiley*. 1989.
- [219] B. MacCluer. "Elementary Functional Analysis". *Springer*, second edition, 2009.
- [220] W. Faris. "Methods of Applied Mathematics. Lecture Notes". *Department of Mathematics, University of Arizona*, 2002. [Online]. Available: <http://math.arizona.edu/faris/>
- [221] K. Atkinson and W. Han. "Theoretical Numerical Analysis: A Functional Analysis Framework". *Springer*, 2009, third edition.
- [222] R. A. Kennedy and P. Sadeghi. "Hilbert Space Methods in Signal Processing". *Cambridge University Press*, 2013.
- [223] G. H. Hardy, J. E. Littlewood and G. Polya. "Inequalities". *Cambridge University Press*, 1952, second edition.

- [224] D. Boyd. “Best constants in a class of integral inequalities”. *Pacific Journal of Mathematics*. 1969, vol. 30, no. 2, pp. 367-383.
- [225] T. X. Wang. “Stability in Abstract Functional Differential Equations. Part II. Applications”. *Journal of Mathematical Analysis and Applications*. 1994, vol. 186, issue 3, pp. 835-861.
- [226] P. M. Fitzpatrick. “Advanced Calculus”. *American Mathematical Society*, 2009, second edition.
- [227] S.S. Dragomir. “Some Integral Inequalities of Grüss type”. *RGMIA Research Report Collection*. 1998, vol. 1, no. 2, pp. 95-111.
- [228] S. Boyd and L. Vandenberghe. “Convex Optimization”. *Cambridge University Press*, 2004. [Online]. Available: <http://stanford.edu/~boyd/cvxbook/>
- [229] S. Krantz and H. Parks. “A Primer of Real Analytic Functions”. *Birkhäuser Advanced Text*, 2002, second edition.
- [230] C. Meyer. “Matrix Analysis and Applied Linear Algebra”. *Society for Industrial and Applied Mathematics*, 2000.
- [231] C.S. Kubrusly. “The Elements of Operator Theory”. *Birkhäuser*, 2010, second edition.
- [232] A. T. Bharucha-Reid. “Random Integral Equations”. *Academic Press*, 1972.
- [233] E. Davies. “Linear Operators and their Spectra”. *Cambridge Studies in Advanced Mathematics*, *Cambridge University Press*, 2007.
- [234] A.N. Kolmogorov and S.V. Fomin. “Introductory Real Analysis”. *Dover Publications Inc.*, 2000.
- [235] H. Garth Dales, P. Aiena, J. Eschmeier, K. Laursen and G. Willis. “Introduction to Banach Algebras, Operators, and Harmonic Analysis”. *Cambridge University Press*, 2003.
- [236] V. S. Sunder. “Functional Analysis: Spectral Theory”. *Birkhäuser Advanced Texts*, 1998.
- [237] P.K. Sahoo and T. Riedel. “Mean Value Theorems and Functional Equations”. *World Scientific Publishing Co Pte Ltd.*, 1998.
- [238] Alin Bostan and Philippe Dumas. “Wronskians and Linear Independence”. *The American Mathematical Monthly*. 2010, vol. 117, no. 8, pp. 722-727.
- [239] J. Michael Steele. “The Cauchy Schwarz Master Class. An Introduction to the Art of Mathematical Inequalities”. *Cambridge University Press*, 2004.

- [240] P. Ascencio, A. Astolfi and T. Parisini, “An Adaptive Observer for a class of Parabolic PDEs based on a Convex Optimization Approach of Backstepping PDE Design,” *American Control Conference*. 2016, pp. 3429-3434.
- [241] R.B. Khosroushahi and H.J. Marquez. “PDE Backstepping Boundary Observer Design for Microfluidic Systems”. *IEEE Transactions on Control Systems Technology*. 2015, vol. 23, no. 1, pp. 380-388.
- [242] S. Moura. “Doyle-Fuller-Newman Electrochemical Battery Model”. Matlab-based programs. [Online]. Available: <https://github.com/scott-moura/dfn>
- [243] L. Jadachowski, T. Meurer and A. Kugi. “An Efficient Implementation of Backstepping Observers for Time-Varying Parabolic PDEs”. *7th Vienna International Conference on Mathematical Modelling (MATHMOD)*, 2012.
- [244] P. Bernard and M. Krstic. “Adaptive output-feedback stabilization of non-local hyperbolic PDEs”. *Automatica*. 2015, vol. 50, pp. 2692-2699.
- [245] P. Ascencio, D. Sbarbaro and S. Feye de Azevedo. “An enhanced model for parameter estimation in bioprocesses”. *UKACC International Conference on Control*, 2000.
- [246] P. Ascencio and D. Sbarbaro. “An EKF based on an enhanced model for parameter estimation in bioprocesses”. *European Control Conference (ECC)*. 2001, pp. 3816-3821.
- [247] C.D. Johnson. “Theory of Disturbance-Accommodating Controllers”. *Advances in Control and Dynamic Systems, Academic Press* (edited by C. T. Leondes). 1976, vol. 12, ch. 7, pp. 387-489.
- [248] S. Tang and C. Xie. “Stabilization for a coupled PDE-ODE control system”. *Journal of the Franklin Institute*. 2011, vol. 348, no. 8, pp. 2142-2155.
- [249] M. Kočvara and M. Stingl. *PENPOT: PENBMI, Version 2.1*. [Online]. Available: www.penopt.com
- [250] D. Henrion, J. Löfberg, M. Kočvara and M. Stingl. “Solving polynomial static output feedback problems with PENBMI”. *Proceedings of the joint IEEE Conference on Decision and Control and European Control Conference*. 2005. LAAS-CNRS Research Report No. 05165.
- [251] J. Barzilai and J. M. Borwein. “Two point step size gradient methods”. *IMA Journal of Numerical Analysis*. 1988, vol. 8, pp. 141-148.
- [252] Ya-xiang Yuan . “Step-Sizes for the Gradient Method”. *Report Institute of Computational Mathematics and Scientific/Engineering Computing*. Chinese Academy of Sciences.

- [253] P. Ascencio, D. Sbarbaro and S. Feyeo de Azevedo “An Adaptive Fuzzy Hybrid State Observer for Bioprocesses”. *IEEE Transactions on Fuzzy Systems*. 2004, vol. 12, no. 5, pp. 641-651.
- [254] Joel C. Forman, Saeid Bashash, Jeffrey L. Stein and Hosam K. Fathy. “Reduction of an Electrochemistry-Based Li-Ion Battery Model via Quasi-Linearization and Padé Approximation”. *Journal of The Electrochemical Society*. 2011, vol. 158, no. 2, pp. A93-A101.
- [255] P. Ascencio, A. Astolfi and T. Parisini, “Backstepping PDE-based adaptive observer for a Single Particle Model of Lithium-Ion Batteries,” *IEEE 56th Conference on Decision and Control*. 2016, pp. 5623-5628.
- [256] M. Doyle, T. F. Fuller and J. Newman. “Modeling of galvanostatic charge and discharge of the Lithium/polymer/insertion cell”. *Journal of the Electrochemical Society*. 1993, vol. 140, no. 6, pp. 1526-1533.
- [257] C. Zou, C. Manzie and D. Nešić. “A Framework for Simplification of PDE-Based Lithium-Ion Battery Models”. *IEEE Transaction on Control System Technology*. To appear.
- [258] A.M. Bizeray, S. Zhao, S.R. Duncan, D.A. Howey. “Lithium-ion battery thermal-electrochemical model-based state estimation using orthogonal collocation and a modified extended Kalman filter”. *Journal of Power Sources*. 2015, vol. 296, pp. 400-412.
- [259] R. Klein, N. A. Chaturvedi, J. Christensen, J. Ahmed, R. Findeisen, and A. Kojic. “Electrochemical Model Based Observer Design for a Lithium-Ion Battery”. *IEEE Transaction on Control System Technology*. 2013, vol. 21, no. 2, pp.2 89-301.
- [260] N. Chaturvedi, R. Klein, J. Christensen, J. Ahmed, and A. Kojic. “Algorithms for Advanced Battery Management Systems: Modeling, estimation, and control challenges for Lithium-ion batteries”. *IEEE Control Systems Magazine*. 2010, vol. 30, no. 3, pp. 49-68.
- [261] S. J. Moura. “Estimation and Control of Battery Electrochemistry Models: A Tutorial”. *IEEE 54th Conference on Decision and Control*. 2015, pp. 3906-3912.
- [262] H. E. Perez, S. J. Moura. “Sensitivity-Based Interval PDE Observer for Battery SOC Estimation”. *American Control Conference*. 2015, pp. 323-328.
- [263] S. Renganathan and R. E. White. “Semianalytical method of solution for solid phase diffusion in Lithium ion battery electrodes: Variable diffusion coefficient”. *Journal of Power Sources*. 2011, vol. 196, pp. 442-448.
- [264] S. Tang, Y. Wang, Z. Sahinoglu, T. Wada, S. Hara and M. Krstic. “State-of-Charge Estimation of Lithium-Ion Batteries via a Coupled Thermal-Electrochemical Model”. *American Control Conference*. 2015, pp. 5871-5877.

- [265] S. Moura, N. Chaturvedi and M. Krstic. “PDE Estimation Techniques for Advanced Battery Management Systems-Part I: SOC Estimation”. *American Control Conference*. 2012, pp. 559-565.
- [266] S. Moura, N. Chaturvedi and M. Krstic. “PDE Estimation Techniques for Advanced Battery Management Systems-Part II: SOH Identification”. *American Control Conference*. 2012, pp. 566-571.
- [267] M. Safari and C. Delacourt. “Modeling of a Commercial Graphite/LiFePO₄ Cell”. *Journal of The Electrochemical Society*. 2011, vol. 158, no. 5, pp. A562-A571.
- [268] John Newman and Karen E. Thomas-Alyea. “Electrochemical Systems”. *Wiley-Interscience*. 2004, third edition.
- [269] J. Marcickia, M. Canovaa, A. T. Conliska and G. Rizzonia. “Design and parametrization analysis of a reduced-order electrochemical model of graphite/LiFePO₄ cells for SOC/SOH estimation”. *Journal of Power Sources*. 2013, vol. 237, pp. 310-324.
- [270] R. P. Brent. “Algorithms for Minimization without Derivatives”. *Prentice-Hall, Englewood Cliffs*, 1973.
- [271] J. Newman. “Fortran Programs for the Simulation of Electrochemical Systems”. *University of California, Berkley, CA*. 1998. [Online]. Available: <http://www.cchem.berkeley.edu/jsngrp/fortran.html>
- [272] D. Ucinski. “Optimal Measurement Methods for Distributed Parameter System Identification”. *CRC Press*, 2005.
- [273] J. B. Lasserre. “Semidefinite Programming vs. LP Relaxations for Polynomial Programming”. *Mathematics of Operations Research*. 2002, vol. 27, no. 2, p.347-360
- [274] P. Parrilo. “Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization”. *PhD thesis, California Institute of Technology*, 2000.
- [275] P. Parrilo. “Semidefinite programming relaxations for semialgebraic problems”. *Mathematical Programming*. 2003, Ser. B 96, pp. 293-320 .
- [276] J. B. Lasserre. “Global optimization with polynomials and the problem of moments”. *SIAM Journal on Optimization*. 2001, vol. 11, no. 3, pp. 796-817.
- [277] J. B. Lasserre. “Sum of Squares Approximation of Polynomials, Nonnegative on a Real Algebraic Set”. *SIAM Journal on Optimization*. 2005, vol. 16, no. 2, pp. 610-628.

- [278] Reza Kamyar and Matthew M. Peet. “Polynomial Optimization with Applications to Stability Analysis and Control - Alternatives to Sum of Squares”. *Journal of Discrete and Continuous Dynamical Systems - Series B*. 2015, vol. 20, no. 8, pp. 2383-2417.
- [279] J. B. Lasserre. “A semidefinite programming approach to the generalized problem of moment”. *Mathematical Programming*. 2008, vol. 112, issue 1, pp. 65-92.
- [280] V. Jeyakumar, J. B. Lasserre, G. Li. “On Polynomial Optimization Over Non-compact Semi-algebraic Sets”. *Journal of Optimization Theory and Applications*. 2014, vol. 163, issue 3, pp. 707-718.
- [281] R. Cominetti, F. Facchinei and J.B. Lasserre. “Modern Optimization Modelling Techniques”, *Birkhäuser Applied Probability and Statistics*, 2012.
- [282] K. Schmüdgen. “The K-moment problem for compact semi-algebraic sets”. *Mathematische Annale*. 1991, vol. 289, no.2, pp. 203-206.
- [283] M. Putinar. “Positive polynomials on compact semi-algebraic sets”. *Indiana University Mathematics Journal*. 1993, vol. 42, pp. 969-984.
- [284] G. Blekherman. “There are significantly more non-negative polynomials than sums of squares”. *Israel Journal of Mathematics*. 2006, vol. 153, no. 1, pp. 355-380.
- [285] H.J. Landau. “Moments in Mathematics”. *Proceedings of Symposia in Applied Mathematics*. 1987, vol. 37.
- [286] J. Shohat and J.D. Tamarkin. “The Problem of Moments”. *American Mathematical Society*, 1943.
- [287] N. I. Akhiezer. “The classical moment problem and some related questions in analysis”, *Oliver and Boyd*, 1965.
- [288] H.J Landau. “The classical moment problem: Hilbertian proofs”. *Journal of Functional Analysis*. 1980, vol. 38, no. 2, pp. 255-272.
- [289] M. Capinski and, E. Kopp. “Measure, Integral and Probability”. *Springer Undergraduate Mathematics Series*, 2004, second edition.
- [290] B. Makarov and A. Podkorytov. “Real Analysis: Measures, Integrals and Applications”. *Springer, Universitext Series*, 2013.
- [291] S. Ovchinnikov. “Measure, Integral, Derivative: A Course on Lebesgue’s Theory”. *Springer, Universitext Series*, 2013.

- [292] E. M. Stein and R. Shakarchi. “Real Analysis: Measure Theory, Integration, and Hilbert Spaces”. *Princeton University Press*, 2005.
- [293] W. Rudin. “Real and Complex Analysis”, *McGraw-Hill*, 1987, third edition.
- [294] E. K. Haviland. “On the momentum problem for distribution functions in more than one dimension II”. *American Journal of Mathematics*. 1936, vol. 58, no. 1, pp. 164-168.
- [295] M. Reed and B. Simon, “Methods of Modern Mathematical Physics, II. Fourier Analysis, Self-Adjointness”. *Academic Press*, 1975.
- [296] V. Powers and C. Scheiderer. “The moment problem for non-compact semialgebraic sets”. *Advances in Geometry*. 2001, vol. 1, pp. 71-88.
- [297] W. Van Assche. “The impact of Stieltjes’ work on continued fractions and orthogonal polynomials”. *Springer-Verlag Berlin, Thomas Jan Stieltjes Oeuvres Complètes*. 1993, vol. I, pp. 5-37.
- [298] K. Schmüdgen. “An example of a positive polynomial which is not a sum of squares of polynomials. A positive, but not strongly positive functional”. *Mathematische Nachrichten*. 1979, vol. 88, no. 1, pp. 385-390.
- [299] C. Berg, J. Christensen and C. Jensen. “A remark on the multidimensional moment problem”. *Mathematische Annalen*. 1979, vol. 243, no. 2, pp. 163-169.
- [300] M. Putinar. “Problème des moments sur les compacts semi-algébriques”. *Comptes rendus de l’Académie des sciences. Série 1, Mathématique*. 1996, vol. 323, issue 7, pp. 787-791.
- [301] M. Laurent. “Revisiting two theorems of Curto and Fialkow on moment matrices”. *Proceedings of the American Mathematical Society*. 2005, vol. 133, no. 10, pp. 2965-2976.
- [302] H. Bauer and R.B. Bruckel. “Probability Theory”. *Walter de Gruyter*, 1996.
- [303] H. Bauer. “Measure and Integration Theory”. *Walter de Gruyter*, 2001.
- [304] G.B. Folland. “Real Analysis: Modern Techniques and Their Applications”. *Wiley*, 1999, second edition.
- [305] M. Krstic, I. Kanellakopoulos and P. Kokotovic. “Nonlinear and Adaptive Control Design”. *Wiley*, 1995.
- [306] D. Colton. “The Solution of Initial-Boundary Value Problems for Parabolic Equations by the Method of Integral Operators”. *Journal of Differential Equations*. 1977, vol. 26, issue 2, pp. 181-190.

- [307] T. I. Seidman. “Two results on exact boundary control of parabolic equations”. *Journal Applied Mathematics and Optimization*. 1984, vol. 11, issue 1, pp. 145-152.
- [308] Andrei D. Polyanin and Alexander V. Manzhirov. “Handbook of Integral Equations”. *Chapman and Hall/CRC*, 2008, second edition.
- [309] A. Balogh and M. Krstic. “Infinite Dimensional Backstepping-Style Feedback Transformations for a Heat Equation with an Arbitrary Level of Instability”. *European Journal of Control*. 2002, vol. 8, no. 2, pp. 165-175.
- [310] Ding-Wen Chung, Martin Ebner, David R Ely, Vanessa Wood and R Edwin García. “Validity of the Bruggeman relation for porous electrodes”. *Modelling and Simulation in Materials Science and Engineering*,. 2013, vol. 21, no.7.

Appendix A

Polynomial Optimization via Sum-of-Squares

As J.B. Lasserre pointed out in [138], optimization problems with linear and quadratic functionals look so familiar that we forgot that they are polynomials. Thanks to remarkable results from Algebraic geometry (positivity certificates), these, amongst other classes of polynomials optimization problems, can be “globally” solved on domains characterized as “semi-algebraic sets” with polynomials terms. In addition, positivity certificates based on “Sum-of-Squares” representations lead to “Convex Optimization” problems which can be solved via semidefinite programming, and allows finding solutions with theoretical arbitrary precision. This provides a general methodology to address a wide class of problems, with a beautiful duality between moment sequences and positive polynomials.

A.1 Polynomial Optimization

In general, the polynomial optimization problem

$$\mathbf{P} : \inf_{\mathbf{x} \in \mathbf{K}} p(\mathbf{x}) \quad \Leftrightarrow \quad \begin{cases} p^* = \sup_{\gamma \in \mathbb{R}} \gamma \\ \text{subj. to: } p(\mathbf{x}) - \gamma \geq 0, \\ \mathbf{x} \in \mathbf{K} \end{cases} \quad (\text{A.1})$$

where $\mathbf{K} \subset \mathbb{R}^n$ and $p \in \mathbb{R}[\mathbf{x}]$, is NP-hard¹[148]. However, this can be approximated by hierarchy of convex relaxations² $(\mathbf{P}_d; p_d^*)$, with p_d^* global optimum for $(p(\mathbf{x}) - \gamma) \in \mathbb{P}_{n,2d}[\mathbf{x}]$, using representations of non-negative polynomials via Sum-of-Squares (SOS) [274, 275, 136] or the dual theory

¹NP-hard: Non-deterministic Polynomial-time hard. See [272] for a glance in complexity theory and NP-problems in control systems design.

²The convex relaxation $\mathbf{P}_d$ grows in size and thus its complexity as d increases. It can be a Linear Programming (LP-)relaxation or SDP-relaxation, depending on the type of certificate of positivity considered [273, 137].

of Moments [276, 277, 137, 138] (see [148] for an excellent survey on these topics).³ These representations can be formulated as a semidefinite programming (SDP) problem [275, 279], namely

$$\begin{aligned} p_d^* &= \sup_{\gamma \in \mathbb{R}} \gamma \\ \text{subj. to: } p(\mathbf{x}) - \gamma &= \sum_J s_J(\mathbf{x}) G_J(\mathbf{x}), \quad s_J \in \Sigma_s, \end{aligned} \quad (\text{A.2})$$

where s_J are SOS polynomials to be determined and G_J denotes a particular combination of polynomial constraints g_j 's. Moreover, if

$$\mathbf{K} = \{\mathbf{x} \in \mathbb{R}^n : g_j(\mathbf{x}) \geq 0, j = 1, \dots, m\} \quad (\text{A.3})$$

is a **compact basic semialgebraic set**⁴ for some polynomials $g_j \in \mathbb{R}[\mathbf{x}]$, the sequence of upper bounds ($p_d^* \geq p^*$) is monotone and converge to the global minimum: $p_d^* \rightarrow p^*$ as $d \rightarrow \infty$ [148, 280, 281]. The central point of this approach is the type of positivity certificate (Positivstellensatz) for $p(x) - \gamma \geq 0, \forall \mathbf{x} \in \mathbf{K}$, which is based on powerful results from Real Algebraic Geometry, and its practical computation^{5, 6}

- **Schmüdgen's Positivstellensatz** [282]⁷: Let $\mathbf{K}$ in (A.3) be a compact basic semialgebraic set. Given $p \in \mathbb{R}[x]$, if $p > 0$ on $\mathbf{K}$, then $p \in P(g_1, \dots, g_m)$. P is called **preordering** on $\mathbb{R}[x]$:

$$P(g_1, \dots, g_m) = \left\{ \sum_{J \subseteq \{1, \dots, m\}} s_J G_J; \quad s_J \in \Sigma_s \right\} \quad (\text{A.4})$$

generated by the polynomials $g_j \in \mathbb{R}[x], j = 1, \dots, m$, $G_J = \prod_{j \in J} g_j$ for $J \subseteq \{1, \dots, m\}$.

- **Putinar's Positivstellensatz** [283]⁸: Let $\mathbf{K}$ in (A.3) be a compact basic semialgebraic set. The set

$$Q(g_1, \dots, g_m) = \left\{ s_0 + \sum_{j=1}^m s_j g_j; \quad (s_j)_{j=0}^m \in \Sigma_s \right\} \quad (\text{A.5})$$

³There are non-SOS representations of positive polynomials on different type of domains, such as representations of Polya's, Bernstein's and Handelman's. Despite these lead to linear and/or semi-definite programming problems, in general, these are restricted to cases with no zeros in the positive orthant. See [278] and references therein.

⁴A set $S \subset \mathbb{R}^n$ defined as $S = \{x \in \mathbb{R}^n; f_i(x) \triangleright_i 0, i = 1, \dots, t\}$, where for each i , $\triangleright_i$ is one of $\geq, >, =, \neq$ and $f_i(x) \in \mathbb{R}[x]$, is called a basic semialgebraic set. A subset of $\mathbb{R}^n$ is a semialgebraic set if it is a finite union of basic semialgebraic sets [149].

⁵There are other representations of positive polynomials in term of SOS, such as of Schweighofer and Stengle. However, these are more difficult to verify in comparison with Schmüdgen's or Putinar's parameterizations.

⁶If $\mathbf{K}$ is non-compact semialgebraic set, Schmüdgen's and Putinar's Positivstellensatz does not hold. However, if for theses cases a particular type of quadratic module satisfies the Archimedean property, the hierarchy of semidefinite programs converge [280].

⁷The main drawback of this remarkable result is the exponential cardinality (2^m) of the set J , which usually implies large SDP problems.

⁸In contrast to Schmüdgen's Positivstellensatz, in this case the number of terms in the representation (A.5) increases linearly, being computationally more tractable.

is denominated **quadratic module** on $\mathbb{R}[x]$ generated by the polynomials $g_j \in \mathbb{R}[x]$, $j = 1, \dots, m$. It is said to be **Archimedean** if there exists a polynomial function $h \in Q(g_1, \dots, g_m)$ such that its level set $L_h = \{\mathbf{x} \in \mathbb{R}^n : h(\mathbf{x}) \geq 0\}$ is a compact set.

Given $p \in \mathbb{R}[x]$ and the quadratic module Q as in (A.5) Archimedean, if $p > 0$ on $\mathbf{K}$, then $p \in Q(g_1, \dots, g_m)$.

The Archimedean property is equivalent to:

$$\forall p \in \mathbb{R}[x], \exists N \in \mathbb{N} \text{ such that } N \pm p \in Q(g_1, \dots, g_m) \Leftrightarrow \quad (\text{A.6})$$

$$\exists N \in \mathbb{N} \text{ such that } N - \sum_{i=1}^n x_i^2 \in Q(g_1, \dots, g_m). \quad (\text{A.7})$$

Particular cases where the Archimedean property is satisfied are: (i) All g_j 's are affine and $\mathbf{K}$ is a polytope.⁹ (ii) The set $\{\mathbf{x} \in \mathbb{R}^n; g_j(\mathbf{x}) \geq 0\}$ is compact for some g_j .

A.1.1 Sum-of-Squares

A polynomial $p \in \mathbb{R}_{n,r}[\mathbf{x}]$ is said to be a **Sum of Square of polynomials** (p is SOS: $p \in \Sigma_s$) if p can be written (decomposed) as:

$$p(\mathbf{x}) = \sum_{j=1}^m q_j^2(\mathbf{x}), \quad (\text{A.8})$$

where $q_j \in \mathbb{R}_{n,d_j}[\mathbf{x}]$. If $p \in \Sigma_s \subset \mathbb{R}_{n,r}[\mathbf{x}]$, then r is even and $\deg(q_j) = d_j \leq (r/2)$, $\forall j := 1, \dots, m$. In the univariate case, a SOS condition is equivalent to global non-negativity. However, in multiple dimensions, this is just a sufficient condition.¹⁰ The set inclusion is commonly referred as¹¹ $p \in \Sigma_s \subset \mathbb{P}_{n,2d}[\mathbf{x}] \subset \mathbb{R}[\mathbf{x}]$ and $\deg(q_j) \leq d$.

Theorem 3. ([136, 137]) Let Φ_r be the standard vector basis of $\mathbb{R}[\mathbf{x}]$ with $z(r) = \binom{n+r}{r}$ monomials in $\mathbf{x}$ with degree $\leq r$. A multivariate polynomial $p \in \mathbb{R}[\mathbf{x}]$ is SOS ($p \in \Sigma_s \subset \mathbb{P}_{n,2d}[\mathbf{x}]$), if and only if there exists a matrix $Q \in \mathbb{S}_+^{z(d)}$ satisfying $p(\mathbf{x}) = \Phi_d^T Q \Phi_d$, $r = 2d$.

The essential feature in Schmüdgen's and Putinar's Positivstellensatz is the condition of existence

⁹A function $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is affine if it is a sum of a linear function and a constant: $f(\mathbf{x}) = A\mathbf{x} + b$, $A \in \mathbb{R}^{n \times m}$ and $b \in \mathbb{R}^m$. A bounded polyhedron is denominated polytope. A polyhedron is the set defined by finitely many linear inequalities or equations (intersection of a finite number of halfspaces and hyperplanes) [228, 136].

¹⁰The converse: $p \geq 0 \Rightarrow p \in \Sigma_s$ is not always true. An example is the Motzkin polynomial (bivariate sextic: $M(x, y) = x^4y^2 + x^2y^4 - 3x^2y^2 + 1$). David Hilbert, more than one century ago, showed its equivalence in three cases: univariate ($n = 1$), quadratic ($2d = 2$) and bivariate quartics ($n = 1, 2d = 4$), i.e., global non-negativity is equivalent to a SOS representation.

¹¹ $\mathbb{P}_{n,2d}[\mathbf{x}]$ is a semialgebraic set but it is not a basic semialgebraic set for $2d \geq 4$. It cannot be described by unquantified inequalities, i.e., the description must include logical operations between sets. From the optimization point of view, this fact leads to non-easily computable barrier functions. However, this computational complexity can be reduced since these sets can be represented or approximated as projections of spectrahedral sets (semialgebraic). To the extent the dimension increases, SOS polynomials are a smaller subset (volume) of the positive polynomials [136, 284].

of some SOS polynomials. Based on these type of representations of p in terms of g_j 's, it can be concluded the positivity of p on $\mathbf{K}$, in a similar way that $p \geq 0$ can be certificated by $p \in \Sigma_s$. As it is stated by Theorem 3, this feature is computationally practical since SOS polynomial condition is equivalent to formulate SDP constrains on polynomial coefficients, namely

$$\begin{aligned} \text{Find } & Q \\ \text{subject to : } & Q = Q^T, Q \succeq 0, \text{ trace}(Q, B_j) = p_j, \forall j = 0, \dots, z(r) - 1 \\ & \Phi_d \Phi_d^T = \sum_{j=0}^{z(r)-1} B_j \mathbf{x}^{\bar{\alpha}_j} \end{aligned} \quad (\text{A.9})$$

where the problem of the size: $z(d) \leq n^d$ is not prohibitive of implementing (solving it in polynomial time). Thus, the polynomial optimization problem (A.1)(equivalent to (A.2)) can be formulated by the relaxation

$$\mathbf{P}_d : \begin{cases} p_d^* = \sup_{Q_k} \gamma \\ \text{subj. to : } \left(p(\mathbf{x}) - \gamma - \sum_{k=1}^N \overbrace{\Phi_{i_k}^\top(\mathbf{x}) Q_k \Phi_{i_k}(\mathbf{x})}^{s_k(\mathbf{x})} G_k(\mathbf{x}) \right) \in \Sigma_s \\ Q_k \in \mathbb{S}_+^{z(i_k)}, i_k = d - d_k, \forall k = 1, \dots, N, \end{cases} \quad (\text{A.10})$$

based on the existence of $s_k \in \Sigma_s$, where¹² $d_k = \lceil \deg(H_k)/2 \rceil$ and $\max_k(\deg(p), \deg(G_k)) \leq 2d$; $N = 2^m$ and $G_k = \prod_{k \in J} g_k$ for $J \subseteq \{1, \dots, m\}$ if Schmüdgen's Positivstellensatz is considered; $N = m$ and $G_k = g_k$ for Putinar's Positivstellensatz (Putinar's refinement requires that the quadratic module generated by $g_1, \dots, g_m$ be Archimedean, which is not very restrictive [138]).

A.1.2 Moments

The moment problem is a relevant topic in functional analysis with several theoretical and practical implications [285]. The classical problem of moments (referred as **full moment problem**)¹³ [286, 287, 288] involves two main questions that can be stated as follows: (i) *Existence*: what are the conditions to find a measure¹⁴ μ on $\mathbf{K}$, given a infinite sequence of real numbers $\mathbf{y} = (y_j)_{j \in \mathbb{N}}$

¹²The function $\lceil a \rceil$, commonly referred to as *ceil* function, rounds to the nearest integer greater than or equal to a .

¹³Initially conducted by Thomas Stieltjes ($\mathbf{K} = [0, \infty)$, 1984), Hans Hamburger ($\mathbf{K} = (-\infty, \infty)$, 1920) and Felix Hausdorff ($\mathbf{K} = [0, 1]$, 1923).

¹⁴In particular, regarding the moment approach for polynomial optimization, this is a Borel measure, i.e., a non-negative ($\mu(\mathcal{B}) \in \mathbb{R}_0^+$) countably additive ($\mu(\mathcal{B}) = \mu(\bigcup_{i=0}^\infty B_i) = \bigcup_{i=0}^\infty \mu(B_i)$, $B_i \cap_{i \neq j} B_j = \emptyset$) set function on Borel sets B_i . Let $\mathcal{A} = \{A_1, \dots, A_i, \dots\}$ be a system of subsets, which is symmetric ($A_i \in \mathcal{A} \Rightarrow A_i^c \in \mathcal{A}, \forall i$), it is called a σ -algebra if $\bigcup_i A_i \in \mathcal{A}$ and $\bigcap_i A_i \in \mathcal{A}, \forall i$. A Borel σ -algebra $\mathcal{B}_\mathcal{A}$ is the minimal σ -algebra containing $\mathcal{A}$. A Borel set is an element of $\mathcal{B}_\mathcal{A}$. Due to its minimality characteristic, $\mathcal{B}_\mathcal{A}$ may not include all subsets of sets of measure zero, i.e., Borel measure is not complete. Lebesgue measure is its extension [289, 290, 291, 292].

such that

$$\int_{\mathbf{K}} \mathbf{x}^{\bar{\alpha}_j} d\mu(\mathbf{x}) = y_j, \quad \forall j \in \mathbb{N} ? \quad (\text{A.11})$$

(if it exists, is said to be a **representing measure** of the sequence $\mathbf{y}$), (ii) *Uniqueness*: under what conditions the measure μ is uniquely determined by the moment sequence $\mathbf{y}$? (if it is unique, is said to be **determinate**). Positiveness is the essential condition to answer these questions. Given a real infinite sequence $\mathbf{y} = (y_j)_{j \in \mathbb{N}}$, and defining the Riesz linear functional:

$$L_{\mathbf{y}} : C_c(\mathbf{K}) \rightarrow \mathbb{R} \\ f \mapsto L_{\mathbf{y}}(f) := \int_{\mathbf{K}} f(\mathbf{x}) d\mu(\mathbf{x}), \quad (\text{A.12})$$

according to the Riesz-Markov representation theorem ([159], Chapter 21; [293], Section 2.14.), if $\mathbf{K}$ is locally compact Hausdorff space¹⁵ and $L_{\mathbf{y}}$ is positive on $C_c(\mathbf{K})$,¹⁶ then there exists a unique positive measure μ on the Borel σ -algebra associated with the topology of $\mathbf{K}$. A particular case is the Haviland's result [294, 137], which considers a closed $\mathbf{K} \subset \mathbb{R}^n$ and a linear functional on non-negative polynomial functions on $\mathbf{K}$:

$$L_{\mathbf{y}} : \mathbb{P}_{n,r}[\mathbf{x}] \rightarrow \mathbb{R} \\ f \mapsto L_{\mathbf{y}}(f) = \int_{\mathbf{K}} \left(\sum_{j=0}^{z(r)-1} f_j \mathbf{x}^{\bar{\alpha}_j} \right) d\mu(\mathbf{x}) = \sum_{j=0}^{z(r)-1} f_j \int_{\mathbf{K}} \mathbf{x}^{\bar{\alpha}_j} d\mu(\mathbf{x}) = \sum_{j=0}^{z(r)-1} f_j y_j. \quad (\text{A.13})$$

For instance, considering a nondecreasing measure μ in the real axis ($\mathbf{K} \subseteq \mathbb{R}$), for the so-called **Stieltjes' moment problem** ($\mathbf{K} = [0, \infty)$), the sufficient and necessary condition of existence is $L_{\mathbf{y}}(p^2 + xq^2) \geq 0, \forall p, q \in \mathbb{R}[x]$. In the case of the so-called **Hamburger's moment problem**¹⁷ ($\mathbf{K} = (-\infty, \infty)$), the condition is $L_{\mathbf{y}}(p^2) \geq 0, \forall p \in \mathbb{R}[x]$ [296]. These conditions are equivalent to the positiveness of the Hankel matrices $H_n \succeq 0$ and $H_{n+1} \succeq 0$, and $H_n \succeq 0, \forall n \in \mathbb{N}$, respectively [297], where

¹⁵A topological space X is called Hausdorff if for any $x, y \in X$ with $x \neq y$ there exists open sets A and B such that $x \in A, y \in B$ and $A \cap B = \emptyset$. Thus, every metric space is Hausdorff, i.e., separable, in particular the Euclidean space $\mathbb{R}^n$. A topological space X is said to be locally compact if for each $x \in X$ there exists an open set A and closed set B with $x \in A \subset B$ and B compact. In particular, the Euclidean space $\mathbb{R}^n$ is locally compact. A subset of the Euclidean space is compact if and only if it is bounded and closed (Heine-Borel theorem) [150, 293].

¹⁶ C_c stands for the collection of all continuous complex functions on $\mathbf{K}$ whose support is compact.

¹⁷A elegant proof of existence and uniqueness of Hamburger's moment problem can be found in [295].

$$H_{n+k}(\mathbf{y})(i, j) = (y_{i+j-2+k}) = \begin{pmatrix} y_{0+k} & y_{1+k} & y_{2+k} & \cdots & y_{n+k} \\ y_{1+k} & y_{2+k} & y_{3+k} & \cdots & y_{n+1+k} \\ y_{2+k} & y_{3+k} & y_{4+k} & \cdots & y_{n+2+k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y_{n+k} & y_{n+1+k} & y_{n+2+k} & \cdots & y_{2n+k} \end{pmatrix}, \quad \forall i, j \in \mathbb{N}, \quad 1 \leq i, j \leq n+1. \quad (\text{A.14})$$

This equivalence is clear on the basis of the existence of positive polynomials. Let

$$\begin{aligned} \mathcal{H}_n(p, q) &= \langle p(x), q(x) \rangle_{\mathbb{R}[x]} = \left\langle \sum_{i=0}^n p_i x^i, \sum_{j=0}^n q_j x^j \right\rangle_{\mathbb{R}[x]} = \sum_{i=0}^n \sum_{j=0}^n p_i q_j \int_{\mathbf{K}} x^{i+j} d\mu(x) = \sum_{i,j=0}^n p_i q_j y_{i+j} \\ &= \langle \bar{p}, H_n \bar{q} \rangle_{\mathbb{R}} = \bar{p}^T H_n \bar{q} \end{aligned} \quad (\text{A.15})$$

be a sesquilinear form¹⁸ in terms of real polynomial coefficients $\bar{p} = [p_0, \dots, p_n]^T$ and $\bar{q} = [q_0, \dots, q_n]^T$ and Hankel matrix H_n as in (A.14). Considering the Hamburger's case, if (A.11) holds, the quadratic Hankel form:

$$\mathcal{H}_n(p, p) = \bar{p}^T H_n \bar{p} = \int_{\mathbf{K}} \left(\sum_{j=0}^n p_j x^j \right)^2 d\mu(x) = \sum_{i,j=0}^n p_i p_j y_{i+j} \geq 0 \quad (\text{A.16})$$

($f(x) = p^2(x)$), which is the necessary and sufficient condition of existence of μ [137], equivalent to $H_n \succeq 0, \forall n \in \mathbb{N}$. $\mathbf{y}$ is referred as a **positive definite sequence of moments**. Therefore, instead of checking $L_{\mathbf{y}}(f) \geq 0, \forall f \geq 0$ on $\mathbf{K}$, this only needs to be checked on polynomials with SOS decomposition.

Due to the generality of $\mathbf{K}$ in the Riesz-Markov's and Haviland's theorems and, in particular, its multi-dimensionality,¹⁹ certificate of positivity is a complex task (the multi-dimensional moment problem for general sets is still unsolved [281]). A significant exception can be stated if $\mathbf{K}$ is a *compact basic semi-algebraic set*. In this case, the remarkable results of K. Schmüdgen (1991)[282] and M. Putinar (1993)[283, 300] lead to a practical implementation, if a **truncated** case²⁰(finite sequence $\mathbf{y} = (y_j)_{j=0}^n$) of the full moment problem is considered.

¹⁸A inner product on a complex scalar field is a sesquilinear form, i.e., it is linear in the first factor and conjugate linear in the second factor: $\langle \cdot, \cdot \rangle : V \times V \rightarrow F$, $\langle v, \alpha w \rangle = \alpha \langle v, w \rangle$, $\langle v, \alpha w_1 + \beta w_2 \rangle = \bar{\alpha} \langle v, w_1 \rangle + \bar{\beta} \langle v, w_2 \rangle$ [218].

¹⁹Schmüdgen [298] and Berg, Christensen and Jensen[299] showed that the Hamburger's result can not be extended to $\mathbb{R}^n$ for $n \geq 2$ based on the fact that there exists polynomials globally non-negative but not SOS.

²⁰Curto and Fialkow have carried out a systematic study of the *truncated moment problem*. Their results have had relevant implications in polynomial optimization (see [301, 148] and references therein).

A.1.2.1 Multi-dimensional Notation

- **Moment Matrix**²¹: Based on the sesquilinear form (A.15) (bilinear in $\mathbb{R}^n$) for a multi-variable case, and the linear functional (A.13), the *Matrix of Moments* is defined as $\mathbf{M}_r(\mathbf{y}) = \mathbf{L}_y(\Phi_r(\mathbf{x})\Phi_r^T(\mathbf{x}))$ with elements $\mathbf{M}_r(\mathbf{y})(\alpha, \beta) = \mathbf{L}_y(\mathbf{x}^\alpha \mathbf{x}^\beta) = y_{\alpha+\beta}$, $\forall \alpha, \beta \in \mathbb{N}_r^n$, namely

$$\mathcal{H}_n(p, q) = \mathbf{L}_y(pq) = \int_{\mathbf{K}} \bar{p}^T \Phi(\mathbf{x}) \bar{q}^T \Phi(\mathbf{x}) d\mu(\mathbf{x}) = \bar{p}^T \int_{\mathbf{K}} \Phi(\mathbf{x}) \Phi(\mathbf{x})^T d\mu(\mathbf{x}) \bar{q} = \langle \bar{p}, \mathbf{M}_r(\mathbf{y}) \bar{q} \rangle_{\mathbb{R}^n}. \quad (\text{A.17})$$

Thus, in similar way to (A.16), for every $p \in \mathbb{R}[\mathbf{x}]$, a SOS decomposition of $f(x) = \sum_{j=0}^{z(r)-1} f_j \mathbf{x}^{\bar{\alpha}_j} = p^2(\mathbf{x})$ leads to

$$\langle \bar{p}, \mathbf{M}_r(\mathbf{y}) \bar{p} \rangle_{\mathbb{R}^n} = \mathbf{L}_y(p^2) = \int_{\mathbf{K}} p^2(\mathbf{x}) d\mu(\mathbf{x}) = \sum_{\alpha \in \mathbb{N}_r^n} f_\alpha y_\alpha \geq 0 \Leftrightarrow \mathbf{M}_r(\mathbf{y}) \succeq 0. \quad (\text{A.18})$$

The moment matrix is a symmetric matrix of dimension $z(r) = \binom{n+r}{r}$ and $\frac{z(r)(z(r)+1)}{2}$ entries in terms of the first $2r$ -order moments. For instance, if $n = 2$ and $r = 2$:

$$\mathbf{M}_0 = (y_{0,0}), \quad \mathbf{M}_1 = \begin{pmatrix} (\mathbf{M}_0) & y_{1,0} & y_{0,1} \\ y_{1,0} & y_{2,0} & y_{1,1} \\ y_{0,1} & y_{1,1} & y_{0,2} \end{pmatrix}, \quad \mathbf{M}_2 = \begin{matrix} & \mathbf{x}^{0,0} & \mathbf{x}^{1,0} & \mathbf{x}^{0,1} & \mathbf{x}^{2,0} & \mathbf{x}^{1,1} & \mathbf{x}^{0,2} \\ \begin{matrix} \mathbf{x}^{0,0} \\ \mathbf{x}^{1,0} \\ \mathbf{x}^{0,1} \\ \mathbf{x}^{2,0} \\ \mathbf{x}^{1,1} \\ \mathbf{x}^{0,2} \end{matrix} & \begin{pmatrix} & & & & & \\ & \mathbf{M}_1 & & & & \\ & & & & & \end{pmatrix} & \begin{pmatrix} y_{2,0} & y_{1,1} & y_{0,2} \\ y_{3,0} & y_{2,1} & y_{1,2} \\ y_{2,1} & y_{1,2} & y_{0,3} \\ y_{4,0} & y_{3,1} & y_{2,2} \\ y_{3,1} & y_{2,2} & y_{1,3} \\ y_{0,2} & y_{1,2} & y_{0,3} & y_{2,2} & y_{1,3} & y_{0,4} \end{pmatrix} \end{matrix}.$$

- **Localizing Matrix**²²: Given a polynomial $g = \sum_{\gamma} g_{\gamma} \mathbf{x}^{\gamma} \in \mathbb{R}[\mathbf{x}]$, where g_{γ} are real coefficients, the *Localizing Matrix* is defined as $\mathbf{M}_r(g\mathbf{y}) = \mathbf{L}_y(g(\mathbf{x})\Phi_r(\mathbf{x})\Phi_r^T(\mathbf{x}))$ with elements:

$$\mathbf{M}_r(g\mathbf{y})(\alpha, \beta) = \mathbf{L}_y(g(\mathbf{x})\mathbf{x}^{\alpha}\mathbf{x}^{\beta}) = \sum_{\gamma} g_{\gamma} y_{\gamma+\alpha+\beta}, \quad \forall \alpha, \beta \in \mathbb{N}_r^n. \quad (\text{A.19})$$

²¹For a one-dimensional problem: $M_n(\mathbf{y}) = H_n(\mathbf{y})$ with the Hankel matrix as in (A.14). In this case, if $\mathbf{y} = (m_j)_{0 \leq j \leq 2n}$ (even case), then it has a representing Borel measure $\mu \in \mathbb{R}$ if and only if $H_n(\mathbf{y}) \succeq 0$ and $\text{rank}(H_n(\mathbf{y})) = \text{rank}(\mathbf{y})$. The rank of the real sequence is the smaller integer $1 \leq i \leq n$ so that $H_n(\cdot, i) \in \text{span}(H_n(\cdot, 1), \dots, H_n(\cdot, i-1))$. Higher order moment matrices are called **flat extension** if their rank does not increase with their size [137].

²²Curto and Fialkow denominated the moment matrix of a shifted vector as *Localizing Matrix* [301].

For example, for $n = 2$, $r = 1$ and $g(\mathbf{x}) = a + x_2 - x_1 x_2^2$:

$$\mathbf{M}_1(g\mathbf{y}) = \begin{pmatrix} ay_{0,0} + y_{0,1} - y_{1,2} & ay_{1,0} + y_{1,1} - y_{2,2} & ay_{0,1} + y_{0,2} - y_{1,3} \\ ay_{1,0} + y_{1,1} - y_{2,2} & ay_{2,0} + y_{2,1} - y_{3,2} & ay_{1,1} + y_{1,2} - y_{2,3} \\ ay_{0,1} + y_{0,2} - y_{1,3} & ay_{1,1} + y_{1,2} - y_{2,3} & ay_{0,2} + y_{0,3} - y_{1,4} \end{pmatrix}.$$

Therefore, if the measure μ has support in the set $\mathbf{K} = \{\mathbf{x} \in \mathbb{R}^n; g(\mathbf{x}) \geq 0\}$, then:

$$\langle \bar{p}, \mathbf{M}_r(g\mathbf{y})\bar{p} \rangle_{\mathbb{R}^n} = \mathbf{L}_y(gp^2) = \int_{\mathbf{K}} g(\mathbf{x})p^2(\mathbf{x})d\mu(\mathbf{x}) = \sum_{\alpha \in \mathbb{N}_r^n} f_\alpha y_\alpha \geq 0 \Leftrightarrow \mathbf{M}_r(g\mathbf{y}) \succeq 0, \quad (\text{A.20})$$

where $f(x) = \sum_{j=0}^{z(r)-1} f_j \mathbf{x}^{\bar{\alpha}_j} = g(\mathbf{x})p^2(\mathbf{x})$, which can be interpreted as the action to *localize* the support of a representing measure for $\mathbf{y}$ [148].

A.1.3 Primal-Dual Perspective

Let $\mathbf{K} \subset \mathbb{R}^n$ be a Borel subset and $\mathcal{M}(\mathbf{K})$ the space of finite signed Borel measures on $\mathbf{K}$, such that $\mu \in \mathcal{M}(\mathbf{K})_+$ (convex non-negative cone). Let $p : \mathbf{K} \rightarrow \mathbb{R}$ be an integrable function with respect to μ , the Generalized Moment Problem (GMP) formulated as a global polynomial optimization problem is:

$$\mathbf{D} : \begin{cases} \rho^* = \inf_{\mu \in \mathcal{M}(\mathbf{K})_+} \int_{\mathbf{K}} p(\mathbf{x})d\mu(\mathbf{x}) \\ \text{subj. to: } \mu(\mathbf{K}) = 1 \end{cases}, \quad (\text{A.21})$$

the constraints of which can be interpreted as an optimization over probability measures ($\mu(\mathbf{K}) = 1$) [137, 289, 302]. The formulation (A.21) and the optimization problem (A.1) are equivalent (this can be easily proved considering for every $\mathbf{x} \in \mathbf{K}$ the Dirac measure²³ $\mu := \delta_{\mathbf{x}} \in \mathcal{M}(\mathbf{K})_+ \Rightarrow \int_{\mathbf{K}} p(\mathbf{x})d\mu(\mathbf{x}) = \int_{\mathbf{K}} p(\mathbf{x})d\delta_{\mathbf{x}}(\mathbf{x}) = f(\mathbf{x})$) [137, 138]. In fact, if the problem (A.1) ($\inf_{\mathbf{x} \in \mathbf{K}} p(\mathbf{x})$) attains its global minimum at $\mathbf{x}^* \in \mathbf{K}$ ($\mathbf{K}$ compact and support of a positive measure²⁴), then the Dirac measure $\mu^* = \delta_{\mathbf{x}^*}$ is the optimal solution of (A.21). This equivalence is also a duality relation:

$$\mathbf{P} : \begin{cases} p^* = \sup_{\gamma \in \mathbb{R}} \gamma \\ \text{subj. to: } p(\mathbf{x}) - \gamma \geq 0 \\ \mathbf{x} \in \mathbf{K} \end{cases} \Leftrightarrow \mathbf{D} : \begin{cases} \rho^* = \inf_{\mu \in \mathcal{M}(\mathbf{K})_+} \int_{\mathbf{K}} p(\mathbf{x})d\mu(\mathbf{x}) \\ \text{subj. to: } \int_{\mathbf{K}} d\mu(\mathbf{x}) = 1 \\ \mu \geq 0 \end{cases} \quad (\text{A.22})$$

²³Let $\mathcal{M}$ be any σ -algebra of a set $\mathbf{X}$. The Dirac measure $\delta_{\mathbf{x}_0}$ is a measure concentrate at one point $\mathbf{x}_0$, such that $\delta_{\mathbf{x}_0}(A) = 1 \Leftrightarrow \mathbf{x}_0 \in A \in \mathcal{M}$ and $\delta_{\mathbf{x}_0}(B) = 0 \Leftrightarrow \mathbf{x}_0 \notin B \in \mathcal{M}$ [159]. The Dirac measure ($\mu = \delta_{\mathbf{x}_0}$) is singular with respect to the Lebesgue measure (η) concentrate on $\mathbb{R}^n \setminus \mathbf{x}_0$. Thus, Dirac measure has no Radon-Nikodym derivative ($d\mu = \rho d\eta$) [303, 304].

²⁴The support of a function is defined to be the closure of the set of points at which the function is nonzero, i.e., closure of the set $\{\mathbf{x} : f(\mathbf{x}) \neq 0\}$ [293, 159].

in the sense of Linear Programming duality,²⁵ with strong duality (there is no duality gap: $\rho^* = p^*$) if $\mathbf{K}$ is compact. For a generic function p , the formulation (A.22) is infinite-dimensional. However, if polynomials $p \in \mathbb{R}_{n,r}[\mathbf{x}]$ are considered, this is a finite-dimensional convex problem:

$$\mathbf{D} : \begin{cases} \rho^* = \inf_{\mathbf{y}} \sum_{\alpha \in \mathbb{N}_r^n} p_{\alpha} \mathbf{y}_{\alpha} \\ \text{subj. to: } \mathbf{y}_0 = 1 \\ \mathbf{y}_{\alpha} = \int_{\mathbf{K}} \mathbf{x}^{\alpha} d\mu(\mathbf{x}), \forall \alpha \in \mathbb{N}_r^n \end{cases} \quad (\text{A.23})$$

but still untractable. This is due to the hard problem of characterization of the convex cone $\mathcal{Y}_{n,r}(\mathbf{K}) = \{(\mathbf{y})_{\alpha} \in \mathbb{N}_r^n : \exists \mu \in \mathcal{M}(\mathbf{K})_+; \mathbf{y}_{\alpha} = \int_{\mathbf{K}} \mathbf{x}^{\alpha} d\mu(\mathbf{x}), \forall \alpha \in \mathbb{N}_r^n\}$, which can be simplified and efficiently implemented if $\mathbf{K}$ is a compact basic semialgebraic set via the following relaxation:

$$\mathbf{D}_d : \begin{cases} \rho_d^* = \inf_{\mathbf{y}} \mathbf{L}_{\mathbf{y}}(p) = \sum_{\alpha \in \mathbb{N}_d^n} p_{\alpha} \mathbf{y}_{\alpha} \\ \text{subj. to: } \mathbf{y}_0 = 1 \\ \mathbf{M}_d(\mathbf{y}) \succeq 0 \\ \mathbf{M}_{d-v_j}(g_j \mathbf{y}) \succeq 0, j = 1, \dots, m \end{cases}, \quad (\text{A.24})$$

dual formulation of (A.10) for the Putinar's Positivstellensatz representation, where $p \in \mathbb{R}_{n,r}[\mathbf{x}]$, $\mathbf{L}_{\mathbf{y}}$ is the Riesz linear functional (A.13), $\mathbf{M}_d(\mathbf{y})$ is a Moment matrix, $\mathbf{M}_{d-v_j}(g_j \mathbf{y})$ is a Localizing matrix, $v_j := \lceil \deg(g_j)/2 \rceil$ and $d \geq \max\{\lceil r/2 \rceil, \max_j v_j\}$.

²⁵Let $\mathcal{B}(\mathbf{K})$ be the space of measurable functions on $\mathbf{K}$. $\mathcal{M}(\mathbf{K})$ and $\mathcal{B}(\mathbf{K})$ constitute a dual pair of spaces with duality bracket $\langle \mu, f \rangle := \int_{\mathbf{K}} f(\mathbf{x}) d\mu(\mathbf{x})$ and the usual scalar product (see [137], appendix C.3 for details). Similarly, the cone of positive polynomials $\mathbb{P}_{n,r}$ and the cone of real sequences with a representing measure $\mathcal{Y}_{n,r}$ are dual each other (see [148], Section 4.4.1).

Appendix B

Backstepping Design for PDEs

B.1 The Fundamental Idea

Backstepping design for PDEs can be seen as an extension of backstepping for nonlinear LPS [305]. This was initiated around 2000 as a methodology for boundary control/observer design for linear PDEs [21, 50, 51, 18, 49]. Its fundamental idea, the *Volterra transformation* [40, 190], traces back to the application of the method of Integral Operators for solving initial-boundary problems [306] derived from the boundary control of parabolic equations [307]. It stands out for its elegant and simple systematic methodology, which: (i) does not involve spatial discretization of the PDE model, (ii) carries out a collective treatment of the system modes instead of a finite analysis of them based on their spectral characteristics, (iii) does not require to formulate the problem in abstract Hilbert spaces, apply semigroup theory, nor solve operator-valued equations.

The central point of this methodology, is the use of the “Volterra Integral Transformation” (coordinate transformation),¹ namely

$$w(t, x) = u(t, x) - \int_0^x K(x, y)u(t, y)dy, \quad (\text{B.1})$$

where the **Kernel** $K = K(x, y)$ is assumed to be continuous on the triangle domain $T(0, b) = \{(x, y); 0 \leq x \leq b, 0 \leq y \leq x\}$, and $k(x, y) \equiv 0 \forall x \leq y$, i.e., the Kernel vanishes above the diagonal of the square $[0, b] \times [0, b]$. Due to the fact that the range of the integration variable y is always limited by x , carrying an analogy with the time variable, this transformation can be seen as *causal in space* [195]. One of the relevant properties of the Volterra transformation is that it is invertible under mild conditions on its Kernel [191], so that stability of the system in $w = w(x, t)$ (target system) translates into stability of the system in $u = u(x, t)$ (original system).

¹Volterra equation of the “first kind”: $\int_a^x K(x, t)f(t)dt = g(x)$; Volterra equation of the “second kind”: $f(t) - \int_a^x K(x, t)f(t)dt = g(x)$, with $f = f(t)$ unknown function and $K(x, t) \equiv 0, \forall t > x$. These equations are particular cases of the Fredholm Integral Equations, where the Kernel does not vanish in the domain ($K(x, t) \neq 0$ in the square $Q(a, b) = \{a \leq x \leq b, a \leq t \leq x\}$)[308].

B.2 Continuous-Time Approach for Observer Design

The main idea is to use the Volterra Integral Transformation (B.1) to map an **Original** PDE into a **Target** PDE. This leads to the so-called “**Kernel-PDE**”, the solution of which constitutes functional gains capable to compensate for the “original” system achieving the stability and convergence properties of the “target” system, selected in advance.² An example of this approach applied to the observer design problem for the *heat equation*, anti-located setup, is summarized as follows:

<u>System (S)</u>			
$u_t = u_{xx} + \lambda u$			
$u_x = 0$			
$u(1, t) = U(t)$	<u>Volterra Transformation (VT)</u>	<u>Kernel PDE (KPDE)</u>	
	$\tilde{u}(x, t) = \tilde{w}(x, t) - \int_0^x K(x, y) \tilde{w}(y, t) dy$	$K_{xx}(x, y) - K_{yy}(x, y) = -\lambda K(x, y)$	
<u>Observer (O)</u>		$K(1, y) = 0$	
$\hat{u}_t = \hat{u}_{xx} + \lambda \hat{u} + h(x) \tilde{u}(0, t)$		$K(x, x) = \frac{\lambda}{2}(x - 1)$	(B.2)
$\hat{u}_x(0, t) = p_0 \tilde{u}(0, t)$	<u>Target Error System (T)</u>		
$\hat{u}(1, t) = U(t)$	$\tilde{w}_t = \tilde{w}_{xx}$	<u>Kernel Solution (KS)</u>	
<u>Original Error System (E)</u>	$\tilde{w}_x(0, t) = 0$	$K(x, y) = -\lambda(1 - x) \cdot \frac{I_1(\sqrt{\lambda(2-x-y)(x-y)})}{\sqrt{\lambda(2-x-y)(x-y)}}$	
$\tilde{u}_t = \tilde{u}_{xx} + \lambda \tilde{u} - h(x) \tilde{u}(0)$	$\tilde{w}(1, t) = 0$		
$\tilde{u}_x(0, t) = -p_0 \tilde{u}(0, t)$			
$\tilde{u}(1, t) = 0$			

where $x \in \Omega = [0, 1]$ is the spatial variable, λ is the reactivity coefficient, U is the control action, $\hat{u}$ is the estimation of the dependent variable u (heat), $\tilde{u} = u - \hat{u}$ is the estimation error, $K \in \mathcal{C}^1(\Omega \times \Omega)$ is the Kernel of the Volterra transformation and $h(x) = K_y(x, 0) = \frac{\lambda(1-x)}{x(2-x)} I_2(\sqrt{\lambda x(2-x)})$ and $p_0 = K(0, 0) = \lambda/2$ are the observer gains.

The observer design consists in two main steps:

- (i) For the PDE system (S) the Luenberger-type Observer (O) is proposed. Subtracting (O) from (S) leads to the observer error (E), which is considered as the “Original” Error System. For (E) the asymptotically stable PDE system (T) is proposed, which is considered as “Target” Error System.
- (ii) The Volterra transformation (VT) is used to map between (E) and (T). Computing $\tilde{u}_t$, $\tilde{u}_{xx}$, $\tilde{u}(0)$ from (VT) and substituting (T) in their resulting expressions yields:

$$\tilde{u}_t(x, t) = \tilde{w}_t(x, t) - \int_0^x K(x, y) \tilde{w}_t(y, t) dy = \tilde{w}_t(x, t) - \underbrace{\int_0^x K(x, y) \tilde{w}_{yy}(y, t) dy}_{T_0}$$

²The transformation alone is not capable of eliminating the undesirable terms of the original system, but it bring these terms to the boundary, where a compensation law eliminates or dominates them.

$$\begin{aligned}
&= \tilde{w}_{xx}(x, t) - \int_0^x K_{yy}(x, y) \tilde{w}(y, t) dy - K(x, x) \tilde{w}_x(x, t) + K(x, 0) \tilde{w}_x(0, t) + \\
&\quad \frac{\partial K}{\partial y}(x, x) \tilde{w}(x, t) - K_y(x, 0) \tilde{w}(0, t), \\
\tilde{u}_x(x, t) &= \tilde{w}_x(x, t) - \int_0^x K_x(x, y) \tilde{w}(y, t) dy - K(x, x) \tilde{w}(x, t), \\
\tilde{u}_{xx}(x, t) &= \tilde{w}_{xx}(x, t) - \int_0^x K_{xx}(x, y) \tilde{w}(y, t) dy - \frac{\partial K}{\partial x}(x, x) \tilde{w}(x, t) - \frac{dK}{dx}(x, x) \tilde{w}(x, t) - K(x, x) \tilde{w}_x(x, t) \\
\tilde{u}(0, t) &= \tilde{w}(0, t),
\end{aligned}$$

where integration by parts has been applied on the term T_0 . According to these expressions the observer error (E) can be written as:

$$\begin{aligned}
\tilde{u}_t(x, t) - \tilde{u}_{xx}(x, t) - \lambda \tilde{u}(x, t) + h(x) \tilde{u}(0) &= \underbrace{\left(h(x) - K_y(x, 0) \right)}_{T_1} \tilde{w}(0, t) + \\
\underbrace{\left(2 \frac{dK}{dx}(x, x) - \lambda \right)}_{T_2} \tilde{w}(x, t) + \int_0^x \underbrace{\left(K_{xx}(x, y) - K_{yy}(x, y) + \lambda K(x, y) \right)}_{T_3} \tilde{w}(y, t) dy &= 0,
\end{aligned}$$

where the identity $\frac{dK}{dx}(x, x) = \frac{\partial K}{\partial x}(x, x) + \frac{\partial K}{\partial y}(x, x)$ has been used. In addition, from the boundary conditions of (E), based on (T), it is obtained:

$$\begin{aligned}
\tilde{u}_x(0, t) &= \tilde{w}_x(0, t) - K(0, 0) \tilde{w}(0, t) = -p_0 \tilde{u}(0, t) = -p_0 \tilde{w}(0, t), \quad \Rightarrow \quad \underbrace{K(0, 0)}_{T_4} = p_0, \\
\tilde{u}(1, t) &= \tilde{w}(1, t) - \int_0^1 K(1, y) \tilde{w}(y, t) dy = 0, \quad \Rightarrow \quad \underbrace{K(1, y)}_{T_5} = 0.
\end{aligned}$$

Thus, applying the coordinate transform $\bar{x} = 1 - y$, $\bar{y} = 1 - x$, $\bar{K}(\bar{x}, \bar{y}) = K(x, y)$, the terms T_2 , T_3 and T_5 constitute the Kernel-PDE (KPDE) and the terms T_1 and T_4 leads to the observer gains. The Kernel K has a closed-form solution given by (KS) [21].

The Kernel-PDE is commonly solved converting this into an *integral equation* and then applying the method of *Successive Approximations*, the expressions of which can be computed recursively³ and evaluated *symbolically*, and in some cases explicitly, as in (B.2) via the first order modified Bessel function I_1 . Except for special cases, the solution of the Kernel-PDE by means of the Successive Approximation method is computationally rather expensive. In addition, if more complex systems are considered, for instance those with spatially and time varying parameters, there could be no direct numerical solution known for the Kernel-PDE [243].

Due to the property of *invertibility* of the Volterra Transformation, and that the “target” system can be selected exponentially stable, the observer shall provide a state estimation error exponentially stable.

³This recursion is also used to compute a quantitative bound of the Kernel, needed for example in inverse optimal control [50].

Appendix C

Pseudo Two-Dimensional Lithium-Ion Battery Model

C.1 Governing Equations

Physics-based electrochemical models provide an accurate prediction of the internal dynamics of Lithium-ion batteries. The most established model of this type, the so-called Doyle-Fuller-Newman (DFN) model [256], includes both macro-scale dynamics in the electrolyte and micro-scale diffusion of Lithium in spherical particles in the porous active material, which are formulated by a set of nonlinear Partial Differential Algebraic Equations (PDAEs). The governing equations are written according to [260, 259] in terms of coordinates shown in Figure C.1.1.

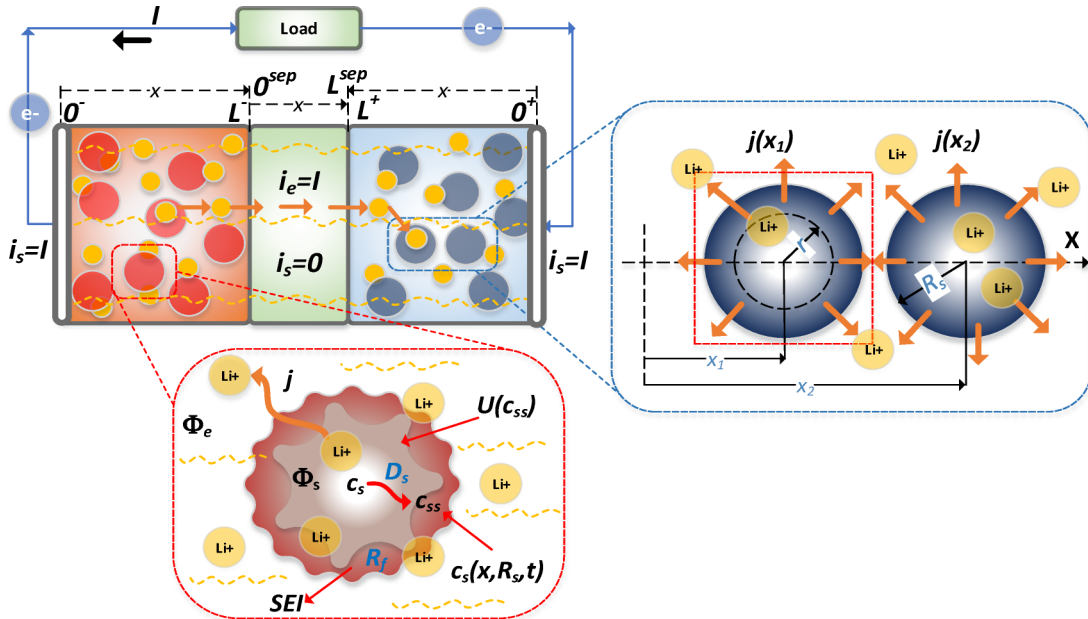


Figure C.1.1: Schematic of an Intercalation Cell.

Figure C.1.1 illustrates a typical sandwich scheme for a Lithium-ion battery (discharging stage). For the sake of notational simplicity, its **three layers** are identified in the governing equations by the **symbol** l as: **Positive Electrode**: $l = +$, **Separator**: $l = \text{sep}$ and **Negative Electrode**: $l = -$. Their spatial domain are: Positive Electrode: $[0, L^+]$, Separator: $[0^{\text{sep}}, L^{\text{sep}}]$ and Negative Electrode: $[0, L^-]$.

C.1.1 Input-Output Signals

For $l \in \{-, \text{sep}, +\}$ and collectors at $0^-, 0^+$:

$$\begin{aligned} V(\tau) &= \phi_s^+(0^+, t) - \phi_s^-(0^-, t), \\ I(\tau) &= i_e^l(x, \tau) + i_s^l(x, \tau), \\ i_s^{\text{sep}}(x, \tau) &= 0 \quad (x \in [0, L^{\text{sep}}]). \end{aligned} \tag{C.1}$$

C.1.2 Potentials in the Solid Phase

For $l \in \{-, +\}$:

$$\begin{aligned} \frac{\partial \phi_s^l}{\partial x}(x, \tau) &= -\frac{1}{\sigma^l} i_s^l(x, \tau) = \frac{i_e^l(x, \tau)}{\sigma^l} - \frac{I(\tau)}{\sigma^l} \\ i_e^-(0^-, \tau) &= 0 = i_e^+(0^+, \tau), \\ i_e^-(L^-, \tau) &= I(\tau) = i_e^+(L^+, \tau), \\ i_e^{\text{sep}}(x, \tau) &= I(\tau), \quad \forall x \in [0, L^{\text{sep}}]. \end{aligned} \tag{C.2}$$

C.1.3 Potentials in the Electrolyte

For $l \in \{-, +\}$:

$$\begin{aligned} \frac{\partial \phi_e}{\partial x}(x, \tau) &= -\frac{i_e^l(x, \tau)}{\kappa^l} + \frac{2RT}{F}(1 - t_0^+) \left(1 + \frac{d \ln f_{c/a}}{d \ln c_e^l}(x, \tau) \right) \frac{\partial \ln c_e^l}{\partial x}(x, \tau), \\ \phi_e(L^-, \tau) &= \phi_e(0^{\text{sep}}, \tau), \\ \phi_e(L^{\text{sep}}, \tau) &= \phi_e(L^+, \tau), \\ \phi_e(0^+, \tau) &= 0. \end{aligned} \tag{C.3}$$

C.1.4 Transport in the Solid Phase

For $l \in \{-, +\}$:

$$\begin{aligned} \frac{\partial c_s^l}{\partial \tau}(x, r, \tau) &= \frac{1}{r^2} \frac{\partial}{\partial r} \left(D_s^l(\cdot) r^2 \frac{\partial c_s^l}{\partial r}(x, r, \tau) \right), \\ \frac{\partial c_s^l}{\partial r}(x, 0, \tau) &= 0, \\ \frac{\partial c_s^l}{\partial r}(x, R_s^l, \tau) &= -\frac{j^l(x, t)}{D_s^l(\cdot)}, \\ c_{ss}^l(x, \tau) &= c_s^l(x, R_s^l, \tau). \end{aligned} \tag{C.4}$$

C.1.5 Transport in the Electrolyte

For $l \in \{-, \text{sep}, +\}$:

$$\begin{aligned}
 \epsilon_e^l \frac{\partial c_e^l}{\partial \tau}(x, \tau) &= \frac{\partial}{\partial x} \left(D_e(\cdot) \frac{\partial c_e^l}{\partial x}(x, \tau) + \frac{(1 - t_0^+)}{F} i_e^l(x, \tau) \right), \\
 \frac{\partial c_e^-}{\partial x}(0^-, \tau) &= 0 = \frac{\partial c_e^+}{\partial x}(0^+, \tau), \\
 D_e(L^-) \frac{\partial c_e^-}{\partial x}(L^-, \tau) &= D_e(0^{\text{sep}}) \frac{\partial c_e^{\text{sep}}}{\partial x}(0^{\text{sep}}, \tau), \\
 D_e(L^{\text{sep}}) \frac{\partial c_e^{\text{sep}}}{\partial x}(L^{\text{sep}}, \tau) &= D_e(L^+) \frac{\partial c_e^+}{\partial x}(L^+, \tau), \\
 c_e(L^-, \tau) &= c_e(0^{\text{sep}}, \tau), \\
 c_e(L^{\text{sep}}, \tau) &= c_e(L^+, \tau).
 \end{aligned} \tag{C.5}$$

C.1.6 Conservation of Charge

For $l \in \{-, +\}$:

$$\frac{\partial i_e^l}{\partial x}(x, \tau) = a^l F j^l(x, \tau) = j^{\text{Li}}(x, \tau) = -\frac{\partial i_s^l}{\partial x}(x, \tau). \tag{C.6}$$

C.1.7 Butler-Volmer Kinetics

For $l \in \{-, +\}$:

$$\begin{aligned}
 j^{\text{Li}}(x, \tau) &= a^l i_o^l(x, \tau) \left(\exp \left(\frac{\alpha^- F}{RT} \eta_s^l(x, \tau) \right) - \exp \left(\frac{-\alpha^+ F}{RT} \eta_s^l(x, \tau) \right) \right), \\
 i_o^l(x, \tau) &= k^l \left(c_{s,ss}^l(x, \tau) \right)^{\alpha^+} \left(c_e^l(x, \tau) \left(c_{s,\text{max}}^l - c_{s,ss}^l(x, \tau) \right) \right)^{\alpha^-}, \\
 \eta^l(x, \tau) &= \phi_s^l(x, \tau) - \phi_e(x, \tau) - U^l(c_{ss}^l(x, \tau)) - F R_f^l j^l(x, \tau).
 \end{aligned} \tag{C.7}$$

C.1.8 Effective Coefficients

In order to account for the tortuosity in the porous media, *effective coefficients* can be used according to the Bruggeman relations.

For $l \in \{-, +\}$:

$$\begin{aligned}
 \kappa^{l,\text{eff}} &= \kappa^l \left(\epsilon_e^l \right)^b, \\
 \sigma^{l,\text{eff}} &= \sigma^l \left(\epsilon_e^l \right)^b, \\
 D_e^{\text{eff}} &= D_e \left(\epsilon^l \right)^b.
 \end{aligned} \tag{C.8}$$

where b is the so-called *Bruggeman exponent* [310].

C.1.9 Variables and Parameters

Symbol	Description	Units
r	Radial Spherical Coordinate $r \in [0, R_s^l]$	m
x	Longitudinal Cartesian Coordinate $x \in [0, L^l]$	m
τ	Time Coordinate	m
c_s^l	Solid Phase Concentration $c_s^l \in [0, c_{s,\max}^l]$	mol/m ³
$c_{s,ss}^l$	Solid Surface Phase Concentration	mol/m ³
c_e	Electrolyte Concentration $c_e \in [0, c_{e,\max}]$	mol/m ³
U^l	Open Circuit Equilibrium Potential (OCP)	V
η^l	Solid-Phase Intercalation Reaction Overpotential	V
ϕ_s^l	Electric Potential in the Solid-Phase	V
ϕ_e	Electric Potential in the Electrolyte	V
i_o^l	Exchange Current Density	A/m ²
j^l	Pore-wall Molar Flux	mol/(m ² s)
j^{Li}	Volumetric Reaction Current	A/m ³
i_s^l	Electronic Current in the Solid Phase	A/m ²
i_e^l	Ionic Current in the Electrolyte	A/m ²
I	Current Density	A/m ²
R	Universal Gas Constant	8.314 J/(mol °K)
T	Standard Ambient Temperature	298 °K
F	Faraday's Constant	96487 C/mol

Table C.1: Variables and constants in the Pseudo Two-Dimensional Lithium-Ion Battery Model

Symbol	Description	Units
$c_{s,0}^l$	Initial Solid Phase Concentration	mol/m ³
$c_{e,0}$	Initial Electrolyte Concentration	mol/m ³
$c_{s,\max}^l$	Maximum Solid Phase Concentration	mol/m ³
$c_{e,\max}^l$	Maximum Electrolyte Concentration	mol/m ³
a^l	Particle Specific Interfacial Area (Surface Area to Volume)	m ² /m ³
D_s^l	Solid Phase Diffusivity Coefficient	m ² /s
D_e	Electrolyte Diffusivity Coefficient	m ² /s
ϵ^l	Volume Fraction or Porosity of Solid Phase	—
$f_{c/a}$	Mean Molar Activity in the Electrolyte	—
ϵ_e	Volume Fraction in the Electrolyte	—
κ^l	Ionic Conductivity of the Electrolyte	S/m
σ^l	Electronic Conductivity of Solid Phase	S/m
R_f^l	Resistance Film Solid Electrolyte Interphase	Ω m ²
t_0^+	Transference Number of Cations in the Electrolyte	—
α^-	Anodic Transference Coefficient	—
α^+	Cathodic Transference Coefficient	—
k^l	Intercalation/deintercalation Reaction Rate Constant	A (m ³ /mol) ^{$\alpha^+ + 2\alpha^-$}
L^l	Layer Thickness	m
R_s^l	Particle Radius	m

Table C.2: Parameters in the Pseudo Two-Dimensional Lithium-Ion Battery Model