Estimation, Soft Sensing and Servo-control of Linear Distributed and Lumped Parameter Systems

by

Junyao Xie

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

PROCESS CONTROL

Department of Chemical and Materials Engineering

University of Alberta

Abstract

State-of-the-art advancements in the realm of industrial process control and monitoring often require accurate descriptions of complex processes and their dynamical behaviours. Usually, many industrial processes are described by partial differential equations (PDE) or ordinary differential equations (ODE) depending on whether their dynamics evolve spatio-temporally or temporally, and thus are classified as distributed parameter systems (DPS) and lumped parameter systems (LPS), respectively. In this thesis, discrete-time estimator, soft sensor, and regulator designs are proposed for linear distributed and lumped parameter systems.

Considering the unavailability of full state information or prohibitive cost for installing spatially-distributed sensors, state estimation is often necessary for the regulation (or control) problems and/or for the monitoring purpose. To estimate the spatio-temporal state, the discrete-time Luenberger observer and Kalman filter are proposed for linear infinite-dimensional systems. Specifically, the discrete-time observer gain can be solved from continuous- or discrete-time Riccati equations numerically. To account for the output and disturbance constraints in the estimation, moving horizon estimation (MHE), as an optimization-based approach, is developed for a rather general class of DPSs, namely regular linear infinite-dimensional systems, by extending the MHE theory of LPSs. Stability and optimality are proved for the proposed MHE design.

Towards the unavailability of accurate model parameters and/or model structures, three estimators are further proposed, ranging from the mode-based MHE for output/state and mode estimation of switching regular linear infinite-dimensional sys-

tems, to the hybrid estimator development for pipeline leak detection, localization, and estimation of distributed pipeline systems based on discrete-time Luenberger observer and support vector machine (SVM), to the entirely model-free (i.e., data-driven) transfer learning (TL) based soft sensor design of linear finite-dimensional systems using variational Bayesian inference. The stability of the proposed advanced MHE design is given.

For the sake of disturbance rejection and reference tracking, the discrete-time output regulator design is developed for linear distributed parameter systems. In particular, the Cayley-Tustin (CT) bilinear transformation is applied to approximate the continuous system by a discrete-time infinite-dimensional system with essential model properties being preserved. Specifically, two types of output regulator designs are presented, namely, state-feedback regulator design and error-feedback regulator design, by exploring the internal model principle. Discrete regulator equations are formulated, and their solvability is proved and linked to the continuous counterparts. To ensure the stability of the discrete-time closed-loop system, the design of stabilizing feedback gain and its dual problem of stabilizing output injection gain design are provided using Riccati equations.

The effectiveness of the developed discrete-time Luenberger observer, Kalman filter, MHE, and advanced MHE methods are demonstrated on the transport-reaction processes, wave, Schrödinger and beam equations, and a heat exchanger system. The proposed soft sensor algorithm (i.e., transfer slow feature analysis) is validated through a numerical example, the Tennessee Eastman (TE) benchmark dataset, and a steam-assisted gravity drainage (SAGD) industrial process. The applicability of the proposed discrete-time output regulator designs is verified through heat equation, transport equation, pipeline networks, and two fluid flow systems (i.e., Kuramoto-Sivashinsky and Ginzburg-Landau equations).

Preface

Chapter 2 of this thesis consists of a revised version of J. Xie, Q. Xu, D. Ni, S. Dubljevic, Observer and Filter Design for Linear Transport-Reaction Systems, *European Journal of Control*, 49 (2019): 26-43, and a revised version of J. Xie, S. Dubljevic, Discrete-Time Kalman Filter Design for Linear Infinite-Dimensional Systems, *Processes*, 7.7 (2019): 451. I was responsible for the formulation, simulation and analysis as well as the manuscript composition. Q. Xu was responsible for the formulation, simulation, and analysis. D. Ni was responsible for formulation and analysis. S. Dubljevic was the supervisory author and was involved with concept formation and manuscript composition.

Chapter 3 of this thesis is a revised version of J. Xie, C. R. Koch, S. Dubljevic, Constrained Receding Horizon Output Estimation of Regular Linear Distributed Parameter Systems, which has been submitted to *IEEE Transaction on Automatic Control*. I was responsible for the formulation, simulation and analysis as well as the manuscript composition. C. R. Koch was responsible for formulation and analysis. S. Dubljevic was the supervisory author and was involved with concept formation and manuscript composition.

Chapter 4 of this thesis is a revised version of J. Xie, C. R. Koch, S. Dubljevic, Moving Horizon Estimation for Switching Regular Linear Systems, which has been submitted to *Automatica*. I was responsible for the formulation, simulation and analysis as well as the manuscript composition. C. R. Koch was responsible for formulation and analysis. S. Dubljevic was the supervisory author and was involved with concept formation and manuscript composition.

Chapter 5 of this thesis is a revised version of J. Xie, B. Huang, S. Dubljevic, Transfer Learning for Dynamic Feature Extraction Using Variational Bayesian Inference, *IEEE Transactions on Knowledge and Data Engineering*, (2021): 1-12. I was

responsible for the formulation, simulation and analysis as well as the manuscript composition. S. Dubljevic was responsible for formulation and manuscript composition. B. Huang was the supervisory author and was involved with concept formation, analysis and manuscript composition.

Chapter 6 of this thesis is a revised version of J. Xie, X. Xu, S. Dubljevic, Long Range Pipeline Leak Detection and Localization Using Discrete Observer and Support Vector Regression, *AIChE Journal*, 65, no. 7 (2019): e16532. I was responsible for the formulation, simulation and analysis as well as the manuscript composition. X. Xu was responsible for problem formulation. S. Dubljevic was the supervisory author and was involved with concept formation and manuscript composition.

Chapter 7 of this thesis is a revised version of J. Xie, C. R. Koch, S. Dubljevic, Discrete Output Regulator Design for Linear Distributed Parameter Systems, *International Journal of Control*, (2020): 1-17. I was responsible for the formulation, simulation and analysis as well as the manuscript composition. C. R. Koch and S. Dubljevic were the supervisory authors and were involved with concept formation and manuscript composition.

Chapter 8 of this thesis is a revised version of J. Xie, C. R. Koch, S. Dubljevic, Discrete-time Model-based Output Regulation of Fluid Flow Systems, *European Journal of Control*, 57 (2021): 1-13. I was responsible for the formulation, simulation and analysis as well as the manuscript composition. C. R. Koch and S. Dubljevic were the supervisory authors and were involved with concept formation and manuscript composition.

Chapter 9 of this thesis is a revised version of J. Xie, S. Dubljevic, Discrete-time Modeling and Output Regulation of Gas Pipeline Networks, *Journal of Process Control*, 98 (2021): 30-40. I was responsible for the formulation, simulation and analysis as well as the manuscript composition. S. Dubljevic was the supervisory author and was involved with concept formation and manuscript composition.

To my family, for their love, encouragement and support.

Acknowledgements

I would like to express my deepest gratitude to those who have helped and supported me during my PhD program over the past four years. First and foremost, I would like to express my sincere gratitude to my supervisor Prof. Stevan Dubljevic for all his efforts, inspiration, guidance, and patience, and for giving me the freedom to explore various research directions during my PhD program. I would like to thank Prof. Dubljevic for many inspiring discussions and for encouraging me to attend academic conferences and interact with industrial engineers through all stages of my PhD program, which has been an invaluable learning experience and I have benefited greatly in this process. Without his guidance and support during my PhD study, this thesis could not be accomplished. I would also like to acknowledge my graduate assistantship supervisors Prof. Biao Huang and Prof. Charles Robert (Bob) Koch. I sincerely thank Prof. Huang for his mentoring, encouragement, and trust when I encountered some difficulties in my research work, and for teaching me to always consider specific process knowledge when building meaningful data-driven models. I would like to appreciate the patience and freedom that Prof. Koch gave me such that I can freely pursue the research topics of my interest, and the time that he invested in guiding my research and revising the papers. It is a great pleasure for me to have them as great advisors, collaborators, and friends in this journey.

Special thanks go to my committee members and examiners Prof. Kirsten Morris, Prof. Tongwen Chen, Prof. Bob Koch, and Prof. Biao Huang for their participation in my doctoral examination and their valuable comments and corrections of my thesis. In addition, I sincerely thank Prof. Zukui Li, Prof. Jinfeng Liu, and Prof. Vinay Prasad for their help in my study and research over the past years.

I am grateful to my fellow friends in the Distributed Parameter Systems group for the friendships and encouragement. I sincerely wish you all the best future and

success in your career development. In addition, I would like to thank Yanjun Ma and Yousef Alipouri for the valuable discussions and advice on the soft sensor design of the SAGD process.

I would like to acknowledge the financial support from Natural Sciences and Engineering Research Council of Canada, Alberta Innovates and Autonomous Systems Initiative, and the scholarships/awards offered by the Faculty of Graduate Studies and Research, University of Alberta.

Last but not least, I would like to express my sincere gratitude to my parents and my sister for their unconditional love, understanding, encouragement, and support on every decision I made. My deep appreciation also goes to my girlfriend for her love, accompany, understanding, and reminding me to enjoy the beautiful life, especially during the difficult pandemic situation.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

In this chapter, the motivations of this thesis are introduced at first. Literature related to distributed parameter systems and the corresponding control and estimation development are reviewed. After that, research objectives, contributions and thesis outline are provided.

1.1 Motivation

Infinite-dimensional systems (or distributed parameter systems, DPS) have been an increasingly active area of research, due to their remarkable capacity of describing ubiquitous spatiotemporal dynamics in chemical, mechanical and civil engineering and applied mathematics. Many existing studies have been devoted to extending and generalizing the infinite-dimensional versions of the linear quadratic control problem [4, 5, 6], Kalman filtering problem [7, 8, 9], internal model control problem [10, 11, 12, 13, 14], and boundary stabilization problem [15, 16, 17]. However, most of the work has been done in a continuous-time setting and one question naturally arises in how to ensure the performance of late-lumping controllers designed for continuous-time PDE models in numerical realizations (namely, in the discrete-time setting) [18]. Considering that system theoretic developments often go in parallel for continuous-time and discrete-time models, it is common practice to derive results for one class of systems and then map these over to the other by using a bilinear transformation in finite-dimensional system theory [19]. Since it is often possible to use some bilinear transform to avoid repeating tedious derivations if results have been obtained for continuous systems and similar results are needed for the discrete ones and vice

versa [19]. As a bilinear transformation, Cayley transform has been widely used for establishing 1-1 correspondences of continuous- and discrete-time PDE systems in terms of conservativity (i.e., energy preserving) [19, 20], reachability, [19], stability [21, 22], stabilizability, controllability and observability [19, 4, 23], and optimality [24]. Based on the preceding considerations, this thesis aims to develop discrete-time servo-controllers and estimators for linear distributed parameter systems using the Cayley-Tustin transformation. Compared to the existing contributions on sampled-data control of infinite-dimensional systems [25, 26, 27, 28], this thesis is concerned with a different approach (namely, Cayley-Tustin transformation) in a late lumping manner (also called direct design).

Compared to the advanced development of control theory of infinite-dimensional systems, the state and parameter estimation theory of such systems are relatively under-explored, especially when it comes to the discrete-time estimator design [29, 30, 31, 32, 33, 34]. On the other hand, the research area of estimation theory of linear and non-linear lumped parameter systems (i.e., finite-dimensional systems) has flourished over the past decades [35, 36]. Among these, a representative approach is moving horizon estimation (also named receding horizon estimation) [37, 38, 39] that can be formulated as a quadratic optimization problem, thus capable of handling disturbances, measurements and input constraints widely presented in distributed or lumped parameter systems and engineering practice. However, constrained state/output estimation for infinite-dimensional systems has not been fully investigated. This thesis is devoted to systematically studying state/output and parameter estimation of deterministic and stochastic linear infinite-dimensional systems by extending the Luenberger observer, Kalman filter, and moving horizon estimator theories and designs developed for continuous-time infinite-dimensional cases and/or discrete-time finite-dimensional cases.

Furthermore, the designed controllers and/or estimators often suffer from time-varying operating conditions and/or abnormal working conditions (such as flow leakage), leading to performance deterioration. Under such conditions, accurate model parameters and/or structures are often not available and hence model re-identification is typically needed. To address these issues, several estimators are further proposed in this thesis, including the advanced moving horizon estimator for state/output and

mode estimation of regular linear infinite-dimensional systems and a hybrid estimator (i.e, a grey box algorithm) for pipeline leak detection, localization and size estimation by combining a model-based Luenberger observer and a data-driven regression algorithm. Considering the extreme cases when the model structure is entirely unknown and under varying working conditions, a pure data-driven method – transfer slow feature analysis algorithm in the form of the linear lumped parameter model with Gaussian noises is proposed for soft-sensor (i.e., predictor) design of quality variables. This thesis covers physical model-based and data-driven methods and their integration for state/output and parameter estimation of linear distributed and lumped parameter systems, and have potential applications in adaptive soft sensor designs for linear PDE systems (including transferring soft sensors designed for one PDE system to another). This is of theoretical and practical significance in model reduction, model adaptation and mismatch handling of distributed and lumped parameter systems as well as transfer learning development from dynamical system point of view.

1.2 Literature review

1.2.1 Distributed parameter systems

A significant amount of industrial systems are "distributed" in space so that their dynamical behaviours depend both on time and space [30]. These systems are named distributed parameter systems and are often modelled by partial different equations and/or delay equations. Compared to lumped parameter systems that are typically described by ordinary differential equations, distributed parameter systems enable us to describe a wide range of industrial processes with spatiotemporal dynamics, such as heat transfer, material processing, fluid dynamics, and structural vibration that are widely presented in chemical, biochemical, mechanical and civil engineering and science [4, 15, 21, 40, 41, 42, 43].

Majority of DPS models can be derived from first principles, which include conservation laws of continuity, momentum, and energy. Given these physical conservation laws, well-defined inputs, states and outputs can be identified with corresponding physical meaning. To complete a physical realization of a distributed parameter system, suitable boundary conditions (BCs) and initial conditions (ICs) should be given

correspondingly. Let us take a general second-order infinite-dimensional model as an example:

$$\frac{\partial^2 x(\zeta,t)}{\partial t^2} = f\left(x(\zeta,t), u(t), \frac{\partial x(\zeta,t)}{\partial \zeta}, \frac{\partial x(\zeta,t)}{\partial t}, \frac{\partial^2 x(\zeta,t)}{\partial t \partial \zeta}, \frac{\partial^2 x(\zeta,t)}{\partial \zeta^2}\right) \tag{1.1}$$

$$y(t) = g(x(\zeta,t), u(t)) \tag{1.2}$$

with initial conditions and boundary conditions given as:

$$x(\zeta,0) = x_0^1(\zeta),\ \frac{\partial x(\zeta,0)}{\partial t} = x_0^2(\zeta) \tag{1.3}$$

$$c_0 \frac{\partial x(0,t)}{\partial \zeta} + c_1 x(0,t) = 0,\ c_2 \frac{\partial x(l,t)}{\partial \zeta} + c_3 \frac{\partial x(l,t)}{\partial t} = 0 \tag{1.4}$$

where $x(\zeta,t)$, $u(t)$ and $y(t)$ denote the state, input and output of interest, and ζ and t stand for spatial and temporal coordinates. In addition, f and g can be two non-linear or linear functions of state x, input u, first-order spatial derivative of state $\frac{\partial x(\zeta,t)}{\partial \zeta}$, second-order spatial derivative of state $\frac{\partial^2 x(\zeta,t)}{\partial \zeta^2}$, first-order temporal derivative of state $\frac{\partial x(\zeta,t)}{\partial t}$, and temporal derivative of first-order spatial derivative $\frac{\partial^2 x(\zeta,t)}{\partial t \partial \zeta}$. In the formulation of lumped parameter systems, there is no spatial dependence of state $x(\zeta,t)$ as in Eq.(1.1). The spatial characteristics make control and estimation design of distributed parameter systems relatively difficult in comparison with lumped parameter systems.

The initial conditions are described by spatial functions as given in Eq.(1.3). In addition, a general mixd Robin boundary condition at $\zeta = 0$ is described as Eq.(1.4), which is more physically realistic for practical processes. When we define the parameters $c_0 = c_2 = 0$, it can lead to the well-known Dirichlet boundary conditions with flux of states equal to zero. On the other hand, one can define the Neumann boundary conditions by having $c_1 = c_3 = 0$.

To classify a linear second-order PDE model, one can linearize Eq.(1.1) as the following expression:

$$a\frac{\partial^2 x(\zeta,t)}{\partial t^2} + 2b\frac{\partial^2 x(\zeta,t)}{\partial t \partial \zeta} + c\frac{\partial^2 x(\zeta,t)}{\partial \zeta^2} + (\text{lower order terms}) = 0 \tag{1.5}$$

Then, the above equation is referred to as:

- Parabolic PDE, if $b^2 - ac = 0$, for diffusion and conduction problems.

- Hyperbolic PDE, if $b^2 - ac > 0$, for vibration and wave motion problems.
- Elliptic PDE, if $b^2 - ac < 0$, for steady-state and potential-type problems.

There is a single state $x(\zeta, t)$ in the PDE model (1.1), and the state varies in one dimensional spatial domain $\zeta \in [0, l]$. In realistic industrial processes, distributed parameter systems can have multiple states that are coupled together in a multi-dimensional spatial domain (such as two dimensions or three dimensions), which poses a significant challenge in state estimation, control and output regulation.

1.2.2 Control and estimation of distributed parameter systems

Generally speaking, there are two main classes of controller and estimator design methodologies proposed for infinite-dimensional systems in literature, namely, *early lumping* and *late lumping* [30]. The basic idea of early lumping approaches is to firstly approximate distributed parameter systems in the spatial variables by utilizing finite difference methods, spectral approaches and/or finite element techniques, and then apply finite-dimensional state estimation and control methods to the resulting early lumped parameter systems. In contrast, late lumping approaches refer to that state estimation and control design steps are carried out directly on the original distributed parameter systems by using infinite-dimensional control theories, and then spatial approximation or discretization is performed in the numerical simulation stage for realization purposes. Both early and late lumping methods have pros and cons when it comes to the theoretical design and numerical performance. For instance, late lumping methods produce accurate estimators and controllers for the original DPSs while relatively complicated mathematical techniques are required as the underlying mathematical models are governed by partial differential equations. Compared to the late lumping methods, early lumping approaches are simpler to be conducted by exploring finite-dimensional control theories and methods, while questions of stability and performance of the designed systems naturally arise [44].

Over the past decades, several central control problems have been focused on in the distributed parameter systems community, namely, optimal control [4, 40, 45, 46, 5, 6, 47, 48, 49, 50, 51], internal model control (i.e., servo-control or output regulation)

[10, 11, 52, 12, 53, 13, 54, 55, 56, 57], boundary control and stabilization [4, 42, 15, 58, 16, 17, 59, 60, 43], predictive control [61, 62, 63, 64, 65], nonlinear and robust control [66, 67, 68], and adaptive control [69, 70]. Regarding these aspects, many important results and methods have been generated, including semigroup methods [4, 42, 21, 43, 71], back-stepping techniques [15, 72, 70], port-Hamiltonian approaches [41, 73, 74], optimization-based methods [61, 64, 65] and etc. Most of these contributions are based on continuous-time infinite-dimensional models or continuous-time finite-dimensional models after applying some forms of model reductions. Hence, designing controllers and estimators for discrete-time infinite-dimensional models are of theoretical and practical significance.

Compared to the aforementioned contributions on controller designs and model stabilization, relevant studies on the state/output and parameter estimation of infinite-dimensional systems have attracted less attentions although the estimation theory (e.g., infinite-dimensional filtering) was initially developed quite early that can be traced back to the 1970s, see [75, 76, 8, 7, 77]. For example, the filtering problem for linear infinite-dimensional continuous-time systems with bounded operators was systematically studied in [78, Cha. 6]. A synthesizing overview on parameter identification in distributed systems was reported in [29], where various techniques and applications were discussed including identification for process control, structural vibrations, agricultural parameter determination, petroleum, gas, and mineral exploration, oceanic temperature and convection current determination and etc. Recently, a comprehensive review on the applications and classifications of observers to chemical process systems was offered in [79]. The state estimation problem of distributed parameter systems based on information from a finite number of sensors was systematically investigated in [40, Cha. 6]. Essentially, unbounded operators and infinite dimensionality of state space are two main difficulties in controller and estimator designs of PDE systems. Moreover, disturbances, constraints and uncertainties issues widely presented in industrial process system engineering pose another layer of technical challenge to the estimation and control of DPSs. On the other hand, the theoretical development of state, parameter estimation and model identification of finite-dimensional systems have been flourished since the 1970s [80, 81, 82, 35]. Numerous methods and algorithms developed therein show great potential to the es-

timation theory development of infinite-dimensional systems. Considering that many important infinite-dimensional control theories and methods are actually extended from finite-dimensional cases (e.g., optimal control, back-stepping and etc), this thesis aims to further extend the finite-dimensional estimation theories to enhance the estimation, soft sensing and servo-control performance of infinite-dimensional systems.

1.3 Research objectives

The overall goal of the thesis is to extend the state and parameter estimation theories of lumped parameter systems into the realm of linear distributed parameter systems and the theories of continuous-time output regulation of linear distributed parameter systems to general discrete-time linear distributed parameter systems. As an alternative of online instrumental measurements, a soft sensor design is proposed for online estimation of key process variables of lumped parameter systems, with the potential to be applied to distributed parameter systems.

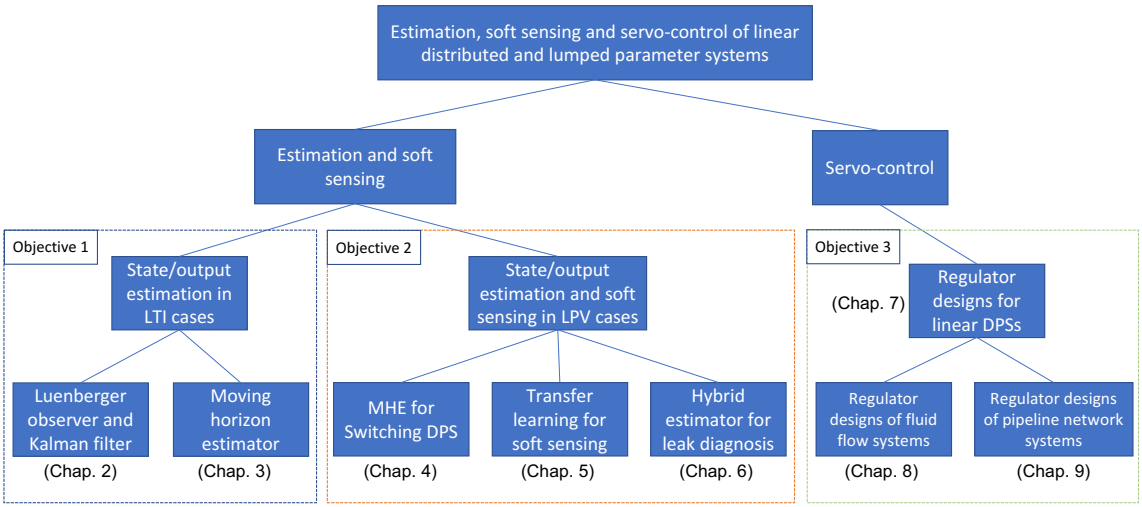


Figure 1.1: Outline of the research objectives

As illustrated in Fig. 1.1, the detailed research objectives are:

- Observer, filter and moving horizon estimator designs for linear distributed parameter systems

  Designs of continuous- and discrete-time infinite-dimensional Luenberger ob-

servers are covered. The theoretical extensions of finite-dimensional discrete-time Kalman filter and moving horizon estimator to linear infinite-dimensional systems are explored. These objectives are considered in Chapters 2-3.

- Hybrid estimator and soft sensor designs for linear distributed and lumped parameter systems

  The design problems range from state/output and mode estimation of switching regular linear systems, to soft sensing of linear lumped parameter systems, to the leak detection, localization and estimation of hyperbolic PDE systems. The design techniques span from the model-based MHE, to the data-driven (i.e., entirely model-free) Bayesian inference, to the hybrid estimator using model-based and data-driven methods. These objectives are considered in Chapters 4-6.

- Discrete output regulation of linear distributed parameter systems

  Two regulator design problems including discrete-time state and error feedback output regulators are considered along with in-domain or boundary disturbance, control, and outputs. The relationship between discrete- and continuous-time regulator design problems is revealed under the Cayley-Tustin transformation. Spectral, non-spectral, and networked PDE systems are investigated. These objectives are considered in Chapters 7-9.

1.4 Contributions and thesis outline

This thesis is concerned with discrete-time observer, estimator, and controller designs using Cayley-Tustin transform in a late lumping manner (sometimes called direct design). Under this transformation, a discrete-time infinite-dimensional system is obtained along with essential system properties being preserved, including stability, observability, etc. The novelty in this work is that the obtained discrete-time system is in a form that facilitates estimator (and controller) designs without any approximation or lumping in the spatial domain. The main contributions of the thesis are the constrained output estimation of distributed parameter systems (and switching DPSs) using moving-horizon strategies and Cayley-Tustin transform.

Chapter 2 extends known finite-dimensional Luenberger observer and Kalman filter designs to the realm of linear transport-reaction systems frequently present in chemical engineering practice. A unified modelling framework for distributed parameter systems that does not induce any type of spatial approximation or model order reduction is developed. Applying the Cayley-Tustin transformation to continuous linear distributed parameter systems leads to structure- and property-preserving discrete distributed parameter models, amenable to observer and filter design developments. Designs presented in this chapter cover well-known state reconstruction methodologies ranging from least square estimation, to continuous- and discrete-time Luenberger observers, to Kalman filter designs. Simple implementation and realization account for the appealing nature of the discrete-time observers and filter designs for linear transport-reaction systems. The simulation scenarios cover the majority of representative examples existing in process system engineering practice.

Chapter 3 addresses constrained output estimation of well-posed regular linear infinite-dimensional systems by using a moving horizon estimation approach. Considering that directly estimating continuous systems modelled by partial differential equations can be challenging due to the unbounded operators (induced by boundary or point-wise disturbance and measurement) and the presence of disturbance and output constraints, the continuous-time system with unbounded operators is transformed into a discrete-time infinite-dimensional model with all bounded operators using Cayley-Tustin transformation without any spatial discretization and/or model reduction. Instead of constructing a continuous-time Riccati equation with unbounded operators, a well-posed discrete-time Riccati equation suitable for numerical computation is shown to have the same minimal nonnegative self-adjoint solution as the continuous-time Riccati equation. The discrete-time model and Riccati equation are utilized in the moving horizon estimator design accounting for estimation efficiency and physical constraints on disturbance and output in an explicit way. The optimality and stability are proved and the corresponding sufficient conditions are presented. The resulting receding-horizon estimator leads to a finite-dimensional constrained quadratic optimization problem easily solvable by standard optimization techniques. Finally, two simulation examples including Schrödinger and wave equations demonstrate the effectiveness of the proposed method.

Chapter 4 addresses moving horizon estimation for a class of switching regular linear infinite-dimensional systems described by partial differential equations, where the system is corrupted with bounded plant (or input) and measurement disturbances and the system mode is regarded as an unknown and unpredictable discrete state to be estimated. To address the issues associated with unbounded operators (induced by boundary or point-wise observation and disturbance) and better suited for discrete-time moving horizon estimator design, the Cayley-Tustin transformation is deployed for model time discretization without any spatial discretization or model reduction while preserving model essential properties. A series of observability concepts along with corresponding properties are proposed and analyzed for the switching linear infinite-dimensional discrete-time systems. Two moving horizon estimation algorithms that accounts for state/output and mode estimation and constraint handling are proposed. Based on the proposed observability properties, the stability of the proposed moving horizon estimators is proved. Two simulation examples are provided to verify the derived results.

In Chapter 5, soft sensor designs for linear lumped parameter systems using variational Bayesian inference are proposed. Data-driven methods have been extensively utilized in establishing predictive models from historical data for process monitoring and prediction of quality variables. However, most data-driven approaches assume that training data and testing data come from steady-state operating regions and follow the same distribution, which may not be the case when it comes to complex industrial processes. To avoid these restrictive assumptions and account for practical implementation, a novel online transfer learning technique is proposed to dynamically learn cross-domain features based on the variational Bayesian inference in this chapter. Stemming from the probabilistic slow feature analysis, a transfer slow feature analysis (TSFA) technique is presented to transfer dynamic models learned from different source processes to enhance prediction performance in the target process. In particular, two weighting functions associated with transition and emission equations are introduced and updated dynamically to quantify the transferability from source domains to the target domain at each time instant. Instead of point estimation, a variational Bayesian inference scheme is designed to learn the parameters under probability distributions accounting for corresponding uncertainties. The effectiveness of

the proposed technique with applications to soft sensor modelling is demonstrated by a simulation example, a public dataset and an industrial case study.

In Chapter 6, a hybrid estimator design for leak detection, localization, and estimation of a realistic pipeline modelled by a nonlinear coupled first-order hyperbolic PDE system is studied. Based on the so-called water hammer equation, a linear distributed parameter system is obtained by linearization. The structure- and energy-preserving Cayley-Tustin time discretization scheme is used to realize a discrete infinite-dimensional hyperbolic PDEs system without spatial approximation or model order reduction. In order to reconstruct pressure and mass flow velocity evolution with limited measurements, a discrete-time Luenberger observer is designed by solving the Riccati equation. Based on this distributed observer system, data associated with different normal and leakage conditions (i.e., various leak amounts and positions) are generated and fed to train a support vector machine model for leak detection, amount and position estimation. Finally, the effectiveness of the developed method on leak detection, amount estimation and localization are verified by a set of simulations.

In Chapter 7, we address discrete-time state and error feedback output regulator designs for a class of linear distributed parameter systems with bounded control and observation operators. By utilizing the Cayley-Tustin bilinear transform, a linear infinite-dimensional discrete-time system is obtained without model spatial approximation or model order reduction. Based on the discrete-time model representation, discrete state and error feedback regulators are designed using the internal model principle. In particular, discrete Sylvester regulator equations are formulated, and their solvability is proved and linked to the solvability of their continuous counterparts. In addition, to ensure the stability of the closed-loop system, the design of stabilizing feedback gain and its dual problem of stabilizing output injection gain design are provided in the discrete-time setting. Finally, three simulation examples including a first-order hyperbolic partial differential equation model and a 1-D heat equation with considerations of step-like, ramp-like and harmonic exogenous signals are shown to demonstrate the applicability of the proposed method.

In Chapter 8, model-based discrete-time output regulator design is proposed for fluid flow systems using a geometric approach. More specifically, a class of vortex

shedding and falling thin film phenomena modelled by the complex Ginzburg-Landau equation (CGLE) and the Kuramoto-Sivashinsky equation (KSE) are considered. Differently from a traditional continuous-time controller design, a novel discrete-time modelling technique is proposed in a general infinite-dimensional state-space setting, which does not pertain any spatial approximation or model reduction, and preserves model intrinsic properties (such as stability, controllability and observability). Based on the time discretized plant model (CGLE and KSE systems) by the Cayley-Tustin method, discrete Sylvester regulation equations are established and facilitated for an output regulator design to achieve fluid flow control and manipulation. To address model instability, a spectrum analysis is utilized in stabilizing continuous-time CGLE and KSE systems, and a link between discrete- and continuous-time closed-loop system stabilizing gains is established. Finally, the proposed methodology is demonstrated through simulation studies, by which the output tracking, disturbance rejection, and model stabilization are achieved for the considered CGLE and KSE systems.

In Chapter 9, a discrete-time output regulator design is proposed for a class of gas pipeline network systems to meet various operating requirements in energy scheduling. Based on the isothermal Euler equations, linearized continuous-time gas pipeline network models with boundary actuation and sensing in the infinite-dimensional space are established, with consideration of Rankine-Hugoniot conditions at junction joints. Cayley-Tustin bilinear transformation is applied for model time discretization without any spatial approximation, by which the continuous-time model with unbounded operators is transformed into an infinite-dimensional discrete-time system with all bounded operators and essential continuous-time properties are invariant under this transformation. Based on the internal model principle, the discrete output regulator is constructed and its solvability conditions are provided. Considering the unavailability of the full state information, observer design methods for state estimation of exogenous and pipeline systems are proposed in order to construct an output feedback regulator. Additionally, a stability analysis of the considered pipeline system is provided. Finally, two simulation examples representing a single gas pipeline and a star-shaped pipeline network are given to demonstrate the applicability of the proposed method.

Chapter 2

Observer and Kalman Filter Design for Linear Infinite-Dimensional Systems

2.1 Introduction

Transport-reaction processes represent the core of relevant first principle based models in chemical engineering practice. The prominent feature of transport-reaction systems is that their models belong to the class of distributed parameter systems (DPS), which are given by partial differential and/or delay equations. Typical distributed parameter system plants are characterized with dominant spatial characteristics along which temporal and spatial evolution of physical properties takes place. In addition, distributed parameter systems are equipped with spatially distributed and/or boundary placed actuators, which are frequently combined with spatially distributed measurement sensor devices. In principle, one cannot measure everywhere in the physical plant either due to the inability to realize spatially distributed sensors and/or due to physical limitations and accessibility limitations associated with the plant design. Therefore, distributed parameter systems are compelling plant models to be explored from the state reconstruction point of view by the estimator, observer and filter designs.

An important issue of monitoring and estimating spatiotemporal states and their evolution in the distributed parameter systems has been explored in the relevant chemical process literature [30, 83] and it has been initiated with seminal contributions of Luenberger [84, 81] and Kalman [85] who introduced basic concepts of state

Luenberger observer and Kalman filter designs. In particular, initial contributions of Harmon Ray [30] and Seinfield [83, 86, 87] explored and extended finite dimensional designs to the framework of partial differential equations. Their work introduced the notion of optimality in the measurement sensors placement and conditions on observability of linear distributed systems that are mainly used to describe models of axial dispersion and tubular reactors [83, 86]. In addition to these efforts, the fundamental theoretical contribution in [4] extended existing finite dimensional concepts of the observer design to the infinite dimensional setting, in a compact and elegant form, in the case of continuous models.

The major theme in the observer design within the DPS setting is to apply either *early lumping* and reduce the system dimensionality by using appropriate spatial discretization methods and/or model approximations suitable for the finite dimensional observer design, or to apply *late lumping* by performing the analysis and observer design in the infinite dimensional setting, and then lump the designed observer along some spatial approximation for the purpose of implementation [30, 88, 89]. Following the line of proposed designs within the area of chemical engineering, advances that addressed the design of state observers for transport reaction systems with unknown kinetics have been developed by Dochain and co-workers and a recent review paper (see [79]) on observers applied in chemical process systems scarcely refers to a possibility that finite dimensional observers can be extended to the models which account for representative examples of axial dispersion and plug flow reactors [90, 91, 92, 79]. Additionally, in some instances the observer design is applied to a fully nonlinear and data based processing setting [93, 94]. However, these contributions always assume some type of model approximation and/or model reduction in the design procedure, or impose requirement on a large number of implemented point measurements (or continuous measurements) in the design procedure [90, 94].

Therefore, the issue of spatial approximation by some modelling technique and/or model reduction is one of the main issues when observer designs for distributed parameters are considered, and addressing the question of the observer design that does not account for spatial approximations nor lumping is of importance. In addition to addressing the design step of model reduction and the drawbacks associated with it, the distributed parameter systems, when endowed with stochastic processes charac-

terized by the noise that exists in every physically realizable setting, bring technical and theoretical complexity in their consideration [95]. For example, the finite dimensional linear state space setting can be endowed easily in the stochastic framework by adding noise arising from the process and measurements, while the distributed parameter systems, due to the infinite-dimensional nature, pose a problem of appropriate setting formulation when it comes to their description. Although discrete design methodologies are very well established and explored for finite dimensional chemical process models [96], the transport-reaction systems in chemical processes are represented ultimately within the continuous infinite-dimensional setting that induces difficulties in obtaining a relevant discrete system description which adequately represents spatially varying transport-reaction continuous models. Along the line of discrete model representations and designs, current modern implementation technologies are ultimately realized as discrete sampled data system realizations, which requires a discrete state space setting in order to provide foundations for easy and robust controllers and estimators design that can be easily reconfigured and/or maintained under a variety of operating conditions in the process plants.

In this chapter, we explore well known discrete estimation techniques within the setting of discrete infinite dimensional models which are obtained from the well-known first principle models of linear transport-reaction systems ranging from axial dispersion to plug flow reactor models. The salient feature of all designs proposed in this chapter is that the models of transport-reaction systems are not approximated by any spatial discretization process and no spatial model reduction is applied in the discrete model generation. This is achieved by the application of Cayley-Tustin discretization which is a symmetric and symplectic discretization approach, such that the physical characteristics (energy) and theoretical properties of considered systems can be preserved. In this way, the continuous distributed parameter systems are transferred to discrete counterparts, which can be utilized and explored within known finite dimensional designs, with well defined corresponding infinite dimensional properties. The chapter presents a standard least square sequential estimation applied to the DPS system, and we point out characteristics, limitations and to what extent this technique can be used in physically relevant DPS systems. Further, we provide the insight in the application and developments associated with the Cayley-Tustin discretization

and associated results that provide foundation for application of this discretization techniques to DPS systems. Along the line of designs, the Luenberger observer design for the continuous transport-reaction systems described by hyperbolic and parabolic PDEs is accomplished by design of converging observer by exploring the corresponding link among Riccati and Lyapunov operator equations. We provide an important insight in the relation among discrete and continuous observer designs which is followed by the proofs of main results. The considered designs account for the boundary and in domain (interval or point) measurement sensor placement providing the output admissibility of the operators settings. Along the same line, the Kalman Filter for the transport-reaction system described by hyperbolic and parabolic PDEs is developed as one-step ahead predictor, and accounts for the process and measurement noise, and the different treatment between the finite-dimensional and infinite-dimensional Kalman filter realizations is clarified. The link between the Cayley-Tustin discretization scheme and Kalman filter design is provided by demonstrating the equality of similar discrete state space realizations which have unique stable input-output (I/O) transfer function realization. Finally, in the numerical simulations section representative examples of hyperbolic and parabolic systems are simulated and the performance of associated least square estimator, discrete and continuous Luenberger observer and Kalman filter are demonstrated through simulation studies.

2.2 Preliminaries

Linear transport-reaction systems are described by the following general form as:

$$\dot{x}(t) = \mathcal{A}x(t) + \mathcal{B}u(t), \quad x(0) = x_0 \tag{2.1}$$

$$y(t) = Cx(t) + \mathcal{D}u(t) \tag{2.2}$$

where spatial state is $x(t) \in \mathcal{X}$, where $\mathcal{X} = L^2((0,l),\mathbb{R})$ is being defined as separable Hilbert space. The input is $u \in L^2_{loc}([0,\infty),U)$, and the U is real Hilbert space, $y \in L^2_{loc}([0,\infty),Y)$ and Y is real Hilbert space. The usual assumption on $\mathcal{A}$ being closed implies that domain $\mathbb{D}(\mathcal{A})$ is also Hilbert space equipped with the graph norm $\|x\|^2_{\mathbb{D}(\mathcal{A})} := \|x\|^2_{\mathcal{X}} + \|\mathcal{A}x\|^2_{\mathcal{X}}$ and with the understanding that resolvent set $\rho(\mathcal{A})$ is nonempty, with the norm $\|x\|_{\mathbb{D}(\mathcal{A})} = \|(\alpha - \mathcal{A})x\|_{\mathcal{X}}$ with an arbitrary $\alpha \in \rho(\mathcal{A})$. Hence,

a different selection of α induces an equivalent norm to $\mathbb{D}(\mathcal{A})$. We denote $\mathcal{X}_1 := \mathbb{D}(\mathcal{A})$ and use the norm $\|(\alpha I - \mathcal{A})x\|_{\mathcal{X}}$, then $(\alpha I - \mathcal{A})^{-1}$ maps $\mathcal{X}$ isometrically to $\mathcal{X}_1$, which enables us to define the space $\mathcal{X}_{-1}$ as the completion of $\mathcal{X}$ with respect to the norm $\|x\|_{\mathcal{X}_{-1}} = \|(\alpha I - \mathcal{A})^{-1}x\|_{\mathcal{X}}$. By iteration of this construction, one can define $\mathcal{X}_j$ for any $j \in \mathbb{Z}$ with $\mathcal{X}_j \subset \mathcal{X}_k$ if $j \leq k$ with a dense inclusion, in other words, the extension or restriction of $\mathcal{A}$ and corresponding semigroup $\mathcal{T}(t)$ to $\mathcal{A}_j \in \mathcal{L}(\mathcal{X}_{j+1}, \mathcal{X}_j)$ and $\mathcal{T}_j(t) \in \mathcal{X}_j$ is feasible. By introducing a block operator called node $\mathcal{S} := \begin{bmatrix} \mathcal{A}\&\mathcal{B} \\ \mathcal{C}\&\mathcal{D} \end{bmatrix}$ as a mapping $\mathcal{X}\times U \to \mathcal{X}\times Y$, one defines the operator node on $(\mathcal{X}, U, Y)$ with the following structure: 1) $\mathcal{A}$ is a closed, densely defined operator on $\mathcal{X}$ with a nonempty resolvent set, 2) $\mathcal{B} \in \mathcal{L}(U, \mathcal{X}_{-1})$, 3) $\mathbb{D}(S) = \left\{ \begin{bmatrix} x \\ u \end{bmatrix} \in \mathcal{X} \times U : \mathcal{A}_{-1}x + \mathcal{B}u \in \mathcal{X} \right\}$, where $\mathcal{A}_{-1}$ is the extension of $\mathcal{A}$ (this is to account for boundary or point actuation) and $\mathbb{D}(S)$ is equipped with the graph norm $\left\| \begin{bmatrix} x \\ u \end{bmatrix} \right\|^2_{\mathbb{D}(S)} := \|\mathcal{A}_{-1}x + \mathcal{B}u\|^2_{\mathcal{X}} + \|x\|^2_{\mathcal{X}} + \|u\|^2_U$, 4) $\mathcal{C}\&\mathcal{D} \in \mathcal{L}(\mathbb{D}(S), Y)$, see [21, 20]. The mapping $\mathcal{S} : u(\cdot) \mapsto y(\cdot)$ is the input/output map of S, and its Laplace transform is the transfer function $\mathcal{G}(s) = \mathcal{C}(sI - \mathcal{A})^{-1}\mathcal{B} + \mathcal{D}$. The operators $\mathcal{B}$ and $\mathcal{C}$ denote input (representing actuation) and output (representing measurements) operators which can be spatially distributed and/or applied at the boundary. In current contribution, we constrain analysis to the case of bounded input and output operators in the following examples, keeping in mind that one needs to revise the space setting as given above, if point or boundary input or measurements are applied. In addition, the setting accounts for majority of linear transport-reaction systems which also includes the systems with spatially varying coefficients.

2.2.1 Cayley-Tustin discretization

Let us consider the following linear systems dynamics given by Eqs. (2.1)-(2.2) and we apply the Crank-Nicolson type of discretization to the continuous system Eqs. (2.1)-(2.2) for any given time discretization interval Δt:

$$\frac{x(k\Delta t) - x((k-1)\Delta t)}{\Delta t} \approx \mathcal{A}\frac{x(k\Delta t) + x((k-1)\Delta t)}{2} + \mathcal{B}u(k\Delta t),\ x(0) = x_0 \tag{2.3}$$

$$y(k\Delta t) \approx \mathcal{C}\frac{x(k\Delta t) + x((k-1)\Delta t)}{2} + \mathcal{D}u(k\Delta t), \quad k \geq 1 \tag{2.4}$$

then by simple approximation of $u(k\Delta t)$ (substitute $u(k\Delta t)$ with $\frac{u(k\Delta t)}{\sqrt{\Delta t}}$), and it can be shown that $\frac{u(k\Delta t)}{\sqrt{\Delta t}}$ converges to $u(k\Delta t)$ as $\Delta t \to 0$ in several different ways, similar for

$y(k\Delta t)$, see [97]. Further, in the engineering literature a finite dimensional discrete time dynamics given by Eqs. (2.3)-(2.4) is frequently called Tustin discretization (it is discovered in 1940s by Tustin and it is referred also as *Tustin transform* in digital and sample-data control literature, see [98, p.27]) and it is given as:

$$\frac{x(k\Delta t) - x((k-1)\Delta t)}{\Delta t} \approx \mathcal{A}\frac{x(k\Delta t) + x((k-1)\Delta t)}{2} + \mathcal{B}\frac{u(k\Delta t)}{\sqrt{\Delta t}},\ x(0) = x_0 \tag{2.5}$$

$$\frac{y(k\Delta t)}{\sqrt{\Delta t}} \approx \mathcal{C}\frac{x(k\Delta t) + x((k-1)\Delta t)}{2} + \mathcal{D}\frac{u(k\Delta t)}{\sqrt{\Delta t}}, \quad k \geq 1 \tag{2.6}$$

In order to address the assumption associated with the discretization of the input signal $\left\{\frac{u(k\Delta t)}{\sqrt{\Delta t}}\right\}$, the following is considered. The discretizing operator is given by $\frac{u(k\Delta t)}{\sqrt{\Delta t}} = \frac{1}{\Delta t}\int_{(k-1)\Delta t}^{k\Delta t} u(t)dt$ which is defined as mean value within a given sampling time. Therefore, for the assumption on piecewise-constant input in the sampled intervals one can obtain the standard input approximation applied in other discretization schemes. Then, the following approximate discrete time counterpart of Eqs. (2.1) -(2.2) is given as:

$$x(k\Delta t) = \mathcal{A}_d x((k-1)\Delta t) + \mathcal{B}_d u(k\Delta t), \quad x(0) = x_0 \tag{2.7}$$

$$y(k\Delta t) = \mathcal{C}_d x((k-1)\Delta t) + \mathcal{D}_d u(k\Delta t), \quad k \geq 1 \tag{2.8}$$

with corresponding discrete-time spatial operators given as follows:

$$\mathcal{S}_d := \begin{bmatrix} \mathcal{A}_d & \mathcal{B}_d \\ \mathcal{C}_d & \mathcal{D}_d \end{bmatrix} = \begin{bmatrix} (\delta I - \mathcal{A})^{-1}(\delta I + \mathcal{A}) & \sqrt{2\delta}(\delta I - \mathcal{A})^{-1}\mathcal{B} \\ \sqrt{2\delta}\mathcal{C}(\delta I - \mathcal{A})^{-1} & \mathcal{G}(\delta) + \mathcal{D} \end{bmatrix} \tag{2.9}$$

and its transfer function is given as $G_d(z) = \mathcal{C}_d z(I - z\mathcal{A}_d)^{-1}\mathcal{B}_d + \mathcal{D}_d$. The well-known bilinear mapping (Mobius transform) which maps the open right-half plain $\mathbb{C}^+ = \{s \in \mathbb{C} : \Re e(s) > 0\}$ into the exterior of the unit disc $\mathbb{D}^+ = \{z \in \mathbb{C} : |z| > 1\}$, given as $z = \frac{\delta+s}{\delta-s}$ and conversely $s = \frac{z-1}{z+1}\delta$. The Cayley transform is defined as minus the Mobius transform and it was introduced by von Neumann [99], so that the mapping between $\mathcal{S} \mapsto \mathcal{S}_d$ is called the Cayley transform of continuous time systems to discrete time systems, so that following folds $G_d(z) = G(\frac{1-z}{1+z}\delta)$ and establishes one-to-one relationship between the continuous and the discrete time system [62, 97]. Further, it can be proved that given $\delta > 0$ and $\mathcal{S}$ to be a system node whose main operator satisfies $\mathbb{R}^+ \subset \rho(\mathcal{A})$, then the system $\mathcal{A}$ is (continuous time) input-output

stable if and only if its Cayley transformation $\mathcal{S}_d$ is (discrete time) input/output stable.

In order to address the issue of dynamical system theoretic preserving properties of Cayley-Tustin discretization we are looking in the energy balances since the conservativity is given by the following expression for node $\mathcal{S}$ as:

$$\frac{d}{dt}\|x(t)\|_{\mathcal{X}}^2 = |u(t)|^2 - |y(t)|^2, \quad \text{for } t \in \mathbb{R}^+ \tag{2.10}$$

and in discrete time for $\mathcal{S}_d$ as:

$$\|x(k\Delta t)\|_{\mathcal{X}}^2 - \|x((k-1)\Delta t)\|_{\mathcal{X}}^2 = |u(k\Delta t)|^2 - |y(k\Delta t)|^2, \quad \text{for } k \in \mathcal{Z}^+ \tag{2.11}$$

Then, the Cayley-Tustin transform $\mathcal{S}_d$ of an energy preserving system node $\mathcal{S}$ is an energy preserving discrete linear system. Moreover, such transform $\mathcal{S}_d$ is conservative (discrete time) if and only if node $\mathcal{S}$ is conservative.

The discrete form of Eqs. (2.1)-(2.2) is dynamical structure preserving since the system theoretic properties (strong stability, approximate controllability and observability, infinite-time input and output admissibility) are preserved under the Cayley-Tustin transform and it is parameterized with the discretization time $\delta = \frac{2}{\Delta t}$ which can be freely chosen in the case of dynamically stable distributed parameter processes, see for details in [23]. Moreover, in [23] the one to one relationship is established among the discrete and continuous time system Lyapunov equations strong stabilizing solutions. In addition, to the conservative systems settings, the framework can be extended to the dissipative distributed parameter system - nodes satisfying that $\dim(U) = 1 = \dim(Y)$, see [97].

In addition, in this chapter we consider distributed parameter systems which are physically realizable. This implies that they are strictly proper and are characterized with $\mathcal{D} = 0$ implying that there is no instantaneous transfer of the input signal to the output, so that the expression in Eq. (2.9) for $\mathcal{G}(\delta) = \mathcal{C}(\delta I - \mathcal{A})^{-1}\mathcal{B}$ is just a transfer function of the model evaluated at $s = \delta$. This implies that whenever one can obtain the closed analytic form of the transfer function of the continuous transport-reaction model it is also feasible to write an exact discrete form of the underlying model, as this will be explored and demonstrated in the ensuing sections. Finally, the benefit of the underlying model's discrete form is also recognized in the ability to

handle boundary/point actuation (if present in the system), as well as boundary or point observation, since discrete formulation provides boundness of input and output operators in the infinite dimensional space setting, see [100].

Remark 1. *The Crank-Nicolson type of discretization setting originates from the well-known implicit midpoint rule, which is a symmetric numerical method since Eq. (2.3) stays unaltered, if* $x(k\Delta t)$ *is exchanged with* $x((k-1)\Delta t)$*, (*$x(k\Delta t) \leftrightarrow x((k-1)\Delta t)$*), and* $\Delta t \leftrightarrow -\Delta t$*, see [101]. This is also the type of discretization referred as reversible in the time discretization method which can guarantee structure-preserving numerical integrations [101]. The integration schemes that preserve the energy equality or more complex dynamics invariants of the system are called symplectic or Hamiltonian integrators. Therefore, the Crank-Nicolson discretization for the linear system is the lowest order symplectic integration scheme from the the family of Gauss quadrature based Runge-Kutta methods [101].*

2.3 Least square estimation of constant distributed parameter state

The finite dimensional least square estimation is well developed and documented as one of the first estimation techniques utilized in systems science [35]. It is of interest to explore how this concept is extended when it comes to the distributed parameter systems setting. First, we consider the time invariant spatially varying steady state $x = x(\zeta)$, that frequently arises in the transport-reaction systems as the physically realizable equilibrium profile. The measurement of the state is given as:

$$y_k = Cx + v_k \tag{2.12}$$

where C is the output measurement operator which determines the place of the output measurement. The noisy measurement y_k is realized as the time invariant spatially varying state $x = x(\zeta)$ measurement augmented with the noise signal v_k given as a random variable with zero mean and covariance R_k (i.e., $v_k \sim \mathcal{N}(0, R_k)$). The realization of the recursive least-square estimator is given as follows:

$$K_k = P_{k-1}C^*(CP_{k-1}C^* + R_k)^{-1} = P_kC^*{R_k}^{-1} \tag{2.13}$$

$$\hat{x}_k = \hat{x}_{k-1} + K_k(y_k - C\hat{x}_{k-1}) \tag{2.14}$$

$$P_k = (I - K_k C)P_{k-1}(I - K_k C)^* + K_k R_k K_k^* \tag{2.15}$$

which is an extension of the well known finite dimensional system least square sequential estimation. In addition, we note that $\hat{x}_0 = E(x_0)$ and $P_0 = E[(x_0 - \hat{x}_0)(x_0 - \hat{x}_0)^*]$.

Proposition 1. *The realization of the recursive least-square estimator for Eq. (2.12) is given by Eqs. (2.13)-(2.15).*

A proof of Proposition 1 and the detailed derivation of Eqs. (2.13)-(2.15) is given in the following proof section.

Proof. The estimated state error mean propagation can be calculated as:

$$\begin{aligned}
&E\left[x\left(\zeta\right) - \hat{x}_k\left(\zeta\right)\right] \\
&= E\left[x - \hat{x}_k\right] \\
&= E\left[x - \hat{x}_{k-1} - K_k(y_k - C\hat{x}_{k-1})\right] \\
&= E\left[x - \hat{x}_{k-1} - K_k(Cx + v_k - C\hat{x}_{k-1})\right] \\
&= E\left[(I - K_k C)\left(x_k - \hat{x}_{k-1}\right) - K_k v_k\right] \\
&= (I - K_k C)\, E\left(x - \hat{x}_{k-1}\right) - K_k E(v_k)
\end{aligned} \tag{2.16}$$

It is obvious to know that for a given initial state estimation $\hat{x}_0 = E(x)$ one can lead to unbiased estimator since $E(v_k) = 0$. Then, the estimation-error covariance propagation can be computed as:

$$\begin{aligned}
P_k &= E\left[\left(x - \hat{x}_k\right)\left(x - \hat{x}_k\right)^T\right] \\
&= E\left[\left[(I - K_k C)\, E\left(x - \hat{x}_{k-1}\right) - K_k E\left(v_k\right)\right]\left[(I - K_k C)\, E\left(x - \hat{x}_{k-1}\right) - K_k E\left(v_k\right)\right]^T\right] \\
&= (I - K_k C)\, E\left[\left(x - \hat{x}_{k-1}\right)\left(x - \hat{x}_{k-1}\right)^T\right](I - K_k C)^* + K_k E\left(v_k {v_k}^T\right){K_k}^* \\
&= (I - K_k C)\, P_{k-1}(I - K_k C)^* + K_k R_k {K_k}^*
\end{aligned} \tag{2.17}$$

In order to determine the gain matrix K_k, one can easily follow recursive least square estimation derivation procedures of finite-dimensional systems (with $J_k = Tr(P_k)$) and obtain:

$$\frac{\partial J_k}{\partial K_k} = \frac{\partial Tr(P_k)}{\partial K_k} = 2\left(I - K_k C\right) P_{k-1}\left(-C^*\right) + 2K_k R_k = 0 \tag{2.18}$$

$$K_k = P_{k-1}C^*(CP_{k-1}C^* + R_k)^{-1} = P_kC^*{R_k}^{-1} \tag{2.19}$$

where J_k is the cost function to be minimized in Eq. (2.18) and with the covariance being defined as $P_k = E\left[(x(\zeta) - \hat{x}_k(\zeta))(x(\eta) - \hat{x}_k(\eta))^T\right]$, with understanding that covariance is given as a nonnegative symmetric operator $P_k(\zeta, \eta)$ which is property preserved by Eq. (2.15). ■

In addition, one needs to resolve how to realize Eq. (2.13), since output operators C and C^* are present and two dimensional covariance P_k is applied in the case of estimation of one dimensional distributed parameter state. An example of a simple transport-reaction system is given in the numerical simulations section to demonstrate aforementioned points and it will be clear that least square estimator fails to estimate the state throughout domain since it does not utilize any distributed parameter system model features. The way to improve to some extent the least square estimation is to enlarge the measurement space by placing more measurements along domain and to extend the space dimensionality of the $y_k \in Y^m$, ($m = 1, \cdots, M_m$ - number of measurements, with $M_m < \infty$). However, there is obviously a limited interest in this, since least square estimation can address the PDE state reconstruction only in approximate sense and under the condition that a large number of measurements is applied.

2.4 Observer design for linear transport-reaction systems

2.4.1 Continuous-time observer design

In order to address the DPS state reconstruction, one can consider the case of an appealing and practically realizable observer design for a transport linear DPS. In particular, this design considers the following representation, without the applied input present in the systems output (i.e., $\mathcal{D}$) given as Eqs. (2.1)-(2.2):

$$\dot{x}(t) = \mathcal{A}x(t) + \mathcal{B}u(t) \tag{2.20}$$

$$y(t) = Cx(t) \tag{2.21}$$

The observer design takes the standard form given as:

$$\dot{z}(t) = \mathcal{A}z(t) + \mathcal{B}u(t) + L(y(t) - Cz(t)) \tag{2.22}$$

such that one can easily define the error as $e(t) = x(t) - z(t)$, and therefore obtain:

$$\dot{e}(t) = (\mathcal{A} - LC)e(t) = A_0 e(t) \tag{2.23}$$

Hence, the design of the observer in the continuous setting is given as the design of an appropriate spatial $L(\zeta)$-operator, and in this case the operator is sought as a spatial function, such that the operator $A_0 = \mathcal{A} - LC$ is stable with desired stability properties. Obviously, this is not a trivial task since the PDE spatial operator $\mathcal{A}$ is characterized by infinite dimensional features and/or different stability characteristics than one used in the finite dimensional theory. Therefore, in order to obtain the desired error decay, we look into the following Lyapunov type of argument that can be applied in the design of the $L(\zeta)$ operator. On one side, this is important design characteristics since we don't apply any type of approximations or lumping of the DPS model utilized for the construction of the $L(\zeta)$, while on the other side the proposed method applies successfully to the spectral and/or non-spectral types of the distributed parameter systems by which accounts for majority of relevant chemical transport-reaction systems.

In order to provide the general theorem that accounts for the observer design, we introduce a following Lemma that embodies the exponential stability of the error dynamics $A_0 = \mathcal{A} - LC$.

Lemma 1. *Let A_0 be an infinitesimal generator of the C_0-semigroup $\mathcal{T}_{\mathcal{A}_o}(t)$ on $\mathcal{X}$ and M be a positive operator on $\mathcal{L}(\mathcal{X})$. Then, $\mathcal{T}_{\mathcal{A}_o}(t)$ is exponentially stable if and only if there exists a nonnegative self-adjoint operator $Q_0 \in \mathcal{L}(\mathcal{X})$ as a solution of the following operator Lyapunov equation such that*

$$\langle Q_0 x, A_0^* x\rangle + \langle A_0^* x, Q_0 x\rangle = -\langle Mx, x\rangle, \text{ for all } x \in D(A_0^*) \tag{2.24}$$

holds and $Q_0(D(A_0^*)) \subset D(A_0)$.

It can be demonstrated that the above equation Eq. (2.24) can be rewritten in the following form:

$$(A_0 Q_0 + Q_0 A_0^* + M)x = 0, \text{ for all } x \in D(A_0^*) \tag{2.25}$$

which is a standard operator Lyapunov equation which given a positive M yields a nonnegative self-adjoint bounded operator Q_0 (that is $\langle x, Q_0 y\rangle = \langle Q_0 x, y\rangle$ for any two

functions $x, y \in \mathcal{X}$) which maps from $D(A_0^*)$ to $D(A_0)$. Based on this Lemma 1, we can provide a description of a general approach of finding observer stabilizing gain $L(\zeta)$ in Eq. (2.23).

Theorem 1. $(\mathcal{A}, C)$ *is exponentially detectable if and only if there exists a nonnegative self-adjoint operator* Q_0 *as a solution to the following operator Riccati algebraic equation:*

$$\mathcal{A}Q_0 + Q_0\mathcal{A}^* + M - 2Q_0C^*CQ_0 = 0, \text{ on } D(\mathcal{A}^*) \tag{2.26}$$

with $Q_0(D(\mathcal{A}^*)) \subset D(\mathcal{A})$*, where* M *is a positive definite design parameter, then the observer gain* $L = Q_0C^*$ *is an exponentially stabilizing output injection gain in* $A_0 = \mathcal{A} - LC$*.*

Proof. In order to demonstrate the stabilizing properties of the obtained observer gain L, one can easily link Eq. (2.24) to Eq. (2.26) by considering:

$$\langle Q_0x, A_0^*x\rangle + \langle A_0^*x, Q_0x\rangle = -\langle Mx, x\rangle \text{ for all } x \in D(A_0^*) \tag{2.27}$$

so that Eq. (2.27) can be easily written in the following inner product form and further transformed as follows:

$$\begin{aligned}
\langle Q_0x, A_0^*x\rangle + \langle A_0^*x, Q_0x\rangle &= -\langle Mx, x\rangle \\
\langle A_0Q_0x, x\rangle + \langle Q_0A_0^*x, x\rangle &= -\langle Mx, x\rangle \\
\langle (\mathcal{A} - LC)Q_0x, x\rangle + \langle Q_0(\mathcal{A} - LC)^*x, x\rangle &= -\langle Mx, x\rangle \\
\langle ((\mathcal{A} - LC)Q_0 + Q_0(\mathcal{A} - LC)^*)x, x\rangle &= -\langle Mx, x\rangle \\
\langle ((\mathcal{A} - Q_0C^*C)Q_0 + Q_0(\mathcal{A} - Q_0C^*C)^*)x, x\rangle &= -\langle Mx, x\rangle \\
\langle (\mathcal{A}Q_0 + Q_0\mathcal{A}^* - Q_0C^*CQ_0 - Q_0C^*CQ_0)x, x\rangle &= -\langle Mx, x\rangle \\
\langle (\mathcal{A}Q_0 + Q_0\mathcal{A}^* - 2Q_0C^*CQ_0 + M)x, x\rangle &= 0
\end{aligned} \tag{2.28}$$

since x is not equal to zero, one obtains Eq. (2.26). Therefore, both equations are satisfied with unique operator Q_0 which implies exponentially stability of $A_0 = \mathcal{A} - LC$.

■

A similar theorem for the control problem is introduced in [102] with the design requirement to find stabilizing gain $K = -\mathcal{B}^*Q_0$ which generates an exponentially stable C_0-semigroup by $(\mathcal{A} + \mathcal{B}K)$ associated with Eq. (2.20).

In order to demonstrate the above design, two most prominent classes of transport-reaction models considered in the numerical simulations section are given by scalar first order hyperbolic and second order parabolic PDE as a majority of chemical and materials process control application are accurately enough described and/or can be approximated with these cases of distributed parameter systems.

Remark 2. *A considered example of transport-reaction linear parabolic PDE in numerical simulations section is given as a model of conductive heat transport in the catalytic metal bar. An identical model as the one given in the Eq. (2.60) in terms of its characteristics, can be considered when it comes to the transport-reaction system modelled by the diffusion process and linear chemical kinetics. Hence, the most important example of the transport-reaction system given by the axial dispersion reactor with given model dynamics as* $x_t = Dx_{\zeta\zeta} - vx_\zeta + kx$*, with Danckwerts boundary conditions can be easily transformed to the parabolic PDE structure given by Eq. (2.60), by applying state transformation* $x(\zeta, t) \to e^{-\frac{v}{2D}\zeta}\bar{x}(\zeta, t)$*. Alternatively, one can directly apply the above design and analysis by considering the convection-diffusion-reaction model belonging to the class of Sturm-Liouville systems with a well defined operator characterized by the point spectrum and the associated eigenfunctions [103, 104, 105]. In other words, this analysis and the design can be applied to other types of distributed parameter systems with a spectral operator of the Riesz type, such as string and/or beam equations in petroleum, structural and mechanical engineering, respectively [4].*

2.4.2 Discrete-time observer design

In a similar manner, as described above, a practical, appealing and easy to realize discrete Luenberger observer design is motivated by the notion that all modern design methodologies in practice assume discrete systems realization and design. In particular, we design an observer based on the following DPS plant model given by Eqs. (2.7)-(2.8):

$$x_k = \mathcal{A}_d x_{k-1} + \mathcal{B}_d u_k \tag{2.29}$$

$$y_k = \mathcal{C}_d x_{k-1} + \mathcal{D}_d u_k \tag{2.30}$$

Then, a discrete Luenberger observer is constructed as follows:

$$\hat{x}_k = \mathcal{A}_d \hat{x}_{k-1} + \mathcal{B}_d u_k + L_d(y_k - C_d \hat{x}_{k-1} - \mathcal{D}_d u_k) \tag{2.31}$$

As the observer design requires us to find L_d such that the error dynamics is given by:

$$e_k = (\mathcal{A}_d - L_d C_d) e_{k-1} = \tilde{\mathcal{A}}_d e_{k-1} \tag{2.32}$$

where $\tilde{\mathcal{A}}_d$ is the operator with the desired decay rate obtained by calculating $L_d(\zeta)$ operator.

However, discrete operators cannot be directly used in calculating the observer gain operator $L_d = Q_d C_d^*$, and therefore one needs to solve the corresponding continuous Lyapunov equation in order to obtain a stabilizing observer gain. In this case, the discrete Lyapunov equation is given as:

$$\langle x, [\tilde{\mathcal{A}}_d^* Q_d \tilde{\mathcal{A}}_d - Q_d] x \rangle = -\langle x, [C_d^* N C_d] x \rangle \tag{2.33}$$

and can be solved by finding a solution to the corresponding continuous Lyapunov equation given as:

$$\langle A_{d0} x, Q_d x \rangle + \langle Q_d x, A_{d0} x \rangle = -\langle Cx, NCx \rangle, \quad x \in D(A_0^*) \tag{2.34}$$

where $A_{d0} = A_0^*$ and N is nonnegative. Let us demonstrate the link between Eq. (2.33) and Eq. (2.34) by considering $\tilde{\mathcal{A}}_d := -I + 2\delta(\delta I - A_{d0})^{-1}$, $\tilde{\mathcal{A}}_d^* := \left[-I + 2\delta(\delta I - A_{d0})^{-1}\right]^*$ and $C_d = \sqrt{2\delta} C \left[\delta I - A_{d0}\right]^{-1}$. Therefore, we have

$$\left[-I + 2\delta(\delta I - A_{d0})^{-1}\right]^* Q_d \left[-I + 2\delta(\delta I - A_{d0})^{-1}\right] - Q_d = -C_d^* N C_d \tag{2.35}$$

by factoring out $(\delta I - A_{d0})^{-1}$, one obtains:

$$\left[\left[-I + 2\delta(\delta I - A_{d0})^{-1}\right]^* Q_d \left[2\delta - (\delta I - A_{d0})\right] - Q_d(\delta I - A_{d0})\right](\delta I - A_{d0})^{-1} = -C_d^* N C_d$$

and in the same way,

$$(\delta I - A_{d0})^{-1*}\left[\left[-(\delta I - A_{d0}) + 2\delta\right]^* Q_d \left[-(\delta I - A_{d0}) + 2\delta\right] - (\delta I - A_{d0})^* Q_d (\delta I - A_{d0})\right] \times (\delta I - A_{d0})^{-1} = -C_d^* N C_d$$

which leads to

$$\begin{aligned}&(\delta I - A_{d0})^{-1*}\left[2A_{d0}^{*}Q_d\delta + 2\delta Q_d A_{d0}\right](\delta I - A_{d0})^{-1}\\&= -\left(\sqrt{2\delta}C\left[\delta I - A_{d0}\right]^{-1}\right)^{*} N\left(\sqrt{2\delta}C\left[\delta I - A_{d0}\right]^{-1}\right)\end{aligned}$$

or

$$\begin{aligned}&(\delta I - A_{d0})^{-1*}\left[2A_{d0}^{*}Q_d\delta + 2\delta Q_d A_{d0}\right](\delta I - A_{d0})^{-1}\\&= -(\sqrt{2\delta})^{2}(\delta I - A_{d0})^{-1*}C^{*}NC(\delta I - A_{d0})^{-1}\end{aligned}$$

hence,

$$A_{d0}^{*}Q_d + Q_d A_{d0} = -C^{*}NC \tag{2.36}$$

which clearly becomes operator Lyapunov equation Eq. (2.34). Therefore, given that the solution of Eq. (2.34) exist and can be found, it is also the solution of the corresponding discrete Lyapunov equation Eq. (2.33), and converse is also true. Further, in order to solve for the operator Q_d, we invoke the Theorem 1. The obtained Q_d operator is the solution of the following operator Riccati equation given by:

$$A_0 Q_d + Q_d A_0^{*} - 2Q_d C^{*}NCQ_d + M = 0 \tag{2.37}$$

which is identical in the setting to the problem already presented and resolved in Eq. (2.26). In general, the issue of necessity to solve for the stabilizing operator Q_d by considering continuous operator Lyapunov or Riccati equation is due to the following facts. Eq. (2.33) does not need to satisfy any conditions when it comes to the integral operators and hence boundary conditions are not required, this is in contrast to the presented continuous operator equations which are always defined with the space of functions and corresponding domain. Along the same line, on the other hand the setting of the expression in Eq. (2.33), $\tilde{\mathcal{A}}_d^{*}Q_d\tilde{\mathcal{A}}_d\psi - Q_d\psi = -C_d^{*}NC_d\psi$ induces the presence of an arbitrary function ψ so that Q_d operator can not be uniquely determined ($\tilde{\mathcal{A}}_d\psi$ and $C_d\psi$ are well defined as integral operators applied to function ψ). As it will be demonstrated in the numerical simulations section the operator Q_d is often found as a solution of boundary value problem obtained as the outcome of constructing a continuous operator equations. Finally, the novelty

associated with discrete Luenberger observer design is that presented methodology does not account for the spatial approximations and/or model reduction, for example by mere discretization of the underlying spatial operators which is a usual procedure in the literature [30, 90, 91, 92, 79].

2.5 Kalman filter design for linear transport-reaction systems

2.5.1 Discrete-time Kalman filter design

Along the line of discrete Kalman filter design for lumped parameter systems, we explore the Kalman filter design for linear distributed parameter systems. Initially, we provide a design procedure for discrete linear infinite-dimensional systems by Cayley-Tustin transformation and case studies of typical linear transport-reaction systems described by hyperbolic and parabolic PDEs are presented. A general linear discrete-time transport-reaction system obtained by Cayley-Tustin time discretization is considered in the following form:

$$\begin{aligned} x_k &= \mathcal{A}_d x_{k-1} + \mathcal{B}_d u_k + w_{k-1} \\ y_k &= \mathcal{C}_d x_{k-1} + \mathcal{D}_d u_k + v_k \end{aligned} \tag{2.38}$$

where the operators $(\mathcal{A}_d, \mathcal{B}_d, \mathcal{C}_d, \mathcal{D}_d)$ are well defined in the rigged spaces $\mathcal{X}_k$, (for $k = 1, 2, \cdots$) constructed from $\mathcal{X}$ by means of the operator $\mathcal{A}$ given in Eq. (2.1), that is $\mathcal{X}_0 = \mathcal{X}$, and $\mathcal{X}_n = \mathcal{D}(\mathcal{A}^k)$ (see 3.6 in [21]). The above equations are solvable since there are no problems induced by unbounded operators, so that the input map is $\mathcal{B}_d : L^2(\mathbb{Z}; U) \to \mathcal{X}$ by control input sequence $\{u_k\}_{k\in\mathbb{Z}^+}$ and augmented by noise signal w_k, the output space $y_k \in R^n = Y$ with n-denoting number of measured outputs is given by $\mathcal{C}_d \& \mathcal{D}_d : \mathcal{X} \times U \to Y$ and is augmented by the measurement noise signal v_k. The process noise w_k is the zero mean multivariate normal distribution with covariance Q_k, and v_k is the measurement noise of having zero mean Gaussian white noise with covariance R_k given as, $w_k \sim \mathcal{N}(0, Q_k)$, $v_k \sim \mathcal{N}(0, R_k)$, $E[w_k w_j^T] = Q_k \delta_{k,j}$, $E[v_k v_j^T] = R_k \delta_{k,j}$, $E[v_k w_j^T] = 0$. The process noise is present in Eq. (2.38) without any spatial characteristics since $w_k = I w_k$, however, one can appropriately model spatial distribution of the noise if there is a prior knowledge of the noise source. We emphasize that in the modelling of discrete infinite dimensional system given by Eq.

(2.38) we augment the state and output noise with the appropriate noise signals in affine manner and our approach is not to define the noise signals in the continuous setting and then to transform to the discrete one by Cayley-Tustin discretization. The compelling reason is that area of stochastic partial differential equations requires completely different functional space setting and different operator definitions which is out of scope of this chapter.

Therefore, having the Eq. (2.38), a priori state estimate is given as $\hat{x}_k^- = E\{x_k|y_1, y_2, \cdots, y_{k-1}\}$ with the corresponding priori estimate covariance $P_k^- = E[(x_k - \hat{x}_k^-)(x_k - \hat{x}_k^-)^*]$. Similarly, a posteriori state estimate is $\hat{x}_k^+ = E\{x_k|y_1, y_2, \cdots, y_k\}$ with the corresponding posteriori estimate covariance $P_k^+ = E[(x_k - \hat{x}_k^+)(x_k - \hat{x}_k^+)^*]$.

In order to guarantee the consistency in the time instants of the discrete transport-reaction system and the standard discrete Kalman filter structure of finite-dimensional systems, one can express y_k in terms of the current state x_k instead of the previous state instance x_{k-1} in Eq. (2.4) and Eq. (2.8), which leads to the following:

$$\begin{aligned} x_k &= \mathcal{A}_d x_{k-1} + \mathcal{B}_d u_k + w_{k-1} \\ y_k &= \bar{C}_d x_k + \bar{\mathcal{D}}_d u_k + v_k \end{aligned} \tag{2.39}$$

where the updated discrete-time spatial operators are denoted as follows:

$$\begin{bmatrix} \mathcal{A}_d & \mathcal{B}_d \\ \bar{C}_d & \bar{\mathcal{D}}_d \end{bmatrix} = \begin{bmatrix} -I + 2\delta\mathcal{R}(\zeta,\delta) & \sqrt{2}\delta\mathcal{R}(\zeta,\delta)B \\ -\sqrt{2}\delta C\mathcal{R}(\zeta,-\delta) & \mathcal{G}(-\delta) \end{bmatrix} \tag{2.40}$$

As it is shown in the above expressions for $\bar{C}_d$ and $\bar{\mathcal{D}}_d$, a resolvent and transfer function of continuous system are evaluated at $s = -\delta$ instead of at $s = \delta$.

Proposition 2. *The noise free discrete systems given by the* $\mathcal{S}_d := \begin{bmatrix} \mathcal{A}_d & \mathcal{B}_d \\ C_d & \mathcal{D}_d \end{bmatrix}$ *and the discrete system given by* $\begin{bmatrix} \mathcal{A}_d & \mathcal{B}_d \\ \bar{C}_d & \bar{\mathcal{D}}_d \end{bmatrix}$ *in Eq. (2.9) have the same transfer function and input/output representation properties.*

Proof. The proof consists of two steps.

(1) State Space Formulation:

From Eqs. (2.7)-(2.8), one has the following:

$$x_k = (\delta I - \mathcal{A})^{-1}(\delta I + \mathcal{A})x_{k-1} + \sqrt{2\delta}(\delta I - \mathcal{A})^{-1}\mathcal{B}u_k$$

It is simple to manipulate above expression as follows:

$$\begin{aligned} x_{k-1} &= \left[(\delta I - \mathcal{A})^{-1}(\delta I + \mathcal{A})\right]^{-1} x_k - \left[(\delta I - \mathcal{A})^{-1}(\delta I + \mathcal{A})\right]^{-1} \sqrt{2\delta}(\delta I - \mathcal{A})^{-1}\mathcal{B}u_k \\ &= (\delta I + \mathcal{A})^{-1}(\delta I - \mathcal{A})x_k - \sqrt{2\delta}(\delta I + \mathcal{A})^{-1}\mathcal{B}u_k \end{aligned}$$

Then, one can substitute this above expression into the output expression Eq. (2.4) and obtain the following

$$\begin{aligned} y_k &= \sqrt{\Delta t}C\frac{1}{2}\left[I + (\delta I + \mathcal{A})^{-1}(\delta I - \mathcal{A})\right] x_k - \sqrt{\Delta t}C\frac{1}{2}\sqrt{2\delta}(\delta I + \mathcal{A})^{-1}\mathcal{B}u_k \\ &= \sqrt{\Delta t}C\frac{1}{2}\left(-2\delta\right)(-\delta I - \mathcal{A})^{-1}x_k + \sqrt{\Delta t}C\frac{1}{2}\sqrt{2\delta}(-\delta I - \mathcal{A})^{-1}\mathcal{B}u_k \\ &= -\sqrt{2\delta}C(-\delta I - \mathcal{A})^{-1}x_k + C(-\delta I - \mathcal{A})^{-1}\mathcal{B}u_k \end{aligned}$$

Hence, the updated discrete operators of $\bar{\mathcal{C}}_d$ and $\bar{\mathcal{D}}_d$ are given by

$$\begin{aligned} \bar{\mathcal{C}}_d &= -\sqrt{2\delta}C(-\delta I - \mathcal{A})^{-1} = -\sqrt{2}\delta C\mathcal{R}(\zeta, -\delta) \\ \bar{\mathcal{D}}_d &= C(-\delta I - \mathcal{A})^{-1}\mathcal{B} = C\mathcal{R}(\zeta, -\delta)\mathcal{B} = \mathcal{G}(-\delta) \end{aligned}$$

(2) Transfer Function Comparison

In order to obtain the discrete transfer function, the $\mathcal{Z}$ transformation is applied to Eq. (2.7) which yields the following:

$$\begin{aligned} x(z) =& \left[I - (\delta I - \mathcal{A})^{-1}(\delta I + \mathcal{A})z^{-1}\right]^{-1}\sqrt{2\delta}(\delta I - \mathcal{A})^{-1}\mathcal{B}u(z) \\ y(z) =& \sqrt{2\delta}C(\delta I - \mathcal{A})^{-1}z^{-1}\left[I - (\delta I - \mathcal{A})^{-1}(\delta I + \mathcal{A})z^{-1}\right]^{-1}\sqrt{2\delta}(\delta I - \mathcal{A})^{-1}\mathcal{B}u(z) \\ &+ C(\delta I - \mathcal{A})^{-1}\mathcal{B}u(z) \end{aligned}$$

Then, one can further derive as follows:

$$\begin{aligned} y(z) =& \sqrt{2\delta}C(\delta I - \mathcal{A})^{-1}z^{-1}\left[I - (\delta I - \mathcal{A})^{-1}(\delta I + \mathcal{A})z^{-1}\right]^{-1}\sqrt{2\delta}(\delta I - \mathcal{A})^{-1}\mathcal{B}u(z) \\ &+ C(\delta I - \mathcal{A})^{-1}\mathcal{B}u(z) \\ =& 2\delta Cz^{-1}(\delta I - \mathcal{A})^{-1}\left[I - (\delta I - \mathcal{A})^{-1}(\delta I + \mathcal{A})z^{-1}\right]^{-1}(\delta I - \mathcal{A})^{-1}\mathcal{B}u(z) \\ &+ C(\delta I - \mathcal{A})^{-1}\mathcal{B}u(z) \\ =& C\left[2\delta z^{-1}\left[(\delta I - \mathcal{A})^2 - (\delta I + \mathcal{A})(\delta I - \mathcal{A})z^{-1}\right]^{-1} + (\delta I - \mathcal{A})^{-1}\right]\mathcal{B}u(z) \\ =& G(z)u(z) \end{aligned}$$

Applying $\mathcal{Z}$ transformation to Eq. (2.39) leads to:

$$\begin{aligned} y(z) =& -\sqrt{2\delta}C(-\delta I-\mathcal{A})^{-1}\left[I-(\delta I-\mathcal{A})^{-1}(\delta I+\mathcal{A})z^{-1}\right]^{-1}\sqrt{2\delta}(\delta I-\mathcal{A})^{-1}\mathcal{B}u(z) \\ &+C(-\delta I-\mathcal{A})^{-1}\mathcal{B}u(z) \end{aligned}$$

Then, one can simplify the above transfer function as follows:

$$\begin{aligned} y(z) =& -\sqrt{2\delta}C(-\delta I-\mathcal{A})^{-1}\left[I-(\delta I-\mathcal{A})^{-1}(\delta I+\mathcal{A})z^{-1}\right]^{-1}\sqrt{2\delta}(\delta I-\mathcal{A})^{-1}\mathcal{B}u(z) \\ &+C(-\delta I-\mathcal{A})^{-1}\mathcal{B}u(z) \\ =&2\delta C(\delta I+\mathcal{A})^{-1}\left[I-(\delta I-\mathcal{A})^{-1}(\delta I+\mathcal{A})z^{-1}\right]^{-1}(\delta I-\mathcal{A})^{-1}\mathcal{B}u(z) \\ &-C(\delta I+\mathcal{A})^{-1}\mathcal{B}u(z)-C(\delta I-\mathcal{A})^{-1}\mathcal{B}u(z)+C(\delta I-\mathcal{A})^{-1}\mathcal{B}u(z) \\ =&C\left[2\delta(\delta I+\mathcal{A})^{-1}\left[I-(\delta I-\mathcal{A})^{-1}(\delta I+\mathcal{A})z^{-1}\right]^{-1}(\delta I-\mathcal{A})^{-1}\mathcal{B}\right]u(z) \\ &-C\left[(\delta I-\mathcal{A})^{-1}(\delta I-\mathcal{A})(\delta I+\mathcal{A})^{-1}+(\delta I-\mathcal{A})^{-1}(\delta I+\mathcal{A})(\delta I+\mathcal{A})^{-1}\right]\mathcal{B}u(z) \\ &+C(\delta I-\mathcal{A})^{-1}\mathcal{B}u(z) \\ =&2\delta C(\delta I+\mathcal{A})^{-1}\left[I-(\delta I-\mathcal{A})^{-1}(\delta I+\mathcal{A})z^{-1}\right]^{-1}(\delta I-\mathcal{A})^{-1}\mathcal{B}u(z) \\ &-2\delta C(\delta I-\mathcal{A})^{-1}(\delta I+\mathcal{A})^{-1}\mathcal{B}u(z)+C(\delta I-\mathcal{A})^{-1}\mathcal{B}u(z) \\ =&2\delta C\left[(\delta I-\mathcal{A})(\delta I+\mathcal{A})-(\delta I+\mathcal{A})(\delta I+\mathcal{A})z^{-1}\right]^{-1}\mathcal{B}u(z) \\ &-2\delta C[(\delta I+\mathcal{A})(\delta I-\mathcal{A})]^{-1}\mathcal{B}u(z)+C(\delta I-\mathcal{A})^{-1}\mathcal{B}u(z) \\ =&2\delta C[(\delta I+\mathcal{A})(\delta I-\mathcal{A})]^{-1}[(\delta I+\mathcal{A})(\delta I-\mathcal{A})] \\ &\times\left[(\delta I-\mathcal{A})(\delta I+\mathcal{A})-(\delta I+\mathcal{A})(\delta I+\mathcal{A})z^{-1}\right]^{-1}\mathcal{B}u(z)+C(\delta I-\mathcal{A})^{-1}\mathcal{B}u(z) \\ =&2\delta C[(\delta I+\mathcal{A})(\delta I-\mathcal{A})]^{-1}\left[(\delta I+\mathcal{A})(\delta I+\mathcal{A})z^{-1}\right] \\ &\times\left[(\delta I-\mathcal{A})(\delta I+\mathcal{A})-(\delta I+\mathcal{A})(\delta I+\mathcal{A})z^{-1}\right]^{-1}\mathcal{B}u(z)+C(\delta I-\mathcal{A})^{-1}\mathcal{B}u(z) \\ =&C\left[2\delta z^{-1}[(\delta I-\mathcal{A})]^{-1}\left[(\delta I-\mathcal{A})-(\delta I+\mathcal{A})z^{-1}\right]^{-1}+(\delta I-\mathcal{A})^{-1}\right]\mathcal{B}u(z) \\ =&C\left[2\delta z^{-1}\left[(\delta I-\mathcal{A})^{2}-(\delta I+\mathcal{A})(\delta I-\mathcal{A})z^{-1}\right]^{-1}+(\delta I-\mathcal{A})^{-1}\right]\mathcal{B}u(z) \\ =&G(z)u(z) \end{aligned}$$

Therefore, the updated Crank-Nicolson discretization setting described in Eq. (2.39) is identical with the original framework shown in Eq. (2.7). ■

The well known configuration for the Kalman filter can be written as a single equation [35], however it is most often realized as the calculation of two distinct computing steps: "Prediction" and "Update". The initial conditions are described as

below:

$$\begin{aligned} \hat{x}_0^+ &= E(x_0) = \hat{x}_0 \\ P_0^+ &= E[(x_0 - \hat{x}_0^+)(x_0 - \hat{x}_0^+)^*] = Q_0 \end{aligned} \tag{2.41}$$

We have the prediction step (also called the prior estimation step) as follows, with measurement up to time $k-1$:

$$\begin{aligned} P_k^- &= \mathcal{A}_d P_{k-1}^+ \mathcal{A}_d^* + Q_{k-1} = \mathcal{A}_d[\mathcal{A}_d P_{k-1}^+]^* + Q_{k-1} \\ \hat{x}_k^- &= \mathcal{A}_d \hat{x}_{k-1}^+ + \mathcal{B}_d u_k \end{aligned} \tag{2.42}$$

and the update step (also called the posterior estimation step) is given as follows, by using additional output measurement y_k at time instance k:

$$\begin{aligned} K_k =& P_k^- \bar{C}_d^* (\bar{C}_d P_k^- \bar{C}_d^* + R_k)^{-1} \\ =& (\bar{C}_d P_k^-)^* [\bar{C}_d (\bar{C}_d P_k^-)^* + R_k]^{-1} \\ P_k^+ =& (I - K_k \bar{C}_d) P_k^- (I - K_k \bar{C}_d)^* + K_k R_k K_k^* \\ =& I P_k^- I^* - K_k \bar{C}_d P_k^- I^* - I P_k^- \bar{C}_d^* K_k^* + K_k \bar{C}_d P_k^- \bar{C}_d^* K_k^* + K_k R_k K_k^* \\ =& I P_k^- I^* - K_k \bar{C}_d P_k^- I^* - I (\bar{C}_d P_k^-)^* K_k^* + K_k \bar{C}_d (\bar{C}_d P_k^-)^* K_k^* + K_k R_k K_k^* \\ =& [(P_k^-)^{-1} + \bar{C}_d^* R_k^{-1} \bar{C}_d]^{-1} \\ =& (I - K_k \bar{C}_d) P_k^- \\ \hat{x}_k^+ =& \hat{x}_k^- + K_k (y_k - \bar{C}_d \hat{x}_k^- - \bar{\mathcal{D}}_d u_k) \end{aligned} \tag{2.43}$$

The proposed framework follows and extends a standard finite-dimensional discrete-time Kalman filter design and realization, and the differences arise in the treatment of discrete spatial operators and consideration of the covariances which admit spatial characteristics. In particular, the priori estimate covariance P_k^- and the posteriori estimate covariance P_k^+ are two-dimensional and self-adjoint, such that $P_k^- = P_k^{-*} = P_k^{-*}(\zeta_1, \zeta_2)$ and $P_k^+ = P_k^{+*} = P_k^{+*}(\zeta_1, \zeta_2)$ provided by Kalman estimation evolution formulas Eqs. (2.42)-(2.43) preserving self-adjoint and positive definiteness due to positive definitness of Q_k and R_k. Finally, the Kalman filter stability is provided by $(C_d, \mathcal{A}_d)$ being detectable, $(\mathcal{A}_d, Q_k^{1/2})$ stabilizable and Q_0 being non-negative definite [9].

2.5.2 One-step discrete Kalman filter design

In addition, we explore a one-step Kalman filter with a single estimation step formulated by combining the prior estimation equation and posterior estimation equation.

In particular, one can manipulate the state and its corresponding covariance prior estimation Eqs. (2.42)-(2.43) as follows:

$$\begin{aligned}
\hat{x}_{k+1}^{-} &= \mathcal{A}_d \hat{x}_k^{-} + \mathcal{B}_d u_{k+1} + L_k \left(y_k - \bar{C}_d \hat{x}_k^{-} - \bar{\mathcal{D}}_d u_k \right) \\
P_{k+1}^{-} &= \mathcal{A}_d P_k^{-} \mathcal{A}_d{}^* - \mathcal{A}_d P_k^{-} C_d^* (C_d P_k^{-} C_d^* + R_k)^{-1} C_d^* P_k^{-} \mathcal{A}_d{}^* + Q_k \\
K_k &= P_k^{-} \bar{C}_d^* (\bar{C}_d P_k^{-} \bar{C}_d^* + R_k)^{-1}
\end{aligned} \tag{2.44}$$

where L_k serves as an one-step ahead Kalman filter prediction gain and its relationship with Kalman gain is given by $L_k = \mathcal{A}_d K_k = \mathcal{A}_d P_k^{-} \bar{C}_d^* (\bar{C}_d P_k^{-} \bar{C}_d^* + R_k)^{-1}$. This state estimation evolution expression is similar to the one provided by Astrom [96] in the finite dimensional state space setting.

Derivation of formula (2.44): By taking Eq. (2.43) into Eq. (2.42), one gets the following expressions:

$$\begin{aligned}
\hat{x}_{k+1}^{-} &= \mathcal{A}_d \hat{x}_k^{+} + \mathcal{B}_d u_{k+1} \\
&= \mathcal{A}_d \left[\hat{x}_k^{-} + K_k (y_k - \bar{C}_d \hat{x}_k^{-} - \bar{\mathcal{D}}_d u_k) \right] + \mathcal{B}_d u_{k+1} \\
&= \mathcal{A}_d \hat{x}_k^{-} + \mathcal{A}_d K_k (y_k - \bar{C}_d \hat{x}_k^{-} - \bar{\mathcal{D}}_d u_k) + \mathcal{B}_d u_{k+1} \\
&= \mathcal{A}_d \hat{x}_k^{-} + \mathcal{B}_d u_{k+1} + L_k \left(y_k - \bar{C}_d \hat{x}_k^{-} - \bar{\mathcal{D}}_d u_k \right)
\end{aligned}$$

The corresponding prior estimated covariance (also called steady-state discrete filtering Riccati matrix) is given as:

$$\begin{aligned}
P_{k+1}^{-} &= \mathcal{A}_d P_k^{+} \mathcal{A}_d{}^* + Q_k \\
&= \mathcal{A}_d \left(P_k^{-} - K_k \bar{C}_d^* P_k^{-} \right) \mathcal{A}_d{}^* + Q_k \\
&= \mathcal{A}_d \left(P_k^{-} - P_k^{-} \bar{C}_d^* (\bar{C}_d P_k^{-} C_d^* + R_k)^{-1} \bar{C}_d^* P_k^{-} \right) \mathcal{A}_d{}^* + Q_k \\
&= \mathcal{A}_d P_k^{-} \mathcal{A}_d{}^* - \mathcal{A}_d P_k^{-} \bar{C}_d^* (C_d P_k^{-} \bar{C}_d^* + R_k)^{-1} \bar{C}_d^* P_k^{-} \mathcal{A}_d{}^* + Q_k
\end{aligned}$$

Remark 3. *The difference between Kalman filter and Luenberger observer estimation is that the one-step ahead Kalman filter prediction gain is updated by recursive estimation of the time varying Riccati equation in terms of the state covariance matrix P_k given above, while the Luenberger observer gain is deterministic and given as constants determined by solving the operator Riccati equation.*

Similarly, one can easily obtain the posterior estimated state and its covariance formulas as follows:

$$\hat{x}_k^{+} = \left(I - K_k \bar{C}_d \right) \left(\mathcal{A}_d \hat{x}_{k-1}^{+} + \mathcal{B}_d u_k \right) + K_k (y_k - \bar{\mathcal{D}}_d u_k)$$

$$
\begin{aligned}
P_k^+ &= \left(I - K_k \bar{C}_d\right)\left(\mathcal{A}_d P_{k-1}^+ \mathcal{A}_d{}^* + Q_{k-1}\right) \\
K_k &= P_k^+ \bar{C}_d^* R_k{}^{-1}
\end{aligned} \tag{2.45}
$$

Remark 4. *In order to describe the state and measurement noise distributions, one can introduce bounded operators G_w and G_v accounting for spatial influence of state noise w_k and different channel measurement noises distributions v_k for a system at each time instance, which yields the following system:*

$$
\begin{aligned}
x_k &= \mathcal{A}_d x_{k-1} + \mathcal{B}_d u_k + G_w w_{k-1} \\
y_k &= \bar{C}_d x_{k-1} + \bar{\mathcal{D}}_d u_k + G_v v_k
\end{aligned} \tag{2.46}
$$

Following the one-step Kalman filter design procedures for distributed parameter systems, one can easily conclude the following posteriori Kalman filtering estimation formulas for the given more general infinite-dimensional system considering the noise distribution characteristics from the state and measurement:

$$
\begin{aligned}
\hat{x}_k^+ &= \left(I - K_k \bar{C}_d\right)\left(\mathcal{A}_d \hat{x}_{k-1}^+ + \mathcal{B}_d u_k\right) + K_k(y_k - \bar{\mathcal{D}}_d u_k) \\
P_k^+ &= \left(I - K_k \bar{C}_d\right)\left(\mathcal{A}_d P_{k-1}^+ \mathcal{A}_d{}^* + G_w Q_{k-1} G_w{}^*\right) \\
K_k &= P_k^+ \bar{C}_d^* (G_v R_k G_v{}^*)^{-1}
\end{aligned} \tag{2.47}
$$

2.6 Numerical simulations

In this section, we provide numerical examples associated with the observer and filter designs. First we consider an example of the least square estimator applied to the typical transport-reaction system characterized by the time invariant spatially nonuniform steady state.

2.6.1 Least square estimator

An example of a simple transport-reaction system is given as:

$$
x_t = -x_\zeta + x, \quad x(0,t) = \bar{x}_0 = 1, \quad x(\zeta, 0) = x_0(\zeta) \tag{2.48}
$$

$$
y(t) = x(l,t) + v(t) \tag{2.49}
$$

where $x_\zeta = \frac{\partial x}{\partial \zeta}$ and with Eq. (2.49) rewritten in the sampled data form given by Eq. (2.12), so that the steady-state solution takes the form of exponential function,

$\lim_{t\to\infty} x(\zeta,t) = x_{ss}(\zeta) = \bar{x}_0 e^{\zeta}$. In this case, the measurement is taken at the reactor's exit (e.g., $l = 1$) and therefore $Cx = \int_0^1 \delta(\zeta-1)x_{ss}(\zeta)d\zeta = x_{ss}(1)$. Since P_k is symmetric in Eq. (2.13), and is realized as $P_{k-1}C^* = (CP_{k-1})^* = \int_0^1 \delta(\zeta-1)P_{k-1}(\zeta,\eta)d\zeta$, so that $CP_{k-1}C^* = C(CP_{k-1})^*$ is meaningful. As it is expected, the least square estimator is capable of proper reconstruction of the state at the boundary of the domain (exit of the plug flow reactor, see Fig. 2.1), but it fails to estimate the state throughout the spatial domain. This is expected, since the least square estimator does not utilize any model of the underlying system dynamics and therefore cannot provide any reliable estimates of the distributed spatial state. In addition, single measurement available to the least square estimator is not sufficient to reconstruct an infinite dimensional spatially distributed state. Possible improvements in the time invariant spatial state estimation could be obtained with a large number of measurements applied along the state domain. This path, however, is not frequently feasible, as distributed parameter system plants are already built in the process industry with no structural feature to add a large number of measurements and change plant design accordingly. In practice, it is possible to add only boundary measurements to already existing plant structure. Hence, the least square state estimation of time invariant spatial state in the distributed parameter systems setting can be achieved only with a large number of point measurements available. Finally, there is obviously a limited interest in this, since least square estimation can address the PDE state reconstruction only in approximate sense and under the condition that a large number of measurements is applied.

2.6.2 Observer design

Example 1: Transport-reaction system modelled by 1^{st} order hyperbolic PDE:
A simple model of the plug flow reactor with constant transport velocity v and spatial function ψ associated with linearized kinetics of the chemical reaction along the reactor length is considered:

$$\begin{aligned} x_t &= -vx_\zeta + \psi x + \frac{1}{2\epsilon}1_{[0-\epsilon,0+\epsilon]}u(t), \quad (\zeta,t) \in [0,1]\times\mathbb{R}^+ \\ x(0,t) &= 0, \quad x(\zeta,0) = f(\zeta) \\ y(t) &= Cx \end{aligned} \tag{2.50}$$

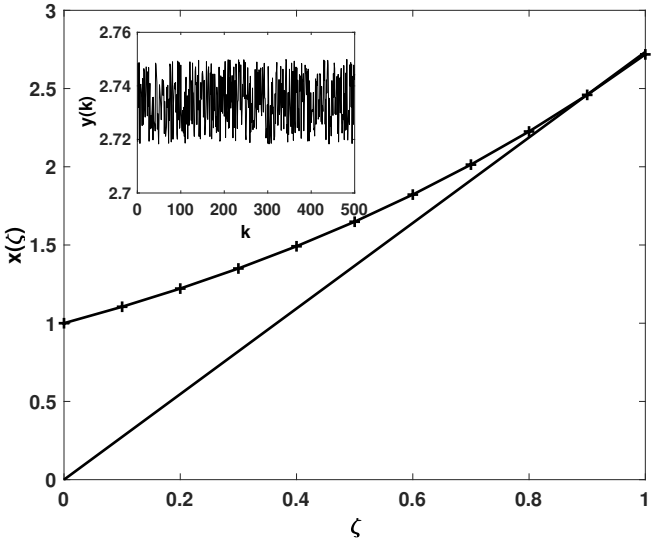


Figure 2.1: Estimated state of the hyperbolic system (solid line), and the steady state $x_{ss} = \bar{x}_0 e^{\zeta}$ given in Eq. (2.48) (line "−+−"), with v_k being a random variable with variance $R_k = Var\{v_k\} = 0.01$ and zero mean, and y_k given by Eq. (2.49) taken at the reactor's exit $\zeta = l$.

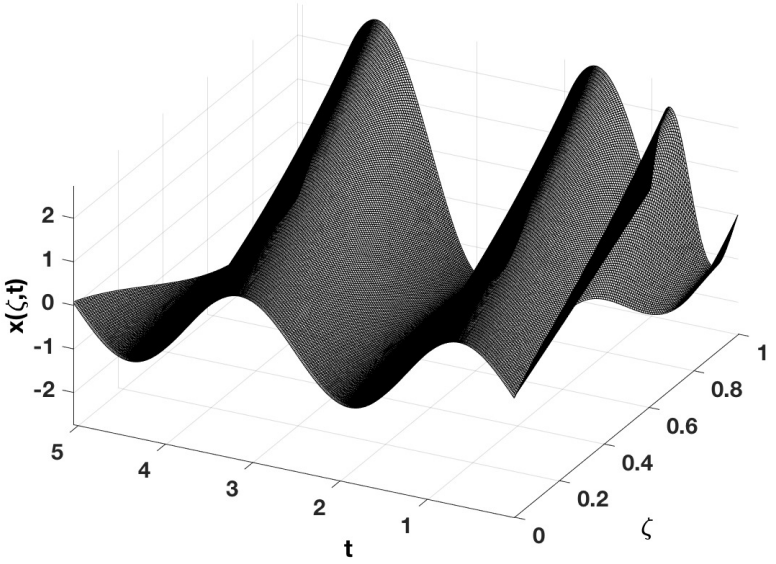


Figure 2.2: State of the hyperbolic system $x(\zeta, t)$ given in Eq. (2.50).

where $x(\cdot, t) \in L^2((0,1), \mathbb{R}^+)$ is a state. The modelling assumption is that measurements are specified with the operator $\mathcal{C}$, while input operator $\mathcal{B}$ approximates boundary actuation with the actuation distribution function $b(\zeta) = \frac{1}{2\epsilon} 1_{[0-\epsilon, 0+\epsilon]}$ which is bounded. In order to apply the above design, one defines the operator $\mathcal{A} = -v\frac{d}{d\zeta} + \psi$

with a domain $\mathcal{D}(\mathcal{A}) = \{\phi(\zeta) \in L^2(0,1), |\phi \text{ is abs. cont.}, \frac{d\phi}{d\zeta} \in L^2(0,1), \phi(0) = 0\}$, and the associated adjoint operator $\mathcal{A}^* = v\frac{d}{d\zeta} + \psi$ with its domain as $\mathcal{D}(\mathcal{A}^*) = \{\phi(\zeta) \in L^2(0,1), |\phi \text{ is abs. cont.}, \frac{d\phi}{d\zeta} \in L^2(0,1), \phi(1) = 0\}$, where "absolutely continuous" is shortly written as "abs. cont." for simplicity. Then, design implies that the expression in Eq. (2.26) is applied, with a free choice of the design parameter $M(\zeta)$ and arbitrary function $\Psi(\zeta)$ in a domain of $\mathcal{D}(\mathcal{A}^*)$, which leads to the following:

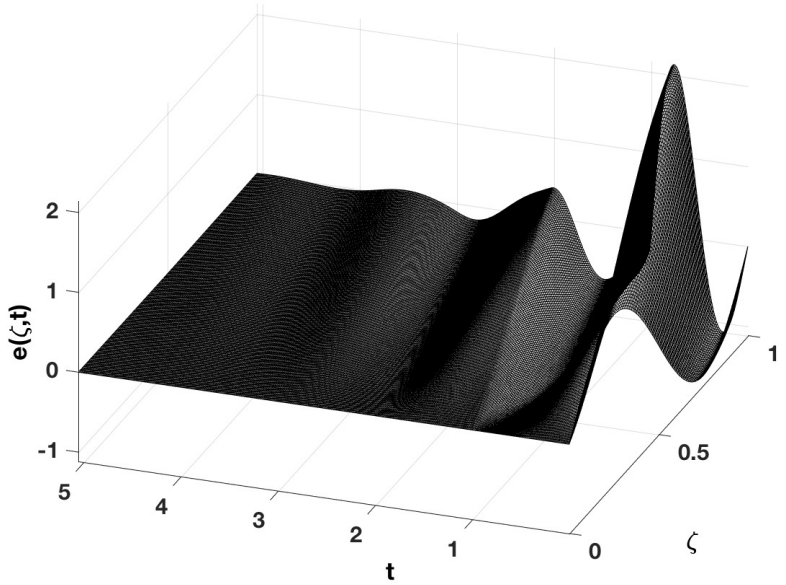


Figure 2.3: Evolution of the observer error $e(\zeta, t)$ given in Eq. (2.23).

$$\mathcal{A}Q_0 + Q_0\mathcal{A}^* + M - 2Q_0C^*CQ_0 = 0, \text{ on } D(\mathcal{A}^*) \tag{2.51}$$

$$\begin{aligned} -v\frac{d}{d\zeta}\left[Q_0\Psi(\zeta)\right] + \psi Q_0\Psi(\zeta) + Q_0 v\frac{d}{d\zeta}\left[\Psi(\zeta)\right] + \psi Q_0\Psi(\zeta) + M\Psi(\zeta) \\ - 2Q_0C^*CQ_0\Psi(\zeta) = 0 \end{aligned} \tag{2.52}$$

$$\begin{aligned} -v\frac{dQ_0}{d\zeta}\Psi(\zeta) - vQ_0\frac{d\Psi(\zeta)}{d\zeta} + Q_0 v\frac{d\Psi(\zeta)}{d\zeta} + 2\psi Q_0\Psi(\zeta) + M\Psi(\zeta) \\ - 2Q_0C^*CQ_0\Psi(\zeta) = 0 \end{aligned} \tag{2.53}$$

where it should be noted that the operator Q_0 is given by a spatial function (also denoted by Q_0 for the ease of notation) multiplied by an identity operator, see [50, The. 5]. Then we can further derive the following

$$\frac{dQ_0}{d\zeta} = \frac{1}{v}\left(2\psi Q_0 - 2Q_0C^*CQ_0 + M\right), \quad Q_0(0) = 0 \tag{2.54}$$

Based on the understanding that the measurement operator C can be defined for different scenarios, we consider several cases of possible solutions to the observer gain $L(\zeta)$.

For example, one can look at the trivial case of having identity operator $C = I$ which implies that the entire state is measured and in the trivial case $C^*C = I$, which leads to the following:

$$\frac{dQ_0}{d\zeta} = \frac{1}{v}\left(-2Q_0^2 + M + 2\psi Q_0\right), \quad Q_0(0) = 0 \tag{2.55}$$

Therefore, one needs to choose $Q_0(0) = 0$ and positive spatial function M to ensure semi-positive definiteness of the function Q_0 and the observer gain is $L = Q_0(\zeta)C^*$. On the other hand, in the case that the output measurement is taken at the exit of the reactor, that is $(y(t) = x(1,t))$, and the operator CQ_0 is defined as $CQ_0 = \int_0^1 \delta(\zeta - 1)Q_0(\zeta)d\zeta = Q_0(1)$. Furthermore, the operator C^* is given as $C^*(\cdot) = \delta(\zeta - 1)\int_0^1 (\cdot)d\eta$. In addition, we provide a quick explanation of the C^* operator features by forming the inner product with any two functions x, y, so that:

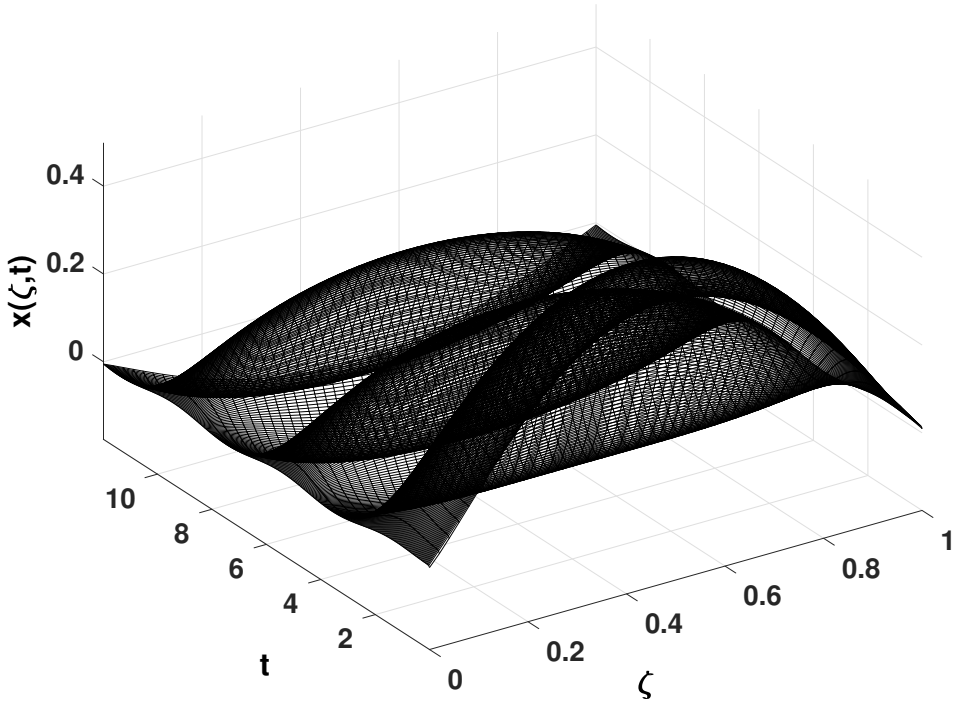


Figure 2.4: State of the parabolic PDE system $x(\zeta, t)$ given in Eq. (2.60).

$$\langle Cx, y\rangle = \int_0^1 \Big[\int_0^1 \delta(\zeta - 1)x(\zeta)d\zeta\Big]y(\eta)d\eta \tag{2.56}$$

$$
\begin{aligned}
&= \int_0^1 [\int_0^1 \delta(\zeta-1)x(\zeta)y(\eta)d\zeta]d\eta \\
&= \int_0^1 x(\zeta)[\delta(\zeta-1)[\int_0^1 y(\eta)d\eta]]d\zeta \\
&= \langle x, C^*y\rangle
\end{aligned}
$$

and therefore $C^*(\cdot) = \delta(\zeta-1)\int_0^1(\cdot)d\eta$, which leads to the following expression of Eq. (2.54),

$$\frac{dQ_0}{d\zeta} = \frac{1}{\nu}\left(-2Q_0C^*CQ_0 + M + 2\psi Q_0\right), \quad Q_0(0) = 0 \tag{2.57}$$

knowing that $CQ_0 = \int_0^1 \delta(\zeta-1)Q_0(\zeta)d\zeta = Q_0(1)$, and by applying the finding of the C^* expression one obtains

$$\frac{dQ_0}{d\zeta} = \frac{1}{\nu}\left(-2Q_0(\zeta)\delta(\zeta-1)Q_0(1) + M + 2\psi Q_0(\zeta)\right), \quad Q_0(0) = 0 \tag{2.58}$$

or

$$\frac{dQ_0}{d\zeta} = \frac{1}{\nu}\left(-2Q_0^2(1) + M + 2\psi Q_0(\zeta)\right), \quad Q_0(0) = 0 \tag{2.59}$$

In this particular case, we consider the following parameters $\psi = 1$, $\nu = 1$, $M = 1$, $f(\zeta) = \sin(2\pi\zeta)$, $\zeta \in [0,1]$ and $Q_0(0) = 0$ in the simulation demonstration. Eq. (2.59) provides the expression for the observer gain and it is solved numerically. For the sake of successful demonstration of the observer design, the case is realized when Eq. (2.22) is augmented with the time varying input applied at the boundary condition $u(t) = \sin(\omega t)$ with $\omega = 0.025$, see Fig. 2.2. Numerical realization of the model Eq. (2.50) and of the associated observer Eq. (2.22) is accomplished with backward finite difference in space $\triangle\zeta = 0.01$ and forward in time with time discretization step $\triangle t = 0.01$, see Fig. 2.3. Finally, as one can observe in the case of observer design for scalar hyperbolic systems, one can also consider spatially varying velocity $\nu(\zeta)$ and $\psi(\zeta)$, since the complexity for numerical integration of Eq. (2.57) is not significantly increased. In Fig. 2.3, one can find that the observer error is more pronounced in the area of the initial condition contribution to the state evolution (that is $x(\zeta,t) \in [0 < \zeta < 1, 0 < t < 1]$, see Figs. 2.2-2.3), as this is due to the fact that the observer is initialized at zero and that the error cannot decay faster than initial data transported throughout the domain by transporting velocity ($\nu = 1$ in example) of the transport-reaction system (for the values $t > 1$ the error evolution is smaller

than the velocity term, $v = 1$).

Example 2: Transport-reaction system modelled by 2^{nd} order parabolic PDE:

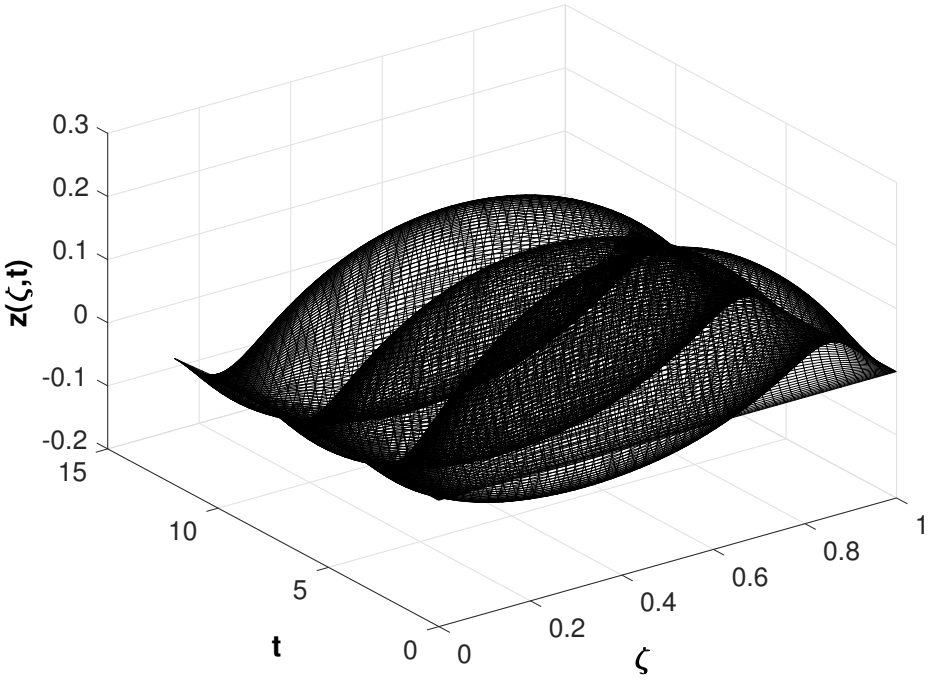


Figure 2.5: State evolution of the observer state $z(\zeta, t)$ given in Eq. (2.22).

A simple model of a heated catalytic metal bar exposed to the time-varying source term applied along the bar length is considered, while the output measurement is applied within the catalytic metal bar domain. Therefore, the model takes the following form of linear parabolic PDEs:

$$x_t = x_{\zeta\zeta} + \alpha x + \beta u(t), \quad (\zeta, t) \in [0, 1] \times \mathbb{R}^+ \tag{2.60}$$

$$x(0, t) = 0 = x(1, t)$$

$$y(t) = Cx(t) = \int_0^1 \frac{1}{2\epsilon} 1_{[\zeta_0 - \epsilon, \zeta_0 + \epsilon]}(\zeta) x(\zeta, t) d\zeta \tag{2.61}$$

where $x(\cdot, t) \in L^2((0, 1), \mathbb{R}^+)$ is a state. An assumption is that output measurements are taken within the domain of the metal bar, $y(t) = Cx$, with $\frac{1}{2\epsilon} 1_{[\zeta_0 - \epsilon, \zeta_0 + \epsilon]}$ being a unit measurement distribution function which is applied at ζ_0 over $[\zeta_0 - \epsilon, \zeta_0 + \epsilon]$ interval. The operator $\mathcal{A} = \frac{\partial^2}{\partial \zeta^2}$ is defined by its domain $\mathcal{D}(\mathcal{A}) = \left\{ \psi \in L^2(0, 1), \psi, \psi' \text{ are abs. cont.}, \frac{d^2\psi}{d\zeta^2} \in L^2(0, 1), \psi(0) = 0 = \psi(1) \right\}$. Along the same line, as in the previous example, the issue of the observer design is given as a systematic

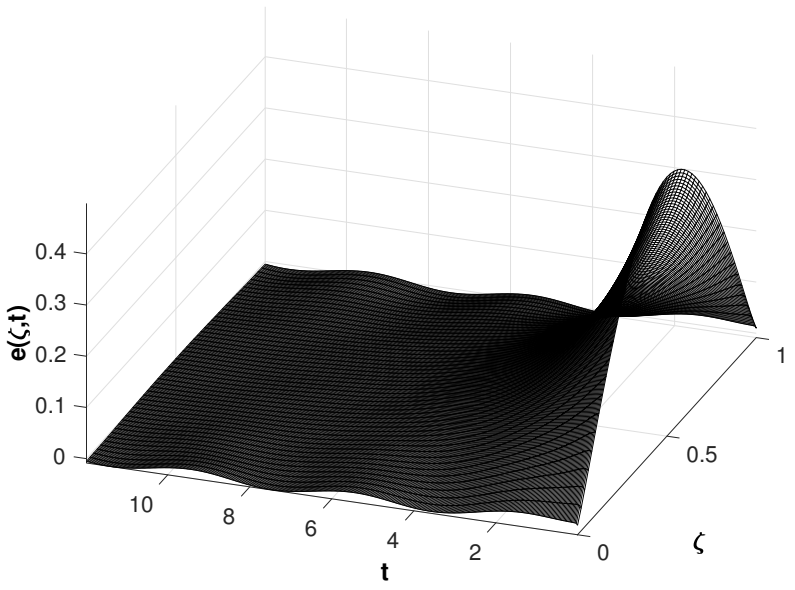


Figure 2.6: Observer error $e(\zeta, t)$ given in Eq. (2.23).

procedure of obtaining the observer gain $L(\zeta)$ in equation Eq. (2.26), so that the following equations are formulated:

$$\mathcal{A}Q_0 + Q_0\mathcal{A}^* + M - 2Q_0C^*CQ_0 = 0, \text{ on } D(\mathcal{A}^*) \tag{2.62}$$

or in the inner product form

$$\langle Q_0x_1, \mathcal{A}^*x_2\rangle + \langle \mathcal{A}^*x_1, Q_0x_2\rangle + \langle Mx_1, x_2\rangle - 2\langle CQ_0x_1, CQ_0x_2\rangle = 0 \tag{2.63}$$

taken that $x_1 = \psi_i$ and $x_2 = \psi_j$, and with the operator being self-adjoint $\mathcal{A} = \mathcal{A}^*$ with λ_i being operator eigenvalues and ψ_i associated eigenfunctions. Therefore, the following expression is obtained:

$$\langle Q_0\psi_i, \mathcal{A}\psi_j\rangle + \langle \mathcal{A}\psi_i, Q_0\psi_j\rangle + \langle M\psi_i, \psi_j\rangle - 2\langle CQ_0\psi_i, CQ_0\psi_j\rangle = 0 \tag{2.64}$$

$$\langle Q_0\psi_i, \lambda_j\psi_j\rangle + \langle \lambda_i\psi_i, Q_0\psi_j\rangle + \langle M\psi_i, \psi_j\rangle - 2\langle CQ_0\psi_i, CQ_0\psi_j\rangle = 0 \tag{2.65}$$

$$\lambda_j\langle Q_0\psi_i, \psi_j\rangle + \lambda_i\langle \psi_i, Q_0\psi_j\rangle + \langle M\psi_i, \psi_j\rangle - 2\langle CQ_0\psi_i, CQ_0\psi_j\rangle = 0 \tag{2.66}$$

Under a mild assumption that the solution Q_0 is given in the form of $Q_0x_1 = \sum_{i,j}^{\infty} q_{ij}\langle x_1, \psi_i\rangle\psi_j$, then $q_{ij} = \langle \psi_i, Q_0\psi_j\rangle$. In addition, M can be taken as a constant, which leads to the following:

$$\lambda_jq_{ij} + \lambda_iq_{ij} + M\delta_{ij} - 2\langle CQ_0\psi_i, CQ_0\psi_j\rangle = 0 \tag{2.67}$$

If the output measurement operator C is defined as $C = \frac{1}{2\epsilon}1_{[\zeta_0-\epsilon,\zeta_0+\epsilon]}$ with $\zeta_0 = 4/5$ and $\epsilon = 1/5$. Then, we obtain the following two expressions,

$$\lambda_j q_{ij} + \lambda_i q_{ij} + M\delta_{ij} - 2\langle Q_0\psi_i, Q_0\psi_j\rangle = 0, \quad \zeta \in [0.6, 1] \tag{2.68}$$

$$\lambda_j q_{ij} + \lambda_i q_{ij} + M\delta_{ij} = 0, \quad \zeta \in [0, 0.6] \tag{2.69}$$

or

$$\lambda_j q_{ij} + \lambda_i q_{ij} + M\delta_{ij} - 2\sum_{l=0}^{\infty} q_{il}q_{jl} = 0, \quad \zeta \in [0.6, 1] \tag{2.70}$$

$$\lambda_j q_{ij} + \lambda_i q_{ij} + M\delta_{ij} = 0, \quad \zeta \in [0, 0.6] \tag{2.71}$$

One can deduce that for $i \neq j$, the expression q_{ij} is the solution of the above equation. And the solution for $i = j$ becomes,

$$2\lambda_i q_{ii} + M - 2q_{ii}^2 = 0, \quad \zeta \in [0.6, 1] \tag{2.72}$$

$$2\lambda_i q_{ii} + M = 0, \quad \zeta \in [0, 0.6] \tag{2.73}$$

which yields

$$\begin{cases} q_{ii} = \frac{\lambda_i \pm \sqrt{\lambda_i^2+2M}}{2}, & \zeta \in [0.6, 1] \\ q_{ii} = -\frac{M}{2\lambda_i}, & \zeta \in [0, 0.6] \end{cases} \quad \text{or} \quad \begin{cases} q_{ii} = \frac{-i^2\pi^2 \pm \sqrt{(i^2\pi^2)^2+2M}}{2}, & \zeta \in [0.6, 1] \\ q_{ii} = \frac{M}{2i^2\pi^2}, & \zeta \in [0, 0.6] \end{cases} \tag{2.74}$$

Since the solution must be non-negative, one obtains:

$$Q_0 = \begin{cases} \sum_{i=0}^{\infty} \frac{-i^2\pi^2 + \sqrt{(i^2\pi^2)^2+2M}}{2}\langle\cdot, \psi_i\rangle\psi_i, & \zeta \in [0.6, 1] \\ \sum_{i=0}^{\infty} \frac{M}{2i^2\pi^2}\langle\cdot, \psi_i\rangle\psi_i, & \zeta \in [0, 0.6] \end{cases} \tag{2.75}$$

In this particular case, we consider a simulation scenario $\alpha = -0.1$, $\beta = 5$, $M = 10$, $f(\zeta) = \sin(\pi\zeta)$, $\zeta \in [0, 1]$. Eq. (2.60) is numerically solved by applying a standard Galerkin numerical solution with 10 eigenfunctions (ψ_i, $i = 1, \cdots, 10$), discretization time $\triangle t = 0.0005$ and an explicit integration scheme with time varying input $u(t) = \sin(\omega t)$ with $\omega = 0.0159$. One can observe that in this case the observer gain is given as an infinite sequence of convergent spatial functions given in Eq. (2.75), and not in the closed form. This implies that the observer gain for the given set of design parameters cannot be exactly determined but it is given in an approximated form.

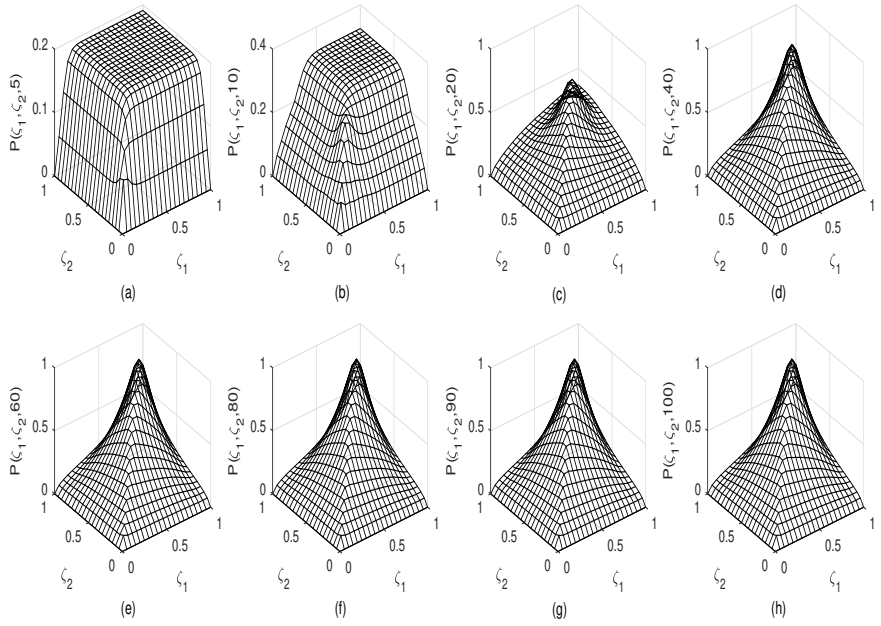


Figure 2.7: Profile of a posteriori estimate covariance P_k^+ described in Eq. (2.42) at time step $k = 5, 10, 20, 40, 60, 80, 90, 100$.

Example 3: Kalman filter design for the transport-reaction system modelled by 1^{st} order hyperbolic PDE:

We are revisiting the convection dominated transport-reaction model described by Eq. (2.50) with the uniformly applied input as in-domain actuation, which accounts for the realization of spatially uniform cooling with the jacket fluid flow. The spatial operator $\mathcal{A} = -v\frac{\partial}{\partial\zeta} + \psi$ is defined on its domain $D(\mathcal{A}) = \{x \in L^2(0,1) | x \text{ is abs. cont. } \frac{dx}{d\zeta} \in L^2(0,1), x(0) = 0\}$ and the operator $C(f(\zeta)) = \int_0^1 f(\zeta)\delta(\zeta - 1)d\zeta = f(1)$, the input is applied along the domain instead of boundary actuation, and ψ is assumed to be a constant spatial function. The discrete linear hyperbolic PDE system corresponding to Eq. (2.50) is obtained by applying Cayley-Tustin transformation and is given in the following form explicitly accounting for time and spatial characteristics:

$$\begin{aligned} x(\zeta,k) &= \mathcal{A}_d x(\zeta,k-1) + \mathcal{B}_d u(k) + w(k-1),\ x(\zeta,0) = x_0 \\ y(k) &= \bar{C}_d x(\zeta,k) + \bar{\mathcal{D}}_d u(k) + v(k),\ k \geq 1 \end{aligned} \tag{2.76}$$

where

$$\begin{aligned} \mathcal{A}_d(\cdot) &= [\delta I - \mathcal{A}]^{-1}[\delta I + \mathcal{A}](\cdot) \\ &= -(\cdot) + 2\delta\Big[\int_0^\zeta \frac{1}{v}(\cdot)e^{-\frac{1}{v}(\psi-\delta)\eta}d\eta\Big]e^{\frac{1}{v}(\psi-\delta)\zeta} \end{aligned}$$

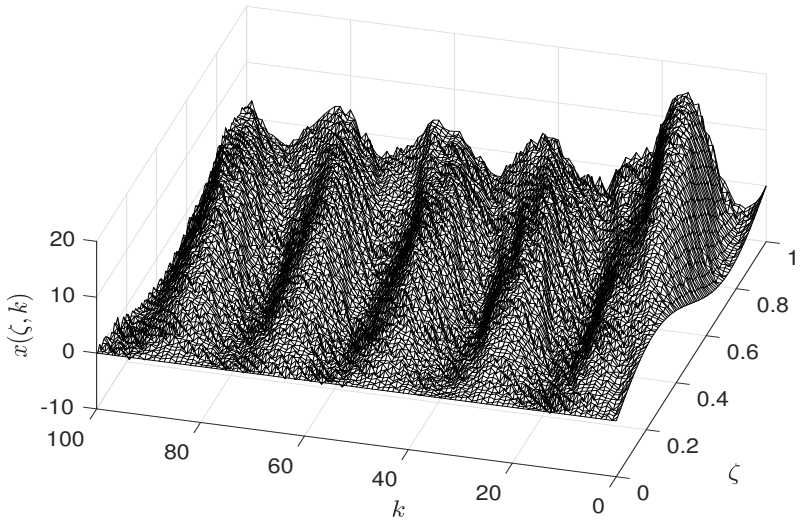


Figure 2.8: Profile of the state with noise $x(\zeta, k)$ described in Eq. (2.76).

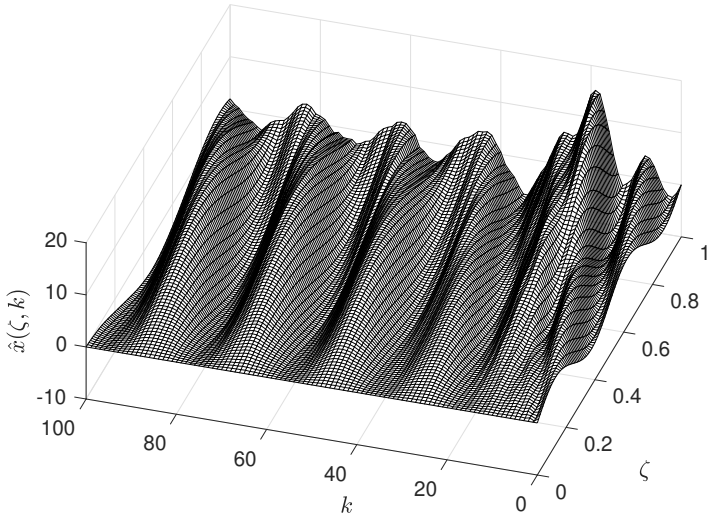


Figure 2.9: Profile of the state estimate $\hat{x}(\zeta, k)$ described in Eq. (2.43).

$$
\begin{aligned}
\mathcal{B}_d &= \sqrt{2\delta}[\delta I - \mathcal{A}]^{-1}B(\zeta) \\
&= \sqrt{2\delta}\Big[\int_0^{\zeta} \frac{1}{\nu}\mathcal{B}(\eta)e^{-\frac{1}{\nu}(\psi-\delta)\eta}d\eta\Big]e^{\frac{1}{\nu}(\psi-\delta)\zeta} \\
\bar{C}_d(\cdot) &= -\sqrt{2\delta}C[-\delta I - \mathcal{A}]^{-1}(\cdot) \\
&= -\sqrt{2\delta}\Big[\int_0^{L} \frac{1}{\nu}(\cdot)e^{-\frac{1}{\nu}(\psi+\delta)\eta}d\eta\Big]e^{\frac{1}{\nu}(\psi+\delta)L} \\
\bar{\mathcal{D}}_d &= C[-\delta I - \mathcal{A}]^{-1}\mathcal{B} + \mathcal{D} \\
&= \Big[\int_0^{L} \frac{1}{\nu}\mathcal{B}e^{-\frac{1}{\nu}(\psi+\delta)\eta}d\eta\Big]e^{\frac{1}{\nu}(\psi+\delta)L}
\end{aligned}
$$

The simulation result of the Kalman filter design given by Eqs. (2.41)-(2.43) for the discrete linear hyperbolic PDE system described in Eq. (2.43) is given in Fig. 2.9. We consider the following parameters: the spatial parameter arising from the linearization of the reaction term in the operator $\mathcal{A}$ is chosen to be $\psi = 0.5$, while the input operator $\mathcal{B}(0 < \zeta < 1) = 1$ represents spatially uniform realized heat transfer across the reactor shell. The time varying input is considered as $u_k = 3\sin(2\pi k)$, the output measurement is taken at $\zeta = 1$ (the exit of the reactor), while initial conditions are $x_0 = 5\sin(2\pi\zeta)$ and $\hat{x}_0 = 3\sin(4\pi\zeta)$. We assume that the process noise is $w_k \in \mathcal{N}(0, Q_k)$ with $Q_k = 0.1$ and measurement noise $v_k \in \mathcal{N}(0, R_k)$ with $R_k = 0.1$. Evolution profile of a posteriori two-dimensional estimate covariance P_k^+ at different time instances is shown in Fig. 2.7 and it converges to the two-dimensional surface given as $P^+(\zeta_1, \zeta_2, 100)$ in Fig. 2.7. Profiles of the state with noise and the estimated state are presented in Fig. 2.8 and Fig. 2.9. In Fig. 2.10 the profiles of the noisy output and the estimated output are presented and it can be seen that in the first $k \approx 25$ instances the Kalman filter estimation is delayed in time by the approximate value of transport residence time, and the mismatch is also influenced by the transients associated with the initial state conditions.

Example 4: The Kalman filter design for transport-reaction system modelled by 2^{nd} order parabolic PDEs:

In this section, the Kalman filter is applied to the representative cases of diffusion dominated model of an axial dispersion reactor described by the parabolic PDE with Dirichlet and Neumann boundary conditions. The linear infinite-dimensional system model of the diffusion dominated transport-reaction system is given by Eq. (2.1) where $\mathcal{A} = \frac{\partial^2}{\partial\zeta^2} + \psi$ is the linear operator defined on its domain $D(\mathcal{A})$ as it is given

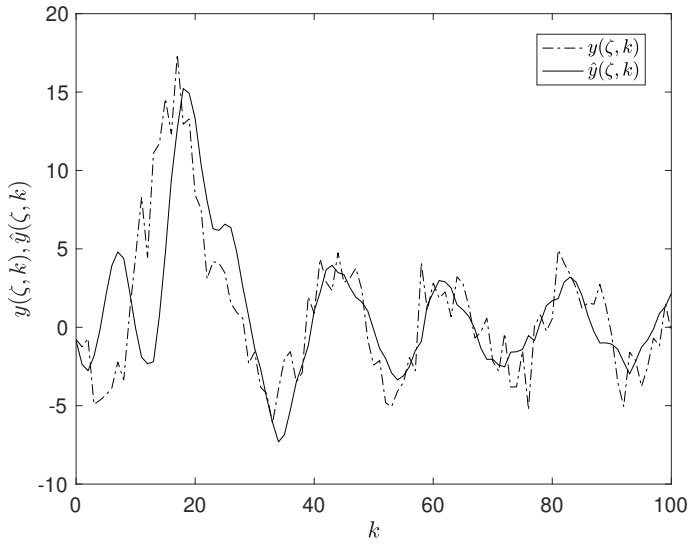


Figure 2.10: Profile of the output with noise $y(k)$ described in Eq. (2.76) and the Kalman filter output estimate $\hat{y}(k)$.

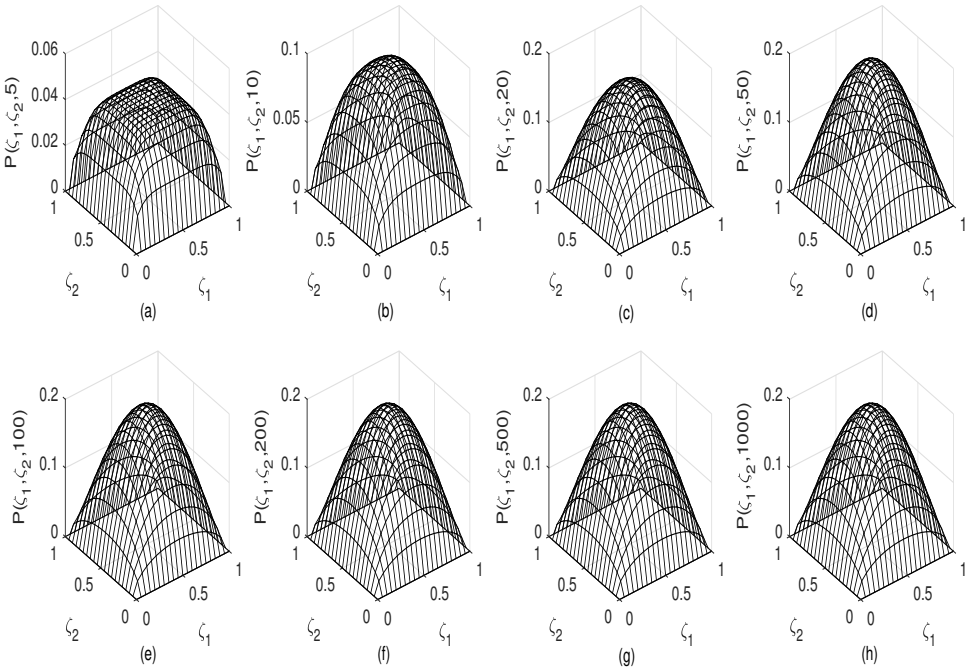


Figure 2.11: Profile of the posteriori estimate covariance P_k^+ described in Eq. (2.42) at time step $k = 5, 10, 20, 50, 100, 200, 500, 1000$.

in Example 2. The output is the state of the PDE at a point within the domain, for example at $\zeta = \zeta_0$ and is obtained by the operator $\mathcal{C}(f(\zeta)) = \int_0^L f(\zeta)\delta(\zeta - \zeta_0)d\zeta = f(\zeta_0)$ and $\mathcal{D} = 0$.

The resolvent operator $R(s, \mathcal{A}) = [sI - \mathcal{A}]^{-1}$ of the operator $\mathcal{A}$ is obtained by applying the Laplace transform and is expressed as follows:

$$\begin{aligned} R(s,\mathcal{A})x(\zeta,0) =& [sI-\mathcal{A}]^{-1}x(\zeta,0) \\ =& \frac{1}{\sqrt{s-\psi}}\frac{\sinh(\sqrt{s-\psi}\zeta)}{\sinh(\sqrt{s-\psi})}\int_0^1 x(\eta,0)\sinh[\sqrt{s-\psi}(1-\eta)]d\eta \\ & - \int_0^\zeta \frac{1}{\sqrt{s-\psi}}x(\eta,0)\sinh[\sqrt{s-\psi}(\zeta-\eta)]d\eta \end{aligned} \tag{2.77}$$

so that $\mathcal{A}_d := \left[-I + 2\delta(\delta I - \mathcal{A})^{-1}\right]$ is evaluated at $s = \delta$ by using the expression in Eq. (2.77). The discrete linear parabolic PDE system with the Dirichlet boundary condition obtained by Cayley-Tustin transformation is given as below:

$$\begin{aligned} x(\zeta,k) &= \mathcal{A}_d x(\zeta,k-1) + \mathcal{B}_d u(k) + w(k-1),\ x(\zeta,0) = x_0 \\ y(k) &= \bar{\mathcal{C}}_d x(\zeta,k) + \bar{\mathcal{D}}_d u(k) + v(k) \end{aligned} \tag{2.78}$$

where the discrete time operators $\mathcal{A}_d$, $\mathcal{B}_d$, $\bar{\mathcal{C}}_d$ and $\bar{\mathcal{D}}_d$ are obtained by using the resolvent operator described in Eq. (2.77) and Eq. (2.40).

The application of the Kalman filter design for parabolic PDE with Dirichlet boundary condition is provided in Figs. 2.12-2.13. In particular, the parameter ψ in the operator $\mathcal{A}$ is taken as $\psi = 5$, and the operator $\mathcal{B}(0 < \zeta < 1) = 1$. In addition, the input is $u_k = 1.5\sin(2\pi k)$, and the output measurement is at $\zeta = 0.5$ while initial conditions are taken as $x_0(\zeta) = 15[-(\zeta - 0.5)^2 + 0.5^2]$ and $\hat{x}_0(\zeta) = 15[-(\zeta - 0.5)^2 + 0.5^2] + w_0$. In simulation, we assume that the process noise is $w_k \in \mathcal{N}(0, Q_k)$ with $Q_k = 0.01$ and measurement noise is $v_k \in \mathcal{N}(0, R_k)$ with $R_k = 0.01$. The simulation results of the Kalman filter design for parabolic PDE with the Dirichlet boundary condition are given in Figs. 2.12-2.13 , while the profiles of a posteriori estimate covariance evolution P_k^+ at different time steps are shown in Fig. 2.11. The profiles of the state with noise and the state estimate are presented in Fig. 2.12 and Fig. 2.13, while Fig. 2.14 shows profiles of the output with noise and estimated output.

In addition to the realization of the Kalman filter for the parabolic PDE with Dirichlet boundary conditions, we provide also the case of a parabolic PDE with

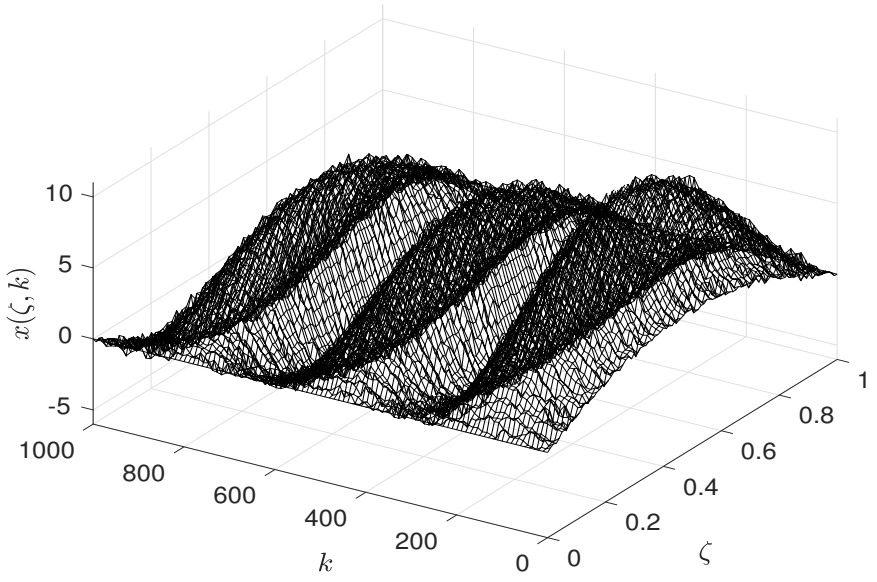


Figure 2.12: Profile of the state with noise $x(\zeta, k)$ described in Eq. (2.78) with Dirichlet boundary condition.

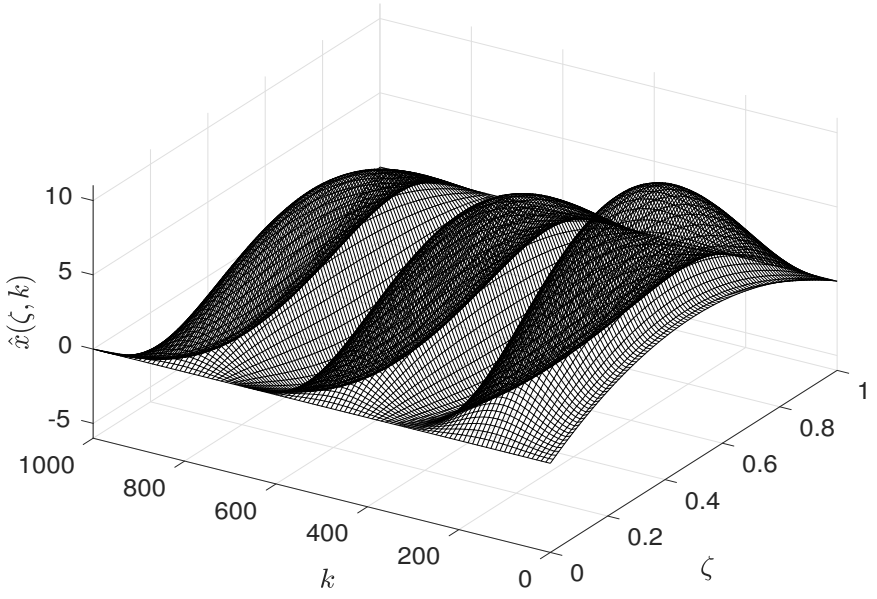


Figure 2.13: Profile of the estimated state $\hat{x}(\zeta, k)$ of parabolic PDE with the Dirichlet boundary condition.

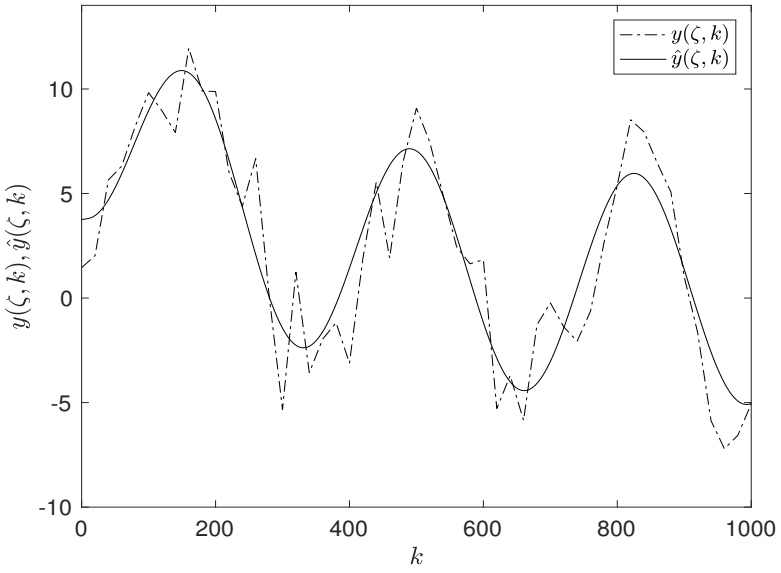


Figure 2.14: Profile of the output with noise $y(k)$ taken at $\zeta = 0.5$ and described in Eq. (2.78) with the Dirichlet boundary condition and the Kalman filter output estimates $\hat{y}(k)$.

natural Neumann boundary conditions. In this case, one needs to provide an insight into the construction of the resolvent operator being the crucial element in the design procedure. Therefore, the resolvent of the operator $\mathcal{A}$ with $D(\mathcal{A}) = \{x \in L^2(0,1) | x \text{ is abs. cont.}, \frac{dx}{d\zeta} \in L^2(0,1), \frac{d^2x}{d\zeta^2} \in L^2(0,1), \frac{dx(0)}{d\zeta} = 0 = \frac{dx(1)}{d\zeta}\}$ is given as:

$$\begin{aligned} R(s,\mathcal{A})x(\zeta,0) =& [sI - \mathcal{A}]^{-1}x(\zeta,0) \\ =& \frac{1}{\sqrt{s-\psi}} \frac{\cosh(\sqrt{s-\psi}\zeta)}{\sinh(\sqrt{s-\psi})} \int_0^1 x(\eta,0)\cosh[\sqrt{s-\psi}(1-\eta)]d\eta \\ & - \int_0^\zeta \frac{1}{\sqrt{s-\psi}} x(\eta,0)\sinh[\sqrt{s-\psi}(\zeta-\eta)]d\eta \end{aligned} \tag{2.79}$$

Finally, the discrete linear parabolic PDE system with the Neumann boundary condition is described in Eq. (2.78) with discrete operators $\mathcal{A}_d$, $\mathcal{B}_d$, $\mathcal{C}_d$ and $\mathcal{D}_d$ according to the resolvent operator $R(s,\mathcal{A})$ described in Eq. (2.79). The simulation conditions are taken so that $\psi = -0.1$ in the operator $\mathcal{A}$, while the input is distributed as follows, $\mathcal{B}(0 < \zeta < 0.2) = 0$, $\mathcal{B}(0.2 < \zeta < 0.8) = 1$, $\mathcal{B}(0.8 < \zeta < 1) = 0$, the input $u_k = \sin(2\pi k)$, the output is taken at $\zeta = 1$, and the initial conditions are $x_0 = 5\cos(2\pi\zeta)$ and $\hat{x}_0 = 5\cos(2\pi\zeta) + w_0$, while the process noise is $w_k \in \mathcal{N}(0, Q_k)$ with $Q_k = 0.01$, and the measurement noise is $v_k \in \mathcal{N}(0, R_k)$ with $R_k = 0.01$. The simula-

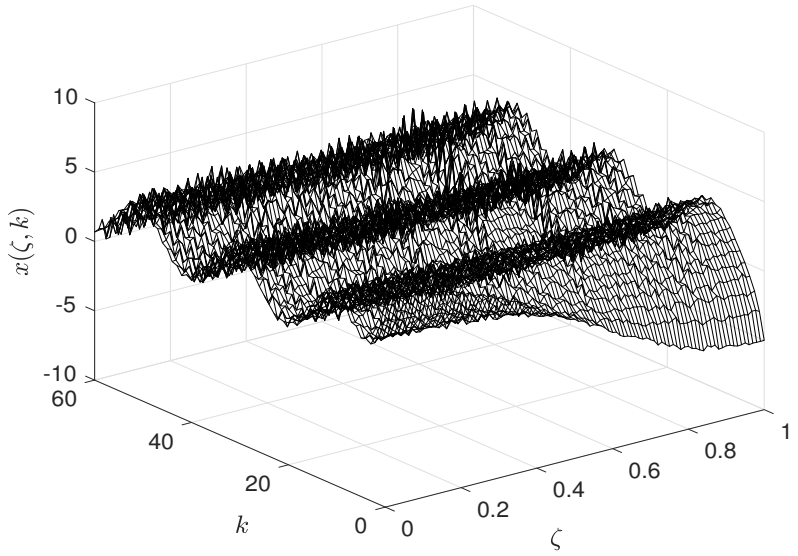


Figure 2.15: Profile of the state with noise $x(\zeta, k)$ described in Eq. (2.78) with the Neumann boundary condition.

tion results of the Kalman filter design for parabolic PDE with Neumann boundary condition are given in Figs. 2.15-2.17. From the simulation results of the linear transport-reaction system, it can be seen that the performance of the Kalman filter is good in estimating the state of the linear infinite-dimensional system described by hyperbolic PDE and parabolic PDE when the output measurements are noisy and when the process disturbances are present. Although the Kalman filter provides good results in estimating the states of linear transport-reaction systems, the issue of naturally present constraints of the state estimates can be adequately addressed through the constrained Kalman filter and/or moving horizon estimator design. As for stability issues of proposed observer and filter designs [106], we will discuss these in our future work.

Remark 5. *The expression for the resolvent $R(s, \mathcal{A})$ is necessary in evaluating the operator $\mathcal{A}_d$. Therefore, for the convection-reaction-diffusion system modelled by $x_t = Dx_{\zeta\zeta} - vx_\zeta + k\zeta$ with Danckwerts boundary conditions $Dx_\zeta(0,t) = vx(0,t)$ and $x_\zeta(1,t) = 0$, expression for resolvent is provided. The resolvent in this case is quite complex,*

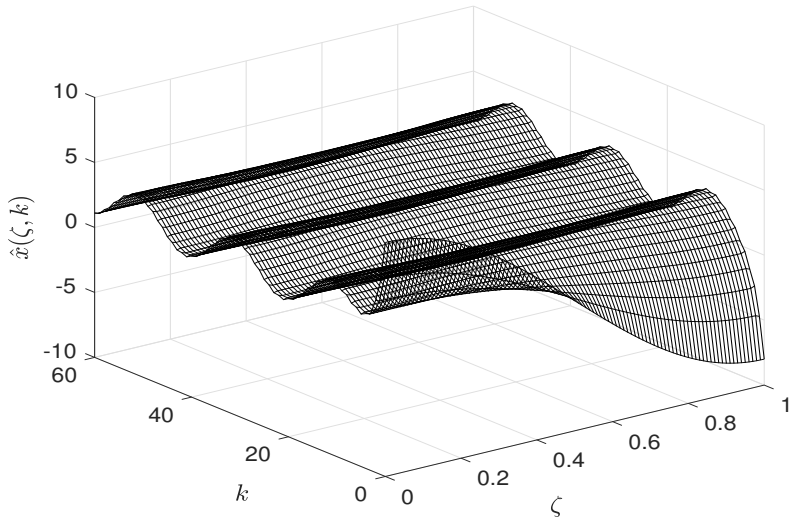


Figure 2.16: Profile of estimated state $\hat{x}(\zeta, k)$ of parabolic PDE with the Neumann boundary condition.

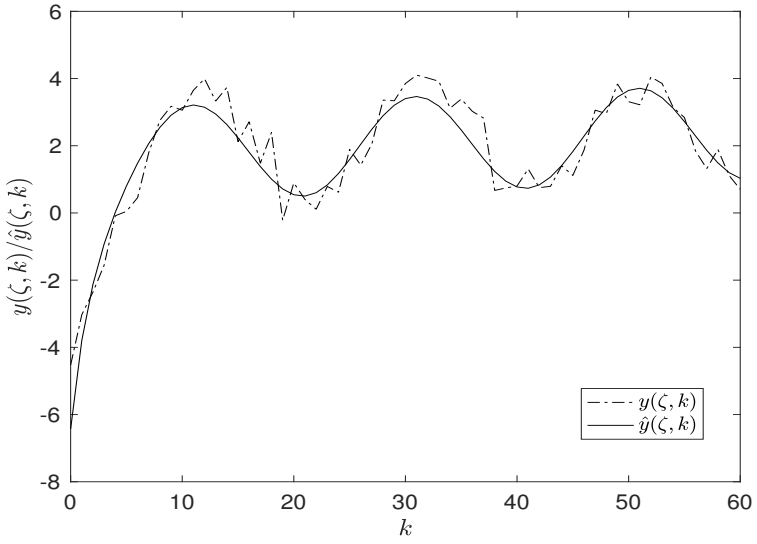


Figure 2.17: Profile of the output with noise $y(k)$ described in Eq. (2.78) with the Neumann boundary condition and estimated output $\hat{y}(k)$.

and it is given as:

$$R(s,\zeta)(\cdot)=\left[\frac{v}{p}P_1+P_2-\frac{2}{1+\frac{D}{p}}\frac{P_1}{p}\right]\int_0^1\frac{P_1}{P_3}(\cdot)d\zeta-\int_0^\zeta\frac{1}{p}\left[e^{\frac{v-p}{2D}(\zeta-\eta)}-e^{\frac{(v+p)}{2D}(\zeta-\eta)}\right](\cdot)d\eta \tag{2.80}$$

where $p=\sqrt{v^2-4\,D\,(s-k)}$, $P_1=e^{\frac{(v-p)}{2D}\zeta}-e^{\frac{(v+p)}{2D}\zeta}$, $P_2=e^{\frac{(v+p)}{2D}\zeta}+e^{\frac{(v-p)}{2D}\zeta}$, *and* $P_3=ve^{\frac{v-p}{2D}}(\frac{v-p}{2D})-v(\frac{v+p}{2D})e^{\frac{v+p}{2D}}+pe^{\frac{v+p}{2D}}(\frac{v+p}{2D})+pe^{\frac{v-p}{2D}}(\frac{v-p}{2D})$. *This can be easily applied in calculating the operators* $(\mathcal{A}_d,\mathcal{B}_d,\bar{\mathcal{C}}_d,\bar{\mathcal{D}}_d)$.

2.7 Summary

This chapter addresses the issue of systematic design of the state reconstruction in the realm of distributed parameter systems in the chemical process industry. The issue of steady state estimation is addressed in the setting of boundary available measurements. Furthermore, design methodology is presented for the systematic observer and the Kalman filter for distributed parameter systems modelled by linear transport-reaction models. The prominent feature of the discrete observer and the Kalman filter is that the modelling framework of a continuous spatially distributed plant is preserved by application of the Cayley-Tustin discretization, while generic properties of stability, controllability and observability are preserved by this transformation. Design methodology accounts for the spectral Riesz (parabolic PDEs with Dirichlet, Neumann or Danckwerts boundary conditions) and non spectral (hyperbolic PDEs) linear distributed parameter systems described in chemical engineering, and provides grounds for addressing other designs originating from the finite dimensional systems theory. In particular, the underlying designs can be easily extended to the systems of different types of coupled PDEs, or along the direction of developing a moving horizon estimator and a robust estimator when plant's uncertainties can be explicitly accounted for in the design.

Chapter 3

Moving Horizon Estimation for Regular Linear Systems

3.1 Introduction

Moving horizon Estimation (MHE) for finite-dimensional systems has been well developed [107, 108, 36]. However, MHE designs for infinite-dimensional systems modelled by partial differential equations (PDE) are significantly under-explored in existing literature. In this chapter, the MHE design for a general class of infinite-dimensional systems, namely regular linear infinite-dimensional systems with unbounded observation and disturbance operators, is considered in the presence of bounded plant and measurement disturbances.

Regular linear systems (RLS) were firstly introduced by Salamon & Weiss in the 1980s [109, 110]. Let X, U and Y be three Hilbert spaces that represent respectively the state, input and output spaces of an infinite-dimensional linear system. This system is described by the following equations:

$$\dot{x}(t) = Ax(t) + Bu(t),\ x(0) = x_0 \in X \tag{3.1a}$$

$$y(t) = Cx(t) + Du(t),\ t \geq 0 \tag{3.1b}$$

where the control operator B and the observation operator C can be unbounded, allowing for actuators and sensors to be supported at isolated points, piecewise, and/or at the boundary. This class of infinite-dimensional systems, although the input and output operators can be unbounded, possesses many properties that parallel finite-dimensional systems in many ways, and hence they are of considerable practical in-

terest. For an introduction of well-posed systems, we refer to the seminal works in [21, 42, 111].

The optimal control problem of well-posed systems has been extensively studied in the literature since the 1990s. Generally, there are two ways of solving the optimal cost operator in linear quadratic optimal control problems on an infinite time horizon, including using algebraic Riccati equations and performing the spectral factorization of the associated Popov functions [5]. For systems with bounded $\boldsymbol{B}$ and $\boldsymbol{C}$ operators, the control algebraic Riccati equation and its corresponding solution parallel these in finite-dimensional systems, while the formulation of Riccati equation is not always well-defined when having unbounded $\boldsymbol{B}$ and $\boldsymbol{C}$ operators [112, 5]. To address this issue, the Riccati equation theory has been extensively investigated in [46, 21] (and the references therein) by proposing complex assumptions and extension operators accounting for admissibility. Another perspective is to find some alternative expressions for Riccati equations, e.g., the Riccati equations associated with the optimal control problems for the reciprocal systems [113, 24], and the discrete-time systems by using Cayley transform [24]. Thus, the control problems of the original systems with unbounded operators can be linked equivalently to the corresponding control problems for the reciprocal (or discrete) systems with all bounded operators and well-posed Riccati equations that are simpler to solve.

As a dual problem of the optimal control problem, the optimal filtering problem of linear infinite-dimensional systems can be traced back to the 1970s, see [75, 7, 8]. The filtering problem for linear infinite-dimensional continuous-time systems with bounded operators was systematically studied in [78, Cha. 6]. Recently, some contributions on this topic include, [114] where consensus filters were designed for a class of infinite-dimensional systems having a network of pointwise measurements, [9] where the solution to an operator Riccati equation minimized the steady-state error variance and the result was further extended to optimal sensor placement by minimizing the trace of the Riccati operator, [115] where an optimal state estimation problem was cast into a nonstandard finite-horizon linear quadratic output tracking problem for non-time invertible evolutionary systems, where disturbance constraints were considered as an upper bound on the objective function. Overall, the constrained optimal filtering problem of well-posed systems with unbounded disturbance and observation

operators is comparatively under-explored. Essentially, the challenges lie in addressing the issues associated with the unboundedness of operators and constraint-handling simultaneously.

The research area of moving horizon estimation with various applications has flourished over the past two decades [107, 108, 36]. The intriguing nature of MHE is threefold, including 1) Unlike Kalman filter using only one single instantaneous measurement for an estimation, MHE deploys N most recent measurements in a moving horizon manner, thus avoiding the so-called short-horizon syndrome and generating more robust and smoother estimation despite disturbances and modelling uncertainties [116]; 2) MHE truncates a full information estimation (FIE) that uses all measurements starting from time instant 0 by a moving horizon containing N most recent measurements to enhance the estimation efficiency of online implementation; and 3) most importantly MHE explicitly accounts for the physical constraints to prevent unrealistic estimates due to spurious measurements [37]. There have been several attempts to tackle MHE design for distributed parameter systems, including [117, 118, 116] by performing early lumping based on model reduction methods. Therefore, how to simultaneously perform the MHE design and treat a linear DPS intact (i.e., without any spatial approximation or model reduction) is the focus of this chapter.

Motivated by the above findings, in this chapter, moving horizon estimation for regular linear systems is developed. In particular, a discrete-time infinite-dimensional model with all bounded operators suitable for MHE design is presented using the Cayley-Tustin (CT) time-discretization approach [97] that does not involve any spatial discretization or model reduction. The CT framework possesses a Crank-Nicolson type of time integration scheme that admits an implicit midpoint integration rule, leading to a symmetric (structure-preserving) and symplectic (energy-preserving) integration scheme [97, 119]. By applying the CT transformation, unbounded disturbance and observation operators in regular linear systems can be transformed to all bounded operators in discrete systems, with better numerical realization ability and essential system properties (e.g., stability) preserved [19, 4, 113, 24, 21, 22, 97]. Given that the system properties and input-output mappings stay invariant under the CT bilinear transformation, one can link the discrete estimation results back to the con-

tinuous regular linear systems, leading to a constrained suboptimal estimation for RLS. Therefore, the contributions of this chapter are listed as follows:

(1) The known moving horizon estimation theory of finite-dimensional systems is extended to regular linear distributed parameter systems to explicitly account for disturbance and output constraints in the estimation algorithm while considering the boundary disturbance and observation. To the best knowledge of the authors, it is the first time that the MHE design is incorporated in the infinite-dimensional setting in a late lumping manner (i.e., spatial discretization is not performed in the design stage). The full information estimator (FIE) and MHE designs are proposed, which provide unified and systematic frameworks for constrained output estimation of regular linear systems with unbounded or bounded input, output and system operators. Theorems 5-6 provide sufficient conditions to prove the strong stability of the proposed FIE and MHE. This chapter extends the previous contribution on model predictive control of regular linear systems [65] into its dual problem, a constrained moving horizon estimation of regular linear systems.

(2) Based on the CT transformation, a well-defined discrete-time algebraic Riccati equation is shown to have the same minimal nonnegative self-adjoint solution as the continuous-time algebraic Riccati equation, and is utilized in the MHE design. The link between discrete- and continuous-time Riccati equations is obtained by considering unconstrained optimal filtering problems on an infinite-time horizon. Theorem 4 shows that a regular linear system is estimatable iff its corresponding discrete system induced by the CT transform is estimatable, and their optimal cost operators are equal. This result extends the optimal control problems of well-posed systems and control algebraic Riccati equations [24] into a filtering problem, where Laguerre transform is utilized as a global transformation between discrete- and continuous-time signals. Considering practical application, average sampling (namely, Crank-Nicolson numerical integration) and zero-order sampling are proved to be capable of providing a good approximation of continuous plant disturbance signal using measurements with a fast sampling rate in Proposition 3 by extending previous work on parameter identification of finite-dimensional systems [120].

The rest of this chapter is organized into 6 sections. Necessary notations and mathematical preliminaries on regular linear systems and the continuous-time fil-

tering problem are provided in Section 3.2. Section 3.3 presents the model time-discretization and the discrete-time Riccati equation. In Section 3.4, the FIE and MHE designs are proposed and the stability analysis is provided. Two examples are given in Section 3.5 to illustrate the effectiveness of the proposed MHE design. Concluding remarks are drawn in Section 3.6.

3.2 Mathematical preliminaries

3.2.1 Notations

Assume that X and Z are two Hilbert spaces and $A : X \mapsto Z$ is a linear operator. Linear bounded operators from X to Z are denoted by $\mathcal{L}(X, Z)$ and $\mathcal{L}(X)$ if $X = Z$. The domain, range, spectrum, resolvent set and resolvent operator of a linear operator A are denoted as: $\mathcal{D}(A)$, $\mathfrak{R}(A)$, $\sigma(A)$, $\rho(A)$, and $\mathcal{R}(s, A) = (sI - A)^{-1}$ with $s \in \rho(A)$, respectively. We define the space X_1 as $\mathcal{D}(A)$ with the norm $\|x\|_1 = \|(\beta I - A)x\|$, and the space X_{-1} as the completion of X with the norm $\|z\|_{-1} = \|(\beta I - A)^{-1}z\|$, where $\forall x \in \mathcal{D}(A)$, $\forall z \in X$, and $\beta \in \rho(A)$. The constructed spaces are related by $X_1 \subset X \subset X_{-1}$, with each inclusion being dense and continuous embedding [42]. The extension of A to X_{-1} is still denoted by A, and the Λ-extension of an operator C is denoted by C_Λ. The inner product is denoted by $\langle \cdot, \cdot \rangle$, and $L^2(0, l)^m$ with a positive integer m denotes a Hilbert space of an m-dimensional vector of the real functions that are square integrable over $[0, l]$ with a spatial length l. Let $\mathbb{R}_0^+ = [0, \infty)$, $\mathbb{R}^+ = (0, \infty)$, $\mathbb{R}^- = (-\infty, 0)$, and $\mathbb{C}_\sigma^+ = \{s \in \mathbb{C} | \operatorname{Re} s > \sigma\}$ delimited by $\sigma \in \mathbb{R}$. $\mathbb{Z}^-$ stands for the set of negative integers, and $\mathbb{Z}^+$ denotes the set of positive integers. For any Hilbert space V, interval $\mathcal{J}$ and $\beta \in \mathbb{R}$, we denote $L^2_\beta(\mathcal{J}; V) = e_\beta L^2(\mathcal{J}; V)$, where $(e_\beta v)(t) = e^{\beta t} v(t)$ with the norm $\|e_\beta v\|_{L^2_\beta} = \|v\|_{L^2}$, and we regard $L^2_{\text{loc}}(\mathbb{R}_0^+, V)$ as a Fréchet space. For any $v \in L^2_{\text{loc}}(\mathbb{R}_0^+, V)$ and $\tau \geq 0$, we regard the time-inversion operator $\boldsymbol{\mathfrak{R}}_\tau$ on $[0, \tau]$ as: $(\boldsymbol{\mathfrak{R}}_\tau v)(t) = v(\tau - t)$ for $t \in [0, \tau]$, $(\boldsymbol{\mathfrak{R}}_\tau v)(t) = 0$ for $t > \tau$, and define $(\boldsymbol{\mathfrak{R}} v)(t) = v(-t)$ for all $t \in \mathbb{R}$. For any Hilbert space V and $\tau \geq 0$, $\mathbf{S}_\tau$ denotes the right shift operator by τ on $L^2(\mathbb{R}_0^+, V)$, and $\mathbf{P}_\tau$ denotes the projection of $L^2(\mathbb{R}_0^+, V)$ onto $L^2([0, \tau], V)$ (by truncation). For any $u, v \in L^2_{\text{loc}}(\mathbb{R}_0^+, X)$ and $\tau \geq 0$, the τ-concatenation of u and v is defined by $u \underset{\tau}{\diamond} v = \mathbf{P}_\tau u + \mathbf{S}_\tau v$, see [121].

3.2.2 Regular linear systems

There are several equivalent ways to define well-posed regular linear systems, and here we shall use the definition from [121] by first introducing global operators $\Sigma_\tau = [\mathrm{T}_\tau, \Phi_\tau; \Psi_\tau, F_\tau]$ for all $\tau \geq 0$.

Definition 1. *Let U, X and Y be Hilbert spaces, $\Omega = L^2(\mathbb{R}_0^+; U)$ and $\Gamma = L^2(\mathbb{R}_0^+, Y)$. A well-posed linear system is a family of operators $\Sigma = (\Sigma_t)_{t\geq 0}$ partitioned as $\Sigma_\tau = [\mathrm{T}_\tau, \Phi_\tau; \Psi_\tau, F_\tau]$, for all $\tau \geq 0$, where (i) $T = (T_t)_{t\geq 0}$ is a strongly continuous semigroup of bounded linear operators on X; (ii) $\Phi = (\Phi_t)_{t\geq 0}$ is a family of bounded linear operators from Ω to X such that $\Phi_{\tau+t}(u \underset{\tau}{\diamond} v) = T_t\Phi_\tau u + \Phi_t v$, for every $u, v \in \Omega$ and all $\tau, t \geq 0$; (iii) $\Psi = (\Psi_t)_{t\geq 0}$ is a family of bounded linear operators from X to Γ such that $\Psi_{\tau+t}x_0 = \Psi_\tau x_0 \underset{\tau}{\diamond} \Psi_t T_\tau x_0$, for every $x_0 \in X$ and all $\tau, t \geq 0$, and $\Psi_0 = 0$; (iv) $F = (F_t)_{t\geq 0}$ is a family of bounded linear operators from Ω to Γ such that $F_{\tau+t}(u \underset{\tau}{\diamond} v) = F_\tau u \underset{\tau}{\diamond} (\Psi_t\Phi_\tau u + F_t v)$, for every $u, v \in \Omega$ and all $\tau, t \geq 0$, and $F_0 = 0$.*

Here the operators Φ, Ψ and F are also called input maps, output maps and input-output maps, i.e., $x(\tau) = \mathrm{T}_\tau x_0 + \Phi_\tau \mathbf{P}_\tau u$ and $\mathbf{P}_\tau y = \Psi_\tau x_0 + F_\tau \mathbf{P}_\tau u$, where $u \in L^2_{\mathrm{loc}}(\mathbb{R}_0^+; U)$, $y \in L^2_{\mathrm{loc}}(\mathbb{R}_0^+; Y)$, $\tau \geq 0$ and $x_0 \in X$. Let Ψ_∞ and F_∞ be the extended output and input-output maps representing strong limits of $\Psi_\tau \in \mathcal{L}(X, L^2_{\mathrm{loc}}(\mathbb{R}_0^+; Y))$ and $F_\tau \in \mathcal{L}(L^2_{\mathrm{loc}}(\mathbb{R}_0^+; U), L^2_{\mathrm{loc}}(\mathbb{R}_0^+; Y))$ as $\tau \to \infty$. Thus, the state and output are expressed as: $x(t) = \mathrm{T}_t x_0 + \Phi_t u$ and $y = \Psi_\infty x_0 + F_\infty u$.

Let U, X and Y be the input, state and output spaces, where X is a separable Hilbert space, and U and Y are finite-dimensional. For a strongly continuous semigroup $T(t)$ on a Hilbert space X, $\forall w > w_0$, there is a constant M_w such that $\forall t \geq 0, \|T(t)\| \leq M_w e^{wt}$, then the constant w_0 is called the growth bound of the semigroup [4, The. 2.1.6]. For a well-posed linear system, we can associate generating operators (A, B, C), where $A : \mathcal{D}(A) \subset X \to X$ denote the generator of T with the growth bound w_0, $B \in \mathcal{L}(U, X_{-1})$ is the control operator, $C \in \mathcal{L}(X_1, Y)$ is the observation operator, such that $\Phi_\tau u = \int_0^\tau T_{\tau-\sigma} B u(\sigma) d\sigma$ and $(\Psi_\infty x_0)(t) = CT_t x_0$ for all $t \geq 0$ and $x_0 \in X_1$. The operator B is called an admissible control operator for T if $\mathfrak{R}(\Phi_\tau) \subset X$, and the operator C is called an admissible observation operator for T if $\int_0^\tau \|CT_t x_0\|^2 dt \leq k_\tau \|x_0\|^2$, with $x_0 \in \mathcal{D}(A)$ and $k_\tau \geq 0$. If there exists a $k > 0$ and $k_\tau \leq k$ such that the above inequality holds as $\tau \to \infty$, then C is called an infinite-time

admissible observation operator for T. The infinite-time admissibility of the observation operator is needed later on for the stability analysis of MHE. Let us denote the Λ-extension of the operator C by C_Λ, where $C_\Lambda x_0 = \lim_{\lambda\to+\infty} C\lambda(\lambda I - A)^{-1}x_0$, and the domain of C_Λ contains all $x_0 \in X$ for which the above limit exists [121]. The well-posed system Σ is regular if $\lim_{s\to+\infty} \mathcal{G}(s)v = Dv$, for all $v \in U$ [121, The. 1.3], where the transfer function $\mathcal{G}(s)$ is defined by:

$$\mathcal{G}(s) := C_\Lambda(sI - A)^{-1}B + D \tag{3.2}$$

Here C_Λ is the Λ-extension of C. To cast the well-posed system into model (3.1), C must be replaced by C_Λ. Define the Laplace transform $\mathfrak{L}$ as $\hat{u}(s) = (\mathfrak{L}u)(s) := \int_0^\infty e^{-st}u(t)dt$ for $s \in \mathbb{C}^+_{\gamma_F}$ where $\gamma_F \in \mathbb{R}$ is the growth bound of F_∞. By Plancherel's theorem, the Laplace transform is a unitary map from $L^2_{\gamma_F}(\mathbb{R}^+, U)$ onto the Hardy space $\mathbf{H}^2(\mathbb{C}^+_{\gamma_F}, U)$. Thus, the output $y \in L^2_{\text{loc}}(\mathbb{R}^+, Y)$ of a well-posed linear system is represented by the initial state x_0 and input $u \in L^2_{\text{loc}}(\mathbb{R}^+_0, U)$ in Laplace domain as:

$$\hat{y}(s) = C(sI - A)^{-1}x_0 + \mathcal{G}(s)\hat{u}(s), \text{ for } s \in \mathbb{C}^+_{w_0} \tag{3.3}$$

where $w_0 \geq \gamma_F$, see [121, Pro. 4.1]. To use the duality between optimal control and filtering problems, the concept of the dual well-posed system is introduced.

Theorem 2. *[111, The. 3.5] Let $\Sigma = [T, \Phi; \Psi, F]$ be a well-posed linear system with input space U, state space X and output space Y. Define Σ^d_τ (the superscript d stands for "dual") as*

$$\Sigma^d_\tau = \begin{bmatrix} \mathrm{T}^d_\tau & \Phi^d_\tau \\ \Psi^d_\tau & F^d_\tau \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & \mathfrak{R}_\tau \end{bmatrix} \begin{bmatrix} \mathrm{T}^*_\tau & \Psi^*_\tau \\ \Phi^*_\tau & F^*_\tau \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & \mathfrak{R}_\tau \end{bmatrix} \tag{3.4}$$

for all $\tau \geq 0$. Then $\Sigma^d = [T^d, \Psi^d; \Phi^d, F^d]$ is a well-posed system with input space Y, state space X and output space U. If A, B and C are the semigroup generator, control operator and observation operator of Σ, then the operators associated with Σ^d are A^*, C^* and B^*. The transfer functions are linked by: $\mathcal{G}^d(s) = \mathcal{G}^*(\bar{s})$, $s \in \mathbb{C}^+_{w_0}$. Both semigroups share the same growth bound, i.e., $w_0 = w^d_0$, where w^d_0 is the growth bound of the semigroup T^d.

3.3 Continuous-time filtering problem

Consider a regular linear system with corrupted measurement:

$$\dot{x}(t) = Ax(t) + Bw(t),\ t \geq 0,\ x(0) = x_0 \in X \tag{3.5a}$$

$$y(t) = Cx(t) + v(t),\ z(t) = Hx(t) \tag{3.5b}$$

where $A : \mathcal{D}(A) \subset X \to X$ is a known infinitesimal generator of a C_0-semigroup $\mathbb{T}(t)$ on a separable Hilbert space X, $w(t)$ denotes the plant (or input) disturbance on a finite-dimensional Hilbert space W, and $v(t)$ is output disturbance (independent with $w(t)$) on a finite-dimensional Hilbert space Y. $y(t), z(t) \in Y$ represent the measured output and the to-be-estimated output, respectively. The operators B, C and H are assumed to be known and admissible: $B \in \mathcal{L}(W, X_{-1})$, $C, H \in \mathcal{L}(X_1, Y)$. As aforementioned, unbounded C and H need to be replaced by their Λ-extension C_Λ and H_Λ in transfer functions and system model (3.5) to ensure well-posedness. The unknown initial state, plant and output disturbances that are generally induced by process noise, sensor noise and/or modelling errors determine the observed output y and the to-be-estimated output z. Instead of estimating the whole state, a linear combination of states Hx is considered, where H may be a state feedback that can be utilized for construction of an output feedback controller [40].

We aim to find an adequate Riccati equation for an infinite-time filtering problem for system (3.5). There are two ways of constructing an infinite-time filtering problem from a finite-time one (on the interval $[t_0, t_1]$), either by letting $t_1 \to +\infty$ or $t_0 \to -\infty$. The latter one is chosen along with $t_1 = 0$ because the (discrete) Kalman filter is most naturally posed on $k \in \mathbb{Z}^-$ [85] and its duality with the optimal control problem on $t \in \mathbb{R}^+$ (e.g. as in [122]) will be used.

Consider system (3.5) on $t \in \mathbb{R}^-$ with measurements $Y(t) = \{y(s), s \in \mathbb{R}^-\}$ and a final state $x_0 \in X$, find the set $\{w\}$ of possible plant disturbance such that the following cost functional is minimized:

$$J(x_0, w) = \int_{-\infty}^{0} \|w(t)\|^2_{Q^{-1}} + \|v(t)\|^2_{R^{-1}}\, dt \tag{3.6}$$

where positive-definite self-adjoint operators $Q^{-1} \in \mathcal{L}(W)$ and $R^{-1} \in \mathcal{L}(Y)$ penalize the uncertainties of plant and output disturbances $w(t)$ and $v(t)$ according to model

(3.5). If there exists a self-adjoint operator $P^{-1} \in \mathcal{L}(X)$ such that the optimal cost is given by:

$$\inf_{(x_0,w)\in X\times L^2(\mathbb{R}^-;W)} J(x_0, w) = \left\langle x_0, P^{-1}x_0 \right\rangle$$

then P^{-1} represents the optimal cost operator. The operator P is the minimal non-negative self-adjoint solution of a continuous-time algebraic Riccati equation (CARE) [123]. For bounded operators B and C, we simply obtain the following CARE [78, Cha. 6 & 7]

$$AP + PA^* - PC^*R^{-1}CP + BQB^* = 0, \text{ on } \mathcal{D}(A^*) \tag{3.7}$$

For unbounded operators B and C, however, it is nontrivial to obtain the corresponding filter Riccati equation since technical treatment needs to be taken into account for admissibility of operators, e.g., by using Λ-extension operators [46, 112, 5], Lur'e form [124], and integral form of Riccati equations [115]. In this chapter, the proposed approach takes a different path by using a discrete-time algebraic Riccati equation that is always well-posed, easily computable, and suitable for MHE design in the discrete-time setting. For completeness, we adopt the optimizability and estimatability concepts from [122, Def. 3.1, Def. 4.1].

Definition 2. *The well-posed system Σ (or the pair (A, B)) is optimizable if for every $x_0 \in X$ there exists a $w \in L^2(\mathbb{R}^+, W)$ such that $x(t) = \mathrm{T}_t x_0 + \Phi_t w \in L^2(\mathbb{R}^+, X)$. The well-posed system Σ (or the pair (A, C)) is estimatable if (A^*, C^*) is optimizable.*

Then, we make the following assumptions:

Assumption 1. *(A, C) is estimatable.*

Assumption 2. *(A, C) is exactly observable in finite time $\tau > 0$, and C is infinite-time admissible for $\mathbb{T}$.*

Assumption 3. *A generates a strongly stable contraction C_0-semigroup, i.e., $\|\mathbb{T}(t)\| \leq 1$ and $T(t)z \to 0$ as $t \to +\infty$ for all $z \in X$.*

Assumption 4. *The disturbance and output spaces are finite-dimensional.*

Assumption 1 ensures that (A^*, C^*) is optimizable and for $z_0 \in X$ there exists a non-empty admissible output set $\mathfrak{Y}(z_0)$ defined by $\mathfrak{Y}(z_0) = \{\mathfrak{R}y \in L^2_{\text{loc}}(\mathbb{R}^+, Y) | \mathfrak{R}w = \Psi^d_\infty z_0 + F^d_\infty \mathfrak{R}y \in L^2_{\text{loc}}(\mathbb{R}^+; W)$ and the cost (3.6) is finite$\}$. Assumptions 2 and 3 are needed to ensure the finite-time exact observability of the discrete-time system (3.8) (defined in the following section). It is clear that if (A, C) is exactly observable, then it is also estimatable. Assumptions 3 and 4 assure that the discrete system (3.8) in the ensuing section provides a good approximation of the continuous system (3.5).

3.4 Discrete-time filtering problem

3.4.1 Model time-discretization

Consider a Tustin time-discretization of model (3.5) for a given time discretization interval h on $t \in (kh, (k+1)h)$:

$$\frac{x_{k+1} - x_k}{h} \approx A\frac{x_{k+1} + x_k}{2} + B\frac{w_k}{\sqrt{h}} \tag{3.8a}$$

$$\frac{y_k}{\sqrt{h}} \approx C\frac{x_{k+1} + x_k}{2} + \frac{v_k}{\sqrt{h}},\ \frac{z_k}{\sqrt{h}} \approx H\frac{x_{k+1} + x_k}{2} \tag{3.8b}$$

where the spatial dependence of x is omitted for brevity. For disturbances and outputs discretization, we consider an "average" sampling operator $\mathbf{T}_h$ as in [97]:

$$\frac{w_k}{\sqrt{h}} = (\mathbf{T}_h w)(k) := \frac{1}{h}\int_{kh}^{(k+1)h} w(t)dt \tag{3.9}$$

and discrete signals v_k, y_k, and z_k are similarly defined. As shown in [97, The. 4.3], the Cayley-Tustin (CT) discretization is a convergent time discretization frame for scalar-valued (dim U=dim Y=1) input-output stable system nodes in the sense that as $h \to 0$, $\sup_{k \le T/h} \|y_k/\sqrt{h} - y(hk)\|$ converges to 0, i.e., that the convergence is uniform in k, where $k \in \mathbb{Z}$ and $0 \le k \le T/h$. Similar convergence condition holds for w_k, v_k and z_k. Direct manipulation of (3.8) leads to the Cayley-Tustin discretization of (3.5) as:

$$x_{k+1} = A_d x_k + B_d w_k,\ x_0 \in X \tag{3.10a}$$

$$y_k = C_d x_k + D_d w_k + v_k = C_d x_k + \bar{v}_k \tag{3.10b}$$

$$z_k = H_d x_k + V_d w_k \tag{3.10c}$$

where $\bar{v}_k = D_d w_k + v_k$ denotes the discrete measurement disturbance, and the discrete operators are given by:

$$\begin{bmatrix} A_d & B_d \\ C_d & D_d \\ H_d & V_d \end{bmatrix} = \begin{bmatrix} -I + 2\delta\mathcal{R}(\delta, A) & \sqrt{2\delta}\mathcal{R}(\delta, A)B \\ \sqrt{2\delta}C\mathcal{R}(\delta, A) & \mathcal{G}(\delta) \\ \sqrt{2\delta}H\mathcal{R}(\delta, A) & \mathcal{T}(\delta) \end{bmatrix} \tag{3.11}$$

where we denote the resolvent operator by $\mathcal{R}(\delta, A) = (\delta I - A)^{-1}$, $\delta = 2/h \in \rho(A) \cap \mathbb{R}^+$, and $\mathcal{T}(\delta) = H_\Lambda \mathcal{R}(\delta, A)B$. By CT transform, we have converted the continuous plant (3.5) with unbounded operators to the discrete system (3.10) having all bounded operators as: $A_d \in \mathcal{L}(X)$, $B_d \in \mathcal{L}(W, X)$, $C_d, H_d \in \mathcal{L}(X, Y)$, and $D_d, V_d \in \mathcal{L}(W, Y)$. Moreover, we note by [125, Exe. 2.26b] that $-\delta$ is not in the point spectrum of A_d if A is the infinitesimal generator of a contraction semigroup. We apply Z-transform to transform a discrete signal w_k from the time domain to the frequency domain: $\hat{w}(z) = \sum_{k\in\mathbb{Z}} z^{-k} w_k$, for those $z \in \mathbb{C}$ for which the sum absolutely converges. We define a transfer function (or characteristic function in [24]) from w_k to y_k for discrete system (3.10) as follows:

$$\mathcal{G}_d(z) = C_d(zI - A_d)^{-1}B_d + \mathcal{D}_d \tag{3.12}$$

for $z \in \rho(A_d)$. Then,

$$\hat{y}(z) = C_d(zI - A_d)^{-1}x_0 + \mathcal{G}_d(z)\hat{w}(z), \text{ for } z \in \rho(A_d)$$

Therefore, once the discrete-time plant disturbance w_k and output z_k are estimated, they can be linked to their continuous counterparts, by applying $\mathbf{T}_h^*$, where the adjoint operator of $\mathbf{T}_h$, $\mathbf{T}_h^*$ is $1/\sqrt{h}$ times the zero-order hold operator on the interval $(kh, (k+1)h)$ [97].

Remark 6. *Based on the average sampling operation (3.9) and the mean value theorem, we note that the constraints on the continuous-time plant disturbance $w(t)$, measurement disturbance $v(t)$, and estimated output $z(t)$ remain identical to the constraints on the discrete-time counterparts $w_k/\sqrt{h}$, $v_k/\sqrt{h}$, and $z_k/\sqrt{h}$, respectively.*

3.4.2 Discrete-time Riccati equation

According to (3.6), a discrete infinite-time filtering problem can be formulated through minimizing: $J_d(x_0, w_d) = \sum_{n=-\infty}^{-1}(\|w_k\|^2_{Q^{-1}} + \|\bar{v}_k\|^2_{\bar{R}^{-1}})$, where for a given final state

$x_0 \in X$ the optimal cost is written in the form $\langle x_0, P_d^{-1} x_0 \rangle$. The optimal cost operator can be found by solving the discrete algebraic Riccati equation (DARE) [123]:

$$A_d P_d A_d^* - P_d + B_d Q B_d^* - L_d S_d L_d^* = 0 \tag{3.13}$$

where the minimal nonnegative self-adjoint solution is $P_d \in \mathcal{L}(X)$, with $L_d = A_d P_d C_d^* S_d^{-1}$, $S_d = C_d P_d C_d^* + \bar{R}$ and $\bar{R} = D_d Q D_d^T + R$, such that $A_d^{cl} = A_d - L_d C_d$ is a strongly stable operator.

The goal of the remainder of this section is to show that the optimal cost operator of functional (3.6) and its discrete-time counterpart are equal under the consideration of unbounded operators B and C (representing point-wise and/or boundary applied actuation and/or measurement). For simplicity, we first consider the case with bounded operators B and C according to Eq.(3.7) and Eq.(3.13), and provide the following lemma through algebraic manipulation based on [23, 113].

Lemma 2. *Let Σ be a regular linear system with bounded operators B and C, and let Σ_c be the corresponding discrete system using the CT transform with parameter δ. Then, i). a nonnegative self-adjoint invertible operator P is a solution of the CARE (3.7) iff it is a solution of the DARE (3.13); ii). The optimal cost operator P^{-1} for the continuous infinite-time filtering problem of Σ equals the optimal cost operator P_d^{-1} for the discrete infinite-time filtering problem of Σ_c.*

In frequency domain, for $\delta > r \geq 0$, the CT transform (i.e., $z = \frac{\delta+s}{\delta-s}$ and $s = \frac{z-1}{z+1}\delta$) maps the right half-plane $\mathbb{C}_r^+$ bijectively onto the exterior disc $\mathbb{D}_r^+$ with center $r/(\delta-r)$ and radius $\delta/(\delta-r)$, and thus induces a unitary transformation between $\mathbf{H}^2(\mathbb{C}_r^+)$ and $\mathbf{H}^2(\mathbb{D}_r^+)$ by $(\mathcal{F}g)(z) = \frac{\sqrt{2\delta}}{z+1} g\left(\frac{z-1}{z+1}\delta\right)$ and $(\mathcal{F}^{-1}f)(s) = \frac{\sqrt{2\delta}}{\delta-s} f\left(\frac{\delta+s}{\delta-s}\right)$, see [21, Cha. 12.1] and [24]. Based on that, we show that the following lemma links model (3.5) to (3.10) in frequency domain under the consideration of unbounded operators B and C.

Lemma 3. *Let Σ be a regular linear system, Σ^d be the dual well-posed system of Σ, and Σ_c^d be the corresponding discrete system of Σ^d using the CT transform with parameter δ. Then*

$$B_d^*(zI - A_d^*)^{-1} = \frac{\delta - s}{\sqrt{2\delta}} B^*(sI - A^*)^{-1} \tag{3.14a}$$

$$B^*(sI-A^*)^{-1}=\frac{z+1}{\sqrt{2\delta}}B_d^*(zI-A_d^*)^{-1},\ \mathcal{G}_d^d(z)=\mathcal{G}^d(s) \tag{3.14b}$$

where $z=\frac{\delta+s}{\delta-s}\in\rho(A_d^*)\backslash\{-1\}$, $s=\frac{z-1}{z+1}\delta\in\mathbb{C}_{w_0}^+\cap\rho(A^*)\backslash\{\delta\}$, $\delta\in\mathbb{R}^+$, *and* $\mathcal{G}^d(s)$ *and* $\mathcal{G}_d^d(z)$ *denote the transfer functions of systems* Σ^d *and* Σ_c^d, *respectively.*

Proof. The proof is similar to [24, Lem. 8] by replacing z by $1/z$ with consideration of real valued $\delta\in\mathbb{R}^+$. ■

The following theorem provides the key connection between the admissible output set $\mathfrak{Y}(x_0)$ of the regular linear system Σ and the admissible output set $\mathfrak{Y}_d(x_0)$ of the corresponding discrete system Σ_c using the CT transform with a suitable parameter δ. Specifically, we introduce the time-domain version the CT transform, Laguerre transform $\mathcal{F}_d$ [21, Def. 12.3.2], which is an isometric isomorphism.

Theorem 3. *Let* Σ *be a regular linear system, and* Σ_c *be the corresponding discrete system using CT transform with parameter* $\delta\in\mathbb{C}_{w_0}^+\cap\mathbb{R}^+$. *Let Assumptions 1 and 4 hold.*

(i). Let w *be the plant disturbance of* Σ *for the final state* x_0 *and output* $y\in\mathfrak{Y}(x_0)$. *Then,* $w_d=\mathcal{F}_d w$ *is the plant disturbance of* Σ_c *for the final state* x_0 *and output* $y_d=\mathcal{F}_d y$.

(ii). Let w_d *be the plant disturbance of* Σ_c *for the final state* x_0 *and output* $y_d\in\mathfrak{Y}_d(x_0)$. *Then,* $w=\mathcal{F}_d^* w_d$ *is the plant disturbance of* Σ *for the final state* x_0 *and the output* $y=\mathcal{F}_d^* y_d$.

Proof. The proof is completed by appealing to the duality between optimal control and filtering problems and by the assumption that both the plant disturbance space W and output space Y are finite dimensional. We first show the proof of (i), then the proof of (ii) will be similar. Let us consider the dual system Σ^d of Σ as:

$$\dot{z}(t)=A^*z(t)+C^*\bar{u}(t),\ z(0)=x_0\in X,\forall t\geq 0 \tag{3.15a}$$

$$\bar{y}(t)=B^*z(t) \tag{3.15b}$$

where $z\in X$, $\bar{u}=\mathfrak{R}y\in L^2_{\mathrm{loc}}(\mathbb{R}_0^+;Y)$, and $\bar{y}=\mathfrak{R}w\in L^2_{\mathrm{loc}}(\mathbb{R}_0^+;W)$ represent the state, control and observation. By Theorem 2, we note that Σ^d is a well-posed system and A^* is a generator of the semigroup T^d with the growth bound w_0. Let Σ_c^d be the discrete system of Σ^d using CT transform with parameter $\delta\in\mathbb{C}_{w_0}^+\cap\mathbb{R}^+$.

Define the quadratic cost functional for system (3.15)

$$\bar{J}(x_0, \bar{u}) = \int_0^\infty \|\bar{u}(t)\|_R^2 + \|\bar{y}(t)\|_Q^2 \, dt \tag{3.16}$$

It can be seen that this problem is same as the Problem 2 in [24] for the system $[T^d, \Phi^d; [\Psi^d; 0], [F^d; I]]$ with the cost operator $J = [Q, 0; 0, R]$. Hence, the positive J-coercivity automatically holds. Under Assumption 1, we note (A^*, C^*) is optimizable. Hence there exists a non-negative self-adjoint operator $\bar{P}$ such that:

$$\inf_{\bar{u} \in \mathfrak{Y}(x_0)} Q(x_0, u) = \left\langle x_0, \bar{P} x_0 \right\rangle_X \tag{3.17}$$

where $\mathfrak{Y}(x_0) = \{\bar{u} \in L^2_{\text{loc}}(\mathbb{R}_0^+; Y) | \bar{y} = \Psi_\infty^d x_0 + F_\infty^d \bar{u} \in L^2_{\text{loc}}(\mathbb{R}_0^+; W)$ and cost (3.16) is finite$\}$ denotes a non-empty set of admissible inputs.

Define $\max\{0, w_0\} \leq r \leq \delta$. Since $\bar{y} \in L^2(\mathbb{R}_0^+, W)$, we have z on the exterior disc $\mathbb{D}_r^+$:

$$\hat{\bar{y}}_d(z) = (\mathcal{F}\hat{\bar{y}})(z) = \frac{\sqrt{2\delta}}{z+1} \hat{\bar{y}}\left(\frac{z-1}{z+1}\delta\right) = \frac{\sqrt{2\delta}}{z+1} \hat{\bar{y}}(s) \tag{3.18}$$

Due to $s = \frac{z-1}{z+1}\delta \in \mathbb{C}_r^+$, we have the following relation of $\bar{y}$ with respect to $\bar{u}$ and x_0 as in (3.3):

$$\hat{\bar{y}}(s) = B^*(sI - A^*)^{-1} x_0 + \mathcal{G}^d(s) \hat{\bar{u}}(s) \tag{3.19}$$

where $\mathcal{G}^d(s)$ is the transfer function of Σ^d. Inserting (3.19) into (3.18) and applying Lemma 3, we then have:

$$\hat{\bar{y}}_d(z) = B_d^*(zI - A_d^*)^{-1} x_0 + \frac{\sqrt{2\delta}}{z+1} \mathcal{G}_d^d(z) \hat{\bar{u}}(s) \tag{3.20}$$

For $\bar{u} \in L^2(\mathbb{R}_0^+, Y)$, we note for $z \in \rho(A_d^*) \backslash \{-1\}$ on the exterior disc $\mathbb{D}_r^+$, $\hat{\bar{u}}_d(z) = \frac{\sqrt{2\delta}}{z+1} \hat{\bar{u}}(s)$. Thus, we can show that $\bar{y}_d$ is indeed the output of the system Σ_c^d for the initial state x_0 and input $\bar{u}_d$. Hence, $w_d = \mathcal{F}_d w$ is the plant disturbance of Σ_c for the final state x_0 and output $y_d = \mathcal{F}_d y$. ■

The following theorem provides an alternative way to find the optimal cost operator of the regular linear system Σ by solving for the minimal nonnegative self-adjoint solution of the discrete-time Riccati equation (3.13) via Cayley-Tustin and Laguerre transforms.

Theorem 4. *Let Σ be a regular linear system, and Σ_c be the discrete system using CT transform with parameter $\delta \in \mathbb{C}^+_{w_0} \cap \mathbb{R}^+$. If Σ is estimatable, then so is Σ_c and the optimal cost operators of Σ and Σ_c are equal. Moreover, the estimated plant disturbances are linked by $w_d = \mathcal{F}_d w$.*

Proof. From Theorem 3, we note that Cayley-Tustin and Laguerre transforms indeed are isometric isomorphisms between stable continuous-time and discrete-time systems and signals, and it follows that the admissible output sets in continuous- and discrete-time settings are isomorphic under these transforms. Thus, the regular linear system Σ is estimatable iff its discrete counterpart Σ_c is estimatable. Moreover, for any final state $x_0 \in X$ and output $y \in \mathfrak{Y}(x_0)$ we have $\inf J(x_0, w) = \inf J_d(x_0, w_d)$. The estimated continuous- and discrete-time disturbances can be linked by $w_d = \mathcal{F}_d w$ following Theorem 3. ■

Although Laguerre transform as a unitary transform inherits most global properties, it does not preserve local properties, i.e., if a finite-part of a continuous-time output signal $y(t)$ is known, then one cannot infer a finite-part of the corresponding discrete-time output signal y_k [21, Cha. 12.3]. Thus, it is not as useful as the average sampling operator $\mathbf{T}_h$ for practical applications. Furthermore, considering that in general y_k may not be sampled by $\mathbf{T}_h$ but rather by a zero-order holder device, the following proposition shows that using the ordinary zero-order holder device with a fast sampling rate ($h \to 0$) can provide a good approximation of plant disturbance $w(t)$.

Proposition 3. *Assume $\mathcal{G}(s)$ is a stable transfer function of regular linear system (3.5), and $\mathfrak{R}y \in L^2(\mathbb{R}^+; Y)$. Then, as $h \to 0$,*

$$||\mathbf{S}_h^* \mathbf{G}_d^d \mathbf{S}_h \mathfrak{R} y - \mathcal{F}_d^* \mathbf{G}_d^d \mathcal{F}_d \mathfrak{R} y||_{L^2(\mathbb{R}^+;W)} \to 0 \tag{3.21a}$$

$$||\mathbf{G}_d^d \mathbf{S}_h \mathfrak{R} y - \mathbf{S}_h \mathcal{F}_d^* \mathbf{G}_d^d \mathcal{F}_d \mathfrak{R} y||_{l^2(\mathbb{Z}^+;W)} \to 0 \tag{3.21b}$$

where $\mathbf{G}_d^d$ denotes the time domain convolution operator corresponding to $\mathcal{G}_d^d(z)$, and $\mathbf{S}_h$ represents an ordinary sampling operator (i.e., by a zero-order holder).

Proof. The main idea follows [120, Pro. 2]. Given that $\mathcal{F}_d$ is a unitary operator, we note that Laguerre transform induces an exact mapping as: $\mathcal{F}_d^* \mathbf{G}_d^d \mathcal{F}_d = F_\infty^d$ [21,

The 12.3.7]. By [97, The. 4.3], we have $||\mathbf{T}_h^* \mathbf{G}_d^d \mathbf{T}_h \mathfrak{R} y - \mathcal{F}_d^* \mathbf{G}_d^d \mathcal{F}_d \mathfrak{R} y||_{L^2(\mathbb{R}^+;W)} \to 0$ as $h \to 0$. Since $\mathbf{T}_h^*$ is isometric, we also have $||\mathbf{G}_d^d \mathbf{T}_h \mathfrak{R} y - \mathbf{T}_h \mathcal{F}_d^* \mathbf{G}_d^d \mathcal{F}_d \mathfrak{R} y||_{L^2(\mathbb{R}^+;W)} \to 0$ as $h \to 0$. Then, we note that $||(\mathbf{T}_h - \mathbf{S}_h) f||_{l^2(\mathbb{Z}^+;V)} \to 0$ with f being defined as $\mathfrak{R} y$ or $F_\infty^d \mathfrak{R} y$ (i.e., $\mathfrak{R}$w), and V being Y or W. By applying [120, Pro. 2] and the triangle inequality, we conclude (3.21). ■

Considering that directly designing MHE for continuous system (3.5) can be challenging due to the presence of unbounded operators and that the MHE is usually designed for discrete-time models (e.g., see [38, 108]), we provide MHE design and the stability analysis for the discrete model (3.10) in the following section. By [97, The. 4.3] and Proposition 3, the estimation results of the discrete model (3.10) can be linked back to those of the continuous model (3.5), with a small time discretization interval.

3.5 Moving horizon estimator design

This section proposes a moving horizon estimator for regular linear systems explicitly accounting for physical constraints. This is done by extending the well-known MHE design for lumped parameter systems [37, 38]. In what follows, we first construct a full information estimator and then link it to the MHE design.

3.5.1 Full information estimation

Consider a constrained linear full information estimation (FIE) problem for system (3.10):

$$\mathbf{\Phi}_T^* = \min_{q, \{\hat{w}_k\}_{k=0}^{T-1}} \mathbf{\Phi}_T(q, \{\hat{w}_k\}) \tag{3.22a}$$

$$\text{s.t. } B_d \hat{w}_k = \hat{x}_{k+1|T} - A_d \hat{x}_{k|T},\ \hat{v}_k = y_k - C_d \hat{x}_{k|T} \tag{3.22b}$$

$$z_{\min} \le H_d \hat{x}_{k|T} + V_d \hat{w}_k \le z_{\max}, \tag{3.22c}$$

$$w_{\min} \le \hat{w}_k \le w_{\max}, v_{\min} \le \hat{v}_k \le v_{\max} \tag{3.22d}$$

where $\hat{w}_k$ and $\hat{v}_k$ denote the estimated discrete plant and measurement disturbances. We consider linear equality constraints (3.22b) by following the system model (3.10) and linear inequality constraints (3.22c)-(3.22d) to prevent physically unrealistic output estimates due to spurious measurements. The cost functional $\mathbf{\Phi}_T$ takes a quadratic

form using all measurements up to time instant T as follows:

$$\Phi_T(q, \{\hat{w}_k\}) = \sum_{k=0}^{T-1} \left\langle \hat{w}_k, Q^{-1}\hat{w}_k \right\rangle_W + \sum_{k=0}^{T} \left\langle \hat{v}_k, \bar{R}^{-1}\hat{v}_k \right\rangle_Y + \left\langle q - \hat{x}_{0|-1}, P_0^{-1}(q - \hat{x}_{0|-1}) \right\rangle_X \tag{3.23}$$

where $\hat{x}_{0|-1}$ denotes a prior estimate of the state at time instant 0, $q = \hat{x}_{0|T}$ represents initial state estimate at time instant T, and P_0^{-1} is a self-adjoint positive-definite operator penalizing the initial state estimation error. The prior estimate $\hat{x}_{0|-1}$ should satisfy the constraints (3.22b)-(3.22d). Q^{-1} and $\bar{R}^{-1}$ (as aforementioned $\bar{R} = D_d Q D_d^T + R$) are symmetric positive-definite penalty matrices on plant and measurement disturbances. Considering directly optimizing (3.22) with respect to infinite-dimensional state $q = \hat{x}_{0|T}$ can be intractable in practical implementation (as the spatial discretization node of the initial state might go to infinity in order to obtain a good estimation result), we deploy the following suitable approximation. By denoting $q - A_d\hat{x}_{0|-1} = B_d\hat{w}_{-1}$ as in [37], we note that optimizing $\Phi_T(q, \{\hat{w}_k\}_{k=0}^{T-1})$ with respect to q and $\{\hat{w}_k\}_{k=0}^{T-1}$ is converted to optimize $\Phi_T(\{\hat{w}_k\}_{k=-1}^{T-1})$ in terms of $\{\hat{w}_k\}_{k=-1}^{T-1}$.

3.5.2 Moving horizon estimation

With the measurements being accumulated, FIE may become intractable. For efficient estimation, moving horizon estimation is formulated by deploying a moving horizon strategy. With a fixed horizon length $N < T$, we formulate the moving horizon estimation as the solution to the following quadratic program problem:

$$\hat{\Phi}_T^* = \min_{q, \{\hat{w}_k\}_{k=T-N}^{T-1}} \hat{\Phi}_T(q, \{\hat{w}_k\}) \tag{3.24}$$

subjected to the constraints (3.22b)-(3.22d). $\hat{\Phi}_T^*$ approximates the optimal cost Φ_T^* in the full information estimation (3.22a) at time T. The cost functional $\hat{\Phi}_T$ is defined as follows:

$$\begin{aligned} \hat{\Phi}_T(q, \{\hat{w}_k\}) =& \hat{\Phi}_T^N(q, \{\hat{w}_k\}_{k=T-N}^{T-1}) + \hat{\Phi}_{T-N}^* \\ =& \sum_{k=T-N}^{T-1} \left\langle \hat{w}_k, Q^{-1}\hat{w}_k \right\rangle_W + \sum_{k=T-N}^{T} \left\langle \hat{v}_k, \bar{R}^{-1}\hat{v}_k \right\rangle_V \\ &+ \left\langle q - \hat{x}_{T-N|T-N-1}, P_\infty^{-1}(q - \hat{x}_{T-N|T-N-1}) \right\rangle_X + \hat{\Phi}_{T-N}^* \end{aligned} \tag{3.25}$$

where $q = \hat{x}_{T-N|T}$, and the cost function $\hat{\Phi}_T$ is divided into two parts, including $\hat{\Phi}^*_{T-N}$ that approximates the optimal cost Φ^*_{T-N} at time $T-N$ and $\hat{\Phi}^N_T(q, \{\hat{w}_k\}^{T-1}_{k=T-N})$ as shown in (3.25). The last two terms in (3.25) are utilized to approximate the arrival cost summarizing all previous information up to time instant $T-N$. For constrained FIE and MHE designs of finite-dimensional systems, it has been shown that using steady-state Kalman filter gain is feasible to approximately summarize previous information in an explicit way [37, 38]. Motivated by that fact, P_∞ solved as the minimal nonnegative self-adjoint solution of Eq.(3.13) denotes a penalty operator on "initial" state (i.e., $\hat{x}_{T-N|T}$) estimation error. Similar to the FIE design, we introduce $q - A_d\hat{x}_{T-N|T-N-1} = B_d\hat{w}_{T-N-1}$ for practical deployment, so optimizing $\hat{\Phi}^N_T(q, \{\hat{w}_k\}^{T-1}_{k=T-N})$ with respect to q and $\{\hat{w}_k\}^{T-1}_{k=T-N}$ is converted to optimize $\hat{\Phi}^N_T(\{\hat{w}_k\}^{T-1}_{k=T-N-1})$ in terms of $\{\hat{w}_k\}^{T-1}_{k=T-N-1}$. Through direct manipulation of (3.25), we obtain the following quadratic program problem:

$$\hat{\Phi}^N_T(\{\hat{w}_k\}^{T-1}_{k=T-N-1}) = \langle W_k, \bar{H}W_k\rangle_{W^{N+1}} + 2\langle W_k, \bar{P}Y_k\rangle_{W^{N+1}} + \langle Y_k, \bar{Q}Y_k\rangle_{Y^{N+1}} \tag{3.26}$$

where $\bar{H} \in \mathcal{L}(W^{N+1})$ is positive-definite and self-adjoint given by

$$\bar{h}_{i,j} = \begin{cases} \sum\limits_{k=0}^{N+1-i} B^*_d(A^k_d)^* C^*_d \bar{R}^{-1} C_d A^k_d B_d + Q^{-1}, & \text{for } i = j \\ \sum\limits_{k=j-i}^{N+1-i} B^*_d(A^k_d)^* C^*_d \bar{R}^{-1} C_d A^{k-j+i}_d B_d, & \text{for } i < j \\ \bar{h}^*_{j,i}, & \text{for } i > j \end{cases}$$

where Q^{-1} should be replaced by $B^*_d P^{-1}_\infty B_d$ when $i = j = 1$. Moreover, we denote $Y_k = \left(y_{T-N+k} - C_d A^k_d \hat{x}_{T-N|T-N-1}\right)^{k=0}_{k=N} \in Y^{N+1}$, and $\bar{P} \in \mathcal{L}(Y^{N+1}, W^{N+1})$ with $\bar{p}_{i,j} = B^*_d(A^*_d)^{j-i}C^*_d\bar{R}^{-1}$ for $i \le j$ and elsewise $\bar{p}_{i,j} = 0$. The cost functional (3.26) is subjected to constraints (3.22b)-(3.22d), i.e., $Z_{\min} \le \hat{Z}_k = T_k\hat{x}_{T-N|T-N-1} + S_k W_k \le Z_{\max}$, $V_{\min} \le V_k = Y_k - \bar{U}_k W_k \le V_{\max}$, $W_{\min} \le W_k \le W_{\max}$, leading to

$$\begin{bmatrix} I \\ -I \\ S_k \\ -S_k \\ \bar{U}_k \\ -\bar{U}_k \end{bmatrix} W_k \le \begin{bmatrix} W_{\max} \\ -W_{\min} \\ Z_{\max} - T_k\hat{x}_{T-N|T-N-1} \\ T_k\hat{x}_{T-N|T-N-1} - Z_{\min} \\ Y_k - V_{\min} \\ V_{\max} - Y_k \end{bmatrix} \tag{3.27}$$

where $W_{\max} = w_{\max} \cdot \mathbf{1}_{N+1}$ with $\mathbf{1}_{N+1}$ being the column vector with all elements 1, and $W_{\min}$, $Z_{\max}$, $Z_{\min}$, $V_{\max}$ and $V_{\min}$ are similarly defined. Moreover, $W_k =$

$(\hat{w}_{T-N+k-1})_{k=N}^{k=0} \in W^{N+1}$, $T_k = H_d(A_d)_{k=N}^{k=0} \in \mathcal{L}(X^{N+1}, Y^{N+1})$, $S_k \in \mathcal{L}(W^{N+1}, Y^{N+1})$ and $\bar{U}_k \in \mathcal{L}(W^{N+1}, Y^{N+1})$ are given by: $s_{i,j} = H_d(A_d)^{i-j}B_d$ for $i \geq j$, $s_{i,j} = V_d$ for $i = j-1$ and elsewise $s_{i,j} = 0$; $\bar{u}_{i,j} = C_d(A_d)^{i-j}B_d$ for $i \geq j$ and elsewise $\bar{u}_{i,j} = 0$. Considering finite-dimensional spaces $W = Y = \mathbb{C}$, the inner products become vector products in the cost functional (3.26), and hence we have a finite-dimensional quadratic optimization problem:

$$\min_{W_k} \hat{\Phi}_T^N(W_k) = \min_{W_k} W_k^T H W_k + 2W_k^T(\bar{P}Y_k) \tag{3.28}$$

Note that the term $\langle Y_k, \bar{Q}Y_k\rangle_{Y^{N+1}}$ can be neglected in the cost functional (3.26) since it is not affected by W_k. Given that all operators utilized in the objective function (3.28) and the constraints (3.27) are bounded under the Cayley-Tustin transformation, the obtained quadratic optimization problem (3.28) is bounded and exactly the same as the ones existing in finite-dimensional cases, e.g., [38, Eq. (5)]. Therefore, the existing MHE theory for finite-dimensional systems can be freely inherited.

3.5.3 Stability analysis

Unlike controller design, the estimator has no control over the state evolution of the closed-loop system and the cost functional (3.22a) may increase without upper bound as time increases [38]. In addition, an improper choice of constraints may prevent the estimated state converging to the true state. To analyze the stability of the proposed estimator in the sense of an observer, we need the following definition and assumption, see [38] for more details.

Definition 3. *The state estimator is an asymptotically stable estimator for the following system:*

$$\tilde{x}_{k+1} = A_d\tilde{x}_k,\ \tilde{y}_k = C_d\tilde{x}_k,\ \tilde{x}_0 = x_0 \in X \tag{3.29}$$

if for any $\varepsilon > 0$, there exist a number $\bar{\delta} > 0$ and a positive integer $\bar{N}$ such that if $\|\tilde{x}_0 - \hat{x}_{0|-1}\|_X \leq \bar{\delta}$ and $z_{\min} \leq H_d\hat{x}_{0|-1} \leq z_{\max}$, then $\|\hat{x}_T - A_d^T\tilde{x}_0\|_X \leq \varepsilon$ for all $T \geq \bar{N}$, and $\lim_{T\to+\infty}\|\hat{x}_T - A_d^T\tilde{x}_0\|_X = 0$.

Assumption 5. *(Finite cost assumption) Suppose that system (3.29) with initial condition $\tilde{x}_0 = x_0$ generates outputs $\{\tilde{y}_k = C_dA_d^k\tilde{x}_0\}_{k=0}^{\infty}$. We assume that there exist*

$\hat{x}_{0|+\infty}$, $\{\hat{w}_{k|+\infty}\}_{k=0}^{+\infty}$, *and a finite number* $\sigma > 0$, *such that*
$\sum_{k=0}^{+\infty} \left\langle \hat{w}_{k|+\infty}, Q^{-1}\hat{w}_{k|+\infty} \right\rangle_W + \left\langle \hat{v}_{k|+\infty}, \bar{R}^{-1}\hat{v}_{k|+\infty} \right\rangle_Y + \left\langle \hat{x}_{0|+\infty} - \hat{x}_{0|-1}, P_0^{-1}(\hat{x}_{0|+\infty} - \hat{x}_{0|-1}) \right\rangle_X \leq$
$\sigma \langle x_0 - \hat{x}_{0|-1}, x_0 - \hat{x}_{0|-1} \rangle_X$
where $z_{\min} \leq H_d\hat{x}_{k|+\infty} + V_d\hat{w}_{k|+\infty} \leq z_{\max}, w_{\min} \leq \hat{w}_{k|+\infty} \leq w_{\max}, v_{\min} \leq \hat{v}_{k|+\infty} \leq v_{\max}$,
$\hat{x}_{k+1|+\infty} = A_d\hat{x}_{k|+\infty} + B_d\hat{w}_{k|+\infty}$, *and* $\hat{v}_{k|+\infty} = \tilde{y}_k - C_d\hat{x}_{k|+\infty}$.

Conceptually, Assumption 5 states that if we consider an infinite amount of data generated by system (3.29), then there exist feasible state and disturbance solutions leading to a finite cost. This assumption can be replaced by the feasibility of optimization problem (3.22) as $T \to +\infty$.

To prove the stability of MHE, we need to show the finite-time exact observability of the discrete system (A_d, C_d) so as to ensure the invertibility of the discrete-time observability Gramian by the following proposition.

Proposition 4. *Consider the regular linear system (3.5) and its discrete-time system (3.10) using the Cayley-Tustin transform with a suitable parameter* δ. *Let Assumptions 2-4 hold. Then* (A_d, C_d) *is exactly observable in finite time* $K > 0$.

Proof. The rational of the proof follows from the continuous-time case [42, Pro. 6.5.2.]. As C is infinite-time admissible for $\mathbb{T}$, there corresponds an infinite-time observability Gramian $Q \in \mathcal{L}(X)$:

$$Qz = \lim_{\tau \to +\infty} \int_0^\tau \mathbb{T}_t^* C^* C \mathbb{T}_t z \, dt, \ \forall z \in \mathcal{D}(A) \tag{3.30}$$

and it is the unique self-adjoint solution of the following Lyapunov equation in the dual space of X_{-1} [42, The. 5.1.1]:

$$A^*Qz + QAz = -CC^*z, \ \forall z \in \mathcal{D}(A) \tag{3.31}$$

or equivalently solved from the discrete-time Lyapunov equation [23, The. 2.4]:

$$A_d^* Q^d A_d x - Q^d x = -C_d C_d^* x, \ \forall x \in X \tag{3.32}$$

By using the CT transform, we note that A_d is strongly stable and a bounded contractive operator since A generates a strongly stable contraction C_0-semigroup [21, The. 3.4.9, The. 12.3.10]. Since A_d is strongly stable and C_d is infinite-time admissible

(induced by the infinite-time admissibility of C) [23, Lem. 2.2, The. 2.4] there exists a unique solution of (3.32) named as the infinite-time discrete observability Gramian $\mathcal{Q}^d \in \mathcal{L}(X)$ defined as:

$$\mathcal{Q}^d x = \lim_{K \to +\infty} \sum_{k=0}^{K} (A_d^*)^k C_d^* C_d (A_d)^k x, \ \forall x \in X \tag{3.33}$$

Hence, $\mathcal{Q}$ and $\mathcal{Q}^d$ are equal and their positive coercivity is invariant under the Cayley-Tustin transform. Thus, the infinite-time exact observability of (A_d, C_d) follows from that of (A, C) under the CT transform. Next, we show that the finite-time exact observability of (A_d, C_d) follows from that of (A, C).

For the infinite-time admissible C and by the exact observability of (A, C) in finite time τ, we have the lower and upper bounds for all $z \in \mathcal{D}(A)$ as:

$$\begin{aligned} \mu^2 \|z\|_X^2 \geq \langle \mathcal{Q} z, z \rangle &= \int_0^{+\infty} \|C T_t z\|_Y^2 \, dt \\ &= \int_0^{\tau} \|C T_t z\|_Y^2 \, dt + \int_{\tau}^{+\infty} \|C T_t z\|_Y^2 \, dt \geq \kappa^2 \|z\|_X^2 \end{aligned}$$

where $\mu > \kappa > 0$. Thus, we have the lower and upper bounds for the discrete observability Gramian $\mathcal{Q}^d$ as:

$$\mu^2 \|z\|_X^2 \geq \left\langle \mathcal{Q}^d z, z \right\rangle = \sum_{k=0}^{+\infty} \left\| C_d (A_d)^k z \right\|_Y^2 \geq \kappa^2 \|z\|_X^2$$

By decomposition in time as in [42, Pro. 6.5.2.], we have

$$\begin{aligned} &\sum_{k=0}^{K} \|C_d (A_d)^k z\|_Y^2 \\ =& \sum_{k=0}^{+\infty} \left\| C_d (A_d)^k z \right\|_Y^2 - \sum_{k=K+1}^{+\infty} \left\| C_d (A_d)^k z \right\|_Y^2 \\ =& \sum_{k=0}^{+\infty} \left\| C_d (A_d)^k z \right\|_Y^2 - \sum_{k=0}^{+\infty} \left\| C_d (A_d)^{k+K+1} z \right\|_Y^2 \\ \geq& (\kappa^2 - \mu^2 \|(A_d)^{K+1}\|^2) \cdot \|z\|_X^2 \end{aligned} \tag{3.34}$$

Since A_d is strongly stable and a bounded contractive operator (i.e., $\|A_d z\| \leq \|z\|$ and $A_d^k z \to 0$ as $k \to +\infty$), the parenthesis in (3.34) becomes positive for some sufficiently large K. For such K, (A_d, C_d) is exactly observable. ■

Theorem 5. *Assume that Q and R are positive definite matrices, and P_0 is a self-adjoint positively coercive operator. Under Assumptions 2-5, the constrained full information estimator (3.22) is an asymptotically stable estimator for the system (3.29).*

Proof. By $\bar{R} = D_d Q D_d^T + R$, we note that $\bar{R}$ is positive definite since Q and R are and $D_d \in \mathcal{L}(W, Y)$ is a matrix by Assumption 4 that supposes the plant disturbance space W and output space Y are finite dimensional. Under Assumptions 2-4 and using Proposition 4, (A_d, C_d) is exactly observable in finite time K. Under Assumptions 2-5, the resulting full information estimator is equivalent to the ones designed in linear finite-dimensional cases (e.g., [38, Eq. (2)]), and hence the results follow from the standard finite-dimensional full information estimation theory, e.g. [126, Pro. 4.4.3]. ■

Theorem 6. *Assume that Q and R are positive definite matrices, and the minimal nonnegative self-adjoint solution P_∞ of (3.13) is invertible. Under Assumptions 2-5, for $N \geq K$, the constrained moving horizon estimator is an asymptotically stable estimator for the system (3.29).*

Proof. By Assumptions 2-4 and using Proposition 4, we note that (A_d, C_d) is exactly observable in finite time K. By the preceding arguments, the quadratic optimization problems (3.24)-(3.28) formulated in the MHE design of discrete-time infinite-dimensional system (3.10) are the same as those in finite dimensional MHE problems (e.g., [38, Eq. (5)]). Under Assumption 4 (having finite-dimensional disturbance and output spaces), the optimality of the proposed MHE follows from the finite-dimensional MHE cases. By using optimality and Assumption 5, a convergence analysis can be done as follows: $\sum_{k=0}^{T-1} \left\langle \hat{w}_k, Q^{-1} \hat{w}_k \right\rangle_W + \sum_{k=0}^{T} \left\langle \hat{v}_k, \bar{R}^{-1} \hat{v}_k \right\rangle_Y \to 0$ as $T \to \infty$ by performing similar derivations as in the proof of [38, Pro. 5]. Based on [38, Lem. 3], the estimation error $\|\hat{x}_T - A_d^T x_0\|_X$ converges to zero as T goes to infinity. Thus, the stability analysis of the proposed constrained moving horizon estimator follows from the finite-dimensional moving horizon estimation theory, e.g. [38, Pro. 5, Cor. 7]. ■

By choosing feasible constraints, the designed discrete-time MHE shall provide a good estimate of the continuous-time regular system using measurements with a fast

sampling rate. In Section 3.6, the MHE design of a damped Schrödinger equation as a stable system is illustrated as an example.

3.5.4 Extension to unstable systems

For unstable systems, we introduce a shifted semigroup generated by $A - \alpha I$, where α is greater than the growth bound w_0 of the semigroup T generated by A, with (B, C, D) being unchanged, resulting a shifted semigroup that is a strongly stable contraction. The admissibility of C and exact observability of (A, C) of the shifted system are equivalent to those of the original system [42, Sec. 6.5]. As in [127, Def. 3.1], we define $\Sigma_\tau^{-\alpha}$ $(\forall \tau \geq 0)$ as the exponential shift of Σ_τ by the amount of $-\alpha \in \mathbb{C}$:

$$\Sigma_\tau^{-\alpha} = \begin{bmatrix} \mathrm{T}_\tau^{-\alpha} & \Phi_\tau^{-\alpha} \\ \Psi_\tau^{-\alpha} & F_\tau^{-\alpha} \end{bmatrix} := \begin{bmatrix} e^{-\alpha\tau}\mathrm{T}_\tau & e^{-\alpha\tau}\Phi_\tau e_\alpha \\ e_{-\alpha}\Psi_\tau & e_{-\alpha}F_\tau e_\alpha \end{bmatrix} \tag{3.35}$$

where $e_{-\alpha}$ denotes the scalar function $(e_{-\alpha}f)(t) = e^{-\alpha t}f(t)$, $t \in \mathbb{R}_0^+$ and $\alpha > w_0$. For the stable shifted system $\Sigma_\tau^{-\alpha}$, we can perform system discretization as in Eqs.(3.10)-(3.11) by replacing $\mathcal{R}(\delta, A)$ by $\mathcal{R}(\delta + \alpha, A)$ and the MHE design. The continuous- and discrete-time plant and measurement disturbances stay unaffected under this shift transform. By using the moving horizon estimator designed for the discrete-time system corresponding to shifted system (3.35), the estimated discrete-time output and disturbance can provide good approximations for the continuous-time output and disturbance of the shifted system with a small time discretization interval. Afterwards, one can freely convert the estimation results of the shifted system to those of the original system by the performing an invertible shift transform e_α since the exponential shift transformation is invertible. The MHE design for an unstable wave equation is given as another example in Section 3.6.

3.6 Examples

3.6.1 Example 1: Schrödinger equation

Consider a stable Schrödinger equation on a 1-D spatial domain $\zeta \in [0, 1]$ with Neumann boundary disturbance and collocated observation:

$$x_t(\zeta, t) = -jx_{\zeta\zeta}(\zeta, t) - vx(\zeta, t) \tag{3.36a}$$

$$x_\zeta(0,t)=0,\ x_\zeta(1,t)=w(t),\ x(\zeta,0)=x_0(\zeta) \tag{3.36b}$$

$$y(t)=x(1,t)+v(t),\ z(t)=x(0,t) \tag{3.36c}$$

where $v>0$ denotes the potential energy of the particle, j is the imaginary unit, and x is the complex-valued state in the state space $X=L^2(0,1)$. We consider the plant disturbance and output spaces $W=Y=\mathbb{C}$. By [42, Rem. 10.1.6], we can formulate system (3.36) in the standard state-space form (3.5) where $Af:=-j\frac{\partial^2 f}{\partial\zeta^2}-vf$ with its domain $\mathcal{D}(A)=\{f\in H^2(0,1;\mathbb{C})|f'(0)=f'(1)=0\}$, and B, C and H can be found as: $B:=j\delta_d(\zeta-1)$, $C:=\langle\delta_d(\zeta-1),\cdot\rangle$ and $H:=\langle\delta_d(\zeta),\cdot\rangle$, where $\delta_d(\cdot)$ denotes the Dirac function. It has been shown in [128, 129] that system (3.36) is well-posed [130]. By [129, The. 2.1], we find system (3.36) (with $v=0$) can be exponentially stabilized under the output feedback $u=-jky$ with $k>0$ which indicates that (A_0,B) is exactly controllable in any finite time $\tau>0$, see [131], where $A_0f:=-j\frac{\partial^2 f}{\partial\zeta^2}$ and $\mathcal{D}(A_0)=\mathcal{D}(A)$. Since $A_0^*=-A_0$ and $C=B^*$, the exact controllability of (A_0,B) is equivalent to the exact observability of (A_0,C) [42, The. 11.2.1], which implies that (A,C) is exactly observable in finite time $\tau>0$ by [42, The. 6.7.2]. By [42, Pro. 3.7.2], we note that A_0 is m-dissipative since it is skew-adjoint, which implies that A is also m-dissipative. Based on the Lumer–Phillips theorem [42, The. 3.8.4], A generates a contraction semigroup. A simple spectral analysis shows that system (3.36) is exponentially stable. Based on that, we note that C is infinite-time admissible [122, Pro. 5.5].

By performing Laplace transform of (3.36a) with the boundary conditions (3.36b), the resolvent operator is determined in the closed analytic form:

$$\begin{aligned}\mathcal{R}(s,A):=&\frac{1}{\sqrt{-w_s}}\times\Big[-\int_0^\zeta\sinh(\sqrt{w_s}(\zeta-\eta))(\cdot)d\eta\\&+\frac{\cosh(\sqrt{w_s}\zeta)}{\sinh(\sqrt{w_s})}\int_0^1\cosh(\sqrt{w_s}(1-\eta))(\cdot)d\eta\Big]\end{aligned} \tag{3.37}$$

where $w_s=-(s+v)j$. Then, a direct calculation leads to the expressions of the discrete operators (A_d,B_d,C_d,D_d) as

$$\begin{aligned}A_d(\cdot)=&-(\cdot)-\frac{2\delta}{\sqrt{-w_\delta}}\int_0^\zeta\sinh(\sqrt{w_\delta}(\zeta-\eta))(\cdot)d\eta\\&+\frac{2\delta\cosh(\sqrt{w_\delta}\zeta)}{\sqrt{-w_\delta}\sinh(\sqrt{w_\delta})}\int_0^1\cosh(\sqrt{w_\delta}(1-\eta))(\cdot)d\eta\end{aligned} \tag{3.38a}$$

$$C_d(\cdot) = -\frac{\sqrt{2\delta}}{\sqrt{-w_\delta}}\int_0^1 \sinh(\sqrt{w_\delta}(1-\eta))(\cdot)d\eta$$
$$+\frac{\sqrt{2\delta}\cosh(\sqrt{w_\delta})}{\sqrt{-w_\delta}\sinh(\sqrt{w_\delta})}\int_0^1 \cosh(\sqrt{w_\delta}(1-\eta))(\cdot)d\eta \tag{3.38b}$$
$$B_d = \frac{\sqrt{2\delta}j\cosh(\sqrt{w_\delta}\zeta)}{\sqrt{-w_\delta}\sinh(\sqrt{w_\delta})},\ D_d = \frac{j\coth(\sqrt{w_\delta})}{\sqrt{-w_\delta}} \tag{3.38c}$$

where $w_\delta = -(\delta+v)j$. We note that $\lim_{s\to+\infty}\mathcal{G}(s) = \lim_{\delta\to+\infty} D_d(\delta) = 0$ so system (3.36) is indeed a well-posed regular system [121]. In the similar manner, it is straightforward to find that $H_d(\cdot) = \frac{\sqrt{2\delta}}{\sqrt{-w_\delta}\sinh(\sqrt{w_\delta})}\int_0^1 \cosh(\sqrt{w_\delta}(1-\eta))(\cdot)d\eta$, and $V_d = \frac{j}{\sqrt{-w_\delta}\sinh(\sqrt{w_\delta})}$. In this case, $v = 1.5$ is taken and we consider constraints on plant and output disturbances as: $w_{\min} = -0.008$, $w_{\max} = 0.008$, $z_{\min} = -0.12$, and $z_{\max} = 0.12$. The plant and output disturbances are chosen to be Gaussian noises with zero mean and variances of $Q = 0.01$ and $R = 0.4$. Using MHE (horizon taken as 3), the estimated state and output (real-part) profiles are illustrated in Fig. 3.1 and Fig. 3.2, where it is apparent that the proposed MHE design is capable of estimating the output of the Schrödinger equation with considered constraints being satisfied. In addition, the time discretization interval is $h = 0.01$ and there are 181 spatial nodes. The initial condition $x_0(\zeta) = \cos(2\pi\zeta)$ is considered in this case.

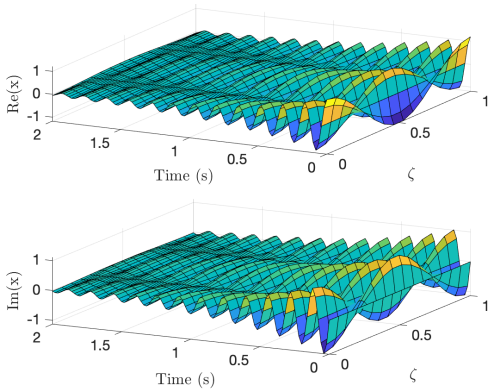


Figure 3.1: Estimated state profiles of the Schrödinger equation

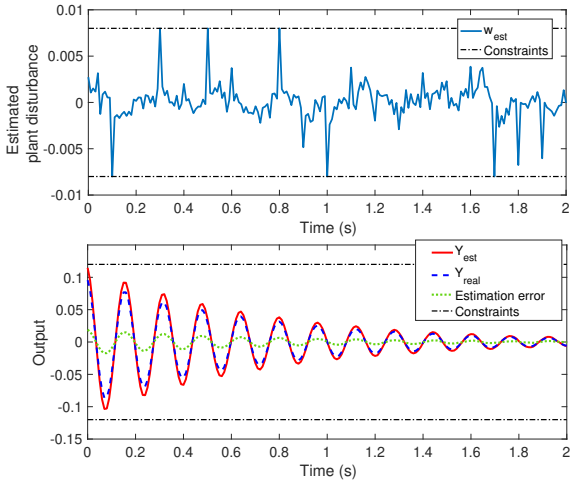


Figure 3.2: Estimated disturbance (above) and output (below) profiles of the Schrödinger equation

3.6.2 Example 2: wave equation

Another example is an unstable 1-D wave equation with boundary disturbance and noncollocated observation:

$$x_{tt}(\zeta,t) = x_{\zeta\zeta}(\zeta,t) \tag{3.39a}$$

$$x(0,t) = 0,\ x_\zeta(1,t) = w(t) \tag{3.39b}$$

$$x(\zeta,0) = x_1^0(\zeta),\ x_t(\zeta,0) = x_2^0(\zeta) \tag{3.39c}$$

$$y(t) = x_\zeta(0,t) + v(t),\ z(t) = x_\zeta(0,t) \tag{3.39d}$$

where $x \in X$ with $X = H_L^1(0,1) \times L^2(0,1)$, $H_L^1(0,1) = \{f | f \in H^1(0,1), f(0) = 0\}$, equipped with the norm $\|(f,g)\|_X^2 = \int_0^1 |f'(\zeta)|^2 + |g(\zeta)|^2 d\zeta$ for any $(f,g) \in X$, and $W = Y = \mathbb{R}$. As in [132], wave equation (3.39) can be formulated in the usual state-space form (3.5), where $A(f,g) = (g, f'')$ with the domain $\mathcal{D}(A) = \{(f,g) \in X | A(f,g) \in X, f'(1) = 0\}$, $B = [0; \delta_d(\zeta - 1)]$, and $C = H = [-\langle \delta_d'(\zeta), \cdot \rangle, 0]$, where $\delta_d(\cdot)$ denotes the Dirac function. By [132, The. 1], we note that C (i.e., H, and B^*) is admissible, (C, A) is exactly observable in finite time $\tau > 2$. By introducing a shifted semigroup generated by $A_s = A - \alpha I$ with $\alpha = 0.1$, the resultant shifted semigroup is exponentially stable and (C, A_s) is exactly observable in finite time $\tau > 2$ and C (i.e., H, and B^*) is still admissible [42, The. 6.7.2]. As in [125, Exa. 2.3.5] or [43, Exa. 2.34], it can be shown that A_s is the infinitesimal generator of a contraction

semigroup.

In this case, we consider input and output constraints as: $w_{\min} = -0.001$, $w_{\max} = 0.001$, $z_{\min} = -1.75$, and $z_{\max} = 1.75$. For simplicity, we consider to-be-estimated and measured outputs at the same location, i.e., $H = C$ in Eq.(3.5). We can compute the resolvent operator and find the discretized operators $(A_d, B_d, C_d, D_d, H_d, V_d)$ and their adjoints. The wave system is a well-posed regular system [121] since $\lim_{s\to+\infty} \mathcal{G}(s) = \lim_{\delta\to+\infty} D_d = \lim_{\delta\to+\infty} \frac{2}{e^{\delta}+e^{-\delta}} = 0$. The plant and output disturbances are chosen to be Gaussian noises with zero mean and variances of $Q = 4\times10^{-4}$ and $R = 0.2$. Using MHE (horizon taken as 3), the estimated state and output profiles are illustrated in Fig. 3.3 and Fig. 3.4, where it can be observed that estimated plant disturbance and output satisfy the presented constraints and the estimation error converges to zero by using the proposed MHE design. In addition, the spatial and time discretization intervals are selected as $\Delta\zeta = 0.001$ and $h = 0.1$. Initial conditions $x_1^0(\zeta) = \sin(1.5\pi\zeta)$ and $x_2^0(\zeta) = 0.2\sin(3.5\pi\zeta)$ are considered. For numerical realization of the resolvent and discrete operators, the trapezoidal rule is applied in both examples, leading to a second-order accuracy in space, which can be enhanced by using advanced numerical integration methods (e.g., fourth-order Runge-Kutta). Alternatively, the spectral methods can be utilized with superior rate of convergence, low dissipation and dispersion errors [133]. To this end, additional efforts need to be made in determining basis functions and the degrees of freedom [134].

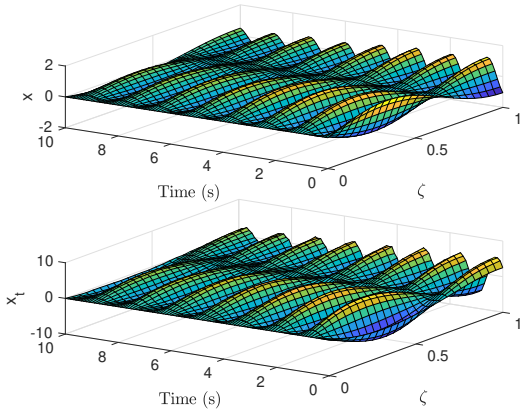


Figure 3.3: Estimated state profiles of the wave equation

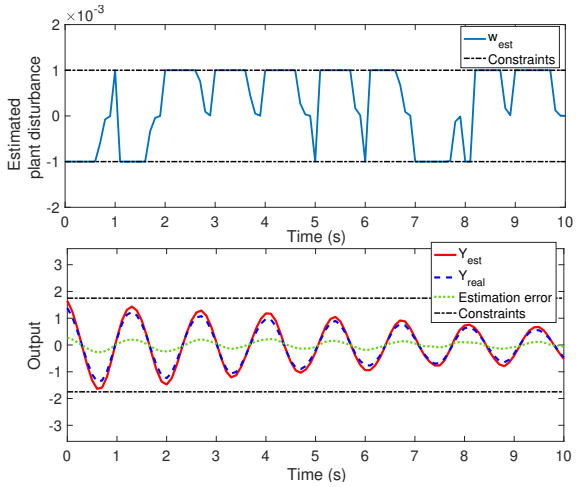


Figure 3.4: Estimated disturbance (above) and output (below) profiles of the wave equation

3.7 Conclusion

In this chapter, a linear moving horizon estimator was designed for constrained output estimation of regular linear distributed parameter systems with unbounded disturbance and observation operators. The salient design points are the application of the Cayley-Tustin transformation to link continuous- to discrete-time regular linear DPSs to address the issues associated with unbounded operators and constraints presence, as well as the proof of Riccati equations and exact observability necessary for the MHE design. By deploying two representative examples (namely, Schrödinger equation and wave equation), the proposed constrained moving horizon output estimation design was successfully demonstrated. The approach developed for the output estimation can be extended to prediction and smoothing problems of regular linear systems, and they will be reported in a future publication. In addition, more changeling DPS examples will be investigated, e.g., Euler-Bernoulli beam equation based on its natural connection with Schrödinger equation [135] and transport equations [65] widely present in process control engineering.

Chapter 4

Moving Horizon Estimation for Switching Regular Linear Systems

4.1 Introduction

Switching infinite-dimensional systems (hybrid distributed parameter systems (DPS)) are of theoretical and practical significance due to the capacity of accounting for multi-mode spatiotemporal dynamics. Many contributions on stability analysis, stabilization and control of switching distributed parameter systems have been made over the past decades [136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 45, 149, 150, 151]. In this chapter, we focus on constrained state/output estimation for a rather general class of switching distributed parameter systems, namely switching regular linear infinite-dimensional systems described by partial differential equations (PDE) with possibly unbounded observation and disturbance operators, in the presence of bounded plant and measurement disturbances, where the system modes switch within a known finite set in an unpredictable way.

The research on switched (linear and non-linear) lumped parameter systems (LPS) that are modelled by ordinary differential equations (ODE) and the associated stability, control and estimation theories have been flourishing since the 1990s (see, for example, [152, 153, 154, 155, 156, 108, 157] and the references therein). For switching infinite-dimensional systems, the corresponding studies on stability, stabilization and control theories have been actively investigated in recent years (e.g., [136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 45, 149, 150, 151]). The initial interest is focused on general switching linear infinite-dimensional systems.

More specifically, the stability of switching linear infinite-dimensional autonomous systems of the form $\dot{x} = A_i x$ with $i = 1, 2$ was given in [137], where the infinitesimal generators A_1 and A_2 commute. Two switching strategies were proposed for stabilizing a class of infinite dimensional switched linear systems with two unstable subsystems whose principal operators (A_1 and A_2) can be non-commutative [144], where the cases of linear autonomous systems and with external perturbations were considered. An optimal actuator-plus-controller switching scheme was formulated and proven for a class of switched DPS subjected to switching bounded control and observation operators (namely, B_p and C_p with model index p) by using the finite horizon LQR optimal control [45]. Apart from these, a special interest has been devoted to switched linear hyperbolic systems (i.e., switched linear systems of conservation laws), for which important results have been reported on stability analysis [151, 148, 142], stabilization [147], and supervisory switching control design [140]. Recently, the boundary stochastic stability of Markov jump linear hyperbolic systems was proved in [138] by using Lyapunov techniques, where the Markov chain deciding the mode switching was assumed to be observable, which may not be the case in practice. For switching well-posed linear systems, an averaging problem for a system that fast switches between two infinite-dimensional well-posed LTI systems was investigated in [149]. In addition, important works have been reported on switching control [141, 146, 139] and sliding-mode control [136, 145] of PDE systems, and boundary observer-based control for cascade hyperbolic PDE-ODE systems with stochastic jumps [143].

However, state/output and parameter estimation of switching DPSs are relatively under-explored, c.f. [158, 159]. In particular, the calculus of variations approach for state and parameter estimation in switched 1D hyperbolic PDEs coupled with an ODE was proposed in [158]. Based on the backstepping method, a boundary observer was constructed to estimate the point-wise expected value of the solution of a randomly switching reaction-diffusion PDE while the random switching was assumed to be governed by an observable Markovian process [159]. As a comparison, estimation of switching (linear and non-linear) LPSs have been extensively studied over the past decades, c.f. [155, 108, 157, 160]. Among these, moving horizon estimation (MHE) or receding horizon estimation, as an optimization-based method, has been proposed for system mode and state estimation of switched (linear or nonlinear) LPSs while

explicitly handling physical constraints [108, 157, 160]. The MHE scheme used in these papers was proposed in [39, 161], where a fixed prior diagonal matrix is used as the arrival cost, and hence the objective function is fully deterministic. This framework is different from the one developed in [38, 162], where MHE is formulated as an approximation of the optimal full information estimator and the arrival cost is introduced to summarize the effect of past measurements $\{y_k\}_{k=0}^{T-N-1}$ on the state x_{T-N}, which might not have general analytic expressions.

In all of the aforementioned papers on MHE for switching systems, the plants are linear or non-linear LPSs modelled by linear or non-linear ODEs. Although for MHE design of non-switching DPSs, there are some related works, including [117, 118, 116] by performing model reduction methods. To the best knowledge of the authors, there is no published work on MHE for the general switching DPSs in literature due to the fact that the state/output estimation problem of general switching PDEs is mathematically challenging and relatively new to the control community.

In this chapter, we focus on constrained state/output estimation for a general class of switching DPSs (i.e., switching regular linear infinite-dimensional systems with possibly unbounded observation and disturbance operators (accounting for point and/or boundary observation and disturbance)), where the system is corrupted with bounded plant and measurement disturbances and switches within a known finite set in an unpredictable way. The main contribution of this chapter is threefold. The first contribution lies in moving horizon estimator design for switching regular linear infinite-dimensional systems. To address unboundedness issues associated with the observation and/or disturbance operators, and better suit the discrete-time MHE design, the Cayley-Tustin bilinear transformation is utilized to map the continuous-time DPS model with unbounded operators to a discrete-time DPS model with all bounded operators, while preserving essential model properties (such as: stability, observability, input-output mapping and etc.), where no model spatial discretization or spatial model reduction is required. Based on the discrete-time infinite-dimensional model with bounded disturbances, an optimal MHE algorithm (Algorithm 1) is proposed for the constrained state/output and mode estimation by extending the moving horizon estimation approach proposed in [161, 157] for switching linear and non-linear discrete-time finite-dimensional systems. Considering that directly optimizing the

infinite-dimensional state may be intractable as the spatial discretization node tends to infinity, a late-lumping MHE algorithm (Algorithm 2) is proposed by converting the estimation of the infinite-dimensional state into an extended finite-dimensional disturbance thus providing a practicable and implementable framework for the optimal MHE algorithm instead of doing early-lumping (spatial discretization or model reduction) in the design stage. It is worth noting that we consider unknown and unpredictable switching modes that are not governed by any observable Markovian process as in [138, 159] and hence more general in practice. The stability properties of the proposed optimal MHE estimator are derived in Theorem 11. The resulting MHE algorithms provide sub-optimal output/state estimation for switching regular linear infinite-dimensional systems due to model time discretization.

The second contribution is the investigation of the essential observability properties of switching regular linear infinite-dimensional systems. The challenges lie in the simultaneous estimation of state/output and mode in the presence of random (unknown and unpredictable) switching pattern and the infinite dimensionality of the state space. Unlike the observability of finite-dimensional systems, there are several different observability concepts often utilized in infinite-dimensional systems, including exact observability, approximate observability and etc [4, 42]. In this chapter, the complete mode observability, complete observability, and simultaneous exact (or approximate) observability are fully explored. This is a generalization of the observability for switching finite-dimensional systems [161, 157] and non-switching infinite-dimensional systems [4, 42]. Based on the simultaneous exact observability, the stability analysis of the proposed MHE is derived in Theorem 11.

The third contribution is the applications to two representative switching regular linear infinite-dimensional systems. In the first example, a counterflow heat changer as a non-spectral system is considered with boundary measurement and spatially distributed disturbance and input. The heat exchanger is modelled by a general class of 2×2 linear coupled hyperbolic PDE system that can describe the packed gas absorber processes, irrigation canals and etc, so the proposed MHE design can be directly applied to such systems. In the second example, the damped Rayleigh beam equation as a spectral system along with boundary disturbance, input and output is considered. The developed MHE design can be extended to other types of beam equations

providing that the models are stable and exact observable. Both examples are of practical and theoretical significance, and a generalization of the non-switching versions [163, 164, 13]. In addition, the moving horizon estimator with perfect switching mode information is investigated and compared with the proposed MHE in both examples. The influences of different levels of disturbances on the estimation performance of the proposed MHE are discussed, which verifies the robustness of the proposed MHE algorithm numerically.

The structure of the chapter is as follows. Section 4.2 presents the switching regular linear systems and the corresponding switching discrete-time infinite-dimensional systems by using Cayley-Tustin transformation. In Section 4.3, a series of observability concepts for the switching discrete-time systems is proposed along with the corresponding analysis. Two MHE algorithms are formulated consisting optimal MHE and late-lumping MHE in Section 4.4. Section 4.5 provides a stability analysis of the proposed estimators. The theoretical results are illustrated in Section 4.6 via two representative PDE models.

Notations: Assume that X and Z are two Hilbert spaces and $A : X \mapsto Z$ is a linear operator. Linear bounded operators from X to Z are denoted by $\mathcal{L}(X, Z)$ and $\mathcal{L}(X)$ if $X = Z$. The domain, range, kernel, spectrum, resolvent set and resolvent operator of a linear operator A are denoted as: $\mathcal{D}(A)$, $\mathrm{Ran}(A)$, $\mathrm{Ker}(A)$, $\sigma(A)$, $\rho(A)$, and $\mathcal{R}(s, A) = (sI - A)^{-1}$ with $s \in \rho(A)$, respectively. We define space X_1 as $\mathcal{D}(A)$ with the norm $\|x\|_1 = \|(\beta I - A)x\|$, and space X_{-1} as the completion of X with the norm $\|z\|_{-1} = \|(\beta I - A)^{-1}z\|$, where $\forall x \in \mathcal{D}(A)$, $\forall z \in X$, and $\beta \in \rho(A)$. The constructed spaces are related by $X_1 \subset X \subset X_{-1}$, with each inclusion being dense and continuous embedding [42]. The extension of A to X_{-1} is still denoted by A, and the Λ-extension of an operator C is denoted by C_Λ. The inner product is denoted by $\langle \cdot, \cdot \rangle$, and $L^2(0, l)^m$ with a positive integer m denotes a Hilbert space of a m-dimensional vector of the real functions that are square integrable over $[0, l]$ with a spatial length l. For a generic time-varying vector v_k, we define $v_{T-N}^T = \mathrm{col}(v_{T-N}, v_{T-N+1}, ..., v_T)$. Moreover, $\prod_{i=1}^N A_i = A_1 A_2 ... A_N$ denotes the ordered product of such operators $A_i \in \mathcal{L}(X)$.

4.2 System description

4.2.1 Switching regular linear systems

Consider a switching regular linear system with corrupted measurement:

$$\dot{x}(t) = A(r_t)x(t) + B(r_t)w(t),\ t \geq 0,\ x(0) = x_0 \in X \tag{4.1a}$$

$$y(t) = C(r_t)x(t) + v(t),\ z(t) = H(r_t)x(t) \tag{4.1b}$$

where $A(r_t) : \mathcal{D}(A(r_t)) \subset X \to X$ is an infinitesimal generator of a strongly stable contraction C_0-semigroup $\mathbb{T}_r(t)$ (i.e., $\|\mathbb{T}_r(t)\| \leq 1$ and $\mathbb{T}_r(t)z \to 0$ as $t \to +\infty$ for all $z \in X$) on a separable Hilbert space X, $r_t \in \mathcal{S} = \{r_1, r_2, ..., r_M\}$ is the discrete state (system mode), $w(t)$ denotes the plant (or input) disturbance on a finite-dimensional Hilbert space W, and $v(t)$ is output disturbance (independent with $w(t)$) on a finite-dimensional Hilbert space Y. In addition, we assume that $\mathcal{D}(A(r_1)) = \mathcal{D}(A(r_2)) = ... = \mathcal{D}(A(r_M))$. $y(t) \in Y$ and $z(t) \in Z$ represent the measured output and the to-be-estimated output, where Z is a finite-dimensional Hilbert space. The operators $B(r_t)$, $C(r_t)$ and $H(r_t)$ are assumed to be admissible, i.e., $B(r_t) \in \mathcal{L}(W, X_{-1})$, $C(r_t) \in \mathcal{L}(X_1, Y)$, and $H(r_t) \in \mathcal{L}(X_1, Z)$. To ensure well-posedness, unbounded $C(r)$ and $H(r)$ must be replaced by their Λ-extension operators $C_\Lambda(r_t)$ and $H_\Lambda(r_t)$, where $C_\Lambda(r_t)x_0 = \lim_{\lambda \to +\infty} C(r_t)\lambda(\lambda I - A(r_t))^{-1}x_0$, and the domain of C_Λ contains all $x_0 \in X$ for which the above limit exists, and $H_\Lambda(r_t)$ is similarly defined. The transfer function associated with each system mode is given by: $\mathcal{G}_r(s) := C_\Lambda(r_t)(sI - A(r_t))^{-1}B(r_t)$, and each system Σ is assumed to be regular in the sense that $\lim_{s \to +\infty} \mathcal{G}_r(s)w = 0$, for all $w \in W$ [121, The. 1.3]. For an introduction of well-posed systems, we refer to [21, 42], and [111].

Moreover, it is assumed that all system operators $(A(r_t), B(r_t), C(r_t), H(r_t))$ are available once the system mode r_t is determined. The unknown initial state, input and output disturbances, and system mode that are generally induced by process noise, sensor noise and/or modelling errors determine the measured output y and the to-be-estimated output z. Instead of estimating the whole state, a linear combination of states $H(r_t)x$ is considered, where $H(r_t)$ maybe a state feedback that can be utilized for construction of an output feedback controller [40].

Considering that direct handling all unbounded operators, constraint presence

and mode detection in a continuous-time setting can be difficult, we transform the continuous model (4.1) into a discrete system using Cayley-Tustin bilinear transformation, which is amenable for discrete MHE design as in [38, 39, 161, 162]. In what follows, we will provide MHE design and analysis for the discrete model. By [97, The. 4.3], the estimation results of the discrete model can be linked back to those of the continuous model (4.1), with a small time discretization interval, as in [65, Sec. 2.3].

4.2.2 Switching linear discrete-time DPSs

Consider a Tustin time-discretization of model (4.1) for a given time discretization interval h on $t \in (kh, (k+1)h)$:

$$\frac{x_{k+1} - x_k}{h} \approx A(r_k)\frac{x_{k+1} + x_k}{2} + B(r_k)\frac{w_k}{\sqrt{h}} \tag{4.2a}$$

$$\frac{y_k}{\sqrt{h}} \approx C(r_k)\frac{x_{k+1}+x_k}{2} + \frac{v_k}{\sqrt{h}},\ \frac{z_k}{\sqrt{h}} \approx H(r_k)\frac{x_{k+1}+x_k}{2} \tag{4.2b}$$

where the spatial dependence of x is omitted for brevity. For the discretization of disturbances and outputs, we consider an "average" sampling operator $\mathbf{T}_h$ as in [97]

$$\frac{w_k}{\sqrt{h}} = (\mathbf{T}_h w)(k) := \frac{1}{h}\int_{kh}^{(k+1)h} w(t)dt \tag{4.3}$$

and discrete signals v_k, y_k, and z_k are similarly defined. As shown in [97, The. 4.3], the Cayley-Tustin discretization (i.e., CT transform in short) is a convergent time discretization frame for scalar-valued (dim U=dim Y=1) input-output stable system nodes in the sense that $\|y_k/\sqrt{h} - y(t)\|$ converges to 0 on the interval $t \in (kh, (k+1)h)$ as $h \to 0$. Direct manipulation of (4.2) leads to the Cayley-Tustin discretization of (4.1) as:

$$x_{k+1} = A_d(r_k)x_k + B_d(r_k)w_k \tag{4.4a}$$

$$y_k = C_d(r_k)x_k + D_d(r_k)w_k + v_k = C_d(r_k)x_k + \bar{v}_k \tag{4.4b}$$

$$z_k = H_d(r_k)x_k + V_d(r_k)w_k \tag{4.4c}$$

where for $x_0 \in X,\ k \in \mathbb{Z}^+$, and $x_k \in X$ denotes discrete distributed state on the infinite-dimensional Hilbert space X. $w_k \in W$, $y_k \in Y$, and $z_k \in Z$ represent discrete plant disturbance, measured and to-be-estimated outputs. For simplicity, we denote the

discrete measurement disturbance by $\bar{v}_k = D_d(r_k)w_k + v_k \in Y$. For ease of notations, we still use v_k to represent $\bar{v}_k$ in what follows and when no confusion arises. Additionally, we denote r_{kh} by r_k for notational simplicity. It is worth noting that all operators in the discrete-time model (4.4) are bounded, i.e., $A_d(r_k) \in \mathcal{L}(X)$, $B_d(r_k) \in \mathcal{L}(W, X)$, $C_d(r_k) \in \mathcal{L}(X, Y)$, $D_d(r_k) \in \mathcal{L}(U, Y)$, $H_d(r_k) \in \mathcal{L}(X, Z)$, $V_d(r_k) \in \mathcal{L}(U, Z)$, which are given by:

$$\begin{bmatrix} A_d(r_k) & B_d(r_k) \\ C_d(r_k) & D_d(r_k) \\ H_d(r_k) & V_d(r_k) \end{bmatrix} = \begin{bmatrix} -I + 2\delta\mathcal{R}(\delta, r_k) & \sqrt{2\delta}\mathcal{R}(\delta, r_k)B(r_k) \\ \sqrt{2\delta}C(r_k)\mathcal{R}(\delta, r_k) & \mathcal{G}_{r_k}(\delta) \\ \sqrt{2\delta}H\mathcal{R}(\delta, r_k) & \mathcal{T}_{r_k}(\delta) \end{bmatrix}$$

where we denote the resolvent operator by $\mathcal{R}(\delta, r_k) = \mathcal{R}(\delta, A(r_k)) = (\delta I - A(r_k))^{-1}$, $\delta = 2/h \in \rho(A(r_k)) \cap \mathbb{R}^+$, $\mathcal{G}_{r_k}(\delta) = C_\Lambda(r_k)\mathcal{R}(\delta, r_k)B(r_k)$, and $\mathcal{T}_{r_k}(\delta) = H_\Lambda(r_k)\mathcal{R}(\delta, r_k)B(r_k)$.

Moreover, we assume that there exists a minimum dwell time bigger than the discretization interval and no switching occurs in the interval $t \in (kh, (k+1)h)$, namely, switching only happens at the sampling time instance. Throughout this chapter, we assume that the system mode is independent of state propagation and can be regarded as an exogenous variable that is governed by some external and unobservable process. This is more general than assuming the switching mode is governed by an observable Markovian process as in [138, 159]. By the preceding discussion, we note that all model operators $(A_d(r_k), B_d(r_k), C_d(r_k), D_d(r_k), H_d(r_k), V_d(r_k))$ are available once the mode identity r_k is determined.

Remark 7. *It is remarkable to note that by the average sampling operation (4.3) and the mean value theorem, the constraints on the estimated output $z(t)$ remain identical to the constraints on its discrete-time counterpart $z_k/\sqrt{h}$, namely, $z_{\min} \le z(t) \le z_{\max} \Leftrightarrow z_{\min} \le z_k/\sqrt{h} \le z_{\max}$. Similar results hold for the constraints on the continuous- and discrete-time disturbances.*

By using the CT transform, the challenges in unbounded operators and constraint handling can be readily tackled by the MHE design for the discrete-time system (4.2). Once the discrete-time disturbance w_k and output z_k are estimated, they can be linked to their continuous counterparts, by applying $\mathbf{T}_h^*$, where the adjoint operator of $\mathbf{T}_h$, $\mathbf{T}_h^*$ is $1/\sqrt{h}$ times the zero-order hold operator on the interval $(kh, (k+1)h)$ [97].

In order to achieve mode detection and analyze the convergence of proposed MHE design, we will establish a series of observability concepts along with corresponding properties in the following section.

4.3 Observability of switching linear DPSs

In this section, we study the observability of system (4.4). For the sake of simplicity and without loss of generality, we will focus on the feasibility of reconstructing the infinite-dimensional state x_0 and the switching pattern r_0^N from a measurements sequence y_0^N over a given time interval $[0, N]$. We first consider the observability of system (4.4) in the absence of disturbances, and then move to the case with disturbances.

4.3.1 Observability in the absence of disturbances

Consider a noise-free version of model (4.4) on the interval $k \in [0, N]$ as:

$$x_{k+1} = A_d(r_k)x_k,\ y_k = C_d(r_k)x_k,\ z_k = H_d(r_k)x_k \tag{4.5}$$

It is straightforward to obtain the measurement sequence as: $y_0^N = \mathcal{C}_d(r_0^N)x_0$, where $\mathcal{C}_d(r_0^N) = [C_d(r_0); C_d(r_1)A_d(r_0); ...; C_d(r_N)\prod_{i=1}^{N} A_d(r_{N-i})] \in \mathcal{L}(X, Y^{N+1})$.

Definition 4. *(Complete Mode Observability). For system (4.5), it is defined that a switching pattern $\Lambda_0^N \in \mathcal{P}_N$ is distinguishable from another switching pattern $\Lambda_0'^N \in \mathcal{P}_N$ if $\mathcal{C}_d(\Lambda_0^N)x_0 \neq \mathcal{C}_d(\Lambda_0'^N)x_0'$ for all $x_0, x_0' \in X$ with $x_0 \neq 0$ or $x_0' \neq 0$. Furthermore, system (4.5) is said to be completely mode observable if in steps $N+1$, for every couple $\Lambda_0^N \neq \Lambda_0'^N \in \mathcal{P}_N$, Λ_0^N is distinguishable from $\Lambda_0'^N$.*

Definition 4 is similar to [161, Def. 1]. According to complete mode observability, it is possible to uniquely determine the switching pattern Λ_0^N as well as the initial state x_0 on the basis of measurements y_0^N, provided that the initial condition is nonzero [161, Lem. 1].

Definition 5. *(Complete Observability). System (4.5) is said to be completely observable in $N+1$ steps if it is completely mode observable in $N+1$ steps and if $\mathcal{C}_d(\Lambda_0^N)x_0 \neq \mathcal{C}_d(\Lambda_0^N)x_0'$ for any $x_0 \neq x_0' \in X$ and any $\Lambda_0^N \in \mathcal{P}_N$.*

Definition 5 is similar to [157, Def. 2], which guarantees that the plant state can be uniquely determined from measurements y_0^N. If $x_0 \neq 0$, both system mode and state can be uniquely determined based on the measurement sequence y_0^N.

Definition 6. *(Simultaneous Approximate Observability). System (4.5) is said to be simultaneously approximately observable in $N+1$ steps if for every couple $\Lambda_0^N \neq \Lambda_0'^N \in \mathcal{P}_N$ and $(x_0, x_0') \in X \bigoplus X$, $C_d(\Lambda_0^N)x_0 - C_d(\Lambda_0'^N)x_0' = 0$ implies that $(x_0, x_0') = (0,0)$, namely, Ker$([C_d(\Lambda_0^N), -C_d(\Lambda_0'^N)]) = [0;0]$.*

Definition 7. *(Simultaneous Exact Observability). System (4.5) is said to be simultaneously exactly observable in $N+1$ steps if for every couple $\Lambda_0^N \neq \Lambda_0'^N \in \mathcal{P}_N$, there exists a finite positive number k_N (depending on N and h) such that for all $z_0, z_0' \in X$, we have $\|C_d(\Lambda_0^N)z_0 - C_d(\Lambda_0'^N)z_0'\|_{Y^{N+1}}^2 \geq k_N^2(\|z_0\|_X^2 + \|z_0'\|_X^2)$.*

Definitions 6 and 7 are a generalization of [42, Def. 6.4.1] for two continuous-time DPSs to the switching discrete-time DPSs where there are finitely many switching modes. The complete observability and simultaneous approximate observability have the following relationship.

Theorem 7. *System (4.5) is said to be completely observable in $N+1$ steps if it is simultaneously approximately observable in $N+1$ steps.*

Proof. On one hand, the complete mode observability immediately follows the simultaneous approximate observability since one has to take $x_0 = 0$ and $x_0' = 0$ in order to ensure $C_d(\Lambda_0^N)x_0 = C_d(\Lambda_0'^N)x_0'$ for every couple $\Lambda_0^N \neq \Lambda_0'^N \in \mathcal{P}_N$. On the other hand, simultaneous approximate observability of $[C_d(\Lambda_0^N), C_d(\Lambda_0'^N)]$ induces approximate observability $C_d(\Lambda_0^N)$ or $C_d(\Lambda_0'^N)$, i.e., one obtain Ker $(C_d(\Lambda_0^N)) = 0$ and Ker $(C_d(\Lambda_0'^N)) = 0$ from Ker $([C_d(\Lambda_0^N), C_d(\Lambda_0'^N)]) = (0,0)$. Therefore, we conclude that $C_d(\Lambda_0^N)x_0 \neq C_d(\Lambda_0^N)x_0'$ for any $x_0 \neq x_0' \in X$ and any $\Lambda_0^N \in \mathcal{P}_N$. ■

Remark 8. *For infinite-dimensional systems, exact observability as a stronger concept indicates approximate observability, while relatively difficult to satisfy. Both concepts ensures that it is possible to reconstruct the initial state from the measurements y_0^N, while the reconstruction operator is bounded for exact observable systems but might be unbounded for approximate observable cases.*

Remark 9. *By the preceding discussion, we note that simultaneous exact observability ⇒ simultaneous approximate observability ⇒ complete observability ⇒ complete mode observability. For switching linear lumped parameter systems (LPSs), stronger results hold: simultaneous exact observability ⇔ simultaneous approximate observability ⇔ complete observability ⇔ complete mode observability, see [161, Lem 1].*

Corollary 1. *System (4.5) is simultaneously exactly (or approximately) observable iff for every couple* $\Lambda_0^N \neq \Lambda_0'^N \in \mathcal{P}_N$ *with* $\Lambda_0^N = \{r_0, r_1, ..., r_k\}$ *and* $\Lambda_0'^N = \{r_0', r_1', ..., r_k'\}$ *the following extended system is exactly (or approximately) observable:*

$$\begin{bmatrix} x_{k+1} \\ x'_{k+1} \end{bmatrix} = \begin{bmatrix} A_d(r_k) & 0 \\ 0 & A_d(r'_k) \end{bmatrix} \begin{bmatrix} x_k \\ x'_k \end{bmatrix} \tag{4.6a}$$

$$y_k^e = \begin{bmatrix} C_d(r_k) & -C_d(r'_k) \end{bmatrix} \begin{bmatrix} x_k \\ x'_k \end{bmatrix} \tag{4.6b}$$

where $k \in [0, N]$, $[x_k; x'_k]$ *denotes the extended infinite-dimensional state, and* y_k^e *is the corresponding measurement.*

From the simultaneous exact (or approximate) observability of the linear infinite-dimensional systems (without switching) (see [42, Def. 6.4.1] and [165, Def. 3.1]), the proof of Corollary 1 is readily obtained and hence omitted here. The simultaneous exact (or approximate) observability concept indicates that by observing the difference in two individual outputs, i.e., $y_k^e = C_d(r_k)x_k - C_d(r'_k)x'_k$, we intend to recover the initial states of the cascaded systems. Based on Corollary 1, one can address the estimation problem of switching DPSs using the idea of diagonalizable semigroups of the extended system.

In what follows, we will analyze the relationship between switching mode distinguishability and output measurements. For switching linear and non-linear LPSs, it has been shown in [161, 157] that it is impossible to distinguish one switching sequence from another if the measurements only differ from the first or the last $[n_x/n_y - 1]$ instants of the observation window $[0, N]$, where the state and output spaces are $X = \mathbb{R}^{n_x}$, $Y = \mathbb{R}^{n_y}$. In this chapter, it is not trivial to obtain such result since we consider infinite-dimensional systems, i.e., $\dim(X) = \infty$, even if n_y is a finite number under the consideration of finite-dimensional space Y. Here, we will derive the results for the switching DPSs.

Theorem 8. *Consider two generic switching patterns $\Lambda_0^N \neq \Lambda_0'^N \in \mathcal{P}_N$ for system (4.5). Then, it is possible for system (4.5) to be simultaneously approximately observable if Λ_0^N and $\Lambda_0'^N$ only differ in the first or in the last m ($m \geq 1$) instants of the observation window $[0, N]$.*

Proof. This proof is completed by induction. We first consider that Λ_0^N and $\Lambda_0'^N$ only differ in the first instant. By defining $\mathrm{Ker}\,([C_d(\Lambda_0^N), C_d(\Lambda_0'^N)]) = (z_0, -z_0')$ and the concatenation of two switching patterns as $\Lambda_0^N = r_0 \otimes \Lambda_1^N$ and $\Lambda_0'^N = r_0' \otimes \Lambda_1^N$ with $\Lambda_1^N = \{r_1, r_2, ..., r_N\}$, we have the following:

$$\begin{bmatrix} C_d(r_0) \\ C_d(r_1)A_d(r_0) \\ \vdots \\ C_d(r_N)\prod_{i=1}^{N} A_d(r_{N-i}) \end{bmatrix} z_0 - \begin{bmatrix} C_d(r_0') \\ C_d(r_1)A_d(r_0') \\ \vdots \\ C_d(r_N)\prod_{i=1}^{N-1} A_d(r_{N-i})A_d(r_0') \end{bmatrix} z_0' = \mathbf{0}$$

which indicates that $C_d(r_0)z_0 - C_d(r_0')z_0' = 0$ and $C_d(r_1)A_d(r_0)z_0 - C_d(r_1)A_d(r_0')z_0' = 0$. From the expression of $C_d(r_1)$ and $A_d(r_0)$ in (4.4), we note that $\sqrt{2\delta}C(r_1)[\mathcal{R}(\delta, r_0)z_0 - \mathcal{R}(\delta, r_0')z_0'] = 0$ and $C_d(r_1)[-(z_0' - z_0') + 2\delta(\mathcal{R}(\delta, r_0)z_0 - \mathcal{R}(\delta, r_0')z_0')] = 0$, which implies that $z_0 = z_0'$. By the second resolvent identity [166], one has $[\mathcal{R}(\delta, r_0) - \mathcal{R}(\delta, r_0')]z_0 = \mathcal{R}(\delta, r_0)(A(r_0') - A(r_0))\mathcal{R}(\delta, r_0')z_0 = 0$. This implies that $z_0 = z_0' = 0$.

Thus, we note that it is possible for system (4.5) to be simultaneously approximately observable if Λ_0^N and $\Lambda_0'^N$ only differ in the first instant of the observation window $[0, N]$. It is immediate that system (4.5) can be simultaneously approximately observable if Λ_0^N and $\Lambda_0'^N$ only differ in the first m ($m \geq 1$) instants of the observation window $[0, N]$.

Similarly, we consider the case that Λ_0^N and $\Lambda_0'^N$ only differ in the last instant, i.e., $\Lambda_0^N = \Lambda_0^{N-1} \otimes r_N$ and $\Lambda_0'^N = \Lambda_0^{N-1} \otimes r_N'$ with $\Lambda_0^{N-1} = \{r_0, r_1, ..., r_{N-1}\}$, then we have

$$\begin{bmatrix} C_d(r_0) \\ C_d(r_1)A_d(r_0) \\ \vdots \\ C_d(r_N)\prod_{i=1}^{N} A_d(r_{N-i}) \end{bmatrix} z_0 - \begin{bmatrix} C_d(r_0) \\ C_d(r_1)A_d(r_0) \\ \vdots \\ C_d(r_N')\prod_{i=1}^{N} A_d(r_{N-i}) \end{bmatrix} z_0' = \mathbf{0}$$

which indicates that $C_d(r_0)z_0 - C_d(r_0)z_0' = 0$ implying that $z_0 = z_0'$ as C_d is not a nilpotent operator in general. From $C_d(r_N)\prod_{i=1}^{N} A_d(r_{N-i})z_0 - C_d(r_N')\prod_{i=1}^{N} A_d(r_{N-i})z_0' = 0$ and $C_d(r_N) \neq C_d(r_N')$ we can further conclude that $z_0 = z_0' = 0$.

Therefore, we conclude that it is possible for system (4.5) to be simultaneously approximately observable if Λ_0^N and $\Lambda_0'^N$ only differ in the first or last m ($m \geq 1$) instants of the observation window $[0, N]$. ■

Theorem 9. *Consider two generic switching patterns $\Lambda_0^N \neq \Lambda_0'^N \in \mathcal{P}_N$ for system (4.5). Then, it is possible for system (4.5) to be simultaneously exactly observable if Λ_0^N and $\Lambda_0'^N$ only differ in the first or in the last m ($m \geq 1$) instants of the observation window $[0, N]$.*

Proof. The proof is similar to that in Theorem 8. We first consider the case that Λ_0^N and $\Lambda_0'^N$ only differ in the last instant, namely, $\Lambda_0^N = \Lambda_0^{N-1} \otimes r_N$ and $\Lambda_0'^N = \Lambda_0^{N-1} \otimes r'_N$ with $\Lambda_0^{N-1} = \{r_0, r_1, ..., r_{N-1}\}$. To show that system (4.5) is simultaneously exactly observable, we derive the following

$$
\begin{aligned}
&\|C_d(r_0)(z_0 - z'_0)\|_Y^2 + \sum_{n=1}^{N-1} \|C_d(r_n) \prod_{i=1}^{n} A_d(r_{n-i})(z_0 - z'_0)\|_Y^2 \\
&\quad + \|C_d(r_N) \prod_{i=1}^{N} A_d(r_{N-i}) z_0 - C_d(r'_N) \prod_{i=1}^{N} A_d(r_{N-i}) z'_0\|_Y^2 \\
\geq & k_{N-1}^2 (\|z_0 - z'_0\|_X^2) + \|C_d(r_N) \prod_{i=1}^{N} A_d(r_{N-i}) z_0 - C_d(r'_N) \prod_{i=1}^{N} A_d(r_{N-i}) z'_0\|_Y^2 \\
\geq & k_{N-1}^2 (\|z_0\|_X^2 + \|z'_0\|_X^2 - 2\|z_0\|_X^2 \|z'_0\|_X^2) \\
&\quad + \|C_d(r_N) \prod_{i=1}^{N} A_d(r_{N-i}) z_0 - C_d(r'_N) \prod_{i=1}^{N} A_d(r_{N-i}) z'_0\|_Y^2
\end{aligned}
$$

where the first inequality is ensured by the finite-time exact observability of $(A_d(r_i), C_d(r_i))$ in terms of the switching pattern Λ_0^{N-1}. The second inequality is obtained by reverse triangular inequality. By having $C_d(r_N)$ and $C_d(r'_N)$ sufficiently different, it is possible to find some positive κ such that $k_{N-1}^2(\|z_0\|_X^2 + \|z'_0\|_X^2 - 2\|z_0\|_X^2\|z'_0\|_X^2) + \|C_d(r_N) \prod_{i=1}^{N} A_d(r_{N-i}) z_0 - C_d(r'_N) \prod_{i=1}^{N} A_d(r_{N-i}) z'_0\|_Y^2 \geq \kappa^2(\|z_0\|_X^2 + \|z'_0\|_X^2)$, which indicates that the simultaneous exact observability of system (4.5). $C_d(r_N)$ and $C_d(r'_N)$ can be made sufficiently different by having sufficiently different modes r_N and r'_N. By induction, we conclude that it is possible for system (4.5) to be simultaneously approximately observable if Λ_0^N and $\Lambda_0'^N$ only differ in the last m ($m \geq 1$) instants of the observation window $[0, N]$.

Then, we consider that Λ_0^N and $\Lambda_0'^N$ only differ in the first instant. Let us define the concatenation of two switching patterns as $\Lambda_0^N = r_0 \otimes \Lambda_1^N$ and $\Lambda_0'^N = r_0' \otimes \Lambda_1^N$ with $\Lambda_1^N = \{r_1, r_2, ..., r_N\}$. To show simultaneous exact observability of system (4.5), we derive the following:

$$\begin{aligned}&\|C_d(r_0)z_0 - C_d(r_0')z_0'\|_Y^2 + \sum_{n=1}^{N} \|C_d(r_n) \prod_{i=1}^{n-1} A_d(r_{n-i})(A_d(r_0)z_0 - A_d(r_0')z_0')\|_Y^2 \\ \geq &\|C_d(r_0)z_0 - C_d(r_0')z_0'\|_Y^2 + \bar{k}_{N-1}^2(\|A_d(r_0)z_0 - A_d(r_0')z_0'\|_X^2)\end{aligned}$$

where the last inequality is indicated by the finite-time exact observability of $(A_d(r_i), C_d(r_i))$ with respect to the switching pattern Λ_1^N. By having $C_d(r_0)$ and $C_d(r_0')$ (or $A_d(r_0)$ and $A_d(r_0')$) sufficiently different, it is possible to find some positive $\bar{\kappa}$ such that $\|C_d(r_0)z_0 - C_d(r_0')z_0'\|_Y^2 + \bar{k}_{N-1}(\|A_d(r_0)z_0 - A_d(r_0')z_0'\|_X^2) \geq \bar{\kappa}^2(\|z_0\|_X^2 + \|z_0'\|_X^2)$. This can be done by having sufficiently different modes r_0 and r_0'. Thus, it is possible for system (4.5) to be simultaneously exactly observable if Λ_0^N and $\Lambda_0'^N$ only differ in the first m ($m \geq 1$ by induction) instants of the observation window $[0, N]$.

Therefore, we conclude that it is possible for system (4.5) to be simultaneously exactly observable if Λ_0^N and $\Lambda_0'^N$ only differ in the first or last m ($m \geq 1$) instants of the observation window $[0, N]$. ■

By Theorems 8-9, we note that for switching DPSs it is possible to have simultaneous exact (or approximate) observability if two switching patterns Λ_0^N and $\Lambda_0'^N$ only differ in the first or in the last m ($m \geq 1$) instants of the observation window $[0, N]$. This is novel when compared with the observability analysis of switching linear or non-linear LPSs.

4.3.2 Observability in the presence of disturbances

Under the influence of state and measurement disturbances, the corrupted measurement sequence on the interval $[0, N]$ is given as:

$$y_0^N = \mathcal{C}_d(r_0^N)x_0 + \mathcal{D}_d(r_0^N)w_0^{N-1} + v_0^N$$

where $\mathcal{D}_d(r_0^N) = [0, 0, ..., 0; C_d(r_1)B_d(r_0), 0, ..., 0; C_d(r_2)A_d(r_1)B_d(r_0), C_d(r_2)B_d(r_1), ..., 0; ...; C_d(r_N)\prod_{i=1}^{N-1} A_d(r_{N-i})B_d(r_0), C_d(r_N)\prod_{i=1}^{N-2} A_d(r_{N-i})B_d(r_1), ..., C_d(r_N)B_d(r_{N-1}]$.

Theorem 10. *Consider two generic switching patterns* $\Lambda_0^N \neq \Lambda_0'^N \in \mathcal{P}_N$ *with* $\Lambda_0^N = \{r_0, r_1, ..., r_N\}$ *and* $\Lambda_0'^N = \{r_0', r_1', ..., r_N'\}$. *Suppose that* W_c *and* V_c *are compact sets and system (4.5) is simultaneously exactly observable in* $N+1$ *steps. If there exist* $(x_0, w_0^{N-1}, v_0^N), (x_0', w_0'^{N-1}, v_0'^N) \in \mathcal{H}^e = X \bigoplus W_c^N \bigoplus V_c^{N+1}$ *satisfying* $\mathcal{C}_d(r_0^N)x_0 + \mathcal{D}_d(r_0^N)w_0^{N-1} + v_0^N = \mathcal{C}_d(r_0'^N)x_0' + \mathcal{D}_d(r_0'^N)w_0'^{N-1} + v_0'^N$, *there must be:* $\|x_0\|^2 \leq \rho_x$ *and* $\|x_0'\|^2 \leq \rho_x$.

Proof. By simple manipulation, it is straightforward that

$$\mathcal{C}_d(r_0^N)x_0 - \mathcal{C}_d(r_0'^N)x_0' = -\mathcal{D}_d(r_0^N)w_0^{N-1} + \mathcal{D}_d(r_0'^N)w_0'^{N-1} - v_0^N + v_0'^N \tag{4.7}$$

Based on the simultaneous exact observability, one can readily obtain the following:

$$\begin{aligned}\|x_0\|_X^2 + \|x_0'\|_X^2 &\leq \frac{1}{k_N^2}\left\|\mathcal{C}_d(r_0^N)x_0 - \mathcal{C}_d(r_0'^N)x_0'\right\|_{Y^{N+1}}^2 \\ &= \frac{1}{k_N^2}\left\|-\mathcal{D}_d(r_0^N)w_0^{N-1} + \mathcal{D}_d(r_0'^N)w_0'^{N-1} - v_0^N + v_0'^N\right\|_{Y^{N+1}}^2 \\ &\leq \frac{2}{k_N^2}[c_0 N\rho_w + (N+1)\rho_v] = \rho_x \end{aligned} \tag{4.8}$$

where we denote $c_0 = \max_{\Lambda_0^N \in \mathcal{P}_N} \|\mathcal{D}_d(r_0^N)\|^2$, $\rho_w = \sup_{w \in W_c} \|w\|^2$, $\rho_v = \sup_{v \in V_c} \|v\|^2$, since all discrete-time operators and disturbances are bounded. This completes the proof. ■

Due to the existence of plant and measurement disturbances, it is possible to have Λ_0^N indistinguishable from another one. Theorem 10 claims that when the estimated initial condition is sufficiently close to 0, then it is impossible to uniquely determine the switching pattern. From (4.8), we can clearly see that the norm bound of $[x_0; x_0']$ is related to the norm bounds of disturbance constraints, which implies that higher disturbance levels will enlarge the norm bound of initial state making it difficult to uniquely determine Λ_0^N from y_0^N, which is consistent with the finite-dimensional cases [161, 157].

4.4 Moving horizon estimation design

In this section, a moving horizon estimator is proposed for output estimation of the discrete-time DPSs with respect to disturbance and output constraints by utilizing

the results developed in Section 4.3. For simplicity, we directly consider constraints on discrete-time disturbances and output as they can be transformed to the constraints on the continuous-time counterparts by Remark 7.

We denote the set of all the admissible switching patterns at time instant T by $\mathcal{P}_T$, namely, the set of all the switching patterns in the observation window $[T-N, T]$ that is consistent with the priori knowledge of the evolution of the discrete state. For brevity, we assume that such a priori knowledge does not diminish with time, i.e., $\mathcal{P}_{T+1} \subseteq \mathcal{P}_T$, $T = N, N+1, ...$, or, less restrictively, $\mathcal{P}_T \subseteq \mathcal{P}_N$. Let us define $\mathcal{Y}_T(\Lambda)$ as the set of all the possible observations vectors associated with the switching pattern $\Lambda \in \mathcal{P}_T$:

$$\begin{aligned}\mathcal{Y}_T(\Lambda) = \{\bar{y}_{T-N}^T \in Y^{N+1} | \bar{y}_{T-N}^T = C_d(\Lambda)\bar{x}_{T-N} + \mathcal{D}_d(\Lambda)\bar{w}_{T-N}^{T-1} \\ + \bar{v}_{T-N}^T,\ \bar{w}_{T-N}^{T-1} \in W^N,\ \bar{v}_{T-N}^T \in V^{N+1}\}\end{aligned}$$

Thus, for a given measurement sequence $\tilde{y}_{T-N}^T$, $\hat{\Lambda}_{T-N}^T \in \mathcal{P}_T$ is referred to be a feasible estimate of the switching pattern Λ_{T-N}^T if $\tilde{y}_{T-N}^T \in \mathcal{Y}_T(\hat{\Lambda}_{T-N}^T)$.

As in [161, 157], we define the set of all feasible switching pattern $\mathcal{P}_T^f$ as:

$$\mathcal{P}_T^f = \{\Lambda | \tilde{y}_{T-N}^T \in \mathcal{Y}_T(\Lambda)\} \tag{4.9}$$

By the preceding discussion, we can formulate the following estimation algorithm: 1). determine the set of feasible switching patterns $\mathcal{P}_T^f$ at time step T; 2). minimize a quadratic cost J_T defined on the interval $[T-N, T]$:

$$\begin{aligned}J_T(\hat{x}_{T-N|T}, \hat{w}_{T-N}^{T-1}, \hat{\Lambda}_{T-N}^T) = &\sum_{k=T-N}^{T-1} \left\langle \hat{w}_{k|T}, Q\hat{w}_{k|T} \right\rangle_W + \sum_{k=T-N}^{T} \left\langle \hat{v}_{k|T}, R\hat{v}_{k|T} \right\rangle_Y \\ &+ \left\langle \hat{x}_{T-N|T} - \bar{x}_{T-N}, P(\hat{x}_{T-N|T} - \bar{x}_{T-N}) \right\rangle_X\end{aligned} \tag{4.10}$$

where $P \in \mathcal{L}(X)$ denotes a self-adjoint positive definite operator that expresses our belief in the "prediction" $\bar{x}_{T-N}$ of the state x_{T-N}, while Q and R as two positive definite matrices representing the penalty weights on the plant and measurement disturbances. Moreover, we note that $\hat{v}_{k|T} = y_k - C_d(\hat{r}_k)\hat{x}_{k|T}$ and $B_d(\hat{r}_k)\hat{w}_{k|T} = \hat{x}_{k+1|T} - A_d(\hat{r}_k)\hat{x}_{k|T}$.

Therefore, the optimal estimates $(\hat{x}_{T-N|T}^o, \hat{w}_{T-N}^{oT-1}, \hat{\Lambda}_{T-N}^{oT})$ can be obtained as follows:

$$(\hat{x}_{T-N|T}^o, \hat{w}_{T-N}^{oT-1}, \hat{\Lambda}_{T-N}^{oT}) = \arg \min_{\hat{x}_{T-N|T} \in X, \hat{w}_{T-N}^{T-1} \in W^N, \Lambda_{T-N}^T \in \mathcal{P}_T} J_T(\hat{x}_{T-N|T}, \hat{w}_{T-N}^{T-1}, \hat{\Lambda}_{T-N}^T) \tag{4.11}$$

According to this framework, we can formulate the following moving horizon estimator algorithm:

Algorithm 1: Optimal MHE Algorithm

(1). Find the set $\mathcal{P}_T^f \in \mathcal{P}_T$ of feasible switching patterns in the moving interval $[T-N, T]$ defined in (4.9);

(2). For every feasible switching pattern $\hat{\Lambda}_{T-N}^T = \{\hat{r}_{T-N}, \hat{r}_{T-N+1}, ..., \hat{r}_T\} \in \mathcal{P}_T^f$, solve the following optimization problem:

$$\min_{\hat{x}_{T-N|T} \in X, \hat{w}_{T-N}^{T-1} \in W^N} J_T(\hat{x}_{T-N|T}, \hat{w}_{T-N}^{T-1}, \hat{\Lambda}_{T-N}^T) \tag{4.12}$$

subjected to the constraints:

$$\begin{aligned}
&w_{\min} \le \hat{w}_{k|T} \le w_{\max}, \text{for } k = T-N, T-N+1, ..., T-1 \\
&v_{\min} \le \hat{v}_{T-N|T} = y_{T-N} - C_d(\hat{r}_{T-N})\hat{x}_{T-N|T} \le v_{\max}, \\
&v_{\min} \le \hat{v}_{k|T} = y_k - C_d(\hat{r}_{T-N}^k)\hat{x}_{T-N|T} - \mathcal{D}_d(\hat{r}_{T-N}^k)\hat{w}_{T-N}^{k-1} \le v_{\max}, \\
&\qquad\qquad \text{for } k = T-N+1, T-N+2, ..., T \\
&z_{\min} \le \hat{z}_{k|T} = \mathcal{H}_d(\hat{r}_{T-N}^k)\hat{x}_{T-N|T} + \mathcal{V}_d(\hat{r}_{T-N}^k)\hat{w}_{T-N}^{k-1} \le z_{\max}, \\
&\qquad\qquad \text{for } k = T-N, T-N+1, ..., T-1
\end{aligned}$$

where

$$\begin{aligned}
C_d(\hat{r}_{T-N}^k) =& C_d(\hat{r}_k) \prod_{i=0}^{k-T+N-1} A_d(\hat{r}_{T-N+i}) \\
\mathcal{D}_d(\hat{r}_{T-N}^k)\hat{w}_{T-N}^{k-1} =& C_d(\hat{r}_k) \times \Big[B_d(\hat{r}_{k-1})\hat{w}_{k-1} \\
&+ \sum_{j=1}^{k-T+N-1} \Big(\prod_{i=j}^{k-T+N-1} A_d(\hat{r}_{T-N+i}) B_d(\hat{r}_{T-N+j-1})\hat{w}(\hat{r}_{T-N+j-1}) \Big) \Big]
\end{aligned}$$

By replacing C_d by H_d in the expressions of C_d and $\mathcal{D}_d$, $\mathcal{H}_d$ and $\mathcal{V}_d$ are similarly defined. Let $\tilde{J}_T(\hat{\Lambda}_{T-N}^T)$ denote the minimum of problem (4.12). Let $\tilde{x}_{T-N|T}$ and $\tilde{w}_{T-N}^{T-1}$ be the estimated state and disturbances that correspond to such a minimum.

(3). The optimal estimate of the switching pattern is solved by

$$\hat{\Lambda}_{T-N}^{oT} = \arg \min_{\hat{\Lambda}_{T-N}^T \in \mathcal{P}_T^f} \tilde{J}_T(\hat{\Lambda}_{T-N}^T)$$

Accordingly, we obtain the optimal estimates $\hat{x}_{T-N|T}^o$ and $\hat{w}_{T-N}^{oT-1}$ that correspond to $\tilde{x}_{T-N|T}$ and $\tilde{w}_{T-N}^{T-1}$ associated to the switching pattern $\hat{\Lambda}_{T-N}^{oT} = \{\hat{r}_{T-N}^o, \hat{r}_{T-N+1}^o, ..., \hat{r}_T^o\}$.

As for the estimated quantities, it should be noted that only $\hat{r}^o_{T-N}$, $\hat{x}^o_{T-N|T}$ and $\hat{w}^o_{T-N}$ are retained in this stage and propagated to the next optimization problem by:

$$\bar{x}_{T-N+1} = A_d(\hat{r}^o_{T-N})\hat{x}^o_{T-N|T} + B_d(\hat{r}^o_{T-N})\hat{w}^o_{T-N} \tag{4.13}$$

where $\bar{x}_{T-N+1}$ denotes a priori prediction of x_{T-N+1}. ■

Apparently, the set $\mathcal{P}_T$ is used when determining $\mathcal{P}_{T+1}$ at time instance $T+1$. Generally, if no a priori assumption on the disturbance and output constraints is given, the computational burden of steps (1)-(2) may be heavy. However, we consider the disturbance and output constraints are bounded polytopes in this chapter, which can simplify the executions of steps (1)-(2). In addition, some computationally efficient numerical solvers might be deployed for online implementation of the proposed MHE algorithm, c.f., [161].

Considering $\hat{x}_{T-N|T}$ is an infinite-dimensional state, direct optimization of (4.12) in step (2) of Algorithm 1 with respect to $\hat{x}_{T-N|T}$ (along with $\hat{w}^T_{T-N}$) can be heavily dependent on the spatial discretization node and may be intractable as the spatial discretization node tends to infinity. To address this issue, we introduce the following implementation strategy that can lead to a fully finite-dimensional MHE problem without doing any spatial discretization in the design stage (in a late lumping manner) and hence easily realizable using standard optimization techniques and finite computing capacity.

As in [37], we define an additional $\hat{w}_{T-N-1|T}$ relating $\hat{x}_{T-N|T}$ to $\bar{x}_{T-N}$ as

$$B_d(\hat{r}^o_{T-N-1|T})\hat{w}_{T-N-1|T} = \hat{x}_{T-N|T} - \bar{x}_{T-N}$$

which converts the estimation of $\hat{x}_{T-N|T}$ to the estimation of a finite-dimensional plant disturbance $\hat{w}_{T-N-1|T}$. By doing so, the cost functional (4.10) can be rewritten as:

$$J_T(\hat{w}^{T-1}_{T-N-1}, \hat{\Lambda}^T_{T-N}) = \sum_{k=T-N-1}^{T-1} \left\langle \hat{w}_{k|T}, \bar{Q}\hat{w}_{k|T} \right\rangle_W + \left\langle \hat{v}_{k+1|T}, R\hat{v}_{k+1|T} \right\rangle_Y$$

where $\bar{Q} = Q$ for $T-N \le k \le T-1$ and $\bar{Q} = B_d(\hat{r}^o_{T-N-1})^* P B_d(\hat{r}^o_{T-N-1})$ when $k = T-N-1$. Along this line, we can propose the following late-lumping MHE algorithm.

Algorithm 2: Late-lumping MHE Algorithm

(1). Identical to step (1) of Algorithm 1;

(2). For every feasible switching pattern $\hat{\Lambda}_{T-N}^{T} = \{\hat{r}_{T-N}, \hat{r}_{T-N+1}, ..., \hat{r}_T\} \in \mathcal{P}_T^f$, solve the following optimization problem:

$$\tilde{J}_T(\hat{\Lambda}_{T-N}^{T}) = \min_{\hat{w}_{T-N-1}^{T-1} \in W^{N+1}} J_T(\hat{w}_{T-N-1}^{T-1}, \hat{\Lambda}_{T-N}^{T}) \tag{4.14}$$

subjected to the constraints identical to step (2) of Algorithm 1, where $\hat{x}_{T-N|T}$ should be expressed by $\hat{x}_{T-N|T} = \bar{x}_{T-N} + B_d(\hat{r}_{T-N-1}^{o})\hat{w}_{T-N-1|T}$. Let $\tilde{x}_{T-N|T}$ and $\tilde{w}_{T-N-1}^{T-1}$ be the estimated state and disturbances that correspond to the minimum $\tilde{J}_T(\hat{\Lambda}_{T-N}^{T})$.
(3). The optimal estimate of the switching pattern is solved by

$$\hat{\Lambda}_{T-N}^{oT} = \arg \min_{\hat{\Lambda}_{T-N}^{T} \in \mathcal{P}_T^f} \tilde{J}_T(\hat{\Lambda}_{T-N}^{T})$$

Accordingly, we obtain the optimal estimates $\hat{w}_{T-N-1}^{oT-1}$ that correspond to $\tilde{w}_{T-N-1}^{T-1}$ associated to the switching pattern $\hat{\Lambda}_{T-N}^{oT}$. The optimal state estimate is given by $\hat{x}_{T-N|T}^{o} = \bar{x}_{T-N} + B_d(\hat{r}_{T-N-1}^{o})\hat{w}_{T-N-1|T}^{o}$. Similarly, only $\hat{r}_{T-N}^{o}$, $\hat{x}_{T-N|T}^{o}$ and $\hat{w}_{T-N}^{o}$ are retained in this step and propagated to the next optimization problem by using (4.13).

■

The Late-lumping MHE Algorithm provides a practicable and implementable framework for the Optimal MHE Algorithm instead of doing early-lumping (spatial discretization or model reduction) in the design stage. In other words, the infinite-dimensionality of the state space remains intact in the Late-lumping MHE Algorithm in the design stage. We will test the performance of Late-lumping MHE Algorithm numerically in the Section 4.6.

Remark 10. *Since stable regular linear systems are focused in this chapter, it is natural to include control actions (i.e., $B^u(r_t)u$, where $u \in U$ and U is assumed to be a finite-dimensional space, and B^u can be either a bounded operator $B^u \in \mathcal{L}(U, X)$ or a unbounded operator $B^u \in \mathcal{L}(U, X_{-1})$ depending on the specific actuations taken place) on the right hand side of system model (4.1a). Consequently, one can derive the counterpart of the discrete-time model (4.4), namely, by adding $B_d^u u_k$, $D_d^u u_k$ and $V_d^u u_k$ to the right hand side of Eqs.(4.4a)-(4.4c), where $B_d^u u_k$, $D_d^u u_k$ and $V_d^u u_k$ are similarly defined as $B_d u_k$, $D_d u_k$ and $V_d u_k$ respectively. In the MHE design, the cost functionals (4.10)-(4.12) and (4.14) are still valid except that we need to update that $\hat{v}_{k|T} = y_k - C_d(\hat{r}_k)\hat{x}_{k|T} - D_d^u(\hat{r}_k)u_k$ and $B_d(\hat{r}_k)\hat{w}_{k|T} = \hat{x}_{k+1|T} - A_d(\hat{r}_k)\hat{x}_{k|T} - B_d^u(\hat{r}_k)u_k$ in the*

formulations of quadratic programming optimization problems and the corresponding constraints.

4.5 Stability analysis

In this section, the stability properties of the proposed optimal MHE estimator in Section 4.4 are derived. To show the convergence properties of such estimator, the following assumptions are needed.

Assumption 6. *System (4.5) is simultaneously exactly observable in $N+1$ steps.*

Assumption 7. *For all $r_i \in \mathcal{S}$, $(A(r_i), C(r_i))$ is exactly observable, and $C(r_i)$ is infinite-time admissible for $\mathbb{T}_r$.*

Assumption 8. *For all $r_i \in \mathcal{S}$, $A(r_i)$ generates a strongly stable contraction C_0-semigroup, and the disturbance and output spaces are finite-dimensional.*

Assumption 9. *Disturbances and output constraints are compact sets.*

Assumptions 6 and 7 guarantees the system mode and state observability. Assumption 8 ensures the infinite-time admissibility of output operators and that the discrete-time systems using CT transformation converges to the continuous-time regular linear systems in the input-output mapping sense as the time discretization interval goes to zero. The stability assumption can be assured by using a shift semigroup approach. Assumption 9 is used to show that disturbances and the estimated output are norm bounded.

Based on the CT transformation, we note that $A_d(r_i)$ is strongly stable and a bounded contractive operator since $A(r_i)$ generates a strongly stable contraction C_0-semigroup [21, The. 3.4.9, The. 12.3.10]. We show that there is a 1-1 correspondence of the exact observability of $(A(r_i), C(r_i))$ and $(A_d(r_i), C_d(r_i))$ by the following proposition.

Proposition 5. *Consider non-switching (i.e., $r_t \equiv r_i \in \mathcal{S}$ is fixed) regular linear system (4.1) and its discrete-time system (4.4) using the Cayley-Tustin transform with a suitable parameter δ. Let Assumptions 7-8 hold. Then $(A_d(r_i), C_d(r_i))$ is exactly observable in finite time $K > 0$.*

Proof. The rational of the proof follows from the continuous-time case [42, Pro. 6.5.2.]. As $C(r_i)$ is infinite-time admissible for $\mathbb{T}_r$, there corresponds an infinite-time observability Gramian $\mathcal{Q}_r \in \mathcal{L}(X)$:

$$\mathcal{Q}_r z = \lim_{\tau \to +\infty} \int_0^\tau (\mathbb{T}_r(t))^* C(r_i)^* C(r_i) \mathbb{T}_r(t) z \, dt \tag{4.15}$$

for all $z \in \mathcal{D}(A(r_i))$, and it is the unique self-adjoint solution of the following Lyapunov equation in the dual space of X_{-1} [42, The. 5.1.1]:

$$A(r_i)^* \mathcal{Q}_r z + \mathcal{Q}_r A(r_i) z = -C(r_i) C(r_i)^* z \tag{4.16}$$

or equivalently solved from the discrete-time Lyapunov equation [23, The. 2.4]:

$$A_d(r_i)^* \mathcal{Q}_r^d A_d(r_i) x - \mathcal{Q}_r^d x = -C_d(r_i) C_d(r_i)^* x \tag{4.17}$$

for all $x \in X$. Since $A_d(r_i)$ is strongly stable (induced by the strong stability of $\mathbb{T}_r$ [21, The. 12.3.10]) and $C_d(r_i)$ is infinite-time admissible (induced by the infinite-time admissibility of $C(r_i)$) [23, Lem. 2.2, The. 2.4] there exists a unique solution of (4.17) named as the infinite-time discrete observability Gramian $\mathcal{Q}_r^d \in \mathcal{L}(X)$ defined as:

$$\mathcal{Q}_r^d x = \lim_{K \to +\infty} \sum_{k=0}^{K} (A_d(r_i)^*)^k C_d(r_i)^* C_d(r_i) (A_d(r_i))^k x \tag{4.18}$$

Hence, $\mathcal{Q}_r$ and $\mathcal{Q}_r^d$ are equal and their positive coercivity is invariant under the Cayley-Tustin transform. Thus, the infinite-time exact observability of $(A_d(r_i), C_d(r_i))$ follows from that of $(A(r_i), C(r_i))$ under the CT transform. Next, we show that the finite-time exact observability of $(A_d(r_i), C_d(r_i))$ follows from that of $(A(r_i), C(r_i))$.

From the infinite-time admissible $C(r_i)$ and the exact observability of $(A(r_i), C(r_i))$ in finite time τ, we have the lower and upper bounds for all $z \in \mathcal{D}(A(r_i))$ as:

$$\begin{aligned} \tilde{k}^2 \|z\|_X^2 \geq \langle \mathcal{Q}_r z, z \rangle &= \int_0^{+\infty} \|C(r_i) \mathbb{T}_r(t) z\|_Y^2 \, dt \\ &= \int_0^\tau \|C(r_i) \mathbb{T}_r(t) z\|_Y^2 \, dt + \int_\tau^{+\infty} \|C(r_i) \mathbb{T}_r(t) z\|_Y^2 \, dt \geq \tilde{\kappa}^2 \|z\|_X^2 \end{aligned}$$

where $\tilde{k} > \tilde{\kappa} > 0$. Thus we have the lower and upper bounds for the discrete observability Gramian $\mathcal{Q}_r^d$ as:

$$\tilde{k}^2 \|z\|_X^2 \geq \left\langle \mathcal{Q}_r^d z, z \right\rangle = \sum_{k=0}^{+\infty} \left\| C_d(r_i) (A_d(r_i))^k z \right\|_Y^2 \geq \tilde{\kappa}^2 \|z\|_X^2$$

By decomposition in time as in [42, Pro. 6.5.2.], we have

$$\begin{aligned}
&\sum_{k=0}^{K} \|C_d(r_i)(A_d(r_i))^k z\|_Y^2 \\
=&\sum_{k=0}^{+\infty} \left\|C_d(r_i)(A_d(r_i))^k z\right\|_Y^2 - \sum_{k=K+1}^{+\infty} \left\|C_d(r_i)(A_d(r_i))^k z\right\|_Y^2 \\
=&\sum_{k=0}^{+\infty} \left\|C_d(r_i)(A_d(r_i))^k z\right\|_Y^2 - \sum_{k=0}^{+\infty} \left\|C_d(r_i)(A_d(r_i))^{k+K+1} z\right\|_Y^2 \\
\geq&(\tilde{\kappa}^2 - \tilde{k}^2 \|(A_d(r_i))^{K+1}\|^2) \cdot \|z\|_X^2
\end{aligned} \tag{4.19}$$

Since A_d is strongly stable and a bounded contractive operator (i.e., $\|A_d(r_i)z\| \leq \|z\|$ and $A_d(r_i)^k z \to 0$ as $k \to +\infty$), the parenthesis in (4.19) becomes positive for some sufficiently large K. For such K, $(A_d(r_i), C_d(r_i))$ is exactly observable. ■

By Proposition 5, we know that Assumptions 7-8 can ensure the stability and observability the non-switching discrete-time system. Next, we prove the stability of the proposed MHE estimator.

Theorem 11. *Under Assumptions 6-9, the square norm of the estimation error is bounded from above as follows:*

$$e_{T-N+k} = \|x_{T-N+k} - \hat{x}^o_{T-N+k|T}\|^2 \leq \xi_{T-N+k}$$

The sequence of ξ_{T-N+k} is shown as

$$\xi_\alpha = \beta_\alpha \tag{4.20a}$$

$$\xi_k = c\,\xi_{k-1} + \beta,\ k = \alpha+1, \alpha+2, \ldots \tag{4.20b}$$

where

$$\begin{aligned}
\beta_\alpha =& \frac{10}{5\underline{\sigma}_P + 2Rc_2} \times \{R[(N+1)\rho_v + 2c_3 N\rho_w + c_1\rho_x] \\
&+ NQ\rho_w + (N+1)R\rho_v + 2\bar{\sigma}_P \|x_0 - \bar{x}_0\|_X^2\}
\end{aligned} \tag{4.21a}$$

$$c = \frac{80\bar{\sigma}_P c_6}{5\underline{\sigma}_P + 2Rc_2} \tag{4.21b}$$

$$\begin{aligned}
\beta =& \frac{10}{5\underline{\sigma}_P + 2Rc_2} \times [2R(N+1)\rho_v + (c_1 R + 8\bar{\sigma}_P c_4)\rho_x \\
&+ (2c_3 NR + NQ + 16\bar{\sigma}_P c_5)\rho_w]
\end{aligned} \tag{4.21c}$$

and the corresponding notations are given by

$$\underline{\sigma}_P = \inf_{x\in X, x\neq 0} \frac{\|Px\|}{\|x\|},\ \bar{\sigma}_P = \sup_{x\in X, x\neq 0} \frac{\|Px\|}{\|x\|}$$

$$c_1 = \sup_{\hat{\Lambda}^T_{T-N}, \Lambda^T_{T-N}\in\mathcal{P}^f_T, x\in X, x\neq 0} \frac{\|[C_d(\hat{\Lambda}^T_{T-N}) - C_d(\Lambda^T_{T-N})]x\|}{\|x\|}$$

$$c_2 = \inf_{\bar{\Lambda}^T_{T-N}\in\mathcal{P}^f_T, x\in X, x\neq 0} \frac{\|C_d(\bar{\Lambda}^T_{T-N})x\|}{\|x\|},\ \rho_w = \sup_{w\in W_c} \|w\|^2$$

$$c_3 = \max_{\bar{\Lambda}^T_{T-N}\in\mathcal{P}^f_T} \|\mathcal{D}_d(\bar{\Lambda}^T_{T-N})\|,\ \rho_v = \sup_{v\in V_c} \|v\|^2$$

$$c_4 = \sup_{\hat{\lambda},\lambda\in\mathcal{S}, x\in X, x\neq 0} \frac{\|[A_d(\hat{\lambda}) - A_d(\lambda)]x\|}{\|x\|}$$

$$c_5 = \sup_{\lambda\in\mathcal{S}, w\in W_c, w\neq 0} \frac{\|B_d(\lambda)w\|}{\|w\|},\ c_6 = \sup_{\lambda\in\mathcal{S}, x\in X, x\neq 0} \frac{\|A_d(\lambda)x\|}{\|x\|}$$

In addition, if P and R are chosen such that $c < 1$, the bounding sequence $\{\xi_k\}$ has the following properties:

(a). $\{\xi_k\}$ converges exponentially to the asymptotic value $e_\infty = \beta/(1-c)$;

(b). if $\xi_k > e_\infty$, then $\xi_{k+1} < \xi_k$, $k = \alpha, \alpha+1, ...$

Proof. The rationale of the proof is inspired by the method presented in [161, The. 2]. Similarly, we seek to find the lower and upper bounds for the optimal cost J^o_T.

For the ease of notations, we drop the second subscript in the estimates of state, disturbances and switching patten. Based on the following equalities:

$$\begin{aligned} y^T_{T-N} &= C_d(\hat{\Lambda}^T_{T-N})\hat{x}_{T-N} + \mathcal{D}_d(\hat{\Lambda}^T_{T-N})\hat{w}^{T-1}_{T-N} + \hat{v}^T_{T-N} \\ &= C_d(\Lambda^T_{T-N})x_{T-N} + \mathcal{D}_d(\Lambda^T_{T-N})w^{T-1}_{T-N} + v^T_{T-N} \end{aligned}$$

one can reformulate the cost functional (4.10) as follows:

$$\begin{aligned} &J_T(\hat{x}_{T-N}, \hat{w}^{T-1}_{T-N}, \hat{\Lambda}^T_{T-N}) \\ &= \|\hat{w}^{T-1}_{T-N}\|^2_Q + \|C_d(\Lambda^T_{T-N})x_{T-N} + \mathcal{D}_d(\Lambda^T_{T-N})w^{T-1}_{T-N} \\ &\quad + v^T_{T-N} - C_d(\hat{\Lambda}^T_{T-N})\hat{x}_{T-N} - \mathcal{D}_d(\hat{\Lambda}^T_{T-N})\hat{w}^{T-1}_{T-N}\|^2_R \\ &\quad + \langle \hat{x}_{T-N} - \bar{x}_{T-N}, P(\hat{x}_{T-N} - \bar{x}_{T-N})\rangle_X \end{aligned} \tag{4.22}$$

Based on the optimality principle, we obtain the upper bound for the cost functional J^o_T, namely, $J^o_T \leq J_T(x_{T-N|T}, w^{T-1}_{T-N}, \Lambda^T_{T-N})$. Combining with the Schwarz inequality

[167, Lem. 3.2-1] and the boundedness of operator P, we obtain:

$$\begin{aligned}
&J_T^o(\hat{x}_{T-N|T}, \hat{w}_{T-N}^{T-1}, \hat{\Lambda}_{T-N}^T) \\
\leq\ & \|w_{T-N}^{T-1}\|_Q^2 + \|v_{T-N}^T\|_R^2 + \|x_{T-N} - \bar{x}_{T-N}\|_X \|P(x_{T-N} - \bar{x}_{T-N})\|_X \\
\leq\ & NQ\rho_w + (N+1)R\rho_v + \bar{\sigma}_P \|x_{T-N} - \bar{x}_{T-N}\|_X^2
\end{aligned} \tag{4.23}$$

Then, we move to show the lower bound on the optimal cost J_T^o. Based on the Schwarz inequality and the triangle inequality [167, Lem. 3.2-1], we can derive the following:

$$\begin{aligned}
&\langle x_{T-N} - \hat{x}_{T-N}, P(x_{T-N} - \hat{x}_{T-N})\rangle_X \\
=&\langle x_{T-N} - \bar{x}_{T-N} + \bar{x}_{T-N} - \hat{x}_{T-N}, P(x_{T-N} - \bar{x}_{T-N} + \bar{x}_{T-N} - \hat{x}_{T-N})\rangle_X \\
=&\langle x_{T-N} - \bar{x}_{T-N}, P(x_{T-N} - \bar{x}_{T-N})\rangle_X + \langle \bar{x}_{T-N} - \hat{x}_{T-N}, P(\bar{x}_{T-N} - \hat{x}_{T-N})\rangle_X \\
&+ 2\langle x_{T-N} - \bar{x}_{T-N}, P(\bar{x}_{T-N} - \hat{x}_{T-N})\rangle_X \\
\leq& 2\|x_{T-N} - \bar{x}_{T-N}\|_P^2 + 2\|\bar{x}_{T-N} - \hat{x}_{T-N}\|_P^2
\end{aligned} \tag{4.24}$$

Thus, the following lower bound for the last term on the right-hand side of (4.22) is obtained:

$$\|\hat{x}_{T-N} - \bar{x}_{T-N}\|_P^2 \geq \frac{1}{2}\|x_{T-N} - \hat{x}_{T-N}\|_P^2 - \|x_{T-N} - \bar{x}_{T-N}\|_P^2 \tag{4.25}$$

As for the second term on the right-hand side of (4.22), we formulate:

$$\begin{aligned}
&\|\mathcal{C}_d(\hat{\Lambda}_{T-N}^T)(x_{T-N} - \hat{x}_{T-N})\|_R^2 \\
=&\|y_{T-N}^T - \mathcal{C}_d(\hat{\Lambda}_{T-N}^T)\hat{x}_{T-N} - \mathcal{D}_d(\hat{\Lambda}_{T-N}^T)\hat{w}_{T-N}^{T-1} \\
&- \mathcal{D}_d(\Lambda_{T-N}^T)w_{T-N}^{T-1} - v_{T-N}^T + \mathcal{D}_d(\hat{\Lambda}_{T-N}^T)\hat{w}_{T-N}^{T-1} \\
&+ [\mathcal{C}_d(\hat{\Lambda}_{T-N}^T) - \mathcal{C}_d(\Lambda_{T-N}^T)]x_{T-N}\|_R^2 \\
\leq& 5\{\|y_{T-N}^T - \mathcal{C}_d(\hat{\Lambda}_{T-N}^T)\hat{x}_{T-N} - \mathcal{D}_d(\hat{\Lambda}_{T-N}^T)\hat{w}_{T-N}^{T-1}\|_R^2 \\
&+ \|\mathcal{D}_d(\Lambda_{T-N}^T)w_{T-N}^{T-1}\|_R^2 + \|\mathcal{D}_d(\hat{\Lambda}_{T-N}^T)\hat{w}_{T-N}^{T-1}\|_R^2 \\
&+ \|v_{T-N}^T\|_R^2 + \|[\mathcal{C}_d(\hat{\Lambda}_{T-N}^T) - \mathcal{C}_d(\Lambda_{T-N}^T)]x_{T-N}\|_R^2\}
\end{aligned}$$

which implies that:

$$\begin{aligned}
&\|y_{T-N}^T - \mathcal{C}_d(\hat{\Lambda}_{T-N}^T)\hat{x}_{T-N} - \mathcal{D}_d(\hat{\Lambda}_{T-N}^T)\hat{w}_{T-N}^{T-1}\|_R^2 \\
\geq& \frac{1}{5}\|\mathcal{C}_d(\hat{\Lambda}_{T-N}^T)(x_{T-N} - \hat{x}_{T-N})\|_R^2 - \|v_{T-N}^T\|_R^2
\end{aligned}$$

$$- \|\mathcal{D}_d(\Lambda_{T-N}^T)w_{T-N}^{T-1}\|_R^2 - \|\mathcal{D}_d(\hat{\Lambda}_{T-N}^T)\hat{w}_{T-N}^{T-1}\|_R^2$$
$$- \|[C_d(\hat{\Lambda}_{T-N}^T) - C_d(\Lambda_{T-N}^T)]x_{T-N}\|_R^2 \tag{4.26}$$

More specifically, we consider the following two cases.

Case I: If $\hat{\Lambda}_{T-N}^T = \Lambda_{T-N}^T$, we note that $\|[C_d(\hat{\Lambda}_{T-N}^T) - C_d(\Lambda_{T-N}^T)]x_{T-N}\|_R^2 = 0$;

Case II: If $\hat{\Lambda}_{T-N}^T \neq \Lambda_{T-N}^T$, we have $\|[C_d(\hat{\Lambda}_{T-N}^T) - C_d(\Lambda_{T-N}^T)]x_{T-N}\|_R^2 \leq c_1\|x_{T-N}\|_R^2 \leq c_1\rho_x$ by Theorem 10.

From Cases I and II, we can derive the following:

$$\|[C_d(\hat{\Lambda}_{T-N}^T) - C_d(\Lambda_{T-N}^T)]x_{T-N}\|_R^2 \leq \max\{0, c_1\rho_x\} = c_1\rho_x$$

By denoting that $e_{T-N} = \|x_{T-N} - \hat{x}_{T-N}^o\|$ and substituting (4.25) and (4.26) into J_T^o, we have the following lower bound:

$$J_T^o(\hat{x}_{T-N}^o, \hat{w}_{T-N}^{oT-1}, \hat{\Lambda}_{T-N}^{oT}) \geq \frac{1}{2}\underline{\sigma}_P e_{T-N} - \bar{\sigma}_P\|x_{T-N} - \bar{x}_{T-N}\|^2 + \frac{1}{5}Rc_2 e_{T-N}$$
$$- R[(N+1)\rho_v + 2c_3N\rho_w + c_1\rho_x] \tag{4.27}$$

In order to derive a bounding sequence on the quadratic estimation error, we combine the lower bound (4.23) and upper bound (4.27) as follows:

$$(\frac{1}{2}\underline{\sigma}_P + \frac{1}{5}Rc_2)e_{T-N} \leq R[(N+1)\rho_v + 2c_3N\rho_w + c_1\rho_x]$$
$$+ NQ\rho_w + (N+1)R\rho_v + 2\bar{\sigma}_P\|x_{T-N} - \bar{x}_{T-N}\|_X^2 \tag{4.28}$$

As for the last term on the right-hand side of (4.28), we derive the following inequality by using (4.4) and (4.13):

$$\begin{aligned}
&\|x_{T-N} - \bar{x}_{T-N}\|_X^2 \\
=&\|A_d(r_{T-N-1})x_{T-N-1} + B_d(r_{T-N-1})w_{T-N-1}^o \\
&- A_d(\hat{r}_{T-N-1}^o)\hat{x}_{T-N-1}^o + B_d(\hat{r}_{T-N-1}^o)\hat{w}_{T-N-1}^o\| \\
\leq&4\{\|A_d(\hat{r}_{T-N-1}^o)(x_{T-N-1} - \hat{x}_{T-N-1}^o)\|^2 \\
&+ \|[A_d(r_{T-N-1}) - A_d(\hat{r}_{T-N-1}^o)]x_{T-N-1}\|^2 \\
&+\|B_d(r_{T-N-1})w_{T-N-1}\|^2 + \|B_d(\hat{r}_{T-N-1}^o)\hat{w}_{T-N-1}^o\|^2\} \\
\leq&4c_6 e_{T-N-1} + 4c_4\rho_x + 8c_5\rho_w
\end{aligned} \tag{4.29}$$

where by considering two cases as previous discussion (i.e., either $\hat{\Lambda}_{T-N-1}^{T-1}$ is equal to Λ_{T-N-1}^{T-1} or not), we obtain that $\|[A_d(r_{T-N-1}) - A_d(\hat{r}_{T-N-1}^o)]x_{T-N-1}\|^2 \leq c_4\rho_x$. In

addition, we note that $e_{T-N-1} = \|(x_{T-N-1} - \hat{x}^o_{T-N-1|T-1})\|^2$. By inserting (4.29) into (4.28), we can readily attain a sequence $\{\xi_k\}$ defined in (4.20) as the upper bound of e_k with the corresponding notations β_α, c and β in (4.21).

Moreover, we note that for $c < 1$, ξ_k converges exponentially to $e_\infty = \beta/(1-c)$ as $k \to +\infty$; if $\xi_k > e_\infty$, then $\xi_{k+1} < \xi_k$, $k = \alpha, \alpha+1, ...$ ■

The condition $c < 1$ can be easily ensured by properly selecting penalty weights P and R. If $\rho_w = 0, \rho_v = 0, \rho_x = 0$, the estimation error converges exponentially to zero under the condition that $c < 1$, which implies that for the noise-free case, we can realize perfect estimation in the sense that $e_k \to 0$ as $k \to +\infty$ provided that we know accurate priori knowledge on the switching patterns. Note that the sequence $\{\xi_k\}$ depends continuously on ρ_w, ρ_v, ρ_x, from which one might conclude that if "small" values of ρ_w, ρ_v, ρ_x are considered, such an upper bound (e_∞) can be "arbitrarily close" to the upper bound in the noise-free cases with switching-patterns being available.

4.6 Examples

4.6.1 Example 1: heat exchanger

Consider the following counterflow heat exchanger model on a 1-D spatial domain $\zeta \in [0, l]$:

$$x_{1t}(\zeta,t) = -a(r_t)x_{1\zeta}(\zeta,t) + b(r_t)(x_2(\zeta,t) - x_1(\zeta,t)) + f_1(\zeta)w(t) \tag{4.30a}$$

$$x_{2t}(\zeta,t) = c(r_t)x_{2\zeta}(\zeta,t) - d(r_t)(x_2(\zeta,t) - x_1(\zeta,t)) + f_2(\zeta)w(t) + g(\zeta)u(t) \tag{4.30b}$$

$$x_1(0,t) = x_2(l,t) = 0, x_1(\zeta,0) = x_{10}, x_2(\zeta,0) = x_{20} \tag{4.30c}$$

$$y(t) = [x_1(l,t); x_2(0,t)] + v(t),\ z(t) = x_1(\frac{l}{2}, t) \tag{4.30d}$$

where $a(r_t), b(r_t), c(r_t), d(r_t)$ are positive constants for every r_t on $t \in \mathbb{R}^+$, and they are randomly switching in a finite set $\mathcal{S} = \{r_1, r_2\}$ and they are known once r_t is identified. x_1 and x_2 represent the transient temperatures of the counter-current fluids. In this case, we consider bounded plant disturbance w equipped with spatial distribution functions $f_1(\zeta)$ and $f_2(\zeta)$, and bounded measurement disturbance v. A spatially distributed control described by $g(\zeta)$ is considered. Non-switching heat exchanger models have been widely studies, e.g., in [168, 163].

We formulate the state space model for the heat exchanger system (4.30). Let us take the Hilbert space $X = L^2(0,l)^2$ to be the state space, and define the input, disturbance and output spaces as $U = Z = W = \mathbb{R}$ and $Y = \mathbb{R}^2$. The unbounded operator $A(r_t) : \mathcal{D}(A(r_t)) \rightarrow X$ is defined by $\mathcal{D}(A(r_t)) = \{\phi = (\phi_1; \phi_2) \in X | \phi_j', \phi_j \in L^2(0,l)$ for $j = 1,2$, $\phi_1(0) = \phi_2(l) = 0\}$ and $A(r_t)(\cdot) = [-a, 0; 0, c]\frac{\partial}{\partial\zeta}(\cdot) + [-b, b; d, -d](\cdot)$. For each r_t on $t \in \mathbb{R}^+$, $A(r_t)$ is the generator of C_0-semigroup of contraction. In addition, we consider bounded control and plant disturbance operators as $B_w = [f_1; f_2]$ and $B_u = [0; g]$, and unbounded observation operators $C(\cdot) = [\int_0^l \delta_l(\cdot)d\zeta, 0; 0, \int_0^l \delta_0(\cdot)d\zeta]$, and $H(\cdot) = \int_0^l \delta_{\frac{l}{2}}(\cdot)d\zeta$, where δ denotes the Dirac delta function. As in [169, 170], one can show that the heat exchanger model is a well-posed regular linear system. By applying linear transformation $x_1 = \sqrt{b}\bar{x}_1$ and $x_2 = \sqrt{d}\bar{x}_2$, it is straightforward to obtain the PDE model corresponding to $(\bar{x}_1, \bar{x}_2)$ as $A_r(\cdot) = [-a, 0; 0, c]\frac{\partial}{\partial\zeta}(\cdot) + [-b, \sqrt{bd}; \sqrt{bd}, -d](\cdot)$. From [171, The. 2, The. 3], we note that the system is exponentially stable and exactly observable in finite-time $l/\min\{a(r_t), c(r_t)\}$ for every r_t. Based on that, we note that $C(r_t)$ is infinite-time admissible [122, Pro. 5.5].

By applying Laplace transformation, we can solve for the resolvent operator analytically as:

$$
\begin{aligned}
R_{11}(s,A) =& \frac{-a_{12}(\zeta)}{aa_{22}(l)}\int_0^l a_{21}(l-\eta)(\cdot)d\eta + \frac{1}{a}\int_0^\zeta a_{11}(\zeta-\eta)(\cdot)d\eta \\
R_{12}(s,A) =& \frac{a_{12}(\zeta)}{ca_{22}(l)}\int_0^l a_{22}(l-\eta)(\cdot)d\eta - \frac{1}{c}\int_0^\zeta a_{12}(\zeta-\eta)(\cdot)d\eta \\
R_{21}(s,A) =& \frac{-a_{22}(\zeta)}{aa_{22}(l)}\int_0^l a_{21}(l-\eta)(\cdot)d\eta + \frac{1}{a}\int_0^\zeta a_{21}(\zeta-\eta)(\cdot)d\eta \\
R_{22}(s,A) =& \frac{a_{22}(\zeta)}{ca_{22}(l)}\int_0^l a_{22}(l-\eta)(\cdot)d\eta - \frac{1}{c}\int_0^\zeta a_{22}(\zeta-\eta)(\cdot)d\eta
\end{aligned}
$$

where

$$
\begin{aligned}
a_{11}(\zeta) =& \frac{1}{2U_s}[(U_s - S_s)e^{\frac{T_s+U_s}{2ac}\zeta} + (U_s + S_s)e^{\frac{T_s-U_s}{2ac}\zeta}] \\
a_{12}(\zeta) =& \frac{bc}{U_s}(-1 + e^{\frac{U_s}{ac}\zeta})e^{\frac{T_s-U_s}{2ac}\zeta},\ a_{21}(\zeta) = -\frac{ad}{bc}a_{12}(\zeta) \\
a_{22}(\zeta) =& \frac{1}{2U_s}[(U_s - S_s)e^{\frac{T_s-U_s}{2ac}\zeta} + (U_s + S_s)e^{\frac{T_s+U_s}{2ac}\zeta}]
\end{aligned}
$$

with $T_s = a(d+s) - c(b+s)$, $S_s = a(d+s) + c(b+s)$, and $U_s = \sqrt{4acs(b+d+s) + T_s^2}$. Along this line, one can substitute the resolvent solution to obtain the analytic forms

of all discrete operators $(A_d, B_d, C_d, D_d, H_d, V_d)$ as follows:

$$A_d(\cdot) = -(\cdot) + 2\delta \begin{bmatrix} R_{11}(\delta, A)(\cdot) & R_{12}(\delta, A)(\cdot) \\ R_{21}(\delta, A)(\cdot) & R_{22}(\delta, A)(\cdot) \end{bmatrix} \tag{4.31a}$$

$$B_d = \sqrt{2\delta} \begin{bmatrix} R_{11}(\delta, A) f_1 + R_{12}(\delta, A) f_2 \\ R_{21}(\delta, A) f_1 + R_{22}(\delta, A) f_2 \end{bmatrix} \tag{4.31b}$$

$$C_d(\cdot) = \sqrt{2\delta} \begin{bmatrix} C_{d11}(\cdot) & C_{d12}(\cdot) \\ C_{d21}(\cdot) & C_{d22}(\cdot) \end{bmatrix}, D_d = \begin{bmatrix} C_{d11} f_1 + C_{d12} f_2 \\ C_{d21} f_1 + C_{d22} f_2 \end{bmatrix} \tag{4.31c}$$

$$\begin{aligned}
C_{d11}(\cdot) &= \frac{1}{a} \int_0^l \Big[\frac{-a_{12}(l)}{a_{22}(l)} a_{21}(l-\eta) + a_{11}(l-\eta) \Big](\cdot) d\eta \\
C_{d12}(\cdot) &= \frac{1}{c} \int_0^l \Big[\frac{a_{12}(l)}{a_{22}(l)} a_{22}(l-\eta) - a_{12}(l-\eta) \Big](\cdot) d\eta \\
C_{d21}(\cdot) &= \frac{-a_{22}(0)}{a a_{22}(l)} \int_0^l a_{21}(l-\eta)(\cdot) d\eta \\
C_{d22}(\cdot) &= \frac{a_{22}(0)}{c a_{22}(l)} \int_0^l a_{22}(l-\eta)(\cdot) d\eta \\
H_d(\cdot) &= \sqrt{2\delta} \begin{bmatrix} H_{d1}(\cdot) & H_{d2}(\cdot) \end{bmatrix}, V_d = H_{d1} f_1 + H_{d2} f_2
\end{aligned} \tag{4.31d}$$

$$\begin{aligned}
H_{d1}(\cdot) &= \frac{-a_{12}(\frac{l}{2})}{a a_{22}(l)} \int_0^l a_{21}(l-\eta)(\cdot) d\eta + \frac{1}{a} \int_0^{\frac{l}{2}} a_{11}(\frac{l}{2} - \eta)(\cdot) d\eta \\
H_{d2}(\cdot) &= \frac{a_{12}(\frac{l}{2})}{c a_{22}(l)} \int_0^l a_{22}(l-\eta)(\cdot) d\eta - \frac{1}{c} \int_0^{\frac{l}{2}} a_{12}(\frac{l}{2} - \eta)(\cdot) d\eta
\end{aligned}$$

In a similar manner, the discrete operators (B_d^u, D_d^u, V_d^u) related to the control input can be determined accordingly. Specifically, we consider bounded spatial functions $f_1(\zeta) = f_2(\zeta) = 1$, $g(\zeta) = \sin(\pi\zeta)$. In this example, we consider two pairs of model parameters as: $(a(r_1), b(r_1), c(r_1), d(r_1)) = (0.5, 6, 0.5, 1.5)$ and $(a(r_2), b(r_2), c(r_2), d(r_2)) = (0.1, 2, 0.75, 6)$. By substituting these into Eq. (4.31), we can determine two pairs of model operators. The initial conditions (i.e., x_{10} and x_{20}), plant and measurement disturbances (namely, w and v) are assumed to be mutually independent random variables. The initial conditions utilized in the example are taken as: $x_{10} = T_{10}(k_1\zeta^2 + 1)e^{-\zeta^2} - T_{10}$ and $x_{20} = T_{20}e^{-(1-\zeta^2)/t_1} - T_{20}$, where $k_1 = \frac{T_{20}}{T_{10}}e^1 - 1$ and $t_1 = 1/\log\frac{T_{20}}{T_{10}}$, see [168]. In the simulations, the values of T_{10} and T_{20} are taken as: $T_{10} \in [18, 22]$ and $T_{20} \in [78, 82]$. Moreover, we consider the control input as $u(t) = 1.8\sin(0.25\pi t)$. The time and space discretization intervals are set as $h = 0.2$ s and $\Delta\zeta = 0.01$ m, with overall length $l = 1$ m. As for the MHE and MHEPI, we select weighting parameters as $P = 0.001$, $Q = 0.1$, and $R = 0.1$, and consider constraints $w_k \in [-r_w, r_w]$, $v_k \in [-r_v, r_v]$, and $z_k \in [-10, 35]$. The priori estimates of x_{10} and

x_{20} for both estimation algorithms are taken as the $0.5x_{10}$ and $0.5x_{20}$ accordingly. Moreover, the values of the minimum dwell time (i.e., the minimum number of time steps between one switch and the next) and the moving horizon length of MHE and MHEPI are taken as 3.

To quantify the estimation performance, we consider two performance indexes including Root Mean Square Error (RMSE) to evaluate the output estimation error and Average Missing Rate (AMR) to assess the mode detection performance as in [157]:

$$\text{RMSE} = \frac{1}{S}\sum_{s=1}^{S}\sqrt{\sum_{k=1}^{T}\frac{\|e_{k,s}\|^2}{T}},\ \text{AMR} = \frac{1}{ST}\sum_{s=1}^{S} f_{d,s}$$

where $e_{k,s}$ denotes the output estimation error at time k in the s-th simulation run and $f_{d,s}$ represents the total number of the false detections of the switching modes in the s-th simulation run. S and T are the number of simulation runs and simulation horizon. In this case, the values of S and T are taken as 200 and 100, respectively.

For brevity, we denote the estimator proposed in the Algorithm 2 by moving horizon estimator (MHE). For the sake of comparison, we investigate one baseline algorithm, namely, moving horizon estimator with perfect mode information (MHEPI) (i.e., by minimizing the cost J_T under the further constraint $\hat{\Lambda}_{T-N}^{T} = \Lambda_{T-N}^{T}$). The quadratic optimization is conducted using Matlab optimization toolbox.

The influences of different levels of disturbances on the performance of considered estimators are investigated and summarized in Table 4.1. It can be seen that MHEPI outperforms MHE in all three cases since MHEPI has access to perfect switching mode information. With the increase of disturbance levels, the estimation errors by using MHE and MHEPI increases as well. Under weak disturbances (i.e., $r_w = r_v = 0.1$), MHE shows a satisfactory performance on mode detection indicated by the average missing rate (AMR) being 0.1003. The AMR of MHE increases as the levels of disturbances increase. It is remarkable that the proposed MHE can achieve comparable output estimation performance to MHEPI when subjected to strong disturbance (i.e., $r_w = r_v = 1.0$). The state and output estimation results under medium level of disturbances (i.e., $r_w = r_v = 0.5$) are shown in Fig. 4.1 and Fig. 4.2, from which one can clearly observe that the proposed MHE is capable of achieving constrained

state/output estimation and mode identification.

Table 4.1: Estimation performance of the proposed estimator on the heat exchanger with different r_w and r_v

r_w		0.1	0.5	1.0
r_v		0.1	0.5	1.0
RMSE	MHE	0.0944	0.1367	0.1971
	MHEPI	0.0789	0.1286	0.1889
AMR	MHE	0.1003	0.1711	0.2432

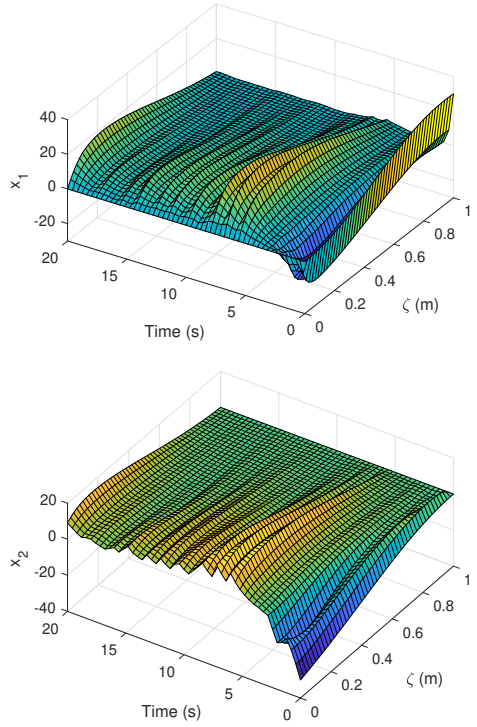


Figure 4.1: Estimated state profiles of the heat exchanger with parameters $r_w = r_v = 0.5$ (the upper one is x_1 and the lower one is x_2)

4.6.2 Example 2: damped Rayleigh beam equation

Consider a damped Rayleigh beam equation on a 1-D spatial domain $\zeta \in [0, \pi]$ and the time $t \geq 0$ [13]:

$$x_{tt}(\zeta,t) - \alpha(r_t)x_{\zeta\zeta tt}(\zeta,t) - a(r_t)x_{\zeta\zeta t}(\zeta,t) + x_{\zeta\zeta\zeta\zeta}(\zeta,t) = 0 \tag{4.32a}$$

$$x(0,t) = x(\pi,t) = x_{\zeta\zeta}(\pi,t) = 0, \tag{4.32b}$$

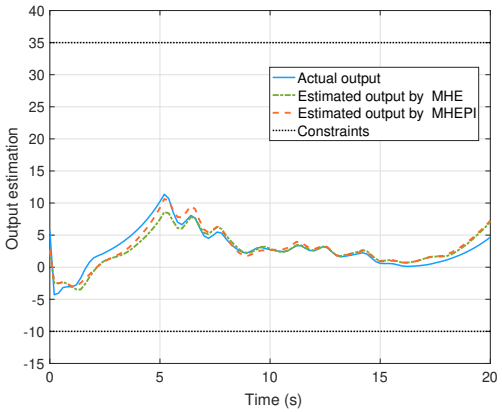


Figure 4.2: Output estimation of the heat exchanger with parameters $r_w = r_v = 0.5$

$$-x_{\zeta\zeta}(0,t) = w(t) + u(t), x(\zeta,0) = x_{10}, x_t(\zeta,0) = x_{20} \tag{4.32c}$$

$$y(t) = x_{\zeta t}(0,t) + v(t),\ z(t) = x_{\zeta t}(\frac{\pi}{5},t) \tag{4.32d}$$

where x denotes the transverse displacement of the beam. Here $\alpha(r_t) > 0$ is proportional to the moment of inertia of the cross section of the beam and $a(r_t) > 0$ is the damping coefficient of the structural damping effect. $\alpha(r_t)$ and $a(r_t)$ are randomly switching in a finite set $\mathcal{S} = \{r_1, r_2\}$ and once r_t is identified $\alpha(r_t)$ and $a(r_t)$ are known, see, [172, 173]. In addition, the torque disturbance w and control input u at $\zeta = 0$ are considered along with collocated angular velocity observation y.

We formulate the state space model for the Rayleigh beam equation as in [13, 164]. We denote $\bar{H} = \mathcal{H}_0^1(0,\pi)$, $V = \mathcal{H}^2(0,\pi) \cap \mathcal{H}_0^1(0,\pi)$, and define the inner product on $\bar{H}$ such that $\langle \phi, \varphi \rangle_{\bar{H}} = \langle (I - \alpha d^2/dx^2)\phi, \varphi \rangle_{L^2}$, $\forall \phi, \varphi \in V$. Let us introduce the operator $\mathcal{R} : L^2[0,\pi] \to V$ as: $\mathcal{R} = (I - \alpha d^2/dx^2)^{-1}$. It has been shown that as a bounded operator, $\mathcal{R}$ is strictly positive [164]. Let us define the operator $A_0 : \mathcal{D}(A_0) \to \bar{H}$ by: $\mathcal{D}(A_0) = \{\phi \in \mathcal{H}^3(0,\pi) | \phi(0) = \phi(\pi) = 0, \phi_{\zeta\zeta}(0) = \phi_{\zeta\zeta}(\pi) = 0\}$, $A_0\phi = (d^4/d\zeta^4)(\mathcal{R}\phi)$, $\forall \phi \in \mathcal{D}(A_0)$. The operator A_0 is self-adjoint, strictly positive, and commutes with $\mathcal{R}$. In what follows, we share use the notations: $\bar{H}_1 = \mathcal{D}(A_0)$, $\bar{H}_{-\frac{1}{2}} = L^2[0,\pi]$, $\bar{H}_{\frac{1}{2}} = V$, and $\bar{H}_{-1} = \mathcal{H}^{-1}(0,\pi)$. Thus we have $A_0 \in \mathcal{L}(\bar{H}_{\frac{1}{2}}, \bar{H}_{-\frac{1}{2}})$ and A_0 commutes with $\mathcal{R}$ (and hence with $\mathcal{R}^{-1}$).

We take transverse displacement x and its velocity x_t as an extended state (i.e., $x^e = [x; x_t]$). Define $X = \bar{H}_{\frac{1}{2}} \times \bar{H}$ and $U = Y = W = Z = \mathbb{C}$ as the state, input, output and disturbance spaces. It has been shown in [13] that A is m-dissipative, hence by

the Lumer–Phillips theorem [42, The. 3.8.4] we note that A is the generator of a contraction semigroup on X as $A : \mathcal{D}(A) = \bar{H}_1 \times \bar{H}_{\frac{1}{2}} \to X$, with $A = [0, I; -A_0, -A_1]$, where $A_1 = -a\mathcal{R}d^2/dx^2 \in \mathcal{L}(V)$, and $A_1 \geq 0$ on V and $\bar{H}$. Denote $X_1 = \mathcal{D}(A)$ and $X_{-1} = \bar{H} \times \bar{H}_{-\frac{1}{2}}$. Define the operator $C_0 \in \mathcal{L}(\bar{H}_{\frac{1}{2}}, \mathbb{C})$ with $C_0\phi = \phi_\zeta(0)$ and the observation operator $C = [0, C_0] \in \mathcal{L}(X_1, \mathbb{C})$. Moreover, we note that $B = C^* = [0; C_0^*] \in \mathcal{L}(\mathbb{C}, X_{-1})$. From [164, p. 649] and [174, Eq.(2.6)], we note that $C_0^* = -\mathcal{R}\frac{d}{d\zeta}\delta_0 = -\frac{1}{\alpha} \times [\coth(\frac{\pi}{\sqrt{\alpha}}) \sinh(\frac{\zeta}{\sqrt{\alpha}}) - \cosh(\frac{\zeta}{\sqrt{\alpha}})]$. Similarly, we define that $H = [0, H_0] \in \mathcal{L}(X_1, \mathbb{C})$, where the operator $H_0 \in \mathcal{L}(\bar{H}_{\frac{1}{2}}, \mathbb{C})$ with $H_0\phi = \phi_\zeta(\frac{\pi}{5})$.

When there is no structural damping (i.e., $a = 0$), it has been shown that the above system is a well-posed regular linear system and it is exactly observable and controllable in [164, Rem. 5.5 and Rem. 6.3]. Since A_1 is a bounded operator, we note that the damped Rayleigh beam system (i.e., $a \neq 0$) is still a well-posed regular system [13] and exactly observable and controllable [175, The. 3.3]. Moreover, it is established that A is exponentially stable in [13] by using [176, Pro. 3.14 and The. 3.18], which implies that C is infinite-time admissible [122, Pro. 5.5].

By performing Laplace transform of (4.32a) with the boundary conditions (4.32c), the resolvent operator $R(s, A) = [R_{11}(s, A), R_{12}(s, A); R_{21}(s, A), R_{22}(s, A)]$ can be determined as in Example 1. By a direct calculation, one can obtain the expressions of the discrete-time operators $(A_d, B_d, C_d, D_d, H_d, V_d)$. In this case, we consider that there are two system modes with parameters $(\alpha(r_1), a(r_1)) = (1 \times 10^{-5}, 1 \times 10^{-5})$ and $(\alpha(r_2), a(r_2)) = (0.05, 0.3)$, respectively. Consequently, one can obtain two pairs of model operators by substituting these parameters. In addition, we assume that the system has a minimum dwell time that is taken as 6. In the parameter setting of MHE and MHEPI, the moving horizon length is taken as $N = 5$, and weighting parameters are selected as $P = 0.001$, $Q = 0.1$, and $R = 1 \times 10^{-6}$. We consider constraints $w_k \in [-r_w, r_w]$, $v_k \in [-r_v, r_v]$, and $z_k \in [-80, 80]$. A periodic input signal is considered as $u(t) = 4\sin(0.2\pi t)$ in this example. The initial conditions are considered to be $x_{10} = x_{20} = A_0 \sin(6\zeta)$, where the value of A_0 is taken as $A_0 \in [1.9, 2.1]$. The priori estimates of x_{10} and x_{20} for MHE and MHEPI are taken as the $0.5x_{10}$ and $0.5x_{20}$, respectively.

Similarly to Example 1, we investigate the proposed MHE algorithm (i.e., Algorithm 2) in comparison with the baseline algorithm MHEPI subjected to different

levels of disturbances in this example. The quadratic optimization is carried out by means of Matlab optimization toolbox. Moreover, the values of simulation runs M and simulation horizon T are taken as 300 and 160, respectively. In each simulation, the time discretization interval and spatial node are taken as $h = 0.25$ s and 100, respectively.

As shown in Table 4.2, MHEPI outperforms MHE in terms of RMSE in all three scenarios. The estimation errors by using MHE or MHEPI increase with the increase of disturbance magnitudes in terms of RMSE. For the considered model switching parameters, the AMR of MHE is around 0.34 and it is remarkable that AMR slightly increases as the level of disturbance increases. More specifically, the estimation performance of MHE subjected to the medium level of disturbances is illustrated in Fig. 4.3 and Fig. 4.4, from which one can observe that the estimated state profiles are quite smooth and the output estimation error of MHE slowly decreases as time increases. Overall, the mode detection for the beam equation is more changeling compared to the heat exchanger case due to the complex model structures (i.e., second-order derivative in time and fourth-order derivative in space). It is expected that by having more distinguishable model switching parameters and weaker disturbances, the estimation performance of MHE can be improved.

Table 4.2: Estimation performance of the proposed estimator on the damped Rayleigh beam equation with different r_w and r_v

r_w		0.1	0.5	1.0
r_v		0.1	0.5	1.0
RMSE	MHE	0.9025	0.9224	0.9251
	MHEPI	0.3560	0.3587	0.3706
AMR	MHE	0.3398	0.3428	0.3453

4.7 Conclusion

In this chapter, moving horizon estimator designs of a class of regular linear infinite-dimensional systems have been accomplished with consideration of bounded disturbances and unknown and unpredictable switching modes. The challenges lie in handling unbounded operators, constraint presence, and simultaneous state/output and mode estimation. To address these, we have applied Cayley-Tustin transformation

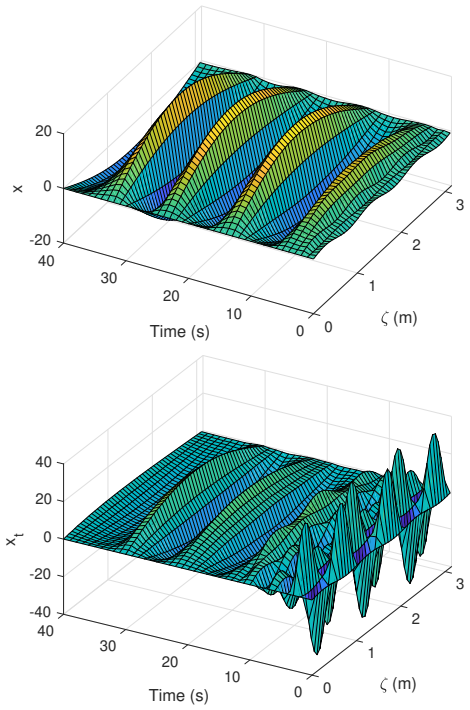


Figure 4.3: Estimated state profiles of the damped Rayleigh beam equation with parameters $r_w = r_v = 0.5$ (the upper one is x and the lower one is x_t)

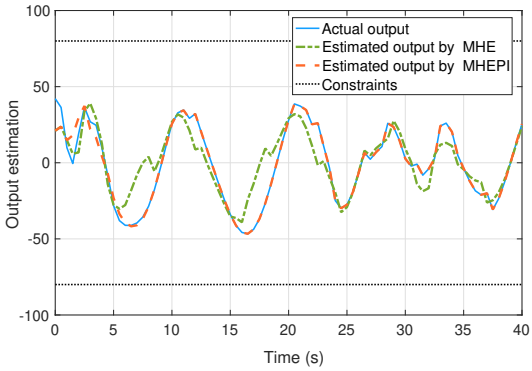


Figure 4.4: Output estimation of the damped Rayleigh beam equation with parameters $r_w = r_v = 0.5$

to the regular linear infinite-dimensional system with unbounded operators leading to a discrete-time infinite-dimensional model with all bounded operators, while preserving essential model properties (e.g., input-output mappings, stability, and etc.). Two discrete-time MHE algorithms have been proposed for simultaneous state/output and mode estimation and constraint handling. Considering that the difficulties of this

problem concerns how to establish appropriate observability conditions, we have proposed simultaneous exact observability concepts and explored their corresponding properties. Based on that, the stability analysis of the proposed MHE algorithm has been provided. Two representative examples have verified the proposed design.

Chapter 5

Soft Sensor Design for Lumped Parameter Systems Using Variational Bayesian Inference

Soft sensors (also called virtual sensors) are widely utilized for online measurements of key process variables [177]. Considering that key process variables might not be measured in an online manner (e.g., concentrations of the quality products in chemical processes are usually measured in laboratories using advanced analyzers), soft sensors are often used to realize online estimation of key process variables and hence supplement online instrument measurements for advanced process monitoring and control [178].

5.1 Introduction

Data-driven methods have been extensively applied to soft sensor development for a wide range of process engineering fields over the past decades [179]. One of the main reasons behind its increasing popularity is that compared to first-principle modelling methods, data-driven soft sensor models are relatively simple to build without the need to know complete knowledge or information on physical models.

With the rapid advance in computing power and multisensory technology, massive data accumulated can be processed by data-driven techniques to learn underlying driving forces and/or hidden patterns for better monitoring and control of industrial processes. Usually, in order to train a data-driven model with satisfactory performance, sizable amounts of labelled ground-truth data are required. However, it is

often expensive, time-consuming and labor-intensive to gather well-labelled data although massive data is being collected. In particular, it becomes more demanding to collect such labelled data for an industrial plant at its early stage. This is referred to as the *cold-start* problem in recommender systems, where recommendations are required for items when almost nothing is known about customer preferences [180]. In the context of the process industry, how to transfer the knowledge learned from source processes with well-labelled data into a related target system that has limited historical labels to establish a satisfactory soft sensor model motives this chapter.

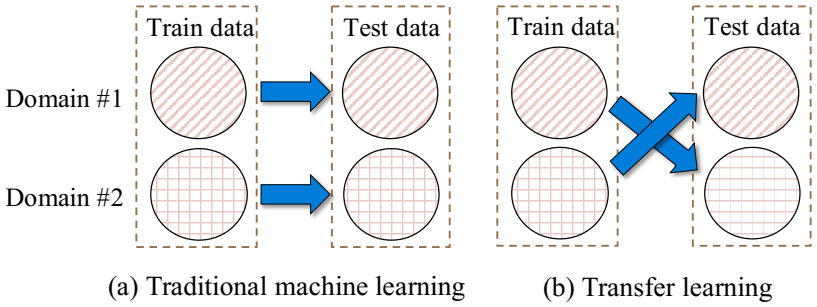


Figure 5.1: Comparison of traditional machine learning and transfer learning

Typically, most multivariate statistical approaches assume that the normal process data come from a single operating region so follow a unimodal distribution, which is not the case when it comes to multi-mode operation processes [181]. Furthermore, traditional machine learning (ML) methods often assume that the training and testing data follow the same distribution and in the same feature space, and hence are suitable for single domain learning. In other words, models learned by traditional ML methods in one domain (e.g. domain # 1) may not be that useful in another domain (e.g. domain # 2) as illustrated in Fig. 5.1. This is the main challenge when applying traditional ML methods to address multimodality of data distribution, cold-start problem, and cross-domain learning tasks. On the other hand, transfer learning (TL) has provided a suitable way to handle cross-domain learning by sharing the knowledge learned from source domains to target domains, and its usefulness and applicability have been demonstrated in various areas, including Web document classification, indoor WIFI localization and machine fault diagnosis [182, 183, 184, 185, 186, 187]. Recently, an instance-based transfer learning method named domain adaptation extreme learning machine was adopted to construct a soft sensor model for multi-grade

chemical processes [188]. A fault description based attribute transfer learning approach was proposed for industrial fault diagnosis with zero-sample of the target faults [189]. Overall, inadequate efforts have been made toward applying transfer learning techniques to soft sensor modelling.

Feature extraction plays a crucial role in soft sensor modelling. To explore the driving forces behind industrial processes from massive historical data, latent variable models (representation learning [190]) have often been deployed to extract representative factors and common causes described by latent variables, such as principal component analysis [191] and partial least squares [192]. The extracted latent variables serve as a bridge connecting inputs and outputs whilst reducing the input dimensionality and redundancy thus facilitating the prediction in soft sensors. However, most data-driven methods in this category make the steady-states assumption by ignoring the underlying process dynamics, which may suppress their potential in dynamic feature extraction greatly. To capture the intrinsic dynamics hidden in the monitoring data, subspace identification methods [193] are often exploited to estimate model structures and parameters from known inputs and outputs and adopt the states estimated through system identification. Moreover, slow feature analysis (SFA), as a novel technique that considers process temporal dynamics, was proposed by finding features with slow varying velocities [194]. Unlike independent component analysis (ICA), SFA aims to learn independent and slow varying features simultaneously. Later on, the optimization-based SFA framework was interpreted to a fully probabilistic version called probabilistic slow feature analysis (PSFA) [195] which was further extended to soft sensor modelling using the expectation maximization (EM) algorithm with respect to fast-rate process data [196] and in a semi-supervised manner [197]. Instead of using EM for point estimation, an interesting work about dynamical slow feature extraction based on variational Bayesian inference (VBI) was accomplished for soft sensing applications by fully considering probability distributions of parameters to account for corresponding uncertainties [198]. However, most of the existing contributions on dynamic feature extraction are about single domain (task, or process) learning so tend to be limited when it comes to varying operating conditions (multimodality of data distributions), model adaption to new processes, and cross-domain learning tasks.

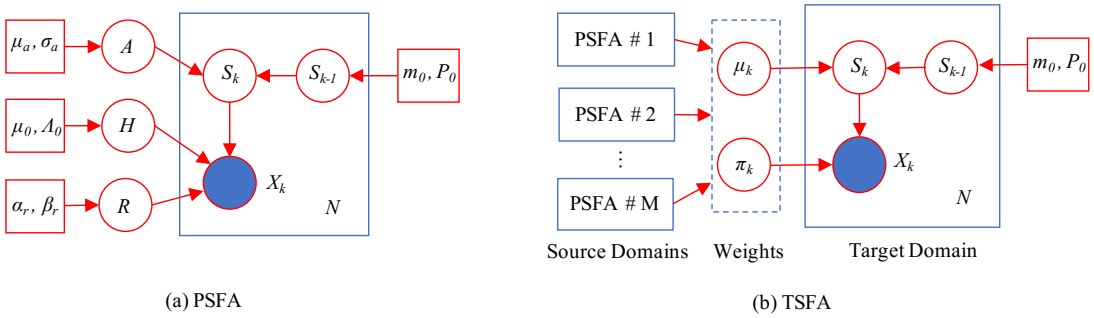
On the other hand, it is of significance to implement designed soft sensors in an online fashion considering practical applications. To account for online realization, some online transfer learning techniques have been proposed to transfer knowledge from multiple homogeneous or heterogeneous source domains [199, 200]. For better probabilistic interpretations, hidden Markov models are utilized to represent multiple source domains and transferred to infer sequential data in target domains using online Bayesian moment matching [201]. In addition, a joint prior probability density function of the source and target model parameters was introduced in the homogeneous transfer learning frame to measure the transferability and then realize a fast optimal Bayesian transfer learning classifier [202].

Motivated by the above considerations, a dynamic homogeneous transfer learning technique named transfer slow feature analysis (TSFA) is proposed in this chapter using variational Bayesian inference (VBI) to expand the applicability of slow feature analysis for cross-domain learning and explore the common driving forces behind similar processes. Unlike instance-based TL, the proposed method possesses model parameter-based TL to transfer feature slowness and it behaves like a hybrid combination of feature-based and parameter-based TL since the feature slowness is characterized by the model parameters (namely transition matrices). The main contributions of this chapter are the following: 1). To learn slow features by VBI, the truncated Gaussian distribution is introduced as a conjugate prior to account for the constrained transition matrix; 2). To quantify model transferability from multiple source domains to the target one, two weighting functions associated with transition and emission equations are introduced and learned dynamically at each time instant.

The remainder of this chapter is organized as follows. Section 5.2 presents the formulation of PSFA and the corresponding parameter estimation are illustrated therein. In Section 5.3, the transfer slow feature analysis technique is developed and parameter learning via VBI is proposed correspondingly. To demonstrate the feasibility and effectiveness of the proposed method, a simulation example, a public dataset and an industrial case study are provided in Section 5.4. Finally, concluding remarks are given in Section 5.5.

5.2 Dynamic model learning

In this section, the probabilistic slow feature analysis is utilized for dynamic model learning with good probability interpretation. As for model parameter learning, conjugate priors are utilized in the variational Bayesian inference scheme for efficient estimation including the truncated Gaussian distribution considered for the transition matrix to account for the constrained supporting range. As illustrated in Fig. 5.2(a), PSFA is presented in this section for dynamic model learning and serves as a preparatory step for the use of TSFA.



(a) PSFA (b) TSFA

Figure 5.2: Probability graphical models of the probabilistic slow feature analysis (PSFA) and the transfer slow feature analysis (TSFA)

5.2.1 Notation

For brevity, a list of notations and the corresponding definitions are summarized in Table 5.1. Let $p_1(V)$ and $p_2(V)$ be two probability density functions. $D_{KL}[p_1(V)||p_2(V)] = \int p_1(V) \ln[p_1(V)/p_2(V)]dV$ and $\mathcal{H}[p_1(V)] = -\int p_1(V) \ln[p_1(V)]dV$ denote the Kullback-Leibler (K-L) divergence from $p_2(V)$ to $p_1(V)$, and the entropy of $p_1(V)$, respectively. Throughout this chapter, if no special note is given, the matrices and vectors are notated with letters in bold fonts, and scalars are represented by regular characters. In particular, $\boldsymbol{I_d}$ is used to denote an identity matrix of dimension d.

5.2.2 Slow feature analysis revisited

As an unsupervised latent variable model, slow feature analysis (SFA) aims at learning a slowness representation scheme as a feature space [194], which is different from other feature learning techniques (such as PCA for maximum variance). Mathematically,

Table 5.1: List of notations

Symbol	Definition
$\boldsymbol{S_k}$	Slow features at time instant k
$\boldsymbol{V}$	Latent variables
S_k^i	i-th slow feature of $\boldsymbol{S_k}$
κ	Element of $\boldsymbol{V}$
$\boldsymbol{S_k^t}$	Slow features at time instant k in the target domain
$\boldsymbol{\Phi}$	Hyper-parameters and slow features of PSFA
$\boldsymbol{X_k}$	Observations or measurements at time instant k
$p(\boldsymbol{X}\|\boldsymbol{\Phi})$	The marginal likelihood of $\boldsymbol{X}$ given $\boldsymbol{\Phi}$
X_k^i	i-th row vector of $\boldsymbol{X_k}$
$q(\boldsymbol{V})$	The proposal distribution of $\boldsymbol{V}$
$\boldsymbol{X_k^t}$	Observations at time instant k in the target domain
$p(\boldsymbol{V}\|\boldsymbol{X},\boldsymbol{\Phi})$	The posterior distribution of $\boldsymbol{V}$ given $\boldsymbol{X}$ and $\boldsymbol{\Phi}$
$\boldsymbol{A}$	The transition matrix of slow features $\boldsymbol{S_k}$
μ_a, σ_a	Hyper-parameters of prior distribution $p(a_i)$
$\boldsymbol{A^{cr}}$	The transition matrix of slow features $\boldsymbol{S_k^t}$
$\boldsymbol{m_0}, \boldsymbol{P_0}$	Mean and variance of prior distribution $p(\boldsymbol{S_1})$
a_i	Diagonal element of $\boldsymbol{A}$
α_r, β_r	Shape and rate of prior (Gamma) distribution $p(r)$
$\boldsymbol{H}$	The emission matrix mapping from $\boldsymbol{S_k}$ to $\boldsymbol{X_k}$
$\boldsymbol{\mu_0}, \boldsymbol{\Lambda_0}$	Mean and precision of prior distribution $p(\boldsymbol{h_i})$
$\boldsymbol{H^{cr}}$	The emission matrix mapping from $\boldsymbol{S_k^t}$ to $\boldsymbol{X_k^t}$
$\hat{\mu}_{a_i}, \hat{\sigma}_{a_i}$	Hyper-parameters of proposal distribution $q^*(a_i)$
$\boldsymbol{h_i}$	i-th row vector of $\boldsymbol{H}$
$\hat{\boldsymbol{m}}_{\boldsymbol{i}}, \hat{\boldsymbol{\Lambda}}_{\boldsymbol{i}}$	Mean and variance of proposal distribution $q^*(\boldsymbol{h_i})$
$\boldsymbol{w_k}$	Gaussian noise involved in slow features $\boldsymbol{S_k}$
$\hat{\alpha}_r, \hat{\beta}_r$	Shape and rate of proposal (Gamma) distribution $q^*(r)$
w_k^i	Gaussian noise involved in slow feature S_k^i
$\boldsymbol{A^i}, \boldsymbol{H^i}, \boldsymbol{B^i}$	Transition and emission matrices learned in i-th source domain
$\boldsymbol{w_k^t}$	Gaussian noise involved in slow features $\boldsymbol{S_k^t}$
$\boldsymbol{Q^i}, \boldsymbol{R^i}$	Variances of noises learned in i-th source domain
$\boldsymbol{Q}$	Variance of Gaussian noise $\boldsymbol{w_k}$
$\boldsymbol{z_k}$, $\boldsymbol{y_k}$	Identity indicators
$\boldsymbol{Q^{cr}}$	Variance of Gaussian noise $\boldsymbol{w_k^t}$
z_{kj}, y_{kj}	j-th elements in $\boldsymbol{z_k}$ and $\boldsymbol{y_k}$
q_i	Diagonal element of $\boldsymbol{Q}$
$\boldsymbol{\pi_k}$, $\boldsymbol{\nu_k}$	Probabilities of having identities $\boldsymbol{z_k}$ and $\boldsymbol{y_k}$
$\boldsymbol{\nu_k}$	Gaussian noise involved in observations $\boldsymbol{X_k}$
π_{kj}, ν_{kj}	Probabilities of having identities z_{kj} and y_{kj}
$\boldsymbol{\nu_k^t}$	Gaussian noise involved in observations $\boldsymbol{X_k^t}$
$\pi_{kj}^{(f)}, \nu_{kj}^{(f)}$	j-th elements of filtered $\boldsymbol{\pi_k}$ and $\boldsymbol{\nu_k}$
$\boldsymbol{R}$	Variance of Gaussian noise $\boldsymbol{\nu_k}$

Table 5.1: List of notations (continued)

Symbol	Definition
$\pi_{kj}^{(p)}, \nu_{kj}^{(p)}$	j-th elements of predicted $\boldsymbol{\pi_k}$ and $\boldsymbol{\nu_k}$
$\boldsymbol{R^{cr}}$	Variance of Gaussian noise $\boldsymbol{v_k^t}$
$\boldsymbol{\Theta}$	Hyper-parameters and slow features of TSFA
r	Parameter of $\boldsymbol{R}$, i.e., $\boldsymbol{R} = r^{-1} \cdot \boldsymbol{I_P}$
θ	Element of $\boldsymbol{\Theta}$
v	Average feature varying velocity
$\boldsymbol{\alpha_\pi}, \boldsymbol{\alpha_\nu}$	Hyper-parameters of prior distributions $p(\boldsymbol{\pi_k})$ and $p(\boldsymbol{\nu_k})$
$\boldsymbol{Y_k}$	The quality outputs at time instant k
$\alpha_\pi^j, \alpha_\nu^j$	j-th elements of $\boldsymbol{\alpha_\pi}$ and $\boldsymbol{\alpha_\nu}$
$\boldsymbol{B}$	The emission matrix mapping from $\boldsymbol{S_k}$ to $\boldsymbol{Y_k}$
$\hat{\boldsymbol{\alpha}}_{\boldsymbol{\pi}}, \hat{\boldsymbol{\alpha}}_{\boldsymbol{\nu}}$	Hyper-parameters of proposal distributions $q^*(\boldsymbol{\pi_k})$ and $q^*(\boldsymbol{\nu_k})$
$\boldsymbol{g_k}$	Gaussian noise involved in outputs $\boldsymbol{Y_k}$
$\hat{\alpha}_\pi^j, \hat{\alpha}_\nu^j$	j-th elements of $\hat{\boldsymbol{\alpha}}_{\boldsymbol{\pi}}$ and $\hat{\boldsymbol{\alpha}}_{\boldsymbol{\nu}}$
$\boldsymbol{T}$	Variance of Gaussian noise $\boldsymbol{g_k}$
$\boldsymbol{m_k^{(f)}}$	Filtered mean of slow feature in the target domain
$\boldsymbol{X_k^e}$	Combined observations of $\boldsymbol{X_k}$ and $\boldsymbol{Y_k}$
$\boldsymbol{P_k^{(f)}}$	Filtered variance of slow feature in the target domain
$\boldsymbol{H^e}$	The emission matrix mapping from $\boldsymbol{S_k}$ to $\boldsymbol{X_k^e}$
$\boldsymbol{m_k^{(p)}}$	Predicted mean of slow feature in the target domain
$\boldsymbol{v_k^e}$	Gaussian noise involved in observations $\boldsymbol{X_k^e}$
$\boldsymbol{P_k^{(p)}}$	Predicted variance of slow feature in the target domain
$\boldsymbol{R^e}$	Variance of Gaussian noise $\boldsymbol{v_k^e}$
$\hat{\boldsymbol{Y}}_{\boldsymbol{k}}$	Predicted quality output in the target domain
$\tilde{\boldsymbol{X}}_{\boldsymbol{k}}$	Augmented measurements with zero vectors
$\tilde{\boldsymbol{A}}, \tilde{\boldsymbol{H}}_{\boldsymbol{k}}$	Fluctuated transition and emission matrices
$\tilde{\boldsymbol{Q}}, \tilde{\boldsymbol{R}}$	Fluctuated variances of noises
$\boldsymbol{F_t}, \boldsymbol{F_e}$	Fluctuation terms in transition and emission equations

the averaged varying velocity for a finite sequence with N discrete samples $\{\boldsymbol{s_k}\}$ with $\boldsymbol{s_k} \in \mathbb{R}^{d\times 1}$ and $1 \leq k \leq N$ is represented as:

$$v(\boldsymbol{s}) = \mathbb{E}[(\boldsymbol{s_k} - \boldsymbol{s_{k-1}})^2] \approx \frac{1}{N-1}\sum_{k=2}^{N}(\boldsymbol{s_k} - \boldsymbol{s_{k-1}})^2 \tag{5.1}$$

To extract slow features in a linear case, an optimization problem is often formulated as follows:

$$\min_{\boldsymbol{F}} \quad v(\boldsymbol{s}) \tag{5.2a}$$

$$s.t. \quad \boldsymbol{s} = \boldsymbol{F}^T \cdot \boldsymbol{X} \tag{5.2b}$$

$$\langle \boldsymbol{s}^{(i)} \rangle_k = 0 \ \text{(zero mean)} \tag{5.2c}$$

$$\langle \boldsymbol{s}^{(i)2} \rangle_k m = 1 \ \text{(unit variance)} \tag{5.2d}$$

$$\langle \boldsymbol{s}^{(i)} \cdot \boldsymbol{s}^{(j)} \rangle_k = 0, \ \forall i \neq j \ \text{(independence)} \tag{5.2e}$$

where the matrix $\boldsymbol{F}^T$ is used to map the measurements $\boldsymbol{X}$ into latent features $\boldsymbol{s}$, and $\langle \cdot \rangle_k$ represents the average over data samples. Specifically, we denote that $\boldsymbol{F} \in \mathbb{R}^{P \times d}$, $\boldsymbol{X} \in \mathbb{R}^{P \times N}$, $\boldsymbol{s} \in \mathbb{R}^{d \times N}$ and $\boldsymbol{s}^{(i)} \in \mathbb{R}^{1 \times N}$. By imposing the constraints (5.2c)-(5.2e), one can attain independent slow features sorted in descending order in terms of slowness by applying the greedy strategy.

5.2.3 PSFA in state space

As a probability framework, probabilistic slow feature analysis (PSFA) can be formulated as a linear Gaussian state-space model by extending the original SFA frame. Specifically, the PSFA model takes the following form:

$$\boldsymbol{S_k} = \boldsymbol{A}\boldsymbol{S_{k-1}} + \boldsymbol{w_k}, \ \ \boldsymbol{w_k} \sim \mathcal{N}(0, \boldsymbol{Q}) \tag{5.3a}$$

$$\boldsymbol{X_k} = \boldsymbol{H}\boldsymbol{S_k} + \boldsymbol{v_k}, \ \ \boldsymbol{v_k} \sim \mathcal{N}(0, \boldsymbol{R}) \tag{5.3b}$$

where $\boldsymbol{S_k} \in \mathbb{R}^{d \times 1}$ and $\boldsymbol{X_k} \in \mathbb{R}^{P \times 1}$ represent the slow features and observations in each individual source domain. Moreover, $\boldsymbol{A} \in \mathbb{R}^{d \times d}$ denotes the transition matrix of slow features and $\boldsymbol{H} \in \mathbb{R}^{P \times d}$ represents the emission matrix mapping slow features into observations, respectively. Additionally, $\boldsymbol{w_k} \in \mathbb{R}^{d \times 1}$ and $\boldsymbol{v_k} \in \mathbb{R}^{P \times 1}$ stand for two Gaussian noises (with variances $\boldsymbol{Q} \in \mathbb{R}^{d \times d}$ and $\boldsymbol{R} \in \mathbb{R}^{P \times P}$, respectively, and zero means having same sizes as $\boldsymbol{w_k}$ and $\boldsymbol{v_k}$, which are notated with regular forms throughout this chapter for brevity.) involved in slow features and observations. Considering the equivalence relationship between SFA and PSFA, applying the unit variance constraint in SFA to model (5.3) leads to $\boldsymbol{Q} = \boldsymbol{I_d} - \boldsymbol{A}^T\boldsymbol{A}$, where $\boldsymbol{I_d}$ denotes the identity matrix of dimension d. By the independence constraint in SFA, the slow features are totally decoupled among each others as follows:

$$S_k^i = a_i S_{k-1}^i + w_k^i, \ \ w_k^i \sim \mathcal{N}(0, q_i) \tag{5.4a}$$

$$a_i^2 + q_i = 1 \tag{5.4b}$$

where $\boldsymbol{A} = \text{diag}(a_1, ..., a_d)$ and $\boldsymbol{Q} = \text{diag}(q_1, ..., q_d)$. S_k^i and w_k^i denote i-th slow feature and the corresponding noise at time instant k. Then, the associated average varying velocity in SFA for latent feature $\boldsymbol{S^i} \in \mathbb{R}^{1\times N}$ can be expressed by:

$$v(\boldsymbol{S^i}) = \mathbb{E}\left[(S_k^i - S_{k-1}^i)^2\right] = 2(1 - a_i) \tag{5.5}$$

Hence, the feature slowness is characterized by the magnitude of a_i, which implies that the feature slowness can be transferred through sharing the model transition matrix $\boldsymbol{A}$.

Due to the positive definiteness of variance q_i in (5.5), a_i is constrained in the range $(-1, 1)$. Considering that noise terms tend to vary faster than meaningful latent features, a_i can be further restricted in the range $(0, 1)$ [198].

For a semi-supervised PSFA regression learning task, we can introduce another emission equation for the quality outputs $\boldsymbol{Y_k} \in \mathbb{R}^{p\times 1}$ as follow:

$$\boldsymbol{Y_k} = \boldsymbol{B}\boldsymbol{S_k} + \boldsymbol{g_k}, \quad \boldsymbol{g_k} \sim \mathcal{N}(0, \boldsymbol{T}) \tag{5.6}$$

where $\boldsymbol{B} \in \mathbb{R}^{p\times d}$ represents the corresponding emission matrix, and $\boldsymbol{g_k} \in \mathbb{R}^{p\times 1}$ denotes the Gaussian noise with variance $\boldsymbol{T} \in \mathbb{R}^{p\times p}$. By combining the process measurements and quality variables together as $\boldsymbol{X_k^e} = [\boldsymbol{X_k}; \boldsymbol{Y_k}] \in \mathbb{R}^{(P+p)\times 1}$, an augmented emission equation accounting for inputs and outputs are obtained as [197]:

$$\boldsymbol{X_k^e} = \boldsymbol{H^e}\boldsymbol{S_k} + \boldsymbol{v_k^e}, \quad \boldsymbol{v_k^e} \sim \mathcal{N}(0, \boldsymbol{R^e}) \tag{5.7}$$

where $\boldsymbol{v_k^e} = [\boldsymbol{v_k}; \boldsymbol{g_k}] \in \mathbb{R}^{(P+p)\times 1}$, and the augmented emission matrix $\boldsymbol{H^e} = [\boldsymbol{H}; \boldsymbol{B}] \in \mathbb{R}^{(P+p)\times d}$ and measurement noise variance $\boldsymbol{R^e} = \text{bdiag}(\boldsymbol{R}, \boldsymbol{T}) \in \mathbb{R}^{(P+p)\times(P+p)}$ (Here bdiag denotes a block diagonal matrix). In the following sections, the model structure (5.3) is inherited by dropping the superscript 'e' in (5.7) for the ease of notation.

5.2.4 Parameter learning via VBI

In this section, the variational Bayesian inference (VBI) approach is utilized for parameter estimation. As shown in Fig. 5.2(a), we denote latent variables by $\boldsymbol{V}(= \boldsymbol{S_{1:N}}, \boldsymbol{A}, \boldsymbol{H}, \boldsymbol{R})$ and the corresponding hyper-parameters by $\boldsymbol{\Phi}(= \mu_a, \sigma_a, \boldsymbol{m_0}, \boldsymbol{P_0}, \alpha_r, \beta_r, \boldsymbol{\mu_0}, \boldsymbol{\Lambda_0})$. In this section, $\boldsymbol{\Phi}$ is fixed to represent a certain model structure so $p(\boldsymbol{X}|\boldsymbol{\Phi})$

is considered as a constant. By marginalization over $\boldsymbol{V}$, we can decompose $p(\boldsymbol{X}|\boldsymbol{\Phi})$ as follows:

$$\begin{aligned}\ln p(\boldsymbol{X}|\boldsymbol{\Phi}) &= \int q(\boldsymbol{V}) \ln \frac{p(\boldsymbol{X},\boldsymbol{V}|\boldsymbol{\Phi})}{q(\boldsymbol{V})} d\boldsymbol{V} + \int q(\boldsymbol{V}) \ln \frac{q(\boldsymbol{V})}{p(\boldsymbol{V}|\boldsymbol{X},\boldsymbol{\Phi})} d\boldsymbol{V} \\ &= \mathcal{L}[q(\boldsymbol{V})] + D_{KL}[q(\boldsymbol{V})||p(\boldsymbol{V}|\boldsymbol{X},\boldsymbol{\Phi})] \end{aligned} \tag{5.8}$$

where $D_{KL}[q(\boldsymbol{V})||p(\boldsymbol{V}|\boldsymbol{X},\boldsymbol{\Phi})]$ represents the Kullback-Leibler (K-L) divergence indicating the probability distance between the proposal distribution $q(\boldsymbol{V})$ and the posterior $p(\boldsymbol{V}|\boldsymbol{X},\boldsymbol{\Phi})$. Considering that K-L divergence is always non-negative, the minimization problem of $D_{KL}[q(\boldsymbol{V})||p(\boldsymbol{V}|\boldsymbol{X},\boldsymbol{\Phi})]$ can be equivalently converted to a maximization problem of the evidence lower bound (LB) $\mathcal{L}[q(\boldsymbol{V})]$. To proceed with the maximization process, the conditional distribution (accounting for measurements, slow features and model parameters) of given hyper-parameters is decomposed as follows:

$$\begin{aligned} p(\boldsymbol{X_{1:N}},\boldsymbol{V}|\boldsymbol{\Phi}) &= p(\boldsymbol{X_{1:N}},\boldsymbol{S_{1:N}}|\boldsymbol{A},\boldsymbol{H},\boldsymbol{R},\boldsymbol{\Phi}) p(\boldsymbol{A},\boldsymbol{H},\boldsymbol{R}|\boldsymbol{\Phi}) \\ &= \prod_{k=1}^{N} p(\boldsymbol{X_k}|\boldsymbol{S_k},\boldsymbol{H},\boldsymbol{R}) \prod_{k=2}^{N} p(\boldsymbol{S_k}|\boldsymbol{S_{k-1}},\boldsymbol{A}) \\ &\quad \times p(\boldsymbol{S_1}|\boldsymbol{\Phi}) p(\boldsymbol{A}|\boldsymbol{\Phi}) p(\boldsymbol{H}|\boldsymbol{\Phi}) p(\boldsymbol{R}|\boldsymbol{\Phi}) \end{aligned} \tag{5.9}$$

where $\boldsymbol{X_{1:N}} \in \mathbb{R}^{(P+p)\times N}$ denotes the input and output observations (i.e., $\boldsymbol{X_k^e}$ in (5.7)) in each source domain, $\boldsymbol{H}$ and $\boldsymbol{R}$ represent $\boldsymbol{H^e}$ and $\boldsymbol{R^e}$ in (5.7), and $S_{1:N} \in \mathbb{R}^{d\times N}$ denotes the time-series slow features. The last equality in (5.9) holds due to the state space model (5.3). To avoid an intractable estimation, a variational factorization is introduced to approximate the posterior distribution as:

$$q(\boldsymbol{V}) = q(\boldsymbol{S_{1:N}}) q(\boldsymbol{A}) q(\boldsymbol{H}) q(\boldsymbol{R}) \rightarrow p(\boldsymbol{V}|\boldsymbol{X_{1:N}},\boldsymbol{\Phi}) \tag{5.10}$$

Let κ be one of these latent variables $\boldsymbol{V}(= \boldsymbol{S_{1:N}},\boldsymbol{A},\boldsymbol{H},\boldsymbol{R})$, and its variational distribution can be updated as follow [203]:

$$\ln q^*(\kappa) \propto \mathbb{E}_{q[\boldsymbol{V}\setminus\kappa]}[\ln\ p(\boldsymbol{X_{1:N}},\boldsymbol{V}|\boldsymbol{\Phi})] \tag{5.11}$$

To facilitate the learning process, Gaussian distribution and inverse Gamma distribution are chosen as conjugate prior distributions of parameters $\boldsymbol{H}$ and $\boldsymbol{R}$ as follows:

$$\boldsymbol{H} = [\boldsymbol{h_1},...,\boldsymbol{h_{P+p}}]^T,\quad p(\boldsymbol{h_i}|\boldsymbol{\mu_0},\boldsymbol{\Lambda_0}) = \mathcal{N}(\boldsymbol{\mu_0},\boldsymbol{\Lambda_0^{-1}}) \tag{5.12a}$$

$$\boldsymbol{R} = r^{-1} \cdot \boldsymbol{I_{P+p}}, \quad p(r|\alpha_r, \beta_r) = \text{Gamma}(\alpha_r, \beta_r) \tag{5.12b}$$

$$p(\boldsymbol{S_1}|\boldsymbol{m_0}, \boldsymbol{P_0}) = \mathcal{N}(\boldsymbol{m_0}, \boldsymbol{P_0}) \tag{5.12c}$$

$$p(\boldsymbol{S_k}|\boldsymbol{S_{k-1}}, \boldsymbol{A}) = \mathcal{N}(\boldsymbol{A S_{k-1}}, \boldsymbol{I_d} - \boldsymbol{A}^T \boldsymbol{A}) \tag{5.12d}$$

$$p(\boldsymbol{X_k}|\boldsymbol{S_k}, \boldsymbol{H}, \boldsymbol{R}) = \mathcal{N}(\boldsymbol{H S_k}, \boldsymbol{R}) \tag{5.12e}$$

where $\boldsymbol{h_i} \in \mathbb{R}^{d\times 1}$, $\boldsymbol{\mu_0} \in \mathbb{R}^{d\times 1}$ and $\boldsymbol{\Lambda_0} \in \mathbb{R}^{d\times d}$ denote i-th row vector of $\boldsymbol{H}$ and its corresponding mean and precision, respectively. $\boldsymbol{S_1} \in \mathbb{R}^{d\times 1}$, $\boldsymbol{m_0} \in \mathbb{R}^{d\times 1}$ and $\boldsymbol{P_0} \in \mathbb{R}^{d\times d}$ denote the initial condition of slow feature and its corresponding mean and variance, respectively. The updating equations for $\boldsymbol{H}$ and $\boldsymbol{R}$ are derived as [198]

$$q^*(\boldsymbol{h_i}) = \mathcal{N}(\hat{\boldsymbol{m}}_i, \hat{\boldsymbol{\Lambda}}_i^{-1}), \; q^*(\boldsymbol{H}) = \prod_{i=1}^{P+p} q^*(\boldsymbol{h_i}) \tag{5.13a}$$

$$q^*(r) = \text{Gamma}(\hat{\alpha}_r, \hat{\beta}_r) \tag{5.13b}$$

where

$$\hat{\boldsymbol{\Lambda}}_i = \boldsymbol{\Lambda_0} + \alpha_r \beta_r \sum_{k=1}^{N} \left\langle \boldsymbol{S_k} \boldsymbol{S_k}^T \right\rangle$$

$$\hat{\boldsymbol{m}}_i = \hat{\boldsymbol{\Lambda}}_i^{-1} \left(\alpha_r \beta_r \sum_{k=1}^{N} \langle \boldsymbol{S_k} \rangle \, X_k^i + \boldsymbol{\Lambda_0 \mu_0} \right)$$

$$\hat{\alpha}_r = \alpha_r + \frac{N(P+p)}{2}$$

$$\hat{\beta}_r = \beta_r + \frac{1}{2} \sum_{i=1}^{P+p} \left[\sum_{k=1}^{N} \left(X_k^i X_k^i - 2 X_k^i \langle \boldsymbol{S_k} \rangle^T \hat{\boldsymbol{m}}_i \right) + tr\left(\sum_{k=1}^{N} \langle \boldsymbol{S_k S_k} \rangle^T \left(\hat{\boldsymbol{m}}_i \hat{\boldsymbol{m}}_i^T + \hat{\boldsymbol{\Lambda}}_i^{-1} \right) \right) \right]$$

where $\langle \cdot \rangle$ denotes the statistical expectation of a random variable under its own distribution, and X_k^i corresponds to i-th row of $\boldsymbol{X_k}$. Moreover, $\hat{\boldsymbol{\Lambda}}_i$ and $\hat{\boldsymbol{m}}_i$ have the same dimensions as $\boldsymbol{\Lambda_0}$ and $\boldsymbol{\mu_0}$, respectively.

To learn the hyper-parameter for updating $q^*(\boldsymbol{A})$, the challenge arises in the constraint issue of $a_i \in (0, 1)$. Compared to sampling methods proposed in [198], a constrained conjugate prior is presented in this chapter considering computational efficiency in practical implementations. Motivated by [204], we consider a truncated Gaussian distribution as a conjugate prior distribution of $p(a_i)$ as follow:

$$p\left(a_i|\mu_a, \sigma_a^2, 0, 1\right) = \mathcal{N}_T\left(\mu_a, \sigma_a^2, 0, 1\right) \tag{5.14}$$

where $\boldsymbol{A} = \text{diag}(a_1, ..., a_d)$ and

$$\mathcal{N}_T\left(\mu_a, \sigma_a^2, 0, 1\right) = \frac{f\left(\frac{a_i - \mu_a}{\sigma_a}\right)}{\sigma_a\left[F\left(\frac{1-\mu_a}{\sigma_a}\right) - F\left(\frac{0-\mu_a}{\sigma_a}\right)\right]}$$

where $f(z) \triangleq \frac{1}{\sqrt{2\pi}} \exp(-\frac{z^2}{2})$ is the probability density function (PDF) of the standard normal distribution, and its cumulative density function (CDF) is given by $F(z) \triangleq \int_{-\infty}^{z} f(t)dt$. The truncated Gaussian distribution can be interpreted as the Gaussian distribution $\mathcal{N}\left(\mu_a, \sigma_a^2\right)$ supported on a bounded interval $(0, 1)$. By recalling (5.11), the updating proposal distribution of a_i is approximated with another truncated Gaussian distribution as:

$$q^*(a_i) = \mathcal{N}_T(\hat{a}_i|\hat{\mu}_{a_i}, \hat{\sigma}_{a_i}^2, 0, 1),\ q^*(\boldsymbol{A}) = \prod_{i=1}^{d} q^*(a_i) \tag{5.15}$$

where

$$\hat{\sigma}_{a_i}^2 = \left(\langle q_i \rangle \sum_{k=2}^{N} \left\langle S_{k-1}^i S_{k-1}^i \right\rangle + \frac{1}{\sigma_a^2}\right)^{-1}$$

$$\hat{\mu}_{a_i} = \hat{\sigma}_{a_i}^2 \left(\langle q_i \rangle \sum_{k=2}^{N} \left\langle S_{k-1}^i S_k^i \right\rangle + \frac{\mu_a}{\sigma_a^2}\right)$$

where $\langle q_i \rangle = \left\langle 1 - a_i^2 \right\rangle$ using (5.14). With the truncated Gaussian distribution being applied, the constraint issue of a_i is resolved in the iterative learning process. Based on the learned proposal distributions of parameters $(\boldsymbol{A}, \boldsymbol{H}, \boldsymbol{R})$, we come to learning the consequent proposal distribution of slow features.

However, it is not straightforward to estimate the slow features due to the uncertainty of the model parameters, which indicates that one cannot simply estimate slow features using expectations of model parameters as:

$$\begin{aligned} &\mathbb{E}_{q(\boldsymbol{A},\boldsymbol{H},\boldsymbol{R})}[\ln p(\boldsymbol{S_{1:N}}|\boldsymbol{X_{1:N}}, \boldsymbol{A}, \boldsymbol{H}, \boldsymbol{R}, \boldsymbol{\Phi})] \\ &\qquad \neq \ln p(\boldsymbol{S_{1:N}}|\boldsymbol{X_{1:N}}, \langle \boldsymbol{A} \rangle, \langle \boldsymbol{H} \rangle, \langle \boldsymbol{R} \rangle) \end{aligned} \tag{5.16}$$

To address this issue, a unified inference frame [205] with the aid of Kalman filter and Kalman-Rauch-Tung-Striebel smoother is formulated. Based on that, we can learn the proposal distributions of slow features $q(\boldsymbol{S_{1:N}})$ through:

$$\ln q^*(\boldsymbol{S_{1:N}}) \propto \mathbb{E}_{q(\boldsymbol{A},\boldsymbol{H},\boldsymbol{R})}[\ln p(\boldsymbol{X_{1:N}}, \boldsymbol{S_{1:N}}, \boldsymbol{A}, \boldsymbol{H}, \boldsymbol{R}|\boldsymbol{\Phi})]$$

$$\propto \ln p(\boldsymbol{S}_{1:N}|\tilde{\boldsymbol{X}}_{1:N}, \tilde{\boldsymbol{A}}, \tilde{\boldsymbol{Q}}, \tilde{\boldsymbol{H}}_{1:N}, \tilde{\boldsymbol{R}}) \tag{5.17}$$

where $\boldsymbol{S}_k \sim \mathcal{N}(\tilde{\boldsymbol{A}}\boldsymbol{S}_{k-1}, \tilde{\boldsymbol{Q}})$ and $\tilde{\boldsymbol{X}}_k \sim \mathcal{N}(\tilde{\boldsymbol{H}}_k\boldsymbol{S}_k, \tilde{\boldsymbol{R}})$. In addition, $\tilde{\boldsymbol{A}} = \langle \boldsymbol{A} \rangle$, $\tilde{\boldsymbol{Q}} = \left\langle \boldsymbol{I_d} - \boldsymbol{A}^T\boldsymbol{A} \right\rangle$, $\tilde{\boldsymbol{X}}_k = [\boldsymbol{X}_k^T, \boldsymbol{0}_{1\times d}, \boldsymbol{0}_{1\times d}]^T$, with $k = 1, ..., N$, $\tilde{\boldsymbol{R}} = \text{bdiag}[\langle \boldsymbol{R} \rangle, \boldsymbol{I_d}, \boldsymbol{I_d}]$, and

$$\tilde{\boldsymbol{H}}_k = \begin{cases} \left[\langle \boldsymbol{H} \rangle^T, \boldsymbol{U_A}, \boldsymbol{U_B}\right]^T, & \text{if } k = 1, ..., N-1 \\ \left[\langle \boldsymbol{H} \rangle^T, \boldsymbol{0}_{d\times d}, \boldsymbol{U_B}\right]^T, & \text{if } k = N \end{cases}$$

$$\boldsymbol{U_A}\boldsymbol{U_A}^T = \boldsymbol{F_t} = \left\langle \boldsymbol{A}^T(\boldsymbol{I_d} - \boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A} \right\rangle - \left\langle \boldsymbol{A}^T \right\rangle \left\langle (\boldsymbol{I_d} - \boldsymbol{A}^T\boldsymbol{A})^{-1} \right\rangle \left\langle \boldsymbol{A} \right\rangle$$

$$\boldsymbol{U_B}\boldsymbol{U_B}^T = \boldsymbol{F_e} = \left\langle \boldsymbol{H}^T\boldsymbol{R}^{-1}\boldsymbol{H} \right\rangle - \left\langle \boldsymbol{H}^T \right\rangle \left\langle \boldsymbol{R}^{-1} \right\rangle \left\langle \boldsymbol{H} \right\rangle$$

Here, $\boldsymbol{F_t}$ and $\boldsymbol{F_e}$ represent two fluctuation terms from the transition and emission equations (5.3). In addition, we note that $\boldsymbol{U_A}, \boldsymbol{U_B}, \boldsymbol{F_t}, \boldsymbol{F_e} \in \mathbb{R}^{d\times d}$, $\tilde{\boldsymbol{H}}_k \in \mathbb{R}^{(P+p+2d)\times d}$, $\tilde{\boldsymbol{R}} \in \mathbb{R}^{(P+p+2d)\times(P+p+2d)}$, and $\tilde{\boldsymbol{A}}$ and $\tilde{\boldsymbol{Q}}$ are of the same dimensions as $\boldsymbol{A}$ and $\boldsymbol{Q}$, respectively.

5.2.5 Lower bound

The variational Bayesian inference not only provides a good probability interpretation but also an effective way for dimensionality determination of latent variables. The original maximization problem of posterior distributions is converted to maximizing a evidence lower bound (LB) as follows [206]:

$$\begin{aligned} \mathcal{L}[q(\boldsymbol{V})] &= \int q(\boldsymbol{V}) \cdot \ln \frac{p(\boldsymbol{X}, \boldsymbol{V}|\boldsymbol{\Phi})}{q(\boldsymbol{V})} d\boldsymbol{V} \\ &= \int q(\boldsymbol{V}) \ln p(\boldsymbol{X}, \boldsymbol{V}|\boldsymbol{\Phi}) d\boldsymbol{V} + \mathcal{H}[q(\boldsymbol{V})] \end{aligned} \tag{5.18}$$

where $\mathcal{H}$ denotes the entropy. For different dimensions of latent variables, selecting the maximum value of LB can lead to an optimal dimension selection of the feature space.

Thus, an efficient probabilistic slow feature analysis method is proposed using VBI for dynamic model learning. A brief summary of the proposed PSFA is given by Algorithm 1.

Algorithm 1. The proposed PSFA

Input: Source domain data $(\boldsymbol{X_{1:N}}, \boldsymbol{Y_{1:N}})$

Output: Source domain model $(\boldsymbol{A}, \boldsymbol{H}, \boldsymbol{R})$

(1). Initialize the hyper-parameters $\boldsymbol{\Phi}$;

(2). Update $q(\boldsymbol{H})$, $q(\boldsymbol{R})$ and $q(\boldsymbol{A})$ using (5.13) and (5.15);

(3). Learn $q(\boldsymbol{S_{1:N}})$ with (5.17);

(4). Iteratively perform steps (2) and (3) until the lower bound (LB) converges.

5.3 Model transfer learning

Considering that there are often insufficient measurements (especially outputs or labels) in the target domain to learn a predictive model, this section proposes an online transfer learning technique to transfer the source domain models $(\boldsymbol{A^j}, \boldsymbol{H^j}, \boldsymbol{R^j})$ ($\boldsymbol{Q^j}$ is omitted since it can be uniquely determined by $\boldsymbol{A^j}$, with $j = 1, ..., M$) to the target domain as depicted in Fig. 5.2(b). As described in (5.5), the feature slowness is characterized by the transition matrix $\boldsymbol{A}$ so it can be naturally shared from source domains to the target one by following this transfer mechanism. To measure the transferability of transition and emission equations from source domains to the target one, two weight functions are introduced and updated dynamically.

5.3.1 Transfer slow feature analysis in state space

To achieve dynamic transfer learning, the transfer slow feature analysis model is constructed in a state space as:

$$\boldsymbol{S_k^t} = \boldsymbol{A^{cr}}\boldsymbol{S_{k-1}^t} + \boldsymbol{w_k^t}, \quad \boldsymbol{w_k^t} \sim \mathcal{N}(0, \boldsymbol{Q^{cr}}) \tag{5.19a}$$

$$\boldsymbol{X_k^t} = \boldsymbol{H^{cr}}\boldsymbol{S_k^t} + \boldsymbol{v_k^t}, \quad \boldsymbol{v_k^t} \sim \mathcal{N}(0, \boldsymbol{R^{cr}}) \tag{5.19b}$$

where $\boldsymbol{S_k^t}$ and $\boldsymbol{X_k^t}$ represent slow features and observations in the target domain with the superscript 't' expressing target quantities in short. Similarly, $\boldsymbol{w_k^t}$ and $\boldsymbol{v_k^t}$ denote the Gaussian noise terms of slow features and observations herein. In addition, we use the superscript 'cr' to represent the cross-domain knowledge learned from different source domains to the target one. More specifically, we consider the same dimensions of quantities used in (5.19) as those in (5.3) or (5.7) depending on whether we use purely input observations or input and output measurements together. In addition,

we note that there are N_t samples (input measurements) in the target domain. Indeed, we propose a three-step slow feature learning mechanism, including a filtering step using only input observations (i.e., $\boldsymbol{X_k^t}$), a filtering step using both input and output observations (i.e., $[\boldsymbol{X_k^t}; \boldsymbol{Y_k^t}]$) as shown in (5.7), and a prediction step using the filtered slow features obtained in the second step. Based on the filtered slow features from the first step, the objective is to predict output ($\hat{\boldsymbol{Y}}_{\boldsymbol{k}}$) in the target domain by merging M learned source domain models. For ease of notation, we drop the superscript 't' when no confusion arises.

To distinguish the impacts between transition and emission equations of the source domain models on the target one, two weight functions are introduced as follows:

$$p(\boldsymbol{z_k}|\boldsymbol{\pi_k}) = \prod_{j=1}^{M} \pi_{kj}^{z_{kj}},\ p(\boldsymbol{y_k}|\boldsymbol{\nu_k}) = \prod_{j=1}^{M} \nu_{kj}^{y_{kj}} \tag{5.20}$$

where z_{kj} and y_{kj} are two binary numbers. Specifically, $\boldsymbol{z_k} \in \mathbb{R}^{M\times 1}$ denotes a M-dimensional binary random variable taking 1-of-M representation. More specifically, $\boldsymbol{z_k} = [z_{k1}, z_{k2}, ..., z_{kM}]^T$ and every element z_{kj} satisfies that $z_{kj} \in \{0, 1\}$ and $\sum_{j=1}^{M} z_{kj} = 1,\ j = 1, ..., M$. Similar notations hold for $\boldsymbol{y_k} \in \mathbb{R}^{M\times 1}$ and y_{kj}. $\boldsymbol{\pi_k} \in \mathbb{R}^{M\times 1}$ and $\boldsymbol{\nu_k} \in \mathbb{R}^{M\times 1}$ denote the corresponding probabilities of having $\boldsymbol{z_k}$ and $\boldsymbol{y_k}$, respectively. Additionally, $z_{kj} = 1$ corresponds to the transition equation of the j-th source model with a probability π_{kj} at time instant k, and $y_{kj} = 1$ indicates that the emission equation of the j-th source model holds with a probability ν_{kj} at time instant k. Indeed, quantities π_{kj} and ν_{kj} are regarded as responsibilities that the transition equation in the source domain z_{kj} and the emission equation in the source domain y_{kj} explain the target observations $\boldsymbol{X_k}$ best. In the proposed transfer learning frame, π_{kj} and ν_{kj} are employed as transferability measures accordingly. Along this line, one can obtain the probabilistic expressions for the transition and emission functions as:

$$p(\boldsymbol{S_k}|\boldsymbol{S_{k-1}}, \boldsymbol{z_k}) = \prod_{j=1}^{M} \left[\mathcal{N}(\boldsymbol{S_k}|\boldsymbol{A^j S_{k-1}}, \boldsymbol{Q^j})\right]^{z_{kj}} \tag{5.21a}$$

$$p(\boldsymbol{X_k}|\boldsymbol{S_k}, \boldsymbol{y_k}) = \prod_{j=1}^{M} \left[\mathcal{N}(\boldsymbol{X_k}|\boldsymbol{H^j S_k}, \boldsymbol{R^j})\right]^{y_{kj}} \tag{5.21b}$$

where $\boldsymbol{Q^j} = \boldsymbol{I_d} - \boldsymbol{A^j}^T \boldsymbol{A^j}$ with $j = 1, 2, ..., M$. It is apparent that once the model identity

indices $(\boldsymbol{z_k}, \boldsymbol{y_k})$ with their corresponding probabilities $(\boldsymbol{\pi_k}, \boldsymbol{\nu_k})$ are determined, the remaining part is a pure state estimation problem $(\boldsymbol{S_k})$ and Kalman filter can provide an optimal estimation [80]. However, the key challenge is how to merge these M source models (with transition and emission equations individually considered) to achieve a good transfer learning result. To tackle this issue, a conditional distribution is adopted from [207]:

$$\ln p(\boldsymbol{X_k}|\boldsymbol{X_{1:k-1}}) = \mathcal{D}_{KL}[q(\boldsymbol{\Theta})||p(\boldsymbol{\Theta}|\boldsymbol{X_{1:k}})] + \mathcal{L}_{q(\boldsymbol{\Theta})}[\boldsymbol{X_k}|\boldsymbol{X_{1:k-1}}] \tag{5.22}$$

where $\boldsymbol{\Theta} = (\boldsymbol{S_k}, \boldsymbol{z_k}, \boldsymbol{\pi_k}, \boldsymbol{y_k}, \boldsymbol{\nu_k})$. Using Bayes' rule, the conditional distribution can be decomposed into two components. The first component represents the K-L divergence that is used to assess the quality of using the proposal distribution $q(\boldsymbol{\Theta})$ to approximate the real posterior distribution $p(\boldsymbol{\Theta}|\boldsymbol{X_{1:k}})$ as follows:

$$\mathcal{D}_{KL}[q(\boldsymbol{\Theta})||p(\boldsymbol{\Theta}|\boldsymbol{X_{1:k}})] = \int q(\boldsymbol{\Theta}) \ln \frac{q(\boldsymbol{\Theta})}{p(\boldsymbol{\Theta}|\boldsymbol{X_{1:k}})} d\boldsymbol{\Theta} \tag{5.23}$$

The second component in (5.22) corresponds to the evidence lower bound (LB):

$$\mathcal{L}_{q(\Theta)}[\boldsymbol{X_k}|\boldsymbol{X_{1:k-1}}] = \int q(\boldsymbol{\Theta}) \ln \frac{p(\boldsymbol{\Theta}, \boldsymbol{X_k}|\boldsymbol{X_{1:k-1}})}{q(\boldsymbol{\Theta})} d\boldsymbol{\Theta} \tag{5.24}$$

The recurrent likelihood $p(\boldsymbol{X_k}|\boldsymbol{X_{1:k-1}})$ is considered as a constant since it is independent of the slow features from the system transfer function perspective. Maximization of LB is an alternative to enhance the quality of approximation (5.25) since K-L divergence is alway non-negative and $p(\boldsymbol{X_k}|\boldsymbol{X_{1:k-1}})$ is a constant. With mean field assumption [206], we consider a factorized variational approximation to the posterior distribution as follow:

$$q(\boldsymbol{\Theta}) = q(\boldsymbol{S_k})q(\boldsymbol{z_k})q(\boldsymbol{\pi_k})q(\boldsymbol{y_k})q(\boldsymbol{\nu_k}) \rightarrow p(\boldsymbol{\Theta}|\boldsymbol{X_{1:k}}) \tag{5.25}$$

where each individual proposal distribution can be updated iteratively by using the coordinate ascent algorithm. Let θ be one of the variables $\boldsymbol{\Theta} = (\boldsymbol{S_k}, \boldsymbol{z_k}, \boldsymbol{\nu_k}, \boldsymbol{y_k}, \boldsymbol{\pi_k})$. By applying VBI, we can update parameters iteratively through:

$$\ln q^*(\theta) \propto \mathbb{E}_{q(\boldsymbol{\Theta}\setminus\theta)}[\ln\ p(\boldsymbol{\Theta}|\boldsymbol{X_{1:k-1}})] \tag{5.26}$$

To start the learning process, we consider Dirichlet distributions as conjugate prior distributions for the two weighting functions:

$$p(\boldsymbol{\pi_k}|\boldsymbol{X_{1:k-1}}) = \mathrm{Dir}(\boldsymbol{\pi_k}|\boldsymbol{\alpha_\pi}) = C(\boldsymbol{\alpha_\pi}) \prod_{j=1}^{M} \pi_{kj}^{\alpha_\pi^j - 1} \tag{5.27a}$$

$$p(\boldsymbol{\nu_k}|\boldsymbol{X_{1:k-1}}) = \mathrm{Dir}(\boldsymbol{\nu_k}|\boldsymbol{\alpha_\nu}) = C(\boldsymbol{\alpha_\nu}) \prod_{j=1}^{M} \nu_{kj}^{\alpha_\nu^j - 1} \tag{5.27b}$$

where we denote the parameters of Dirichlet distributions by $\boldsymbol{\alpha_\pi}, \boldsymbol{\alpha_\nu} \in \mathbb{R}^{M\times 1}$, and $\boldsymbol{\alpha_\pi} = [\alpha_\pi^1, ..., \alpha_\pi^M]^T$ with the constraint $\alpha_\pi^j > 0,\ j = 1, ..., M$, ensuring that the distribution (5.27a) can be normalized [203]. $\boldsymbol{\alpha_\nu}$ and its component α_ν^j are similarly defined. Moreover, $C(\boldsymbol{\alpha_\pi})$ and $C(\boldsymbol{\alpha_\nu})$ are two normalization constants for the considered Dirichlet distributions.

5.3.2 Learning proposal distributions by VBI

To learn the hyper-parameters and slow features, the following joint distribution is considered as:

$$\begin{aligned} p(\boldsymbol{\Theta}, \boldsymbol{X_k}|\boldsymbol{X_{1:k-1}}) &= p(\boldsymbol{X_k}, \boldsymbol{S_k}|\boldsymbol{X_{1:k-1}}, \boldsymbol{z_k}, \boldsymbol{y_k}) \\ &\quad \times p(\boldsymbol{z_k}, \boldsymbol{y_k}|\boldsymbol{\pi_k}, \boldsymbol{\nu_k}) p(\boldsymbol{\pi_k}, \boldsymbol{\nu_k}|\boldsymbol{X_{1:k-1}}) \end{aligned} \tag{5.28}$$

where $p(\boldsymbol{\pi_k}, \boldsymbol{\nu_k}|\boldsymbol{X_{1:k-1}}) = \mathrm{Dir}(\boldsymbol{\pi_k}|\boldsymbol{\alpha_\pi}) \cdot \mathrm{Dir}(\boldsymbol{\nu_k}|\boldsymbol{\alpha_\nu})$ can be obtained from the prior information (5.27). In addition, one can decompose $p(\boldsymbol{X_k}, \boldsymbol{S_k}|\boldsymbol{X_{1:k-1}}, \boldsymbol{z_k}, \boldsymbol{y_k})$ by marginalizing over $\boldsymbol{S_{k-1}}$ as:

$$\int_{S_{k-1}} p(\boldsymbol{X_k}|\boldsymbol{S_k}, \boldsymbol{y_k}) p(\boldsymbol{S_k}|\boldsymbol{S_{k-1}}, \boldsymbol{z_k}) p(\boldsymbol{S_{k-1}}|\boldsymbol{X_{1:k-1}})\ d\boldsymbol{S_{k-1}}$$

where $p(\boldsymbol{S_{k-1}}|\boldsymbol{X_{1:k-1}}) = \mathcal{N}(\boldsymbol{m_{k-1}^{(f)}}, \boldsymbol{P_{k-1}^{(f)}})$. For clarification, the superscripts '(f)' and '(p)' denote the filtered and predicted quantities. With the transition and emission likelihood functions (5.21), the joint distribution (5.28) is rewritten as

$$\begin{aligned} p(\boldsymbol{\Theta}, \boldsymbol{X_k}|\boldsymbol{X_{1:k-1}}) = &\prod_{j=1}^{M} \left[\nu_{kj}^{(p)} \mathcal{N}(\boldsymbol{X_k}|\boldsymbol{H^j S_k}, \boldsymbol{R^j}) \right]^{y_{kj}} \\ &\cdot C(\boldsymbol{\alpha_\nu}) \prod_{j=1}^{M} \left[\nu_{kj}^{(p)} \right]^{\alpha_\nu^j - 1} \int_{\boldsymbol{S_{k-1}}} \left\{ \prod_{j=1}^{M} \left[\pi_{kj}^{(p)} \mathcal{N}(\boldsymbol{S_k}|\boldsymbol{A^j S_{k-1}}, \boldsymbol{Q^j}) \right]^{z_{kj}} \right. \\ &\left. \cdot \mathcal{N}\left(\boldsymbol{S_{k-1}}|\boldsymbol{m_{k-1}^{(f)}}, \boldsymbol{P_{k-1}^{(f)}}\right) \right\} d\boldsymbol{S_{k-1}} \cdot C(\boldsymbol{\alpha_\pi}) \prod_{j=1}^{M} \left[\pi_{kj}^{(p)} \right]^{\alpha_\pi^j - 1} \\ = &\prod_{j=1}^{M} \left[\nu_{kj}^{(p)} \mathcal{N}(\boldsymbol{X_k}|\boldsymbol{H^j S_k}, \boldsymbol{R^j}) \right]^{y_{kj}} \cdot C(\boldsymbol{\alpha_\nu}) \prod_{j=1}^{M} \left[\nu_{kj}^{(p)} \right]^{\alpha_\nu^j - 1} \end{aligned}$$

$$\cdot \prod_{j=1}^{M} \left[\pi_{kj}^{(p)} \mathcal{N}(\boldsymbol{S_k} | \boldsymbol{m_k^{j(p)}}, \boldsymbol{P_k^{j(p)}}) \right]^{z_{kj}} \cdot C(\boldsymbol{\alpha_\pi}) \prod_{j=1}^{M} \left[\pi_{kj}^{(p)} \right]^{\alpha_\pi^j - 1} \tag{5.29}$$

The second equality above is due to the binary property of z_{kj}, which permits the exchange of integral and product operators [207]. Then the predicted parameters are readily drawn as:

$$\boldsymbol{m_k^{j(p)}} = \boldsymbol{A^j} \cdot \boldsymbol{m_{k-1}^{(f)}} \tag{5.30a}$$

$$\boldsymbol{P_k^{j(p)}} = \boldsymbol{A^j} \cdot \boldsymbol{P_{k-1}^{(f)}} \cdot \boldsymbol{A^{j^T}} + \boldsymbol{Q^j} \tag{5.30b}$$

Applying parameter learning formula (5.26), we can update the variational distribution as follows:

$$q^*(\boldsymbol{z_k}) = \prod_{j=1}^{M} [\pi_{kj}^{(f)}]^{z_{kj}},\ q^*(\boldsymbol{y_k}) = \prod_{j=1}^{M} [\nu_{kj}^{(f)}]^{y_{kj}} \tag{5.31a}$$

$$q^*(\boldsymbol{\pi_k}) = \mathrm{Dir}(\boldsymbol{\pi_k} | \hat{\boldsymbol{\alpha}}_{\boldsymbol{\pi}}),\ q^*(\boldsymbol{\nu_k}) = \mathrm{Dir}(\boldsymbol{\nu_k} | \hat{\boldsymbol{\alpha}}_{\boldsymbol{\nu}}) \tag{5.31b}$$

where the hyper-parameters are updated as

$$\begin{aligned}
\pi_{kj}^{(f)} &\propto \pi_{kj}^{(p)} \cdot \mathcal{N}(\boldsymbol{S_k} | \boldsymbol{m_k^{j(p)}}, \boldsymbol{P_k^{j(p)}}) \\
\nu_{kj}^{(f)} &\propto \nu_{kj}^{(p)} \cdot \mathcal{N}(\boldsymbol{X_k} | \boldsymbol{H^j m_k^{j(p)}}, \boldsymbol{R^j} + \boldsymbol{H^j P_k^{j(p)} H^{j^T}}) \\
\hat{\alpha}_\pi^j &= \alpha_\pi^j + \pi_{kj}^{(f)},\ \hat{\alpha}_\nu^j = \alpha_\nu^j + \nu_{kj}^{(f)}
\end{aligned}$$

where $\hat{\boldsymbol{\alpha}}_{\boldsymbol{\pi}} = [\hat{\alpha}_\pi^1, ..., \hat{\alpha}_\pi^M]^T \in \mathbb{R}^{M \times 1}$ and $\hat{\boldsymbol{\alpha}}_{\boldsymbol{\nu}} = [\hat{\alpha}_\nu^1, ..., \hat{\alpha}_\nu^M]^T \in \mathbb{R}^{M \times 1}$ with constraints $\hat{\alpha}_\pi^j, \hat{\alpha}_\nu^j > 0$, and $j = 1, ..., M$. Based on the updated parameters, the joint distribution (5.29) can be approximated with a single Gaussian distribution by combining the first- and second-order moments of $\boldsymbol{S_k}$, which leads to

$$q^*(\boldsymbol{S_k}) = \mathcal{N}(\boldsymbol{S_k} | \boldsymbol{m_k^{(f)}}, \boldsymbol{P_k^{(f)}}) \tag{5.32}$$

where

$$\boldsymbol{P_k^{(f)}} = \left[\sum_{j=1}^{M} \left[\nu_{kj}^{(f)} \boldsymbol{H^{j^T}} \boldsymbol{R^{j^{-1}}} \boldsymbol{H^j} + \pi_{kj}^{(f)} \boldsymbol{P_k^{j(p)^{-1}}} \right] \right]^{-1}$$

$$\boldsymbol{m_k^{(f)}} = \boldsymbol{P_k^{(f)}} \sum_{j=1}^{M} \left[\nu_{kj}^{(f)} \boldsymbol{H^{j^T}} \boldsymbol{R^{j^{-1}}} \boldsymbol{X_k} + \pi_{kj}^{(f)} \boldsymbol{P_k^{j(p)^{-1}}} \boldsymbol{m_k^{j(p)}} \right]$$

With the filtered slow features in (5.32), one can predict the outputs (labels) in the target domain as follows:

$$\hat{\boldsymbol{Y}}_{\boldsymbol{k}} = \sum_{j=1}^{M} \nu_{kj}^{(f)} \boldsymbol{B}^{j} \boldsymbol{m}_{\boldsymbol{k}}^{(f)} \tag{5.33}$$

The procedures for implementing the proposed transfer slow feature analysis technique are summarized as Algorithm 2.

Algorithm 2. The proposed TSFA

Input: M source domain data $(\boldsymbol{X}_{\boldsymbol{1:N}}^{\boldsymbol{j}}, \boldsymbol{Y}_{\boldsymbol{1:N}}^{\boldsymbol{j}})$, and target domain data $(\boldsymbol{X}_{\boldsymbol{k}})$.

Output: Predicted outputs $\hat{\boldsymbol{Y}}_{\boldsymbol{k}}$ in the target domain

Offline learning

For $j = 1:1:M$

(1). Initialize the hyper-parameters $\boldsymbol{\Phi}$;

(2). Update $q(\boldsymbol{H}^{\boldsymbol{j}})$, $q(\boldsymbol{R}^{\boldsymbol{j}})$ and $q(\boldsymbol{A}^{\boldsymbol{j}})$ using (5.13) and (5.15);

(3). Learn $q(\boldsymbol{S}_{\boldsymbol{1:N}})$ with (5.17);

(4). Iteratively perform the above steps (2) and (3) until the LB converges.

end

Obtain M source domain models $(\boldsymbol{A}^{\boldsymbol{j}}, \boldsymbol{H}^{\boldsymbol{j}}, \boldsymbol{R}^{\boldsymbol{j}})$, with $j = 1, ..., M$

Online learning

For $k = 1:1:N_t$

(1). Predict $\boldsymbol{m}_{\boldsymbol{k}}^{\boldsymbol{j}(\boldsymbol{p})}$ and $\boldsymbol{P}_{\boldsymbol{k}}^{\boldsymbol{j}(\boldsymbol{p})}$ using (5.30);

(2). Learn model identities $(\boldsymbol{z}_{\boldsymbol{k}}, \boldsymbol{y}_{\boldsymbol{k}})$ and weights $(\boldsymbol{\pi}_{\boldsymbol{k}}, \boldsymbol{\nu}_{\boldsymbol{k}})$ by (5.31) using input observations $(\boldsymbol{X}_{\boldsymbol{k}})$;

(3). Estimate slow features by (5.32) based on input observations $(\boldsymbol{X}_{\boldsymbol{k}})$;

(4). Predict the quality outputs $(\hat{\boldsymbol{Y}}_{\boldsymbol{k}})$ with (5.33);

(5). Repeat steps (2) and (3) using input and output observations (namely $[\boldsymbol{X}_{\boldsymbol{k}}; \boldsymbol{Y}_{\boldsymbol{k}}]$) if $\boldsymbol{Y}_{\boldsymbol{k}}$ is available.

end

5.4 Simulation and validation

In this section, a numerical example, a public dataset and an industrial case study are utilized to demonstrate the applicability of the proposed transfer slow feature analysis method in soft sensor modelling. All data samples used in the three cases are time-series sequences. The well labeled source domain data (including all inputs

and outputs) are utilized for the offline training of the TSFA model in order to learn source domain models. The time-series data with limited labels (namely all inputs and limited outputs) in the target domain are used to infer the TSFA model and quality outputs in an online learning manner. The time difference between two available outputs (labels) in the target domain is set as two (i.e., one of every two consecutive labels is missing) in all cases. In the proposed TSFA model, hyper-parameters are chosen with the preference for process slowness as: $\{\mu_a, \sigma_a, \boldsymbol{m_0}, \boldsymbol{P_0}, \alpha_r, \beta_r, \boldsymbol{\mu_0}, \boldsymbol{\Lambda_0}\} = \{0.5, 1, \mathbf{1}_{\boldsymbol{d\times 1}}, \boldsymbol{I_d}, 0.1, 100, \mathbf{1}_{\boldsymbol{d\times 1}}, 200 \times \boldsymbol{I_d}\}$, where $\mathbf{1}_{\boldsymbol{d\times 1}}$ denotes a d dimensional vector with all elements 1.

As for comparison, six state-of-the-art methods including probabilistic slow feature analysis (PSFA) [195], probabilistic principal component analysis (PPCA) [208], linear transfer component analysis (TCA) [209], metric transfer learning framework (MTLF) [210], multi-component transfer metric learning (MCTML) [211], and domain-adversarial neural network (DANN) [212, 213] are investigated as baseline techniques in three cases. The hyper-parameters are selected via 5-fold cross-validation on the training data for baseline methods. In addition, the DANN utilized here considers the mean squared error as the objective function for regression tasks instead of the binary cross-entropy for classification tasks in the original scheme [212, 213] and the trade-off parameter λ is selected from $[2^{-9}, 2^{-8}, ..., 1]$ by cross-validation. For the implementation of the DANN algorithm, a linear discriminative network is used as the last layer for regression tasks. Considering the small data sizes, a one hidden layer architecture is considered as the DANN framework and the neuron numbers are selected from 50, 100, and 150 by cross-validation. Principal component analysis (PCA) [191] is taken as a preprocessing method for the MTLF and the MCTML when dimension reduction is needed for dealing with high-dimensional data.

5.4.1 Numerical simulation example

To verify the effectiveness of the proposed transfer slow feature analysis method, a synthetic dataset containing 4 domains are generated based on system (5.3). More specifically, 4 models are constructed with the following transition matrices: $\boldsymbol{A^1} = \text{diag}(0.99, 0.95, 0.92, 0.87)$, $\boldsymbol{A^2} = \text{diag}(0.99, 0.94, 0.91, 0.88)$, $\boldsymbol{A^3} = \text{diag}(0.80, 0.40, 0.30, 0.20)$, and $\boldsymbol{A^4} = \text{diag}(0.25, 0.24, 0.23, 0.12)$, where it is apparent that the first two tran-

sition matrices (namely $\boldsymbol{A^1}$ and $\boldsymbol{A^2}$) correspond to slow features, while $\boldsymbol{A^3}$ accounts for both fast and slow features, and $\boldsymbol{A^4}$ represents fast varying features. Using (5.4b), the corresponding variance matrices $\boldsymbol{Q^j}$ $(j = 1, ..., 4)$ can be readily computed. As for emission matrices, 4 random $\boldsymbol{H^j} \in \mathbb{R}^{5\times4}$ and $\boldsymbol{R^j} = r_j^{-1} \cdot \boldsymbol{I_5}$ with $j = 1, ..., 4$ (Here $\boldsymbol{H^j}$ and $\boldsymbol{R^j}$ represent $\boldsymbol{H^e}$ and $\boldsymbol{R^e}$ in (5.7)) are introduced for synthetic data generation. For demonstration purpose, we further choose $\boldsymbol{H^1} = \boldsymbol{H^2}$, while $\boldsymbol{H^3}$ and $\boldsymbol{H^4}$ are different random matrices.

In this case, we set the first model as the target domain and the remaining 3 models as the source domains. Along this line, 150 samples from 3 source domains are used for training and 50 samples from the target domain are utilized for testing. Based on the generated measurements of inputs and outputs from domain 2 to domain 4, 3 linear Gaussian state-space models are learned as source domain models. Through the evidence lower bound, the dimension of latent variables of 3 source domain models are determined as 3 that is quite close to the real dimension value 4. As shown in Fig. 5.3(a), the learned transition matrices can basically reflect the real values. Especially, the learned source model 1 has a clear preference for slowness, which shows the potential to be transferred to the target model. In addition, each source domain model has a distinct emission matrix and measurement noise variance as illustrated in Fig. 5.3(b) and Fig. 5.4.

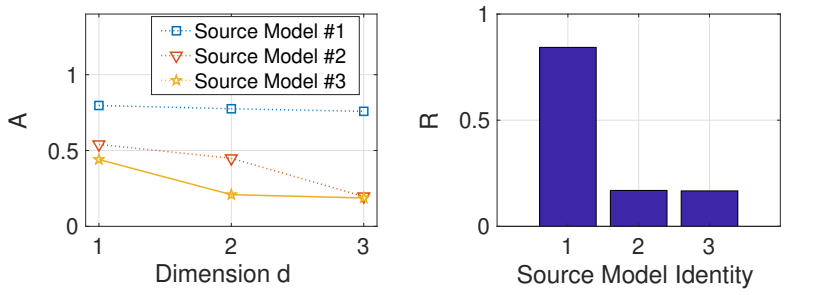


Figure 5.3: Transition matrices and measurement noise variances learned for source domain models by TSFA

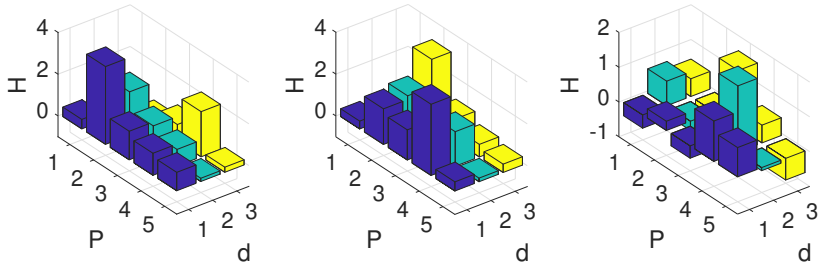


Figure 5.4: Emission matrices learned for source domain models by TSFA

Table 5.2: Comparison of different dimensions of latent variables in the synthetic data in terms of correlation coefficient

d (N)	PPCA	PSFA	TCA	MTLF	MCTML	DANN	TSFA
2 (50)	0.526	0.167	0.411	0.393	0.502	0.261	0.619
3 (100)	0.317	0.519	0.378	0.401	0.435	0.315	**0.631**
4 (150)	0.316	-0.058	0.362	0.473	0.419	0.318	0.628

Table 5.3: Comparison of different dimensions of latent variables in the synthetic data in terms of root mean square error

d (N)	PPCA	PSFA	TCA	MTLF	MCTML	DANN	TSFA
2 (50)	1.452	1.723	1.602	1.282	1.237	1.187	1.173
3 (100)	1.447	1.338	1.597	1.284	1.271	1.194	**1.155**
4 (150)	1.447	1.888	1.594	1.228	1.227	1.193	1.166

To quantify the soft sensing performance, Pearson correlation coefficient (PCC) and Root Mean Square Error (RMSE) are calculated for TSFA and six competing algorithms under consideration of latent variables with different dimensions (namely $d = 2, 3, 4$) or neuron numbers (i.e., $N = 50, 100, 150$) in the DANN method. As observed in Table 5.2 and Table 5.3, the proposed TSFA outperforms two traditional machine learning methods (namely PPCA and PSFA) and four transfer learning algorithms (including TCA, MTLF, MCTML, and DANN). Additionally, TSFA performs better than PSFA since it reduces the effects of negative transfer from unrelated source domains by reducing the corresponding weights. The possible reason why TCA, MTLF, MCTML, and DANN are not as good as TSFA in terms of regression

performance (especially PCC) might be because TCA, MTLF, MCTML, and DANN perform a static transfer learning mechanism, which ignores model dynamics in cross-domain transfer learning. It can be noted that MTLF, MCTML and DANN achieve smaller prediction errors than those of TCA, and MCTML and MTLF achieve better prediction performance than TCA and DANN with respect to PCC. MCTML outperforms MTLF in terms of PCC and RMSE, owing to its capacity of automatically learning distinct components from source domains and mitigate the negative influence of unrelated samples. Moreover, DANN shows comparable performance to TSFA in terms of RMSE. Graphically, the soft sensor performance of TSFA and PSFA with $d = 3$ is shown in Fig. 5.5, where both TSFA and PSFA can learn the main trend of the reference output, while PSFA shows obvious local divergence especially at time instants 20~25. In addition, the learned weighting functions of transition and emission equations corresponding to the 1st source domain (namely model 2) converge to 1 rapidly as depicted in Fig. 5.6. This validates the fact that the 1st source domain resembles the target domain as expected. Hence, the learned weights can be utilized to dynamically measure the cross-domain transferability.

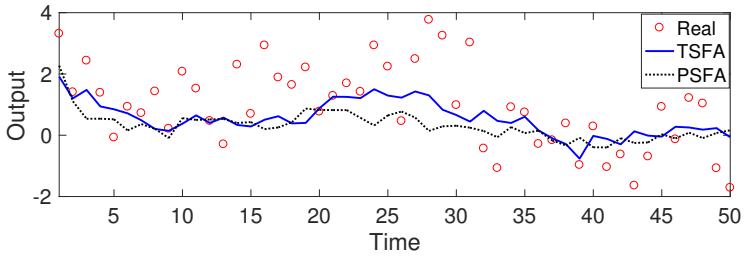


Figure 5.5: Prediction performance of TSFA on simulated data

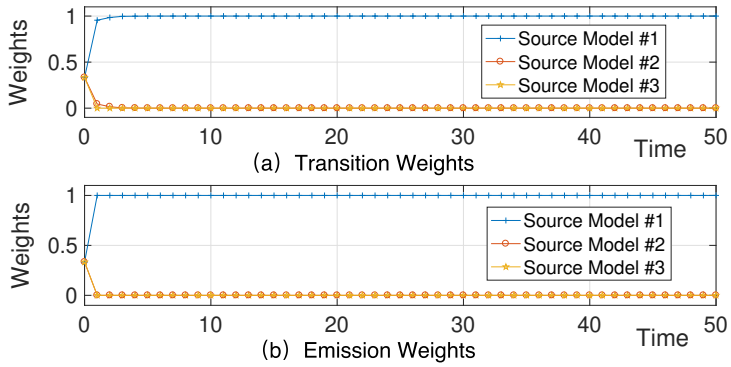


Figure 5.6: Learning to update transition and emission weights by TSFA

Table 5.4: Comparison of different target domains in the synthetic data in terms of correlation coefficient

#	PPCA	PSFA	TCA	MTLF	MCTML	DANN	TSFA
2	-0.491	0.122	0.488	0.412	0.312	0.222	**0.514**
3	0.187	0.102	-0.078	0.212	**0.252**	0.052	0.238

Table 5.5: Comparison of different target domains in the synthetic data in terms of root mean square error

#	PPCA	PSFA	TCA	MTLF	MCTML	DANN	TSFA
2	1.616	1.321	1.140	1.121	1.141	1.245	**1.064**
3	2.660	2.701	2.721	2.639	2.621	**1.662**	2.641

To further investigate the transfer learning performance of the propose TSFA, we test model 2 and model 3 with slow varying features respectively as the target domain and the remaining 3 domains as source domains. The performance of soft sensors designed by TSFA and six baseline methods is evaluated and compared in Table 5.4 and Table 5.5 with respect to PCC and RMSE. It is shown that the proposed TSFA outperforms most baseline methods with respect to regression performance since it aims at learning cross-domain slow features. When model 3 is selected as the target domain, DANN shows the smallest RMSE among other methods, while having smaller PCC than most methods including the proposed TSFA, and MCTML achieves the best prediction performance in terms of PCC and smaller RMSE than the proposed TSFA. Using variational Bayesian inference, the proposed transfer slow feature analysis is able to integrate the process knowledge and historical data in the source domains and achieves model parameters transfer and slow features learning in the target domain, which leads to a better online prediction of quality variables in the target domain compared to the competing algorithms.

5.4.2 Simulation study on TE process

In this section, the Tennessee Eastman (TE) process [214] as a public simulation benchmark is utilized to test the applicability and effectiveness of the proposed method. TE process has been extensively deployed to validate and evaluate control, identification and estimation techniques [198, 215, 216, 217, 218, 219]. Based on the

simulator and decentralized control strategies proposed in [217, 218], we generate data from 41 measurement variables and 12 manipulated variables under six normal operating modes as shown in Table 5.6. Although the proposed transfer slow feature analysis aims at dynamic feature extraction, it should be applicable for static or weak dynamic process analysis under similar operating conditions [198].

Table 5.6: Operating modes of TE process

Mode	Mass Ratio (G/H)	Production Rate (m^3/h)
1	50/50	22.89
2	40/60	22.89
3	60/40	22.89
4	50/50	19.45
5	40/60	19.45
6	60/40	19.45

Table 5.7: Validation results of estimating component C in the TE process in terms of correlation coefficient

#	PPCA	PSFA	TCA	MTLF	MCTML	DANN	TSFA
1	0.184	0.437	0.016	0.809	0.748	0.751	**0.835**
2	0.362	-0.459	-0.039	0.426	0.685	0.703	**0.763**
3	0.093	0.016	-0.179	0.543	0.654	0.705	**0.897**
4	0.556	0.167	-0.323	0.689	0.566	0.719	**0.724**
5	0.458	-0.184	-0.279	0.655	0.503	0.623	**0.853**
6	-0.251	0.157	-0.316	0.765	0.510	0.565	**0.806**

Table 5.8: Validation results of estimating component C in the TE process in terms of root mean square error

#	PPCA	PSFA	TCA	MTLF	MCTML	DANN	TSFA
1	3.244	2.967	3.350	1.956	2.197	2.034	**1.873**
2	2.314	3.874	2.569	2.460	1.838	1.852	**1.667**
3	1.691	2.079	1.770	1.449	1.272	1.562	**0.791**
4	1.986	2.241	2.310	1.590	1.798	1.732	**1.551**
5	2.473	3.398	2.820	2.067	2.364	1.857	**1.449**
6	2.759	2.619	2.617	1.620	2.180	1.810	**1.485**

Based on 31 commonly used online measurements that consist of 9 manipulated variables and 22 process measurements, we intend to establish soft sensing models for

predicting a quality variable (usually measured offline) under six different operating conditions. With prior knowledge [216], 13 representative features are selected out of 31 features and fed into feature learning algorithms for building predictive models to infer a key product (Component C) in an intermediate stream (Stream 6) [214]. Specifically, 13 representative features include feed flow, purge flow, reactor feed rate, reactor pressure and etc., and we refer to [216] for detailed physical meanings of all 13 features. There are 205 input-output samples collected under each operating condition. In the training process, 1,025 samples from five operating modes (source domains) are used for predictive model learning, while 205 samples from the remaining operating mode (target domain) are utilized for validation.

The testing results are evaluated by PCC and RMSE as shown in Table 5.7 and Table 5.8, respectively. It is apparent that the proposed TSFA is capable of extracting transferrable slow features and thus providing better prediction results in comparison to six baseline methods in all cases. In addition, MTLF, MCTML, and DANN as the-state-of-art transfer learning techniques outperform other baseline methods with higher correlation coefficients and smaller prediction errors in most cases, and achieve comparable performance to TSFA in some cases (such as case #1 where mode 1 is taken as the target domain). It is noted that the performance of MCTML is close to (or worse than) that of MTLF in some cases, because only independent components are extracted in the source domains [211].

5.4.3 Validation through industrial process

In order to further investigate the usage of the proposed method, an industrial process with potential latent slow features is considered for a soft sensor modelling exercise. This process is about steam-assisted gravity drainage (SAGD), which is a widely utilized in-situ technology for heavy oil recovery using horizontal well-pairs. In SAGD operations, hot steam is injected to underground heavy oil or bitumen reservoirs through injector wells and the mixture of emulsion (including crude oil, water and etc) is retrieved through producer wells using submersible pumps. The injected hot steam heats up heavy crude oil and reduces its viscosity, making it easier to be pumped out. In this process, water is inevitably mixed in the crude oil emulsion which is undesired and needs to be separated for the recycling purpose. However, current engineering

practice relies on the off-line measurement of the emulsion water content, which is not efficient. Therefore, real-time estimation is preferable for efficient process control and operation. Considering the similarities of well-pairs in working mechanisms and operating conditions, it is potential to transfer the knowledge learned from source well-pairs to enhance soft sensing performance of target well-pairs of interest.

In this study, 11 well-pairs are considered and each well-pair has 19 process variables for predicting the quality variable (water content), including steam temperatures, emulsion flow rate and etc. As stated above, the 19 process variables are measured online while the water content is measured offline with a slow sampling rate by a device called test separator. For proprietary reasons, all data have been normalized. In addition, the fast-rate data is averaged within an acceptable time range to align them with the slow-rate water content data. Consequently, there are 49 sample measurements of the water content for each well-pair within a one-year working period (one averaged sample per week). Based on the prior process knowledge, 6 features including steam temperature, total flow rate, produced emulsion pressure and temperature, casing gas pressure and temperature are selected for further feature extraction.

Table 5.9: Validation results of estimating water content in the SAGD process in terms of correlation coefficient

#	PPCA	PSFA	TCA	MTLF	MCTML	DANN	TSFA
1	0.322	0.131	-0.311	-0.158	0.046	-0.294	**0.467**
2	0.283	0.121	0.228	-0.208	-0.007	0.116	**0.537**

Table 5.10: Validation results of estimating water content in the SAGD process in terms of root mean square error

#	PPCA	PSFA	TCA	MTLF	MCTML	DANN	TSFA
1	**0.911**	0.962	0.954	0.941	0.937	1.149	0.961
2	0.883	0.926	0.876	0.899	0.899	1.567	**0.759**

To demonstrate the performance of the proposed transfer learning technique, we consider well-pair 1 and well-pair 2 as the target domain individually and the remaining 10 well-pairs are treated as source domains. In this study, process variables and the water content variable are considered available for learning each source model,

while process variables and limited labels are considered known in the target domain. In total, 490 samples from 10 source domains are used for model training and 49 samples from the target domain are utilized for model testing. Thus 10 source domain models are learned and transferred to the target domain for soft sensor design by TSFA. As shown in Table 5.9 and Table 5.10, the proposed method greatly improves the prediction performance in comparison to the baseline methods in terms of PCC and RMSE. This indicates that the proposed method is able to capture the slowness characteristics in source well-pairs and transfer them to the target well-pair. By calculating lower bounds, the latent variable dimensions in both cases are determined as 5 for PSFA and TSFA, while the latent variable dimensions of PPCA and TCA are selected as 1, 1 and 3, 2 based on the accumulated eigenvalues. In both cases, the dimensions of latent space of MTLF and MCTML are chosen as 5 by using PCA (based on the accumulated eigenvalues), and the neuron numbers of 1 hidden layer DANN are selected as 100 and 150 respectively (from 50, 100, and 150) according to cross-validation. In this case, TSFA shows much better performance than TCA, MTLF, MCTML, and DANN due to its ability to account for dynamical transfer learning by the use of variational Bayesian inference. Although PPCA, as the best competitor, shows the smallest prediction error in the second case, it is not capable of achieving better PCC than the proposed method. Hence, PPCA is not suitable for such dynamic transfer learning task because it can only predict some global trends while fails to reflect the local variations that can be well estimated by using the proposed TSFA method as illustrated in Fig. 5.7.

5.5 Conclusion

In this chapter, an online transfer slow feature analysis (TSFA) was developed for dynamically transferring the multiple source models to the target domain that has limited output labels. More specifically, the probabilistic slow feature analysis method was adopted for multiple source domain models learning (with a preference for slow features) using input and output measurements from different source domains based on variational Bayesian inference. To address the constraint issue associated with the transition matrix, the truncated Gaussian distribution was introduced as a conjugate

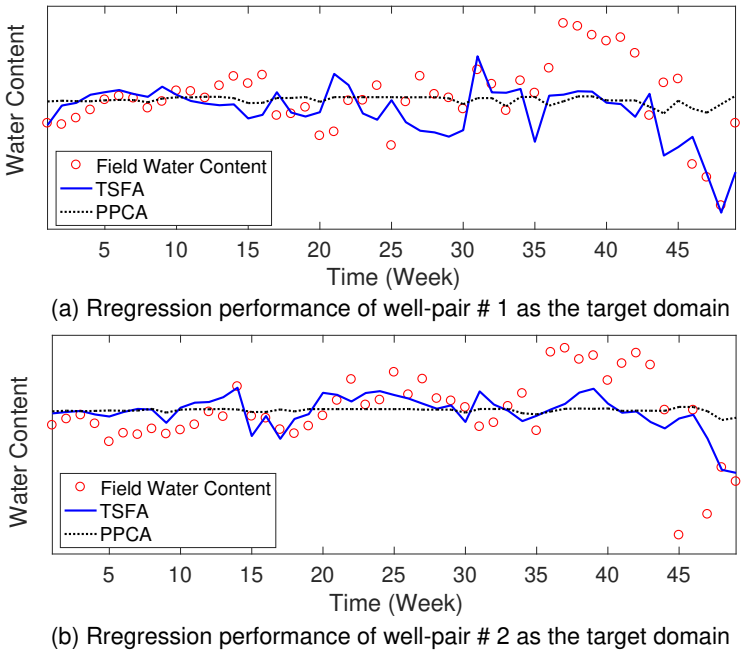


Figure 5.7: Prediction performance of TSFA in the SAGD process

prior for efficient learning. In addition, two weighting functions corresponding to transition and emission matrices of source models were dynamically updated to measure the transferability from each source domain model to the target domain model at each time instant. The effectiveness of the proposed technique with application in soft sensor modelling is validated through a simulation example, a public dataset, and an industrial SAGD process.

Chapter 6

Hybrid Estimator Design for Long Range Pipeline Leak Detection and Localization

6.1 Introduction

As one of the widely utilized infrastructures for material transportation, numerous pipelines have been laid onshore and offshore for crude and refined oil and gas products delivering [220]. The practical difficulty, however, is that the operational status of buried pipeline is rarely directly observable. On the other hand, harsh and changeable operational and environmental conditions (extreme pressure and temperature, etc.) may induce pipeline defects (such as leakage, breakage, bending, or other failure modes) which, in turn, may affect its reliability and transportation efficiency, and even result in severe damage, such as economic loss, ecological contamination and human accidents [221, 222, 223]. Among these defects, leakage is considered to be one of the most damaging defects affecting a pipeline system [224]. It is crucial, therefore, to propose reliable and effective leak detection and localization methods to guarantee pipeline integrity.

To address this, numerous methods have been proposed. Mostly, they fall under three categories: 1) non-technical (manual) methods, 2) hardware based (semi-automated) methods and 3) software based (automated) methods [221]. Furthermore, non-technical methods and hardware based methods can be classified as direct methods (or external methods), while software based techniques can be referred as indirect methods (or internal methods). Most of the direct methods are usually realized by

routine survey (offline monitoring) rather than continuous online monitoring, which can be done by software based techniques. Another difference is that software based methods could be updated in existing software systems, while hardware based techniques are difficult to be reinstalled once the hardware system is built. As stated in the literatures [221, 225], non-technical methods involve patrolling by skilled workers along pipelines to detect leaks by relying on natural senses (seeing, smelling or hearing). Despite being accurate in detecting large leaks, these methods are not suitable for detection of small leaks and in particular the leaks at their early stages. On top of it, they require well-trained operators or qualified agents in the process, and heavily depend on the patrolling frequency and the surrounding environment. Particularly, it is hard to utilize these approaches when it comes to subsea or onshore buried pipes.

As for hardware based methods, they mostly rely on physical sensors and devices, such as swimming robots, intelligent pigs, cable sensors, optical fibres, infrared cameras, vapour/liquid sensing tubes, etc [226, 227, 228]. The sensors and devices are used to monitor working status of pipeline systems to detect any irregular working condition, such as leakage, and then locate a specific leakage position. One inevitable drawback of these methods is prohibitive costs of installation and maintenance of numerous sensors along entire pipeline systems. On the other hand, software based methods comprise model-based approaches and data-driven techniques. With limited measurement of process variables at pipeline ends provided by SCADA (Supervisory Control and Data Acquisition) systems, including pressure, mass flow rate and temperature, model-based methods aim at finding the relationship between measurements and the process states, by using first principle modelling. Then, the estimated outputs and associated measurements are compared, to diagnose and pinpoint leakage [225].

The advantage of model-based methods is that physical meaning is clearly described which enables us to capture the intrinsic nature of hydraulic dynamics, whereas model adaption may fail when it comes to varying working conditions. Typical models in this category are: mass/volume balance model, real time transient model, negative pressure wave method and pressure point analysis. On the other hand, data-driven techniques, as black box modelling approaches (such as support vector machine [229, 230, 231], neural networks [232], wavelet transform [233], etc.), have

become increasingly popular in the past decades because of the advances in sensor technology and computing capacity. However, these techniques fail to provide a link between the estimation variables and pipeline physical characteristics, if they are employed by pure application of statistic algorithms without guidance on physical model features.

As a representative of model-based approaches, real time transient model has attracted a lot of interest from academia and industry. In particular, this model has been widely used in water-pipeline and oil/gas-pipeline transient analysis and condition monitoring due to its cost-effective performance, noninvasive installation and wide operational ranges [225]. Specifically, the presence of leakage in pipelines can alter hydraulic transient behaviour by changing the spatiotemporal profiles of pressure, mass flow velocity and/or temperature. Hence, abnormal phenomena can be captured by comparing the measured and estimated process variables obtained by real time transient model based on mass, energy and momentum conservation laws.

In the past decades, some enhanced transient model-based leak-detection techniques have been proposed, including inverse transient analysis (ITA) [228], frequency method [227] and direct transient analysis (DTA) [225]. However, there are still some drawbacks that prevent their further application in pipeline leak detection. For instance, in order to achieve better performance, ITA has to entail a wide range of measurements at different locations along a pipeline system to calibrate its model in terms of deviation analysis between measured values and calculated ones. Moreover, when it comes to larger scale pipeline networks, numerical simulation error of ITA may be amplified by solving the inverse problem [225]. Likewise, the frequency domain method depends on persistent excitation at one location for transient flow generation to actuate leak-induced system frequency responses, while the relationship between leak-induced frequencies and leak characteristics is still unclear. As an alternative to ITA, direct transient analysis does not utilize numerous spatial measurements but relies on full pressure trajectory estimation by pipeline boundary measurements. Since it is pressure point analysis, it may not be successful in capturing leak-induced features on other process variables and estimating of varying parameters is quite challenging, caused by system uncertainty, process noise and changeable working conditions [225].

To reduce the cost of numerous sensors placed along pipelines, researchers and

engineers aim at designing an adaptive observer or estimator to monitor working condition of pipeline systems and diagnose various faults such as leakage and breakage [234] with limited measurements. Instead of using four boundary conditions (pressure and velocity at both upstream and downstream ends) as measurements, an extended Kalman observer is proposed in the reference [235] for pipeline leak detection and localization with two measurements, i.e., pressure at the pump inlet and pressure at the system downstream end. In Aamo's paper [236], leak size and position can be estimated independently by introducing a loss term to the original water-hammer equations. A probability density function is defined in the loss term to describe the leak spatial distribution. By formulating a continuous-time Luenberger observer, both single and multiple leaks can be estimated effectively by this method. In order to address real-time deployment of the observer under varying operating conditions, an adaptive online observer is design by Wang [237]. To enhance local leakage detectability, a consensus algorithm based on $\mathcal{H}_\infty$ estimator is designed for water pipeline system leak detection using water hammer equation [238]. However, most of these Luenberger or Kalman observers are designed within a continuous-time system setting, which prevents direct applications to digital devices, such as SCADA system or data-processing centres. Although continuous-time designs can be realized in discrete settings by applying Euler finite difference method in time and space, simulation accuracy cannot be often guaranteed, especially when dealing with inverse transient analysis [239], which requires more accurate and robust simulation techniques. On the other hand, most of the model-based diagnosis approaches are based on recursive estimation, which tends to take more time to render estimation results compared to pattern recognition methods.

Due to these drawbacks, a hybrid method in terms of discrete Luenberger observer of a direct transient model and support vector machine model is developed in this chapter. More specifically, a linear infinite-dimensional pipeline hydraulic system is obtained by applying linearization to the original nonlinear water hammer equation. As for model discretization, the Crank-Nicolson approach realized by Cayley-Tustin time discretization is utilized for the pipeline infinite-dimensional system representation in this chapter. There are several main advantages of using Crank-Nicolson discretization setting. Among these, the first one is that no spatial discretization or

model reduction technique is applied. In addition, the Crank-Nicolson method, as an implicit mid-point integration, has second-order discretization accuracy in time, while explicit Euler and implicit Euler methods have first-order discretization accuracy in time. Moreover, the Crank-Nicolson method is symmetric or time-reversible given by the implicit mid-point rule, which ensures numerical stability which might be lost by applying the explicit Euler discretization method [62, 97, 119]. In particular, the discrete hydraulic system is attained by solving for the corrsponding resolvent operator. Based on that, a discrete Luenberger observer is designed to reconstruct spatiotemporal process states with limited boundary measurements. By means of varying boundary conditions of mass flow velocity, different normal conditions and leakage scenarios are simulated, including different leak amounts and leak positions. Based on the database generated by the proposed observer, a machine learning (ML) model - support vector machine (SVM) is utilized to reveal the nonlinear relationship between leak-induced patterns and process variables. Unlike Kalman filter, SVM can convert this recursive estimation process into a pattern recognition problem, which speeds up leak detection, amount estimation and localization process substantially. Finally, a long range pipeline is studied in this chapter, and different leakage amounts and positions can be accurately diagnosed and pinpointed intelligently by combining discrete Luenberger observer and SVM.

The main contributions of this chapter are shown as follows: 1) a closed-form solution of pressure, velocity and density with respect to time and space of distributed pipeline systems is derived and obtained; 2) a distributed discrete Luenberger observer is designed; 3) based on the developed discrete infinite-dimensional model and its corresponding analytical solution, numerous datasets are generated under normal working conditions and different leakage scenarios via limited external measurements, and one-class support vector machine (OCSVM) and support vector regression (SVR) models are exploited to detect leak existence, estimate leak amounts and expose leak locations with high accuracy. The rest of this chapter is constructed as follows: following the description of the water hammer equation, steady state analysis, determination of resolvent operator and discrete-time pipeline hydraulic model are realized in Section 6.2. Based on these, details of the discrete Luenberger observer design are discussed in Section 6.3. Different working conditions, such as normal conditions and

various leak situations including different leak magnitudes and positions, are simulated, detected and diagnosed by means of OCSVM model and SVR model in Section 6.4. Finally, conclusions are drawn in Section 6.5.

6.2 Problem formulation

In this section, the so-called water hammer equation is introduced for pipeline hydraulic modelling. In particular, steady state profiles of pressure, density and velocity are determined by linearization of the original first-order coupled nonlinear hyperbolic PDEs system and then a distributed parameter system setting is given to describe this linearized model. By determining the resolvent operator, one obtains a discrete-time state space realization of the pipeline infinite-dimensional system without spatial approximation or model reduction.

6.2.1 Model description

As one of the most representative and cost-effective ways for material transportation, pipelines have been widely utilized in oil, gas and water distribution industry. Fundamental conservation laws are utilized in pipeline hydraulic modelling, which include conservation laws of mass, momentum and energy.

In the ensuing section, we consider that: 1) a rigid buried pipe is assumed and pipeline parameters are considered to be constants as shown in Table 6.1; 2) the flow is single-phase liquid flow; 3) the flow is viscous (viscosity causes shear stresses in the moving fluid); 4) the flow is one-dimensional flow meaning that mass flow velocity, density and pressure vary only along pipeline axial direction; 5) small density changes in liquid flow imply local compressibility which induces interesting phenomena in piping systems, such as water hammer and leakage [240]; 6) the flow is isothermal. The last assumption implies that temperature changes due to pressure changes and friction effects can be neglected in the flow system. This is motivated by the fact that majority of the pipeline infrastructure is buried underground, and for example, in the case of a buried gasoline pipeline, the heat flux between the flow and surrounding soil is negligible [241].

In the following model development, the conservation of energy is not considered

due to the isothermal flow assumption. Then, a distributed parameter system is provided by applying conservation laws of mass and momentum, which are augmented with the equation of the speed of sound for a liquid flow transport pipeline (physical properties of incoming fluid are given as velocity of 2.05 m/s, pressure of 1.68×10^6 Pa and density of 680 kg/m^3), as follows:

$$\rho \frac{dv}{d\zeta} + \frac{d\rho}{dt} = 0 \tag{6.1a}$$

$$\frac{dv}{dt} + \frac{1}{\rho}\frac{dp}{d\zeta} + g\sin\alpha + \frac{\lambda v|v|}{2D} = 0 \tag{6.1b}$$

$$\frac{dp}{d\rho} = a^2 \tag{6.1c}$$

with boundary conditions:

$$v(0,t) = v_0, p(L_d,t) = p_0, \rho(L_d,t) = \rho_0 \tag{6.2}$$

where v represents flow velocity, p is pressure, ρ stands for liquid density, ζ is spatial position along the pipeline, t is time, the quantity of $g\sin\alpha$ is the ζ-component of the original gravity acceleration g, a is the speed of sound in the fluid, λ is a dimensionless friction coefficient described in Eq.(6.3a) for laminar flow (with Reynolds number smaller than 2320) and the Colebrook equation Eq.(6.3b) for turbulent flow (for larger values of the Reynolds number) [240].

$$\lambda = \frac{64}{\mathrm{Re}} \tag{6.3a}$$

$$\frac{1}{\sqrt{\lambda}} = -2\log_{10}\left(\frac{2.51}{\mathrm{Re}\sqrt{\lambda}} + 0.27\frac{k_R}{D}\right) \tag{6.3b}$$

where D is the pipeline inner diameter, and $\frac{k_R}{D}$ is relative pipe roughness and Re is the Reynolds number. The above system of Eq.(6.1a)-(6.1c) is usually referred as water hammer equation [242]. In the ensuing section, numerical value of λ is directly adopted from the reference [242] and it is shown in Table 6.1. From Eq.(6.1c), one can easily get the relationship between the following derivatives of pressure and density with respect to time and space:

$$\frac{dp}{dt} = a^2\frac{d\rho}{dt}, \quad \frac{dp}{d\zeta} = a^2\frac{d\rho}{d\zeta} \tag{6.4}$$

In fact, considering different boundary conditions of pressure and density as shown in Eq.(6.2), the actual relationship for pressure and density becomes Eq.(6.4) instead of Eq.(6.1c). From full derivatives of $\frac{dv}{dt}$, $\frac{dp}{dt}$ and $\frac{d\rho}{dt}$, one obtains:

$$\frac{dv}{dt} = \frac{\partial v}{\partial t} + v\frac{\partial v}{\partial \zeta} \tag{6.5a}$$

$$\frac{dp}{dt} = \frac{\partial p}{\partial t} + v\frac{\partial p}{\partial \zeta} \tag{6.5b}$$

$$\frac{d\rho}{dt} = \frac{\partial \rho}{\partial t} + v\frac{\partial \rho}{\partial \zeta} \tag{6.5c}$$

Substituting these derivatives in Eq.(6.5a)-(6.5c) into the original distributed parameter system described by Eq.(6.1a)-(6.1c), then the system model becomes as follows:

$$\frac{\partial \rho}{\partial t} + v\frac{\partial \rho}{\partial \zeta} + \rho\frac{\partial v}{\partial \zeta} = 0 \tag{6.6a}$$

$$\rho\frac{\partial v}{\partial t} + \rho v\frac{\partial v}{\partial \zeta} + \frac{\partial p}{\partial \zeta} + \rho g\sin\alpha + \rho\frac{\lambda v|v|}{2D} = 0 \tag{6.6b}$$

$$\frac{\partial p}{\partial t} + v\frac{\partial p}{\partial \zeta} + a^2\rho\frac{\partial v}{\partial \zeta} = 0 \tag{6.6c}$$

Plugging the relationship of pressure and density Eq.(6.4) into Eq.(6.6a)-(6.6c), the density state can be eliminated and then the system can be simplified as the following two-state coupled nonlinear first-order hyperbolic system with the same boundary conditions given by Eq.(6.2):

$$p\frac{\partial v}{\partial t} + pv\frac{\partial v}{\partial \zeta} + a^2\frac{\partial p}{\partial \zeta} + pg\sin\alpha + p\frac{\lambda v|v|}{2D} = 0 \tag{6.7a}$$

$$\frac{\partial p}{\partial t} + v\frac{\partial p}{\partial \zeta} + p\frac{\partial v}{\partial \zeta} = 0 \tag{6.7b}$$

Remark 11. *The above section is devoted to the model development of a single phase pipeline fluid flow model, however, without lost of generality, one can replace Eq.(6.1c) with the equation of state and develop a hydraulic model for single-phase gas flow. A similar type of derivation is presented in Reddy's work [234] where the gas pipeline flow system is considered with mass flow rate and pressure as system states.*

6.2.2 Steady states analysis

In this chapter, we are focused on the pipeline hydraulic model around steady states. Hence, by applying the linearization procedure, the steady states v_{ss}, p_{ss} and ρ_{ss} can

be determined by setting all derivatives with respect to time equal to zero ($\partial/\partial t := 0$) in Eq.(6.6a)-(6.6c) which leads to Eq.(6.8a)-(6.8c).

$$v_{ss}\frac{\partial \rho_{ss}}{\partial \zeta} + \rho_{ss}\frac{\partial v_{ss}}{\partial \zeta} = 0 \tag{6.8a}$$

$$\rho_{ss}v_{ss}\frac{\partial v_{ss}}{\partial \zeta} + \frac{\partial p_{ss}}{\partial \zeta} + \rho_{ss}g\sin\alpha + \rho_{ss}\frac{\lambda v_{ss}|v_{ss}|}{2D} = 0 \tag{6.8b}$$

$$v_{ss}\frac{\partial p_{ss}}{\partial \zeta} + a^2\rho_{ss}\frac{\partial v_{ss}}{\partial \zeta} = 0 \tag{6.8c}$$

Based on this, three corresponding steady states profiles are illustrated in Fig. 6.1 by using finite difference methods. Pipeline parameters used for simulation are shown in Table 6.1 [242].

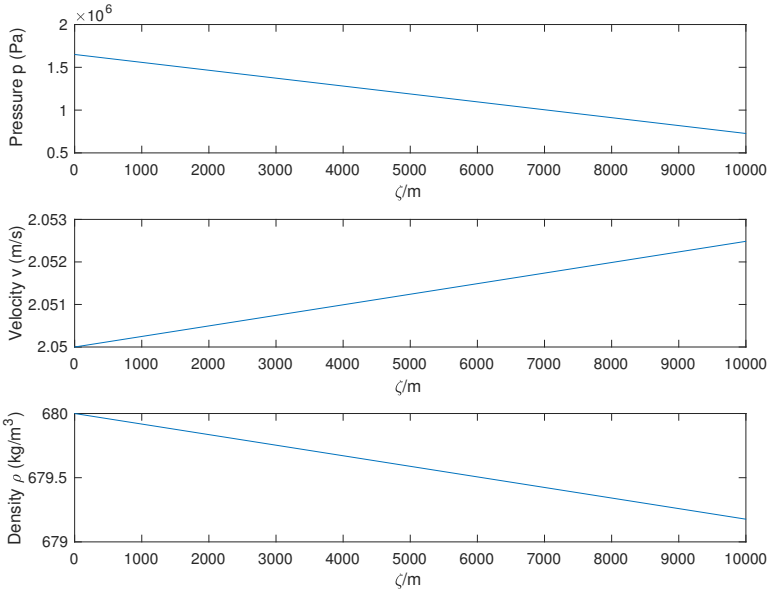


Figure 6.1: Profiles of steady states in Eq.(6.8a)-(6.8c).

Table 6.1: Pipeline parameters

	Notations	Numerical Values
Length	L_d	$10,000\ m$
Speed of Sound	a	$1,059\ m/s$
Friction Coefficient	λ	0.0158
Gravity Acceleration	g	$9.81\ m/s$
Diameter	D	$0.2065\ m$
Inclination Angle	α	-0.00256

Remark 12. *Steady states v_{ss}, p_{ss} and ρ_{ss} must satisfy their corresponding boundary conditions Eq.(6.2). By subtracting the steady state models Eq.(6.8a)-(6.8c), the original system Eq.(6.1a)-(6.1c) can be linearized around steady states via Taylor series expansion.*

By using the change of variables, new states are introduced as follows:

$$\bar{p}(\zeta,t) = p(\zeta,t) - p_{ss}(\zeta,t) \tag{6.9a}$$

$$\bar{v}(\zeta,t) = v(\zeta,t) - v_{ss}(\zeta,t) \tag{6.9b}$$

$$\bar{\rho}(\zeta,t) = \rho(\zeta,t) - \rho_{ss}(\zeta,t) \tag{6.9c}$$

with new boundary conditions given as:

$$\bar{v}(0,t) = 0, \bar{p}(L_d,t) = 0, \bar{\rho}(L_d,t) = 0 \tag{6.10}$$

Substituting new state variables shown in Eq.(6.9a)-(6.9c) and steady states Eq.(6.8a)-(6.8c) into Eq.(6.6a)-(6.6c) yields two models since v can be positive or negative. In the following derivation, we mainly consider the case of velocity with positive sign. In this case, the nonlinear distributed parameter system in Eq.(6.6a)-(6.6c) can be linearized as follows:

$$\frac{\partial \bar{p}}{\partial t} + v_{ss}\frac{\partial \bar{p}}{\partial \zeta} + \bar{v}\frac{\partial p_{ss}}{\partial \zeta} + a^2\rho_{ss}\frac{\partial \bar{v}}{\partial \zeta} + \bar{p}\frac{\partial v_{ss}}{\partial \zeta} = 0 \tag{6.11a}$$

$$\frac{\partial \bar{v}}{\partial t} + \left(\frac{\bar{p}v_{ss}}{a^2\rho_{ss}} + \bar{v}\right)\frac{\partial v_{ss}}{\partial \zeta} + v_{ss}\frac{\partial \bar{v}}{\partial \zeta} + \frac{1}{\rho_{ss}}\frac{\partial \bar{p}}{\partial \zeta} + \left(\frac{g\sin\alpha}{a^2\rho_{ss}} + \frac{\lambda v_{ss}^2}{2Da^2\rho_{ss}}\right)\bar{p} + \frac{\lambda \bar{v} v_{ss}}{D} = 0 \tag{6.11b}$$

with the same boundary conditions shown in Eq.(6.10). Given that $\bar{p} = \bar{\rho}a^2$, one can further reduce the system to a two-state system (pressure and velocity) as shown in Eq.(6.11a)-(6.11b), which can also be linearized directly from the aforementioned system Eq.(6.7a)-(6.7b). For the sake of simplicity, the linearized infinite-dimensional model is given in the following matrix form:

$$\frac{\partial}{\partial t}\begin{bmatrix} \bar{p}(\zeta,t) \\ \bar{v}(\zeta,t) \end{bmatrix} + \begin{bmatrix} L & N \\ G & L \end{bmatrix}\frac{\partial}{\partial \zeta}\begin{bmatrix} \bar{p}(\zeta,t) \\ \bar{v}(\zeta,t) \end{bmatrix} + \begin{bmatrix} T & R \\ Q & F \end{bmatrix}\begin{bmatrix} \bar{p}(\zeta,t) \\ \bar{v}(\zeta,t) \end{bmatrix} = 0 \tag{6.12}$$

where all notations are defined as follows:

$$L = v_{ss}, \quad R = \frac{\partial p_{ss}}{\partial \zeta}, \quad N = a^2\rho_{ss}, \quad T = \frac{\partial v_{ss}}{\partial \zeta}, \quad G = \frac{1}{\rho_{ss}} \tag{6.13a}$$

$$Q = \frac{v_{ss}}{a^2\rho_{ss}}\frac{\partial v_{ss}}{\partial \zeta} + \frac{g\sin\alpha}{a^2\rho_{ss}} + \frac{\lambda v_{ss}^2}{2Da^2\rho_{ss}}, \quad F = \frac{\partial v_{ss}}{\partial \zeta} + \frac{\lambda v_{ss}}{D} \tag{6.13b}$$

From the steady state profiles of pressure and velocity in Fig. 6.1, the spatial derivative of velocity steady state $\frac{\partial v_{ss}}{\partial \zeta}$ can be neglected due to its near-zero value, which implies $T = 0$. Considering the tremendous contribution from the coefficient $N = a^2\rho_{ss}$, relatively smaller values of parameters v_{ss} and $\frac{\partial p_{ss}}{\partial \zeta}$ can be ignored consequently, which infers $L = 0$ and $R = 0$ such that Q and F can be simplified as:

$$Q = \frac{g\sin\alpha}{a^2\rho_{ss}} + \frac{\lambda v_{ss}^2}{2Da^2\rho_{ss}}, F = \frac{\lambda v_{ss}}{D} \tag{6.14}$$

Then, the distributed parameter pipeline system Eq.(6.12) can be rewritten in a simpler form as follow:

$$\frac{\partial}{\partial t}\begin{bmatrix} \bar{p}(\zeta,t) \\ \bar{v}(\zeta,t) \end{bmatrix} + \begin{bmatrix} 0 & N \\ G & 0 \end{bmatrix}\frac{\partial}{\partial \zeta}\begin{bmatrix} \bar{p}(\zeta,t) \\ \bar{v}(\zeta,t) \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ Q & F \end{bmatrix}\begin{bmatrix} \bar{p}(\zeta,t) \\ \bar{v}(\zeta,t) \end{bmatrix} = 0 \tag{6.15}$$

As shown in Eq.(6.15), simplified first-order coupled hyperbolic partial differential equations (PDEs) are obtained to represent the continuous-time hydraulic system in terms of pressure and velocity as process states. Then, one can directly apply Laplace transformation to Eq.(6.15), and rewrite it in the following form:

$$\begin{bmatrix} 0 & N \\ G & 0 \end{bmatrix}\frac{\partial}{\partial \zeta}\begin{bmatrix} \bar{p}(\zeta,s) \\ \bar{v}(\zeta,s) \end{bmatrix} = -\begin{bmatrix} s & 0 \\ Q & F+s \end{bmatrix}\begin{bmatrix} \bar{p}(\zeta,s) \\ \bar{v}(\zeta,s) \end{bmatrix} + \begin{bmatrix} \bar{p}(x,0) \\ \bar{v}(x,0) \end{bmatrix} \tag{6.16}$$

In order to simplify the derivation, one can make the following notations:

$$A_0 = \begin{bmatrix} 0 & N \\ G & 0 \end{bmatrix}, B_0 = \begin{bmatrix} s & 0 \\ Q & F+s \end{bmatrix}, M = -{A_0}^{-1}B_0 \tag{6.17a}$$

$$Q = \frac{g\sin\alpha}{a^2\rho_{ss}} + \frac{\lambda v_{ss}^2}{2Da^2\rho_{ss}}, F = \frac{\lambda v_{ss}}{D} \tag{6.17b}$$

Through further calculation, a frequency-domain solution of the distributed parameter pipeline system is finally generated as follows:

$$\begin{bmatrix} \bar{p}(\zeta,s) \\ \bar{v}(\zeta,s) \end{bmatrix} = e^{M\zeta}\begin{bmatrix} \bar{p}(0,s) \\ \bar{v}(0,s) \end{bmatrix} + \int_0^{\zeta} e^{M(\zeta-\eta)}{A_0}^{-1}\begin{bmatrix} \bar{p}(\eta,0) \\ \bar{v}(\eta,0) \end{bmatrix} d\eta \tag{6.18}$$

By checking the nonzero elements off the diagonal of matrix A_0, it is easy to prove that A_0 is invertible, which guarantees that the solution shown in Eq.(6.18) is well posed.

Remark 13. *As for negative velocity in Eq.(6.6a)-(6.6c), the same linearization procedure can be adopted and the analytical solution in Laplace domain is identical as Eq.(6.18) with the same notations as in Eq.(6.17a) but different expressions of* Q *and* F *described as follows:*

$$Q = \frac{g\sin\alpha}{a^2\rho_{ss}} - \frac{\lambda v_{ss}^2}{2Da^2\rho_{ss}}, F = -\frac{\lambda v_{ss}}{D} \tag{6.19}$$

6.2.3 Resolvent operator

In order to utilize the Crank-Nicolson discretization framework, the resolvent operator need to be determined beforehand. In general, for any given continuous-time PDE system, there is a link between its analytical solution in Laplace-domain and the associated resolvent operator. For example, for a given first-order hyperbolic PDE, $\dot{x}(t,\zeta) = \mathcal{A}x(t,\zeta)$ with $\mathcal{A} := \frac{\partial}{\partial x}$, one can obtain the resolvent operator by applying the Laplace transform as $\mathfrak{R}(s,\mathcal{A})(\cdot) = [sI - \mathcal{A}]^{-1}(\cdot)$ being related to $\mathcal{A}$. Then, the frequency-domain solution can be further expressed by the resolvent operator as:

$$x(s,\zeta) = \mathfrak{R}(s,\mathcal{A})x(0,\zeta) \tag{6.20}$$

Along this line, one can determine the resolvent operator for the pipeline system according to the aforementioned frequency-domain solution in Eq.(6.18). For simplicity, one can then define these following notations in order to determine the resolvent operator:

$$e^{M\zeta} = \begin{bmatrix} M_{11}(\zeta,s) & M_{12}(\zeta,s) \\ M_{21}(\zeta,s) & M_{22}(\zeta,s) \end{bmatrix} \tag{6.21a}$$

$$A_0{}^{-1} = H(\zeta,s) = \begin{bmatrix} 0 & \frac{1}{G(\zeta,s)} \\ \frac{1}{N(\zeta,s)} & 0 \end{bmatrix} = \begin{bmatrix} 0 & H_{12}(\zeta,s) \\ H_{21}(\zeta,s) & 0 \end{bmatrix} \tag{6.21b}$$

Through further manipulation, closed-form analytical solutions of the state evolutional matrix can be shown as follows:

$$\begin{cases} M_{11}(\zeta,s) = [\cosh(Y\zeta) - \frac{Z}{Y}\sinh(Y\zeta)][\cosh(Z\zeta) - \sinh(Z\zeta)] \\ M_{12}(\zeta,s) = \frac{(F+s)N}{X}[\sinh((Z+Y)\zeta) - \cosh((Z+Y)\zeta)][\cosh(2Y\zeta) + \sinh(2Y\zeta) - 1] \\ M_{21}(\zeta,s) = \frac{2Gs}{X}\sinh(Y\zeta)[\sinh(Z\zeta) - \cosh(Z\zeta)] \\ M_{22}(\zeta,s) = [\cosh(Y\zeta) + \frac{Z}{Y}\sinh(Y\zeta)][\cosh(Z\zeta) - \sinh(Z\zeta)] \end{cases} \tag{6.22}$$

where the notations X, Y, Z are defined as follows:

$$X = \sqrt{N(NQ^2 + 4Gs^2 + 4FGs)}, \quad Y = \frac{\sqrt{N(NQ^2 + 4Gs^2 + 4FGs)}}{2GN}, \quad Z = \frac{Q}{2G} \tag{6.23}$$

Finally, the distributed parameter pipeline system can be manipulated in the following compact form:

$$\begin{array}{l} \left[\begin{array}{c} \bar{p}(\zeta, s) \\ \bar{\nu}(\zeta, s) \end{array} \right] = \left[\begin{array}{cc} M_{11}(\zeta, s) & M_{12}(\zeta, s) \\ M_{21}(\zeta, s) & M_{22}(\zeta, s) \end{array} \right] \left[\begin{array}{c} \bar{p}(0, s) \\ \bar{\nu}(0, s) \end{array} \right] + \\ \int_0^{\zeta} \left[\begin{array}{cc} M_{11}(\zeta - \eta, s) & M_{12}(\zeta - \eta, s) \\ M_{21}(\zeta - \eta, s) & M_{22}(\zeta - \eta, s) \end{array} \right] \left[\begin{array}{cc} 0 & H_{12}(\eta, s) \\ H_{21}(\eta, s) & 0 \end{array} \right] \left[\begin{array}{c} \bar{p}(\eta, 0) \\ \bar{\nu}(\eta, 0) \end{array} \right] d\eta \end{array} \tag{6.24}$$

In order to complete this infinite-dimensional linearized model, one needs to facilitate boundary conditions in Eq.(6.10) as follows:

(a) At $\zeta = 0$, one can plug $\bar{\nu}(0, s) = 0$ into Eq.(6.24) which leads to:

$$\left[\begin{array}{c} \bar{p}(0, s) \\ \bar{\nu}(0, s) \end{array} \right] = \left[\begin{array}{cc} M_{11}(0, s) & M_{12}(0, s) \\ M_{21}(0, s) & M_{22}(0, s) \end{array} \right] \left[\begin{array}{c} \bar{p}(0, s) \\ \bar{\nu}(0, s) \end{array} \right] \tag{6.25}$$

From the above Eq.(6.25), it is easy to show $M_{21}(0, s) = 0$ which can be derived via Eq.(6.22).

(b) At $\zeta = L_d$, one can substitute $\bar{p}(L_d, s) = 0$ into Eq.(6.24) which yields:

$$\begin{array}{l} \left[\begin{array}{c} \bar{p}(L_d, s) \\ \bar{\nu}(L_d, s) \end{array} \right] = \left[\begin{array}{cc} M_{11}(L_d, s) & M_{12}(L_d, s) \\ M_{21}(L_d, s) & M_{22}(L_d, s) \end{array} \right] \left[\begin{array}{c} \bar{p}(0, s) \\ \bar{\nu}(0, s) \end{array} \right] + \\ \int_0^{L_d} \left[\begin{array}{cc} M_{11}(L_d - \eta, s) & M_{12}(L_d - \eta, s) \\ M_{21}(L_d - \eta, s) & M_{22}(L_d - \eta, s) \end{array} \right] \left[\begin{array}{cc} 0 & H_{12}(\eta, s) \\ H_{21}(\eta, s) & 0 \end{array} \right] \left[\begin{array}{c} \bar{p}(\eta, 0) \\ \bar{\nu}(\eta, 0) \end{array} \right] d\eta \end{array} \tag{6.26}$$

From this, one can solve for:

$$\begin{array}{c} \bar{p}(0, s) = -\frac{1}{M_{11}(L_d, s)} \int_0^{L_d} \left[M_{12}(L_d - \eta, s) H_{21}(\eta, s) \bar{p}(\eta, 0) \right. \\ \left. + M_{11}(L_d - \eta, s) H_{12}(\eta, s) \bar{\nu}(\eta, 0) \right] d\eta \end{array} \tag{6.27}$$

Boundary conditions of this distributed parameter pipeline system are bi-directional, as shown in Eq.(6.10), so one needs to convert them to be both at $\zeta = 0$ utilizing the system Eq.(6.24) and Eq.(6.27) in order to determine the resolvent operator in the following form:

$$\left[\begin{array}{c} \bar{p}(\zeta, s) \\ \bar{\nu}(\zeta, s) \end{array} \right] = \left[\begin{array}{cc} \mathfrak{R}_{11}(s, M) & \mathfrak{R}_{12}(s, M) \\ \mathfrak{R}_{21}(s, M) & \mathfrak{R}_{22}(s, M) \end{array} \right] \left[\begin{array}{c} \bar{p}(\eta, 0) \\ \bar{\nu}(\eta, 0) \end{array} \right] \tag{6.28}$$

Following this, the frequency-domain solution of the infinite-dimensional pipeline system can be expressed as Eq.(6.28) with the associated resolvent operator as:

$$\begin{cases} \mathfrak{R}_{11}(s,M)(\cdot) = -\frac{M_{11}(\zeta,s)}{M_{11}(L_d,s)} \int_0^{L_d} M_{12}(L_d-\eta,s)H_{21}(\eta,s)(\cdot)d\eta \\ \qquad\qquad + \int_0^{\zeta} M_{12}(\zeta-\eta,s)H_{21}(\eta,s)(\cdot)d\eta \\ \mathfrak{R}_{12}(s,M)(\cdot) = -\frac{M_{11}(\zeta,s)}{M_{11}(L_d,s)} \int_0^{L_d} M_{11}(L_d-\eta,s)H_{12}(\eta,s)(\cdot)d\eta \\ \qquad\qquad \int_0^{\zeta} M_{11}(\zeta-\eta,s)H_{12}(\eta,s)(\cdot)d\eta \\ \mathfrak{R}_{21}(s,M)(\cdot) = -\frac{M_{21}(\zeta,s)}{M_{11}(L_d,s)} \int_0^{L_d} M_{12}(L_d-\eta,s)H_{21}(\eta,s)(\cdot)d\eta \\ \qquad\qquad \int_0^{\zeta} M_{22}(\zeta-\eta,s)H_{21}(\eta,s)(\cdot)d\eta \\ \mathfrak{R}_{22}(s,M)(\cdot) = -\frac{M_{21}(\zeta,s)}{M_{11}(L_d,s)} \int_0^{L_d} M_{11}(L_d-\eta,s)H_{12}(\eta,s)(\cdot)d\eta \\ \qquad\qquad \int_0^{\zeta} M_{21}(\zeta-\eta,s)H_{12}(\eta,s)(\cdot)d\eta \end{cases} \tag{6.29}$$

6.2.4 Discrete-time pipeline model realization

Based on the resolvent operator proposed in the previous section, one can get the discrete-time pipeline hydraulic system by evaluating s as a real value $\delta = \frac{2}{h}$ (h is the time increment) in the resolvent operator Eq.(6.28) and Eq.(6.29), which can be directly utilized for digital model simulation. Considering that there are two main drawbacks of the Euler finite difference discretization methods, which may hinder leak detection and localization performance, we utilize the Cayley-Tustin discretization technique to achieve a higher order (second-order) discretization accuracy in time. More specifically, Euler finite difference methods could not avoid spatial discretization, and they only have first-order discretization accuracy in time [119]. On the contrary, by exploiting the Cayley-Tustin discretization framework, one can map the original continuous distributed pipeline system into its corresponding discrete version without spatial discretization or model order reduction, which ensures realization of accurate leak localization.

In particular, the continuous state evolutional operator $\mathcal{A}$ can be utilized to generate its associated discrete evolutional operator $\mathcal{A}_d$ using $\mathcal{A}_d = [\delta - \mathcal{A}]^{-1}[\delta + \mathcal{A}] = -I + 2\delta[\delta I - \mathcal{A}]^{-1} = -I + 2\delta\mathfrak{R}(\delta, \mathcal{A})$ where I is an identity operator [62]. From the formulation of the resolvent operator shown in Eq.(6.28) and Eq.(6.29), one can obtain

a discrete-time system realization as follows:

$$\begin{aligned}\begin{bmatrix} \bar{p}(\zeta,kh) \\ \bar{v}(\zeta,kh) \end{bmatrix} &= \mathcal{A}_d \begin{bmatrix} \bar{p}(\zeta,(k-1)h) \\ \bar{v}(\zeta,(k-1)h) \end{bmatrix} \\ &= (-I + 2\delta\mathfrak{R}(\delta,M)) \begin{bmatrix} \bar{p}(\zeta,(k-1)h) \\ \bar{v}(\zeta,(k-1)h) \end{bmatrix} \\ &= \left(-I + 2\delta \begin{bmatrix} \mathfrak{R}_{11}(\delta,M) & \mathfrak{R}_{12}(\delta,M) \\ \mathfrak{R}_{21}(\delta,M) & \mathfrak{R}_{22}(\delta,M) \end{bmatrix}\right) \begin{bmatrix} \bar{p}(\zeta,(k-1)h) \\ \bar{v}(\zeta,(k-1)h) \end{bmatrix}\end{aligned} \tag{6.30}$$

Combining Eq.(6.9a)-(6.9c), a discrete-time transient distributed parameter pipeline system is formulated as follows:

$$\begin{bmatrix} p(\zeta,kh) \\ v(\zeta,kh) \end{bmatrix} = \begin{bmatrix} \bar{p}(\zeta,kh) \\ \bar{v}(\zeta,kh) \end{bmatrix} + \begin{bmatrix} p_{ss}(\zeta) \\ v_{ss}(\zeta) \end{bmatrix} \tag{6.31}$$

where k stands for time instance and one can easily implement this framework into digital device settings such as computers or SCADA systems.

6.3 Discrete observer design

In general, one cannot realize leak detection and localization by spatially distributed sensors along the entire long-range pipeline system either due to physical limitations of installation and/or prohibitive expenses if possible to be realized. In these cases, a model-based observer design serves a purpose of reducing the target number of external measurements to be taken. In this section, a discrete Luenberger observer for this linearized coupled water hammer equation under discrete-time infinite-dimensional setting is designed. In order to simplify the derivation process, one introduces the following change of variables:

$$\begin{cases} \bar{x}(\zeta,t) = \frac{\bar{p}(\zeta,t) - \sqrt{\frac{N}{G}}\bar{v}(\zeta,t)}{2} \\ \bar{y}(\zeta,t) = \frac{\bar{p}(\zeta,t) + \sqrt{\frac{N}{G}}\bar{v}(\zeta,t)}{2} \end{cases} \tag{6.32}$$

Then, plugging the Eq.(6.32) into the pipeline system Eq.(6.15) yields:

$$\begin{aligned}\tfrac{\partial}{\partial t}\begin{bmatrix} \bar{x}(\zeta,t) \\ \bar{y}(\zeta,t) \end{bmatrix} + \begin{bmatrix} -\sqrt{NG} & 0 \\ 0 & \sqrt{NG} \end{bmatrix} \tfrac{\partial}{\partial \zeta}\begin{bmatrix} \bar{x}(\zeta,t) \\ \bar{y}(\zeta,t) \end{bmatrix} \\ + \begin{bmatrix} \frac{-\sqrt{(\frac{N}{G})}Q+F}{2} & -\frac{\sqrt{(\frac{N}{G})}Q+F}{2} \\ \frac{\sqrt{(\frac{N}{G})}Q-F}{2} & \frac{\sqrt{(\frac{N}{G})}Q+F}{2} \end{bmatrix} \begin{bmatrix} \bar{x}(\zeta,t) \\ \bar{y}(\zeta,t) \end{bmatrix} = 0\end{aligned} \tag{6.33}$$

For simplicity, one can define an expanded state $x^e(\zeta,t) = [\bar{x}(\zeta,t), \bar{y}(\zeta,t)]'$, and make the following notations:

$$\kappa = \begin{bmatrix} -\sqrt{NG} & 0 \\ 0 & \sqrt{NG} \end{bmatrix}, \quad \varphi = \begin{bmatrix} \frac{-\sqrt{(\frac{N}{G})}Q+F}{2} & -\frac{\sqrt{(\frac{N}{G})}Q+F}{2} \\ \frac{\sqrt{(\frac{N}{G})}Q-F}{2} & \frac{\sqrt{(\frac{N}{G})}Q+F}{2} \end{bmatrix} = \begin{bmatrix} -\Delta & -\Pi \\ \Delta & \Pi \end{bmatrix} \tag{6.34}$$

Then, the model Eq.(6.33) can be rewritten as:

$$\frac{\partial x^e(\zeta,t)}{\partial t} = \mathcal{A}_1 x^e(\zeta,t) = -\kappa \frac{\partial x^e(\zeta,t)}{\partial \zeta} - \varphi x^e(\zeta,t) \tag{6.35}$$

where $\mathcal{A}_1(\cdot) = -\kappa \frac{\partial}{\partial \zeta}(\cdot) - \varphi(\cdot)$ with the domain as $\mathcal{D}(\mathcal{A}_1) = \{x^e = [\phi_1(\zeta), \phi_2(\zeta)]' \in L_2(0, L_d), \phi_1(L_d) + \phi_2(L_d) = 0, \phi_2(0) - \phi_1(0) = 0,\ \phi_1(\zeta)$ and $\phi_2(\zeta)$ are absolutely continuous$\}$. Then, the distributed parameter pipeline system is completed with an output defined as:

$$\begin{cases} \dot{x}^e(t) = \mathcal{A}_1 x^e(t) \\ y^e(t) = C_1 x^e(t) \end{cases} \tag{6.36}$$

where the output operator C_1 is defined as follows:

$$C_1 := \begin{bmatrix} \int_0^{L_d} \delta(\zeta - 0)(\cdot) d\zeta & 0 \\ 0 & \int_0^{L_d} \delta(\zeta - L_d)(\cdot) d\zeta \end{bmatrix} \tag{6.37}$$

where $\delta(\zeta)$ is the Dirac function.

6.3.1 Stabilization observer gain

In order to monitor the distributed pipeline system with limited boundary measurements, a Luenberger observer is formed directly by applying infinite-dimensional system theory [4] as follows:

$$\begin{cases} \dot{\hat{x}}^e(t) = \mathcal{A}_1 \hat{x}^e(t) + \mathcal{L}_1 [y^e(t) - \hat{y}^e(t)] \\ \hat{y}^e(t) = C_1 \hat{x}^e(t) \end{cases} \tag{6.38}$$

By introducing an observed error variable $\tilde{x}^e(t) = x^e(t) - \hat{x}^e(t)$, the observer error dynamic is then obtained by combining the original system Eq.(6.36) and the observer system Eq.(6.38) as follows:

$$\dot{\tilde{x}}^e(t) = \dot{x}^e(t) - \dot{\hat{x}}^e(t) = (\mathcal{A}_1 - \mathcal{L}_1 C_1)\tilde{x}^e(t) \tag{6.39}$$

Theorem 12. *An observer system ($\mathcal{A}_1$, $\mathcal{C}_1$) is exponentially detectable with a distributed parameter setting, if there exists a nonnegative self-adjoint operator Q_0 guaranteeing that the following operator Riccati equation holds.*

$$\mathcal{A}_1 Q_0 + Q_0 \mathcal{A}_1^* + M - 2Q_0 \mathcal{C}_1^* \mathcal{C}_1 Q_0 = 0, \quad on \quad \mathcal{D}(\mathcal{A}_1^*) \tag{6.40}$$

where M is a positive definite matrix which serves as a design parameter, such that $Q_0(\mathcal{D}(\mathcal{A}_1^*)) \subset \mathcal{D}(\mathcal{A}_1)$, and $\langle \phi_1, Q_0 \phi_2 \rangle = \langle Q_0 \phi_1, \phi_2 \rangle$, and the observer gain is given by $\mathcal{L}_1 = Q_0 \mathcal{C}_1^*$. Then, the stability of the observer error system is ensured, such that $(\mathcal{A}_1 - \mathcal{L}_1 \mathcal{C}_1)$ generates an exponentially stable C_0-semigroup.

A similar theorem for the control problem is introduced in reference [102]. To solve the operator Riccati equation, adjoint operators of $\mathcal{A}_1$ and $\mathcal{C}_1$ need to be determined beforehand. For the system, it is easy to find the adjoint operator $\mathcal{A}_1^*$ using condition of $\langle \mathcal{A}_1 \phi_1, \phi_2 \rangle = \langle \phi_1, \mathcal{A}_1^* \phi_2 \rangle$, as follows:

$$\mathcal{A}_1^*(\cdot) = \kappa \tfrac{\partial}{\partial \zeta}(\cdot) - \varphi^T(\cdot) \tag{6.41}$$

with the domain defined as $\mathcal{D}(\mathcal{A}_1^*) = \{x = [\phi_1(\zeta), \phi_2(\zeta)]' \in L_2(0, L_d), \phi_1(L_d) + \phi_2(L_d) = 0, \phi_1(0) - \phi_2(0) = 0,\ \phi_1(\zeta)$ and $\phi_2(\zeta)$ are absolutely continuous$\}$.

Similarly, $\mathcal{C}_1^*$ is determined as follows, by applying $\langle \mathcal{C}_1 \phi_1, \phi_2 \rangle = \langle \phi_1, \mathcal{C}_1^* \phi_2 \rangle$:

$$\mathcal{C}_1^* := \begin{bmatrix} \delta(\zeta - 0) \int_0^{L_d} (\cdot) d\eta & 0 \\ 0 & \delta(\zeta - L_d) \int_0^{L_d} (\cdot) d\eta \end{bmatrix} \tag{6.42}$$

Remark 14. *Different spatial functions of M can lead to different Q_0, so we can define $M = [M_{11}(\zeta), M_{12}(\zeta); M_{12}(\zeta), M_{22}(\zeta)]$ without loss of generality. Since Q_0 is a symmetrical operator, one can define it as a diagonal form for simplicity.*

One is free to choose arbitrary functions $[\phi_1(\zeta), \phi_2(\zeta)]'$ in $\mathcal{D}(\mathcal{A}_1^*)$ for determination of operator Riccati equation (6.40) as:

$$\begin{aligned} &\begin{bmatrix} \sqrt{NG}\frac{\partial}{\partial \zeta} + \Delta & \Pi \\ -\Delta & -\sqrt{NG}\frac{\partial}{\partial \zeta} - \Pi \end{bmatrix} \begin{bmatrix} Q_{11}(\zeta) & 0 \\ 0 & Q_{22}(\zeta) \end{bmatrix} \begin{bmatrix} \phi_1(\zeta) \\ \phi_2(\zeta) \end{bmatrix} \\ &\quad + \begin{bmatrix} Q_{11}(\zeta) & 0 \\ 0 & Q_{22}(\zeta) \end{bmatrix} \begin{bmatrix} -\sqrt{NG}\frac{\partial}{\partial \zeta} + \Delta & -\Delta \\ \Pi & \sqrt{NG}\frac{\partial}{\partial \zeta} - \Pi \end{bmatrix} \begin{bmatrix} \phi_1(\zeta) \\ \phi_2(\zeta) \end{bmatrix} \\ &\qquad + \begin{bmatrix} M_{11}(\zeta) & M_{12}(\zeta) \\ M_{12}(\zeta) & M_{22}(\zeta) \end{bmatrix} \begin{bmatrix} \phi_1(\zeta) \\ \phi_2(\zeta) \end{bmatrix} + 2Q_0 \mathcal{C}_1^* \mathcal{C}_1 Q_0 \begin{bmatrix} \phi_1(\zeta) \\ \phi_2(\zeta) \end{bmatrix} = 0 \end{aligned} \tag{6.43}$$

where

$$Q_0 C_1^* C_1 Q_0 \begin{bmatrix} \phi_1(\zeta) \\ \phi_2(\zeta) \end{bmatrix} = \begin{bmatrix} L_d {Q_{11}}^2(0) & 0 \\ 0 & L_d {Q_{22}}^2(L_d) \end{bmatrix} \begin{bmatrix} \phi_1(\zeta) \\ \phi_2(\zeta) \end{bmatrix} \tag{6.44}$$

Given by the special structures of $\mathcal{A}_1$ and $\mathcal{A}_1^*$, $\frac{\partial \phi_1}{\partial \zeta}$ and $\frac{\partial \phi_2}{\partial \zeta}$ can be naturally eliminated such that one can easily simplify the above operator Riccati equation as follows:

$$\begin{cases} \sqrt{NG}\frac{dQ_{11}(\zeta)}{d\zeta} + 2Q_{11}\Delta + M_{11}(\zeta) + 2L_d Q_{11}^2(0) = 0 \\ \Pi Q_{22}(\zeta) - \Delta Q_{11}(\zeta) + M_{12}(\zeta) = 0 \\ -\sqrt{NG}\frac{dQ_{22}(\zeta)}{d\zeta} - 2Q_{22}\Pi + M_{22}(\zeta) + 2L_d Q_{22}^2(L_d) = 0 \end{cases} \tag{6.45}$$

Remark 15. *To determine $Q_{11}(\zeta)$, $Q_{22}(\zeta)$, $M_{11}(\zeta)$, $M_{12}(\zeta)$ and $M_{22}(\zeta)$, the necessary condition to ensure is the positive definiteness. In addition, another condition shown in Theorem 12 is to guarantee that $Q_0(\mathcal{D}(\mathcal{A}_1^*)) \subset \mathcal{D}(\mathcal{A}_1)$ as follows:*

$$Q_{11}(L_d) + Q_{22}(L_d) = 0, \quad Q_{11}(0) - Q_{22}(0) = 0 \tag{6.46}$$

For simplicity, one can further fix $Q_{11}(L_d) = 0$ and $Q_{22}(L_d) = 0$, and suppose $M_{11}(\zeta) + 2L_d Q_{11}^2(0)$ and $M_{22}(\zeta) + 2L_d Q_{22}^2(L_d)$ are just known spatial functions. Then, one gets the profiles of observer gain Q_0 and design parameter M by using the finite difference method, as shown in the Figs. 6.2-6.3.

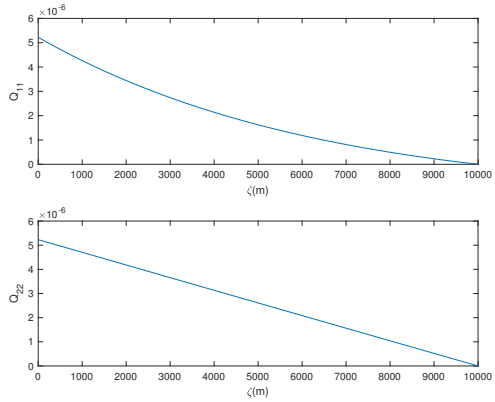


Figure 6.2: Plots of observer gain Q_{11} and Q_{22} in Eq.(6.45) with boundary conditions Eq.(6.46).

By applying the similarity transformation, one can easily find the observer gain corresponding to the original states $[\bar{p}(\zeta,t), \bar{v}(\zeta,t)]'$ as follows:

$$T_s^{-1}\mathcal{A}T_s Q_0 + Q_0(T_s^{-1}\mathcal{A}T_s)^* + M - 2Q_0(CT_s)^* CT_s Q_0 = 0, \text{ on } \mathcal{D}(\mathcal{A}^*) \tag{6.47}$$

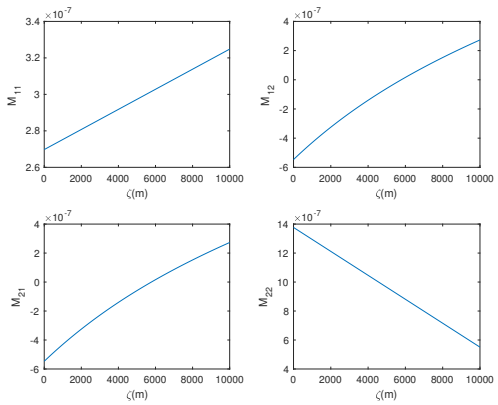


Figure 6.3: Plots of design parameter M matrix in Eq.(6.45) with boundary conditions Eq.(6.46).

where $\begin{bmatrix} \bar{p}(\zeta,t) \\ \bar{v}(\zeta,t) \end{bmatrix} = T_s \begin{bmatrix} \bar{x}(\zeta,t) \\ \bar{y}(\zeta,t) \end{bmatrix}$ with $T_s = \begin{bmatrix} 1 & 1 \\ -\sqrt{\frac{G}{N}} & \sqrt{\frac{G}{N}} \end{bmatrix}$.

By combining Eq.(6.15) and Eqs.(6.32)-(6.33), one can obtain the observer gain $\tilde{Q}_0 = T_s Q_0 T_s^T$ corresponding to the original states. Finally, the observer error dynamics is shown as:

$$e(t) = \begin{bmatrix} \bar{p}(\zeta,t) \\ \bar{v}(\zeta,t) \end{bmatrix} - \begin{bmatrix} \bar{p}_0(\zeta,t) \\ \bar{v}_0(\zeta,t) \end{bmatrix} \tag{6.48a}$$

$$\dot{e}(t) = (\mathcal{A} - \mathcal{L}C)e(t) = (\mathcal{A} - \tilde{Q}_0 C^* C)e(t) = (\mathcal{A} - T_s Q_0 T_s^T C^* C)e(t) \tag{6.48b}$$

Remark 16. *The continuous observer gain Q_0 is identical to its associated discrete observer gain Q_d, which yields the corresponding discrete observer error system as follows:*

$$e(\zeta,kh) = \begin{bmatrix} \bar{p}(\zeta,kh) \\ \bar{v}(\zeta,kh) \end{bmatrix} - \begin{bmatrix} \bar{p}_0(\zeta,kh) \\ \bar{v}_0(\zeta,kh) \end{bmatrix} \tag{6.49a}$$

$$\begin{aligned} e(\zeta,kh) &= (\mathcal{A}_d - \mathcal{L}_d C_d)e(\zeta,(k-1)h) = (\mathcal{A}_d - \tilde{Q}_0 C_d^* C_d)e(\zeta,(k-1)h) \\ &= (\mathcal{A}_d - T_s Q_0 T_s^T C_d^* C_d)e(\zeta,(k-1)h) \end{aligned} \tag{6.49b}$$

where $\bar{p}_0(\zeta,kh)$ and $\bar{v}_0(\zeta,kh)$ correspond to the discrete linearized states. From the Crank-Nicolson discretization setting, the resolvent operator can be utilized as a link to connect the discrete-time output operator C_d and continuous one C, as follows [62]:

$$C_d(\cdot) = \sqrt{2\delta} C[\delta - \mathcal{A}]^{-1}(\cdot) = \sqrt{2\delta} C \mathfrak{R}(\delta, M)(\cdot) \tag{6.50}$$

Until now, a discrete-time Luenberger observer for the distributed parameter pipeline system is completely constructed and can be directly implemented for the

pipeline condition monitoring, leak detection and diagnosis. With limited measurements of pressure at upstream and mass flow velocity at downstream, the entire spatiotemporal states evolution can be reconstructed by the well-designed discrete Luenberger observer. Along this line, velocities at upstream and downstream are collected for feature selection, and are then used for support vector machine training and testing later on.

6.4 Leak simulation and detection

In this section, the dynamical evolution of this derived distributed parameter pipeline system is analyzed around the steady state profiles primarily based on a set of physical parameters [242] illustrated in Table 6.1. Subsequently, normal working conditions and different single individual leakage scenarios, including different leak magnitudes and positions are generated by varying initial conditions and boundary conditions of pressure and velocity at upstream and downstream. Based on these datasets, a one-class support vector machine (OCSVM) is used for leak detection and a support vector regression (SVR) model [243] is trained and exploited for leak amount and position estimation.

6.4.1 Simulation of normal conditions

In this section, a long range pipeline system under normal working conditions is simulated based on the proposed discretized-time infinite-dimensional system. In particular, different normal working conditions are generated by varying the model initial conditions. Considering the limited measurements of pressure at upstream and velocity at downstream, the presented Luenberger observer is further utilized to reconstruct the entire state evolution profiles along with the corresponding error evolution profile. In particular, one adopts pipeline parameters as shown in Table 6.1. Then, simulation parameters are set as spatial interval $\Delta\zeta = 500\ m$, temporal interval $h = 2$ seconds, based on which, the profiles of linearized pressure and velocity and their associated transients are determined and illustrated in the following Figs. 6.4-6.7 over 120 seconds of simulation.

From these profiles, it is apparent that this distributed parameter pipeline system

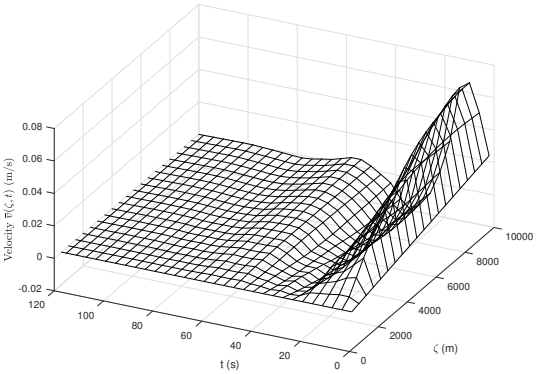


Figure 6.4: Profile of linearized velocity $\bar{v}(\zeta, t)$ evolution given by Eq.(6.15) and numerically realized by Eq.(6.30).

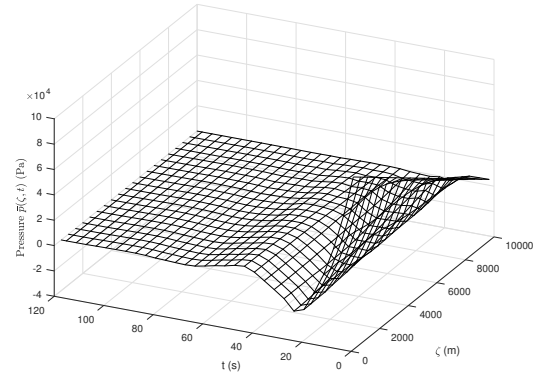


Figure 6.5: Profile of linearized pressure $\bar{p}(\zeta, t)$ evolution given by Eq.(6.15) and numerically realized by Eq.(6.30).

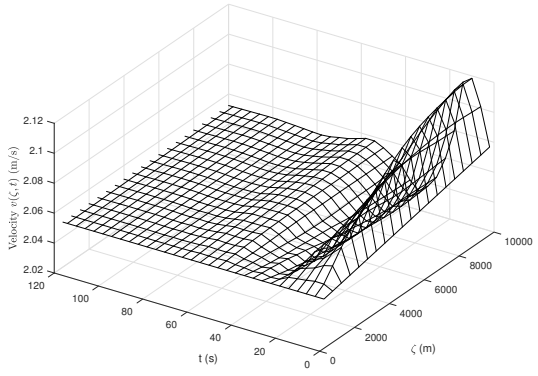


Figure 6.6: Profile of transient velocity $v(\zeta, t)$ evolution numerically realized by Eq.(6.31).

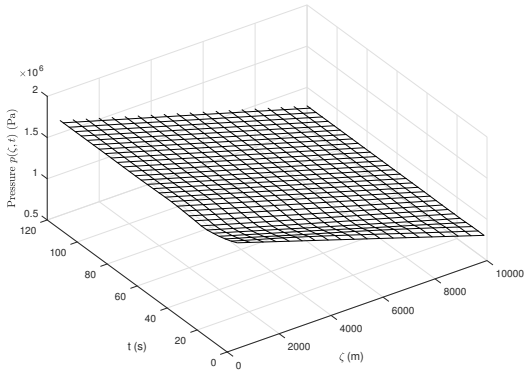


Figure 6.7: Profile of transient pressure $p(\zeta, t)$ evolution numerically realized by Eq.(6.31).

is dynamically stable (small perturbation in the states at a given operating condition will not induce transients to another operating state) and the change of the velocity direction does not affects its stability. Considering the space limitation, the simulation results with negative velocity are not depicted here. The initial conditions of system states and observer error states for this case are shown as follows. Adding the steady state profiles into this linearized profiles leads to the real transient diagrams that can be harnessed for the pipeline condition monitoring and leak detection in the next section. From Figs. 6.8-6.9, one can see that observed errors converge to zero with time increase, which implies that the estimated states are capable of converging to real states even with complex trigonometric initial error conditions Eqs.(6.51a)-(6.51b), which further verifies the effectiveness of the aforementioned observer gain.

$$\begin{cases} \bar{p}_0(\zeta) = -0.1\frac{dp_{ss}}{d\zeta}(L_d - \zeta), \bar{v}_0(\zeta) = 10\frac{dv_{ss}}{d\zeta}\zeta; \quad (v > 0) \\ \bar{p}_0(\zeta) = -0.1\frac{dp_{ss}}{d\zeta}(L_d - \zeta), \bar{v}_0(\zeta) = -10\frac{dv_{ss}}{d\zeta}\zeta; \quad (v \leq 0) \end{cases} \tag{6.51a}$$

$$\begin{cases} \tilde{p}_0(\zeta) = 0.001\frac{dp_{ss}}{d\zeta}L_d \sin(0.002\pi\zeta) \\ \tilde{v}_0(\zeta) = -0.002\sin(0.002\pi\zeta) \end{cases} \tag{6.51b}$$

6.4.2 Simulation of leakage scenarios

In this section, different leakage scenarios considering single individual leak are simulated, including different leak amounts from 1% to 20% of the corresponding local mass flow velocity at the leak point (before a leak occurs) with an increment of 1%

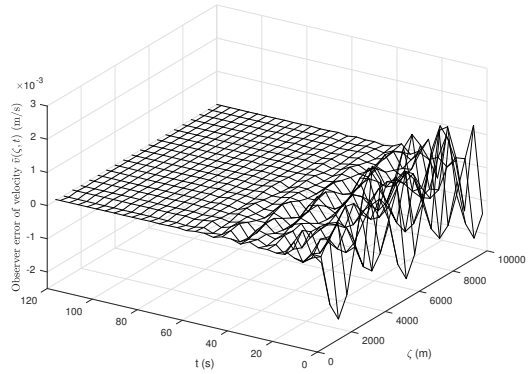


Figure 6.8: Observer error evolution profile of velocity $\tilde{v}(\zeta, t)$ in Eq.(6.49b) with initial conditions Eq.(6.51b).

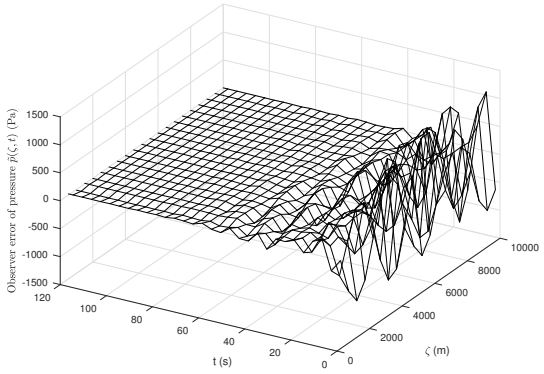


Figure 6.9: Observer error evolution profile of pressure $\tilde{p}(\zeta, t)$ in Eq.(6.49b) with initial conditions Eq.(6.51b).

and different leak positions from 500 m to $9,500$ m with spatial increment of 500 m. In order to emulate internal process dynamics when a leak occurs, one can adjust the mass flow velocity at the upstream and downstream of the pipeline system. More specifically, the distributed pipeline can be separated into two independent systems once a leak occurs. However, the pressure values at upstream and downstream will not change, since they rely on their respective upstream and downstream systems. Take the reservoir-pipeline-reservoir system as an example, the pipeline pressure at both ends will depend on two associated reservoir heights only, regardless the leakage in the pipeline.

In order to simulate the velocity profiles of the pipeline after a leak occurs, the shooting method is utilized for simulation of new steady states. More specifically, the

leak amount is pre-determined as the change of mass flow velocity at the leak position. As shown in Fig. 6.10, pressure variation at the leak point can be determined based on the leak amount (i.e., leak size). Considering the downstream part of the pipeline (i.e., the pipe segment from the leak position to the downstream end of the pipeline), the velocity at the downstream end can be determined using the unchanged pressure at downstream and velocity at the leak point. Similarly, to match the unchanged pressure at the upstream end, one can calculate the velocity of the upstream part of the pipeline (i.e., the pipe segment from the upstream end of the pipeline to the leak position) by using the shooting method. Thus, steady state profiles under different leakage scenarios are generated, and can be added to linearized states to formulate the transient model under different leakage conditions including different leakage amounts and locations.

The simulated leak occurrence time for all cases is at 100 seconds (≈ 1.67 minutes), and the entire simulation time period is 160 seconds (≈ 2.67 minutes). Since this chapter does not focus on the leak propagation process, and also this period is very short (within 10 seconds for this pipeline setting) due to the high propagation speed (1,069 m/s), priority is then given to the discrepancy of the steady state profile once a leak is present. In particular, one leak case with a 2% leakage amount and location at $4,000\ m$ from the left end as the leak position is shown in the Figs. 6.10-6.14.

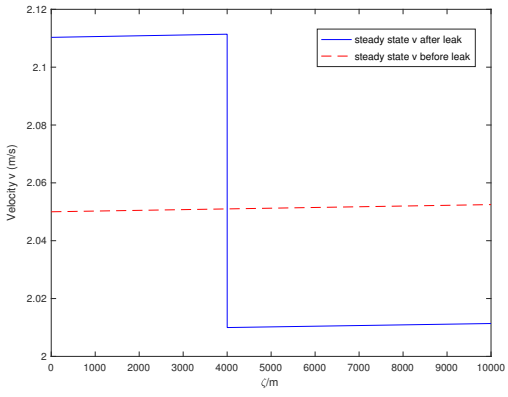


Figure 6.10: Comparison of velocity steady state before and after leakage occurs.

From Figs. 6.10-6.11, one can clearly see the discrepancy of steady states profiles before and after a leak occurs. In particular, a polyline appears instead of a straight line after the existence of leak, and the knee point corresponds to the leak position

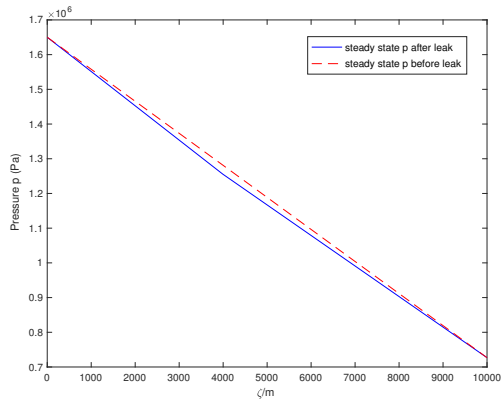


Figure 6.11: Comparison of pressure steady state before and after leakage occurs.

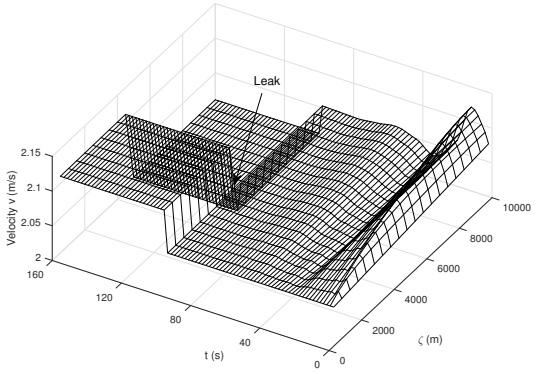


Figure 6.12: 3D profile of velocity when leakage is present.

directly, while the deviation distance from the original red dashed line implies the leak magnitude in Fig. 6.10. In addition, there is an ascending jump at upstream and a descending drop at the downstream in the velocity profile. As for the pressure change, based on the comparison of the steady state lines before and after the leak, we can conclude that the jumping point is located at the leak position (in this case it is at $4,000\ m$ from the left end), and the slope changes of the upstream and downstream profiles can be utilized to calculate leak magnitude.

When compared with the state profiles under normal working condition Figs. 6.6-6.7, there is a corresponding change in three-dimensional profiles of pressure and velocity, as shown in Figs. 6.12-6.13 with the existent leakage. From these reconstructed 3-D profiles calculated by the pre-designed observer, velocity profiles both at upstream and downstream can be collected and illustrated, as shown in Fig. 6.14.

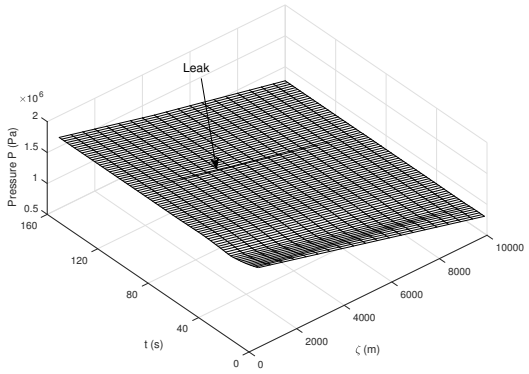


Figure 6.13: 3D profile of pressure when leakage is present.

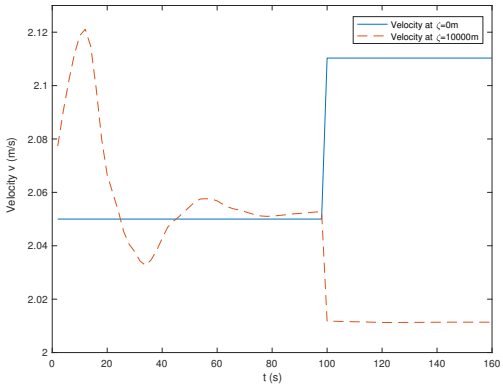


Figure 6.14: Velocity change when leakage is present.

In particular, these changes in velocity profiles are consistent with the above Fig. 6.10. With limited measurements of pressure at upstream and mass flow velocity at downstream, one can ultimately rebuild the entire spatiotemporal profiles of pressure and velocity. In the next section, only measured and estimated velocity profile at both ends will be utilized for leak detection and localization, while one can directly extend this proposed method by combining the pressure profiles.

6.4.3 Leak detection and localization

Stemming from the mass flow velocity data generated in the previous section, a machine learning (ML) model- support vector machine is trained and validated for leak detection, amount estimation and localization.

The support vector machine (SVM) as a classical machine learning method has

been widely utilized for pattern recognition and trend analysis (also known as classification and regression problems) [243]. In particular, the one-class SVM is a well-known method for anomaly detection [244]. When it comes to solving for regression problem, SVM can be referred as support vector regression (SVR). The main idea behind this technique is to maximize the margin between different datasets by constructing hinge loss functions. Compared to other machine learning algorithms, such as neural networks, it has good generalization ability and the sparseness of the solution with the aid of less training samples. In particular, in order to address nonlinear characteristics of datasets, nonlinear kernel functions are introduced to the original formulation of SVM such that a low-dimensional Euclidean space can be mapped into a higher-dimensional Hilbert space where originally difficult-to-classify data can be separated. Thus, it attracts more and more attention from academia and industry, especially when it comes to small training samples.

However, the classification and regression performances of SVM highly depend on the kernel parameters and penalty parameters. To address this, the grid search method is utilized for hyper-parameters tuning and a one-class SVM model and an epsilon support vector regression model are utilized for pipeline leak detection, amount estimation and localization. To evaluate the model performance, classification/regression accuracy, standard deviation (STD) and root mean square error (RMSE) are utilized. In particular, penalty factor C and kernel function parameter γ are searched in the range of $[10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 1, 10^{1}, 10^{2}, 10^{3}, 10^{4}, 10^{5}]$ and the loss function parameter epsilon is optimized from $[10^{-4}, 10^{-3}, 10^{-2}]$. To avoid overfitting, a five-fold cross-validation method is exploited in the training stage and the estimation accuracy evaluation follows a hold-out cross validation procedure. As for SVM realization, the *libsvm* library is utilized in this chapter [245].

Feature selection plays an important role when it comes to training an accurate and robust support vector machine model. In this chapter, velocity profiles at upstream and downstream are taken into account, then representative features are extracted including mean value, maximum value, minimum value of the first part, second part of the pipeline and maximum difference value of the entire pipeline (totally 14 features extracted from two velocity profiles). The entire datasets contain leakage conditions with leakage locations at $[500\ m, 1000\ m, \cdots, 9500\ m]$ and leak

amounts of $[0.01, 0.02, \cdots, 0.2]$. For the sake of leak localization, the features corresponding to leak positions at $[1000\ m, 2000\ m, \cdots, 9000\ m]$ and leak amounts of $[0.06, 0.10, 0.14, 0.18]$ are fed into the SVR model for testing, and the remaining data is used for model training. In order to estimate leak magnitudes, the features associated with leak amounts of [0.03, 0.05, ..., 0.19] and leak positions at $[1000\ m, 3000\ m, \cdots, 9000\ m]$ and are fed into the SVR model for testing and the remaining data serves as training data. It should be emphasized that testing datasets are completely different leakages scenarios which are not used for model training process. More details about the training and testing datasets segregation are shown in Table 6.2.

Table 6.2: Datasets segregation

Datasets		Leak Locations (m)	Leak Amounts
Total Simulation Dataset		[500, 1000, ..., 9500]	[0.01, 0.02, ..., 0.2]
Leak Position Estimation	Training Dataset	—	—
	Testing Dataset	[1000, 2000, ..., 9000]	[0.06, 0.10,0.14,0.18]
Leak Amount Estimation	Training Dataset	—	—
	Testing Dataset	[1000, 3000, ..., 9000]	[0.03, 0.05, ..., 0.19]

Table 6.2: Datasets segregation (continued)

Datasets		Sets	Samples
Total Simulation Dataset		S	380×14
Leak Position Estimation	Training Dataset	TR_{PO} (s.t. $TR_{PO} \cup TE_{PO} = S$; $TR_{PO} \cap TE_{PO} = \emptyset$)	344×14
	Testing Dataset	TE_{PO}	36×14
Leak Amount Estimation	Training Dataset	TR_{AM} (s.t. $TR_{AM} \cup TE_{AM} = S$; $TR_{AM} \cap TE_{AM} = \emptyset$)	335×14
	Testing Dataset	TE_{AM}	45×14

To realize leak detection, the one-class SVM model is utilized here using datasets under normal working conditions generated by the developed observer systems with different initial conditions. In particular, the one-class SVM model with Gaussian radial basis function is trained using 180 data points under normal working situations. In the testing process, 90 data samples (randomly chosen 45 points under leakage conditions and 45 points under normal conditions) are utilized to validate the trained one-class SVM model and run five times. As shown in classification results, the average leak detection accuracy is about 96.00% with small standard deviation and RMSE values of 3.00% and 2.69%, respectively.

As for leak amount estimation and localization, SVR models with consideration of four typical kernel functions, including: Gaussian radial basis function (RBF), linear kernel function, sigmoid kernel function and polynomial kernel function, are investigated as shown in Table 6.3 and Table 6.4. From these tables, it is apparent that SVR with Gaussian radial basis function outperforms other kernel functions in terms of more accurate estimated positions and smaller standard deviation values. Except that, SVR with polynomial kernel function has better leakage localization performance, compared to linear function and sigmoid function. As for leak amount estimation, SVR models with four different kernel functions all have good estimation performance in terms of expectation and standard deviation values. Hence, the RBF kernel is demonstrated be the best kernel among the four kernel functions.

The leakage amount estimation and localization of SVR with Gaussian radial basis function are further depicted in Figs. 6.15-6.18. As shown in Figs. 6.15-6.16, the estimated leak positions perfectly match the actual ones with total diagnosis accuracy of $98.12\% - 99.80\%$. In particular, the mean localization accuracies reach above 99% except leak positions at $4,000\ m$ and $9,000\ m$ with relatively lower estimation accuracy. Considering other assessment criteria, the feasibility of the proposed method is proved by high localization accuracy and small RMSE values. As for the leak amount estimation, the accuracy is even higher, i.e., it is $99.89\%-99.98\%$ as illustrated in Figs. 6.17-6.18. In terms of estimation accuracy and RMSE, one can clearly see that the estimation discrepancy is small enough when compared with actual leakage amounts, which further verifies the effectiveness and reliability of the proposed method.

To investigate the pipeline parameter sensitivity of the developed observer model,

Table 6.3: Leak localization results by SVR with different kernel functions

Actual Leak Position (m)	Estimated Position by RBF SVR (m)		Estimated Position by Linear SVR (m)		Estimated Position by Sigmoid SVR (m)		Estimated Position by Polynomial SVR (m)	
	Mean	STD	Mean	STD	Mean	STD	Mean	STD
1000	996.92	88.48	2551.33	1169.37	2553.30	1171.52	1030.15	362.16
2000	2027.09	61.31	5458.08	293.40	5460.05	295.53	1771.57	311.73
3000	2967.55	22.00	6616.33	103.10	6618.29	101.04	2708.28	115.24
4000	3939.76	47.74	7250.87	322.74	7252.83	320.57	3846.64	173.17
5000	4952.52	32.46	7654.09	470.61	7656.05	468.44	4954.17	161.04
6000	5976.28	73.53	7933.81	576.42	7935.78	574.25	5953.06	77.49
7000	6986.13	67.51	8139.53	656.03	8141.50	653.86	6831.09	56.58
8000	7946.40	17.27	8297.30	718.17	8299.27	716.00	7597.55	189.93
9000	8830.49	93.88	8422.18	768.07	8424.14	765.90	8266.99	330.27

Table 6.4: Leak amount estimation results by SVR with different kernel functions

Actual Leak Amount	Estimated Amount by RBF SVR		Estimated Amount by Linear SVR	
	Mean	STD	Mean	STD
0.03	0.030033	0.000045	0.029979	0.000019
0.05	0.050034	0.000037	0.049969	0.000031
0.07	0.070012	0.000027	0.069957	0.000042
0.09	0.089976	0.000024	0.089944	0.000053
0.11	0.109942	0.000027	0.109929	0.000063
0.13	0.129922	0.000024	0.129913	0.000072
0.15	0.149921	0.000018	0.149896	0.000080
0.17	0.169930	0.000027	0.169877	0.000088
0.19	0.189926	0.000033	0.189858	0.000095

we vary the friction coefficient λ as one of the representative parameters with $\pm 5\%$ and repeat the Luenberger observer design, and SVR based leak localization and amount estimation process. The simulation results are shown as Table 6.5 and Table 6.6. The notations are made as: Lambda_No_SVR stands for original λ without any variation, Lambda_Ne_SVR represents original λ with minus 5%, and Lambda_Po_SVR stands for original λ with plus 5%. It is apparent that the position and amount estimation using PDE models with different friction coefficients preserve good performance. In terms of mean and standard deviation values, one can observe that there are negligible differences brought by variation of friction coefficient λ, which demonstrates

Table 6.4: Leak amount estimation results by SVR with different kernel functions (continued)

Actual Leak Amount	Estimated Amount by Sigmoid SVR		Estimated Amount by Polynomial SVR	
	Mean	STD	Mean	STD
0.03	0.029978	0.000015	0.030026	0.000025
0.05	0.049969	0.000024	0.050003	0.000034
0.07	0.069960	0.000032	0.069981	0.000040
0.09	0.089950	0.000040	0.089961	0.000044
0.11	0.109938	0.000048	0.109944	0.000048
0.13	0.129925	0.000055	0.129931	0.000051
0.15	0.149911	0.000061	0.149924	0.000054
0.17	0.169896	0.000067	0.169924	0.000057
0.19	0.189880	0.000072	0.189931	0.000061

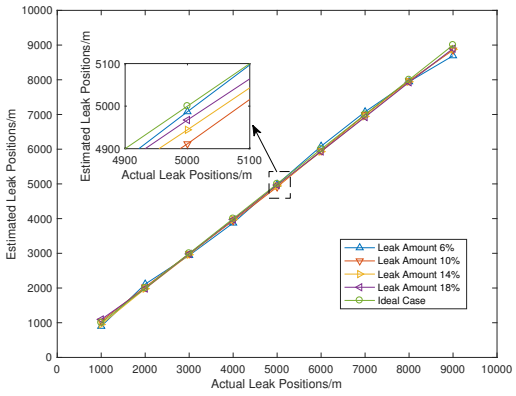


Figure 6.15: Comparison of estimated leak positions and actual ones.

the developed SVR model is not sensitive to the parameter uncertainty. In this case, the kernel function is set as Gaussian radial basis function.

For the sake of comparison, the pipeline leak detection case from Guillen's work [235] is considered. In particular, a 85 m pipeline is considered in Guillen's work [235], while the pipeline length considered in this chapter is a long range pipeline ($L_d = 10,000\ m$) as shown in Table 6.7. In addition, instead of solving a recursive filtering problem, leak diagnosis is based on SVR method, which is essentially a pattern recognition problem. It reduces the computing time to approximately 0.016 second for a single testing based on MacBook Pro (processor is 3.1GHz, Intel Core i5, and memory is 8 GB 2133 MHz LPDDR3 with operation system macOS Sierra

Table 6.5: Leak localization results by SVR with variations of friction coefficient λ

Actual Leak Position (m)	Estimated Position by Lambda_No_SVR(m)		Estimated Position by Lambda_Ne_SVR(m)		Estimated Position by Lambda_Po_SVR(m)	
	Mean	STD	Mean	STD	Mean	STD
1000	996.92	88.48	995.83	89.25	999.55	89.24
2000	2027.09	61.31	2027.13	62.85	2027.22	60.44
3000	2967.55	22.00	2966.19	22.54	2965.37	21.18
4000	3939.76	47.74	3936.80	49.30	3938.56	47.00
5000	4952.52	32.46	4949.99	32.08	4951.82	33.45
6000	5976.28	73.53	5974.85	71.63	5975.65	73.80
7000	6986.13	67.51	6985.62	65.70	6985.73	67.33
8000	7946.40	17.27	7946.34	16.62	7946.32	15.71
9000	8830.49	93.88	8830.59	93.93	8830.87	94.21

Table 6.6: Leak amount estimation results by SVR with variations of friction coefficient λ

Actual Leak Amount	Estimated Amount by Lambda_No_SVR		Estimated Amount by Lambda_Ne_SVR		Estimated Amount by Lambda_Po_SVR	
	Mean	STD	Mean	STD	Mean	STD
0.03	0.030033	0.000045	0.029940	0.000013	0.029925	0.000008
0.05	0.050034	0.000037	0.049919	0.000009	0.049911	0.000038
0.07	0.070012	0.000027	0.069935	0.000008	0.069900	0.000038
0.09	0.089976	0.000024	0.089945	0.000008	0.089888	0.000019
0.11	0.109942	0.000027	0.109939	0.000016	0.109880	0.000007
0.13	0.129922	0.000024	0.129921	0.000024	0.129875	0.000015
0.15	0.149921	0.000018	0.149902	0.000028	0.149868	0.000016
0.17	0.169930	0.000027	0.169892	0.000031	0.169857	0.000032
0.19	0.189926	0.000033	0.189879	0.000034	0.189831	0.000049

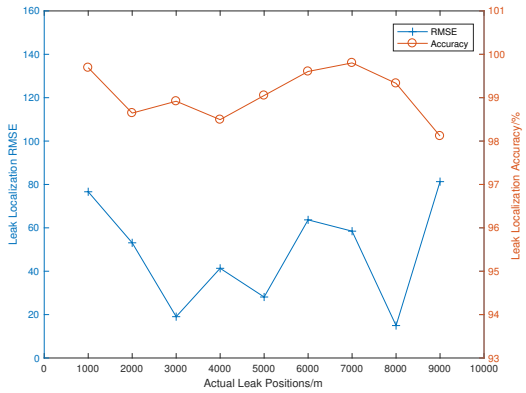


Figure 6.16: RMSE and accuracy of leak localization.

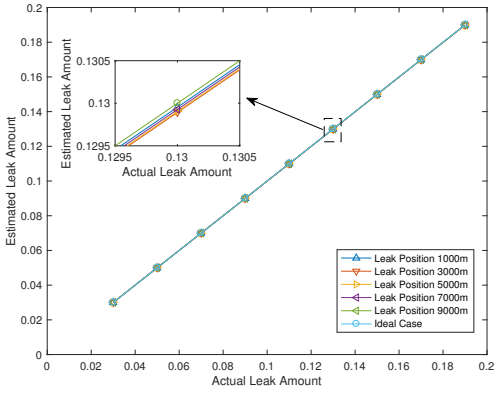


Figure 6.17: Comparison of estimated leak amount and actual ones (leak amount refers to the percentage of local mass flow velocity at the leak point).

10.12.06). As for diagnosis accuracy, it can reach 99% meaning that it outperforms the filtering method, where the accuracy is around 95% [235], and waiting time for one estimation result is around 60 seconds.

Remark 17. *Since the spatial discretization interval is* 500 m *and leak amount is discretized by 0.01 (the loss proportion of mass flow velocity at a given leak point) in the simulation of leakage scenarios, the ensuing training and testing datasets of SVM are generated based on these. However, this chapter can be extended for leak detection, leak amount and position estimation of long-range pipeline systems with leakage occurring other positions along the pipeline and/or other leakage amounts. The discrete observer and support vector machine model design can preserve high accuracy by setting the desired spatial discretization interval and specifying desired*

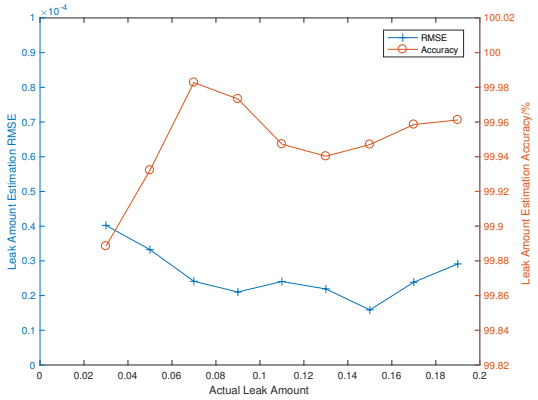


Figure 6.18: RMSE and accuracy of leak amount estimation (leak amount refers to the percentage of local mass flow velocity at the leak point).

Table 6.7: Comparison of simulation results with Guillen's paper

Methods	Guillen's Method	Proposed Method
Mechanism	Recursive estimation	Pattern recognition
Pipeline Length	85 m	$10,000$ m
Computing Time	60 s	$\leq$1 s
Diagnosis Accuracy	$\approx$ 95 %	$\approx$99 %

leakage amount in pipeline simulation.

Remark 18. *Considering the extrapolation performance, it can be seen that the SVR model deteriorates when it comes to localization close to both ends as shown in Fig. 6.15. This is due to the fact that in the simulation process the current spatial discretization is relatively coarse while on the other hand, Gaussian radial basis function is a typical local kernel with good performance in interpolation but not in extrapolation. Hence, one can enhance the extrapolation performance through decreasing the spatial discretization interval and/or proposing mixed kernel functions for SVM, such as by combing Gaussian radial basis function and polynomial kernel function [246].*

6.5 Conclusions

State estimation plays an important role for pipeline condition monitoring and leak diagnosis. This chapter presents a discrete Luenberger observer and investigates its utilization together with support vector machine models in leakage detection, amount estimation and localization based on infinite-dimensional nonlinear coupled

first-order hyperbolic equations. The major conclusions are drawn as: 1) the closed-form solution of the pipeline dynamics is mathematically formulated to reveal the spatiotemporal representation of the key process states, such as pressure, mass flow velocity and density; 2) the design of a discrete Luenberger observer is realized by utilizing the resolvent operator and solving the operator Riccati equation to obtain the stabilization observer gain for this class of the distributed parameter system; 3) based on the proposed infinite-dimensional discrete Luenberger observer, normal working situations and different leakage cases are simulated, including various leakage amounts and positions. And finally, two support vector machine models (OCSVM and SVR) are used to expose the presence of leakage and diagnose the corresponding leakage location and its amount. Simulation results show that the presented discrete observer and support vector machine models achieve satisfactory diagnosis accuracy (around 99%) with limited external measurements. The ultimate benefit of this design formulation is that the proposed design can be further integrated in discrete SCADA and embedded control systems.

Chapter 7

Discrete Output Regulator Design for Linear Distributed Parameter Systems

7.1 Introduction

Output regulation has been an active research and practical application driven topic during the past decades. The main idea of output regulation is to design regulators capable of stabilizing unstable systems, tracking desired output references, and rejecting undesired disturbance signals simultaneously. In general, there are two distinct design problems associated with output regulation, i.e., the state feedback and output/error feedback regulator designs. In the case of the state feedback regulator design the full state information is assumed to be available while only the output or error is known for the output/error feedback regulator design problems. Since the pioneering work of [247], the internal model principle has initiated a plethora of important contributions in the output regulation theory of various linear and nonlinear finite-dimensional systems [248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258].

When it comes to complex dynamical systems in engineering applications which are often modelled by partial differential equations (PDE) and/or partial integral-differential equations (PIDE), the major challenge in regulator design is to account for the characteristics of infinite-dimensional systems and incipient efforts were made toward extending the output regulation theory in finite-dimensional systems to infinite-dimensional systems. In particular, the non-model based PI controllers were naturally explored as early regulators for stable distributed parameter systems with constant

disturbance by [259], and later [260] introduced the concepts of the state feedback and the output feedback in the design of regulators based on state-space representations. [251] studied the output regulator design problem of finite-dimensional nonlinear systems by solving certain output regulator equations, which were generalized to infinite-dimensional systems with bounded control and observation operators in [10], and recently further extended to an important class of regular linear systems by [13]. Within the regulator design approaches the distinct realizations can be grouped in the cases when the exo-system is finite-dimensional [261, 262, 10, 74] or infinite-dimensional [262, 52, 12, 53], which implies that reference and disturbance signals can take finite- or infinite-dimensional representation.

Except the existing contributions on geometric regulation [57], the backstepping approach [15] was introduced to solve output regulation problems leading to systematic regulator design methods. In particular, [263] applied a backstepping method for disturbance rejection of a boundary controlled linear 2×2 hyperbolic system with co-located sensing and actuation, which was extended for the same type of system with interior domain disturbance by [264], $n+1$ systems by [265] and more general n+m heterodirectional first-order hyperbolic systems by [266] and [267]. Recently, [60] introduced an integral transform into hyperbolic PDEs and proposed a backstepping-forwarding controller and observer for this class of hyperbolic PDEs with Fredholm integrals, which was further extended to the output regulation problem by [268]. As for parabolic systems, [54] first developed a backstepping-based regulator design approach for a scaler PDE system and the results were extended to output regulation problems of a 1-D Schrödinger equation by [269] and PIDE systems by [270]. Recently, there are intense interests in cascade PDE systems by the use of backstepping approaches, including cascaded parabolic PDEs by [271], hyperbolic-parabolic PDE-PDE cascade by [272], ODE-PDE-ODE cascade systems by [273] and etc.

However, most of the existing work on the output regulation problem is conducted on the continuous-time setting, and relatively limited references are available on output regulation of discrete-time systems. Among these, discrete-time output regulation were considered for linear lumped parameter systems with input saturation by [274], piecewise-linear systems by [275], and linear finite-dimensional multi-agent systems by [276]. Nevertheless, there are even more scattered contributions when it comes to

discrete-time output regulator design in infinite-dimensional systems. Among these, a simple sampled-data low-gain controller was proposed for approximate tracking and disturbance rejection of a class of exponentially stable well-posed infinite-dimensional systems [277]. A sampled-data control problem of output tracking and disturbance rejection for unstable well-posed linear infinite-dimensional systems was considered with respect to constant disturbance and reference signals [278], where a frequency-domain technique based on coprime factorizations approach was employed. A lifting technique was used to design a discrete-time feedback controller that achieves approximate robust output tracking and disturbance rejection in [279]. Motivated by the fact that digital controllers and discrete-time systems are of great practical and theoretical interest, this chapter addresses discrete-time output regulator design problem for linear distributed parameter systems with state-space models and consideration of general exogenous signals (including step-like, ramp-like and harmonic signals).

In particular, we consider discrete-time output regulation design problem for linear distributed parameter systems driven by a finite-dimensional exo-system and extend key results of [10] and [102]. More specifically, novel contributions of this chapter lie in the following aspects: (1) state and error feedback discrete regulators are designed for linear discrete-time distributed parameter systems by employing the Cayley-Tustin bilinear transform which preserves model structure and properties of linear infinite-dimensional continuous-time systems; (2) the discrete regulator equations are formulated and proved in the design of state and error feedback regulators, and the discrete state and error feedback regulator design problems are solvable if and only if the discrete regulator equations can be solved; (3) A 1-1 correspondence between the solutions of discrete regulator equations and the corresponding continuous regulator equations is established, implying that one can solve for the continuous Sylvester equations and utilize the results in a discrete regulator design and vice versa; (4) the non-resonance solvability conditions of the discrete regulator equations are provided and linked to the corresponding continuous regulator equations; (5) a novel way of determining discrete-time stabilizing feedback gain (and its dual problem) is provided for the infinite-dimensional discrete-time systems using discrete-time Lyapunov and Riccati equations.

The rest of this chapter is organized as follows: in Section 7.2, continuous-time

infinite-dimensional plant and exogenous system are described and discretized in time by using the Cayley-Tustin transform. In Section 7.3, after revisiting some main results in the continuous-time state feedback regulator, the discrete-time state feedback regulator is designed and the solvability of the discrete Sylvester equations is proved and linked to its continuous counterpart. To ensure the stability of the closed-loop system, continuous- and discrete-time Lyapunov and Riccati equations are introduced to determine the discrete stabilizing feedback gain along with their continuous parallels. In the same manner, a discrete error feedback regulator is formulated and the solvability of the corresponding regulator equations is proved, along with its continuous counterpart, and the dual problem of solving output injection gain is studied in Section 7.4. Finally, the results are shown to be applicable to a first-order hyperbolic PDE model and a heat equation model in Section 7.5, and conclusions are offered in Section 7.6.

We use the following notations in this chapter. Assume that $\mathcal{X}$ and $\mathcal{V}$ are two Hilbert spaces and $\mathcal{A} : \mathcal{X} \mapsto \mathcal{V}$ is a linear operator from $\mathcal{X}$ to $\mathcal{V}$. $\mathcal{L}(\mathcal{X}, \mathcal{V})$ denotes the set of linear bounded operators from $\mathcal{X}$ to $\mathcal{V}$. If $\mathcal{X} = \mathcal{V}$, we simply write $\mathcal{L}(\mathcal{X})$. The domain, spectrum, resolvent set and resolvent operator of a linear operator $\mathcal{A}$ are denoted as: $\mathcal{D}(\mathcal{A})$, $\sigma(\mathcal{A})$, $\rho(\mathcal{A})$, and $\mathcal{R}(s, \mathcal{A}) = (sI - \mathcal{A})^{-1}$ with $s \in \rho(\mathcal{A})$, respectively. We denote the space $\mathcal{X}_1$ as the space $\mathcal{D}(\mathcal{A})$ with the norm $\|x\|_1 = \|(\beta I - \mathcal{A})x\|$, and the space $\mathcal{X}_{-1}$ as the completion of $\mathcal{X}$ with the norm $\|z\|_{-1} = \|(\beta I - \mathcal{A})^{-1} z\|$, where $\forall x \in \mathcal{D}(\mathcal{A})$, $\forall z \in \mathcal{X}$, and $\beta \in \rho(\mathcal{A})$. The constructed space are linked by $\mathcal{X}_1 \subset \mathcal{X} \subset \mathcal{X}_{-1}$, with each inclusion being dense and continuous embedding [42]. In addition, the inner product is denoted by $\langle \cdot, \cdot \rangle$, and $L^2(0, l)^m$ with a positive integer m denotes a Hilbert space of a m-dimensional vector of the real functions that are square integrable over $[0, l]$ with a spatial length l.

7.2 Preliminaries

The plant - The following linear infinite-dimensional continuous-time system is considered:

$$\dot{x}(t) = \mathcal{A}x(t) + \mathcal{B}u(t) + \Xi d(t), \; x(0) = x_0 \in \mathcal{X} \tag{7.1a}$$

$$y_c(t) = Cx(t) \tag{7.1b}$$

where the spatial state $x(\cdot,t) \in \mathcal{X}$, with $\mathcal{X} = L^2((0,l),\mathbb{C})$ is being defined as a complex separable Hilbert space. $\zeta \in [0,l] \subset \mathbb{R}$ and $t \in [0,\infty)$ represent temporal and spatial coordinates. We denote the input $u(t) \in L^2_{loc}([0,\infty),U)$, the disturbance $d(t) \in L^2_{loc}([0,\infty),U_d)$, and the controlled output $y_c(t) \in L^2_{loc}([0,\infty),Y)$, where U, U_d and Y are finite-dimensional Hilbert spaces. $\mathcal{A} : \mathcal{D}(\mathcal{A}) \subset \mathcal{X} \mapsto \mathcal{X}$ is an infinitesimal generator of a C_0-semigroup $\mathbb{T}_{\mathcal{A}}(t)$ on $\mathcal{X}$. The operators $\mathcal{B} \in \mathcal{L}(U,\mathcal{X})$, $\Xi \in \mathcal{L}(U_d,\mathcal{X})$, and $\mathcal{C} \in \mathcal{L}(\mathcal{X},Y)$ are assumed to be bounded operators. We remark that our approach can be extended to the more general class of well-posed linear systems which include unbounded input and output operators in the sense of [121, 280]. To address that, we need to introduce the spaces $\mathcal{X}_1$ and $\mathcal{X}_{-1}$ so as to define unbounded control, observation and disturbance operators as in [42, Pro. 2.10.3] and [281].

Naturally, one can express the transfer functions as follows:

$$\mathcal{G}_c(s) = \mathcal{C}\mathcal{R}(s,\mathcal{A})\mathcal{B},\ s \in \rho(\mathcal{A}) \tag{7.2a}$$

$$\mathcal{T}_c(s) = \mathcal{C}\mathcal{R}(s,\mathcal{A})\Xi,\ s \in \rho(\mathcal{A}) \tag{7.2b}$$

where $\mathcal{G}_c(s)$ and $\mathcal{T}_c(s)$ stand for transfer functions from $u(t)$ to $y_c(t)$ and from $d(t)$ to $y_c(t)$, respectively.

The plant discretization in time - In order to address the issue of time discretization, the energy and structure preserving Cayley-Tustin bilinear transform is applied to the linear infinite-dimensional continuous-time system (7.1) for a given time discretization interval $\triangle t > 0$ as follows:

$$\frac{x(k\Delta t) - x((k-1)\Delta t)}{\Delta t} \approx \mathcal{A}\frac{x(k\Delta t) + x((k-1)\Delta t)}{2} + \mathcal{B}u(k\Delta t) + \Xi d(k\Delta t) \tag{7.3a}$$

$$y_c(k\Delta t) \approx \mathcal{C}\frac{x(k\Delta t) + x((k-1)\Delta t)}{2},\ x(0) = x_0,\ k \geq 1 \tag{7.3b}$$

As shown in Eq.(7.3), the discretization is performed based on the implicit mid-point integration rule, which does not rely on any spatial discretization or model reduction, leading to a symmetric (in time) and symplectic discretization scheme [119]. Through simple algebraic manipulation of Eq.(7.3), one can obtain the infinite-dimensional discrete-time state-space model as:

$$x_k = \mathcal{A}_d x_{k-1} + \mathcal{B}_d u_k + \Xi_d d_k,\ k \geq 1 \tag{7.4a}$$

$$y_{ck} = \mathcal{C}_d x_{k-1} + \mathcal{D}_d u_k + \Upsilon_d d_k \tag{7.4b}$$

where the discrete state, disturbance and output are denoted by x_k, d_k and y_{ck}. Additionally, the discrete input is given by $\frac{u_k}{\sqrt{\Delta t}} = \frac{1}{\Delta t}\int_{(k-1)\Delta t}^{k\Delta t} u(t)dt$, and it can be shown that $\frac{u_k}{\sqrt{\Delta t}}$ converges to $u(t)$ on the interval $t \in ((k-1)\Delta t, k\Delta t)$ as $\Delta t \to 0$ [97]. Similar expressions hold for the discrete time d_k and y_{ck}. The associated discrete-time operators are given as follows:

$$\begin{bmatrix} \mathcal{A}_d & \mathcal{B}_d & \Xi_d \\ \mathcal{C}_d & \mathcal{D}_d & \Upsilon_d \end{bmatrix} = \begin{bmatrix} -I + 2\delta\mathcal{R}(\delta,\mathcal{A}) & \sqrt{2\delta}\mathcal{R}(\delta,\mathcal{A})B & \sqrt{2\delta}\mathcal{R}(\delta,\mathcal{A})\Xi \\ \sqrt{2\delta}\mathcal{C}\mathcal{R}(\delta,\mathcal{A}) & \mathcal{G}_c(\delta) & \mathcal{T}_c(\delta) \end{bmatrix} \tag{7.5}$$

where $\mathcal{R}(\delta,\mathcal{A})$, $\mathcal{G}_c(\delta)$ and $\mathcal{T}_c(\delta)$ denote the resolvent operator $\mathcal{R}(s,\mathcal{A}) = (sI-\mathcal{A})^{-1}$, transfer functions $\mathcal{G}_c(s)$ and $\mathcal{T}_c(s)$ with s evaluated at $s = \delta = 2/\Delta t \in \rho(\mathcal{A})$. In addition, there are feedforward operators $\mathcal{D}_d$ and Υ_d appearing in the discrete-time setting (7.4) after applying Cayley-Tustin discretization, which is not necessarily present in the continuous model (7.1b). The transfer functions of the discrete-time system (7.4) are given by:

$$\mathcal{G}_d(z) = C_d(zI-\mathcal{A}_d)^{-1}\mathcal{B}_d + \mathcal{D}_d,\ z \in \rho(\mathcal{A}_d)\backslash\{-1\} \tag{7.6a}$$

$$\mathcal{T}_d(z) = C_d(zI-\mathcal{A}_d)^{-1}\Xi_d + \Upsilon_d,\ z \in \rho(\mathcal{A}_d)\backslash\{-1\} \tag{7.6b}$$

where $\mathcal{G}_d(z)$ and $\mathcal{T}_d(z)$ are transfer functions from u_k to y_{ck} and from d_k to y_{ck}, respectively. Based on the well-known bilinear mapping $z = \frac{\delta+s}{\delta-s}$ and $s = \frac{z-1}{z+1}\delta$ [23] (taken as: $\delta = 1$), the discrete- and continuous-time transfer functions are linked as:

$$\mathcal{G}_d(z) = \mathcal{G}_c(\frac{z-1}{z+1}\delta),\ \mathcal{T}_d(z) = \mathcal{T}_c(\frac{z-1}{z+1}\delta),\ z \in \rho(\mathcal{A}_d)\backslash\{-1\} \tag{7.7a}$$

$$\mathcal{G}_c(s) = \mathcal{G}_d(\frac{\delta+s}{\delta-s}),\ \mathcal{T}_c(s) = \mathcal{T}_d(\frac{\delta+s}{\delta-s}),\ s \in \rho(\mathcal{A})\backslash\{\delta\} \tag{7.7b}$$

By the Cayley-Tustin bilinear transformation, the open right-half plane $\mathbb{C}^+ = \{s \in \mathbb{C} : \Re(s) > 0\}$ is mapped into the exterior of the unit disc $\mathbb{D}^+ = \{z \in \mathbb{C} : |z| > 1\}$ and vice versa. Based on that, a 1-1 correspondence of stability, admissibility, controllability and observability between continuous- and discrete-time systems has been established in terms of Lyapunov and Riccati equations by [23]. In this chapter, we will explore the 1-1 equivalence of discrete- and continuous-time regulation problems in terms of Sylvester equations.

Remark 19. *In the discrete-time model (7.4), all discrete-time operators are bounded and defined as:* $\mathcal{A}_d \in \mathcal{L}(\mathcal{X})$, $\mathcal{B}_d \in \mathcal{L}(U,\mathcal{X})$, $\Xi_d \in \mathcal{L}(U_d,\mathcal{X})$, $\mathcal{C}_d \in \mathcal{L}(\mathcal{X},Y)$, $\mathcal{D}_d \in \mathcal{L}(U,Y)$, $\Upsilon_d \in \mathcal{L}(U_d,Y)$, *see details in the reference [279].*

Exo-system - In order to generate disturbance and reference signals, a finite-dimensional exogenous system (exo-system in short) is introduced as follows:

$$\dot{q}(t) = Sq(t),\ q(0) = q^0 \in \mathbb{C}^{n_q} \tag{7.8a}$$

$$d(t) = Fq(t),\ y_r(t) = Qq(t) \tag{7.8b}$$

where q, d and y_r represent the state, disturbance and reference signals of the continuous-time exo-system. In addition, $q(t) \in \mathbb{C}^{n_q}$, S, F and Q have compatible dimensions.

Assumption 10. *$S : \mathcal{D}(S) : \mathbb{C}^{n_q} \to \mathbb{C}^{n_q}$ is a matrix with all eigenvalues on the imaginary axis and S has two candidates S_m and S_n. S_m has distinct eigenvalues and is diagonalizable with dimension of n_q, while S_n is a nilpotent matrix of dimension 2 as $S_n = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$. Hence, this design of S accounts for the modelling of step-like, ramp-like, and harmonic exogenous signals.*

Exo-system discretization - A discrete exo-system is formulated to generate discrete-time disturbance and output reference signals as follows:

$$q_k = S_d q_{k-1},\ q_0 = q^0 \in \mathbb{C}^{n_q} \tag{7.9a}$$

$$d_k = F_d q_k,\ y_{rk} = Q_d q_k,\ k \geq 1 \tag{7.9b}$$

where q_k, d_k and y_{rk} are the state, disturbance and output reference signals in the discrete-time setting. Specifically, S_d is the discrete state evolution matrix and obtained by discretizing the corresponding continuous evolution matrix S using the Cayley-Tustin transform as:

$$S_d = -I + 2\delta_0(\delta_0 I - S)^{-1},\ \delta_0 \in \rho(S) = \mathbb{C}\backslash\sigma(S) \tag{7.10a}$$

$$F_d = F,\ Q_d = Q \tag{7.10b}$$

where $\delta_0 = 2/\Delta T$ with ΔT defined as the discretization time for the exo-system and we assume $\Delta T = \Delta t$ for simplicity, which implies that $\delta = \delta_0$.

By the light of Assumption 10, we have two candidates for S_d as $S_d = S_m^d$ or $S_d = S_n^d$, which implies that all the eigenvalues of S_d are on the unit circle boundary on the complex plane. In particular, $S_i^d = -I + 2\delta(\delta I - S_i)^{-1}$, $i = m, n$, and S_m^d has all distinct

eigenvalues and $S_n^d = \begin{bmatrix} 1 & 0 \\ \frac{2}{\delta} & 1 \end{bmatrix}$. Hence, S_d is capable of generating step-like, ramp-like, and harmonic signals in the discrete-time setting. Thus, the correspondence between the continuous- and discrete-time exogenous systems is established in order to link the solvability of discrete Sylvester equations and the continuous counterparts.

Remark 20. *We need discretize F and Q as below using Cayley-Tustin discretization approach in order to ensure the corresponding relationship of discrete- and continuous-time error feedback regulators.*

$$S_d = -I + 2\delta(\delta I - S)^{-1},\ \delta \in \rho(S) = \mathbb{C}\backslash\sigma(S) \tag{7.11a}$$

$$F_d = \sqrt{2\delta}F(\delta I - S)^{-1},\ Q_d = \sqrt{2\delta}Q(\delta I - S)^{-1} \tag{7.11b}$$

Corollary 1. *With diagonalizable S and S_d (i.e., $S = S_m$ and $S_d = S_m^d$), for each eigenpair (λ_i^s, ϕ_i^s) of S, the associated eigenpair (λ_i^d, ϕ_i^d) of S_d is given by $\lambda_i^d = -1 + 2\delta(\delta - \lambda_i^s)^{-1}$ and $\phi_i^d = \phi_i^s$ where $\delta \in \rho(S)$; With non-diagonalizable S and S_d (i.e., $S = S_n$ and $S_d = S_n^d$), we have the multiplicity of eigenvalues $\lambda^s = 0$ and $\lambda^d = 1$ (for simplicity we drop the subscript i in this case), which induces a standard eigenvector $S\phi_1^s = \lambda^s\phi_1^s$ or $S_d\phi_1^d = \lambda^d\phi_1^d$ and a generalized eigenvector as $S\phi_2^s = \lambda^s\phi_2^s + \phi_1^s$ or $S_d\phi_2^d = \lambda^d\phi_2^d + \phi_1^d$ by using the chain rule. Furthermore, we suppose $\lambda_i^s \neq \delta$ and $\lambda_i^d \neq -1$, which can always be ensured by Assumption 10 and a proper choice of the discretization interval.*

For clarification, we introduce the following stability concepts.

Definition 8. *The C_0-semigroup $\mathbb{T}_{\mathcal{A}}(t)$ on $\mathcal{X}$ is exponentially stable if there exist positive constants M and α such that:*

$$\|\mathbb{T}_{\mathcal{A}}(t)\| \leq Me^{-\alpha t},\ \forall t \in \mathbb{R}^+$$

and it is strongly stable if $\|\mathbb{T}_{\mathcal{A}}(t)x\| \to 0$ as $t \to \infty$ for all $x \in \mathcal{X}$. $\mathcal{A}_d$ is power stable if there exist positive constants M and $\gamma < 1$ such that:

$$\|\mathcal{A}_d^k\| \leq M\gamma^k,\ \forall k \in \mathbb{N}$$

and $\mathcal{A}_d$ is strongly stable if $\mathcal{A}_d^k x \to 0$ as $k \to \infty$ for all $x \in \mathcal{X}$.

Theorem 13. *[23, The. 2.9] Suppose that* $\Sigma(\mathcal{A},\mathcal{B},C,\mathcal{D})$ *and* $\Sigma_d(\mathcal{A}_d,\mathcal{B}_d,C_d,\mathcal{D}_d)$ *are continuous- and discrete-time analogues. Then* $\Sigma_d(\mathcal{A}_d,\mathcal{B}_d,C_d,\mathcal{D}_d)$ *is strongly stabilizable (detectable) if and only if* $\Sigma(\mathcal{A},\mathcal{B},C,\mathcal{D})$ *is strongly stabilizable (detectable).*

Normally, it is more favourable to show the 1-1 correspondence between the power stability of $\mathcal{A}_d$ and exponential stability of $\mathbb{T}_{\mathcal{A}}(t)$. However, it needs more rigorous assumption on the boundedness of $\mathcal{A}$ on $\mathcal{X}$ [282, Lem. 4.4] or the semigroup generated by $\mathcal{A}$ and $\mathcal{A}^{-1}$ [22, The. 4.4], which needs more technical treatments and it will not be considered in this chapter.

Throughout this chapter, we make some general assumptions as below:

Assumption 11. *The spectrum of* S *is included in the resolvent set of* $\mathcal{A}$*, i.e.,* $\sigma(S)\subset\rho(\mathcal{A})$.

By applying the Cayley-Tustin transform with some proper time discretization interval, we can ensure that the spectrum of S_d is contained in the resolvent set of $\mathcal{A}_d$, i.e., $\sigma(S_d)\subset\rho(\mathcal{A}_d)$.

Assumption 12. *The pair* $(\mathcal{A},\mathcal{B})$ *is exponentially stabilizable.*

By Theorem 13 and Assumption 12, we note that the pair $(\mathcal{A},\mathcal{B})$ is strongly stabilizable, which further implies that $(\mathcal{A}_d,\mathcal{B}_d)$ is strongly stabilizable.

Assumption 13. *The pair* $\left(\begin{bmatrix}\mathcal{A}&P\\0&S\end{bmatrix},\begin{bmatrix}C&-Q\end{bmatrix}\right)$ *is exponentially detectable and there exists* $\mathcal{G}_2=[G_1;G_2]\in\mathcal{L}(Y,\Omega)$ *where* Ω *is a Hilbert space with* $\Omega=\mathcal{X}\bigoplus\mathbb{C}^{n_q}$ *such that* $\begin{bmatrix}\mathcal{A}&P\\0&S\end{bmatrix}-\mathcal{G}_2\begin{bmatrix}C&-Q\end{bmatrix}=\begin{bmatrix}\mathcal{A}-G_1C&P+G_1Q\\-G_2C&S+G_2Q\end{bmatrix}$ *generates an exponentially stable* C_0*-semigroup, where* $P=\Xi F$.

By Theorem 13 and Assumption 13, we can induce that the pair $\left(\begin{bmatrix}\mathcal{A}_d&P_d\\0&S_d\end{bmatrix},[C_d\ \Theta_{cd}-Q_d]\right)$ is strongly detectable and there exists $\mathcal{G}_{2d}\in\mathcal{L}(Y,\Omega)$, such that $\begin{bmatrix}\mathcal{A}_d&P_d\\0&S_d\end{bmatrix}-\mathcal{G}_{2d}\begin{bmatrix}C_d&\Theta_{cd}-Q_d\end{bmatrix}=\begin{bmatrix}\mathcal{A}_d-G_{1d}C_d&P_d-G_{1d}(\Theta_{cd}-Q_d)\\-G_{2d}C_d&S_d-G_{2d}(\Theta_{cd}-Q_d)\end{bmatrix}$ is a strongly stable operator, where $\mathcal{G}_{2d}=[G_{1d};G_{2d}]$, $P_d=\Xi_dF_d$ and $\Theta_{cd}=\Upsilon_dF_d$. It can proved by showing that $\left(\begin{bmatrix}\mathcal{A}_d&P_d\\0&S_d\end{bmatrix},\begin{bmatrix}C_d&\Theta_{cd}-Q_d\end{bmatrix}\right)$ corresponds to the

discrete-time operator $\left(\begin{bmatrix} \mathcal{A} & P \\ 0 & S \end{bmatrix}, \begin{bmatrix} C & -Q \end{bmatrix}\right)$ by using the Cayley-Tustin transform.

In order to show the solvability of the regulator equations, we introduce the concept of transmission zero. Under the consideration of finite-dimensional input and output spaces, we make the following definition as in [10].

Definition 9. *$s_0 \in \mathbb{C}$ is a transmission zero of continuous-time plant (7.1) if* $\det \mathcal{G}_c(s_0) = 0$*, and $z_0 \in \mathbb{C}$ is a transmission zero of discrete-time plant (7.4) if* $\det \mathcal{G}_d(z_0) = 0$.

7.3 State feedback regulation

In this section, with the full state information being provided, a discrete-time state feedback regulator is designed for the discrete system (7.4) and it is presented in parallel with its continuous analogue for comparison. Based on the Cayley-Tustin discretization, we prove the solvability of the discrete output regulator equations and provide a 1-1 correspondence between the solutions of discrete- and continuous-time regulator equations.

7.3.1 Continuous-time state feedback regulator

To proceed with the discrete-time state-feedback regulator design, we briefly revisit the corresponding continuous-time counterpart in this section.

For simplicity, the continuous-time state feedback regulator design problem is reviewed as follows. A continuous-time state-feedback regulator is designed for the system (1) by finding a control law having the following form:

$$u(t) = Kx(t) + Lq(t) \tag{7.12}$$

where $K \in \mathcal{L}(\mathcal{X}, U)$, $L \in \mathcal{L}(\mathbb{C}^{n_q}, U)$ such that the following conditions hold.

[c1]: The closed-loop system operator $\mathcal{A} + \mathcal{B}K$ generates an exponentially stable C_0-semigroup;

[c2]: For the closed-loop system, the output tracking error $e(t) = y_c(t) - y_r(t) \to 0$ with $t \to +\infty$ for any given initial conditions of $x_0 \in \mathcal{X}$ and $q^0 \in \mathbb{C}^{n_q}$.

To determine the control law (7.12), the following theorem is often utilized:

Theorem 14. *Let Assumptions 11-12 hold. The continuous-time state feedback regulation problem is solvable if and only if there exist mappings* $\Pi \in \mathcal{L}(\mathbb{C}^{n_q}, \mathcal{X})$ *with* $\Pi D(S) \subset D(\mathcal{A})$ *and* $\Gamma \in \mathcal{L}(\mathbb{C}^{n_q}, U)$ *such that the following Sylvester equations hold [10, The. IV.1]:*

$$\Pi S = \mathcal{A}\Pi + \mathcal{B}\Gamma + P \tag{7.13a}$$

$$C\Pi = Q \tag{7.13b}$$

where $P = \Xi F$*, and* $L = \Gamma - K\Pi$ *can be utilized for computing the control input* $u(t) = Kx(t) + Lq(t)$*.*

7.3.2 Discrete-time state feedback regulator

A discrete state feedback regulator is designed for the discrete system (7.4) in a discrete-time setting by satisfying the following conditions.

[C1]: The closed-loop system operator $\mathcal{A}_d + \mathcal{B}_d K_d$ is strongly stable;

[C2]: For the closed-loop system, the output tracking error $e_k = y_{ck} - y_{rk} \to 0$ with $k \to +\infty$ for any given initial conditions of $x_0 \in \mathcal{X}$ and $q^0 \in \mathbb{C}^{n_q}$.

Discrete time regulator design - a full state feedback: With full state information of plant and exo-system being available, the discrete state feedback regulator design problem is addressed by finding a discrete regulator in the following form:

$$u_k = K_d x_{k-1} + L_d q_k \tag{7.14}$$

where $K_d \in \mathcal{L}(\mathcal{X}, U)$, $L_d \in \mathcal{L}(\mathbb{C}^{n_q}, U)$ such that [C1] and [C2] hold.

To address the discrete-time state feedback regulator design problem, we propose the following theorem.

Theorem 15. *Under Assumptions 11-12, the discrete state feedback regulation problem is solvable if and only if there exist mappings* $\Pi_d \in \mathcal{L}(\mathbb{C}^{n_q}, \mathcal{X})$ *and* $\Gamma_d \in \mathcal{L}(\mathbb{C}^{n_q}, U)$ *such that the following discrete Sylvester equations hold:*

$$\Pi_d S_d = \mathcal{A}_d \Pi_d + (\mathcal{B}_d \Gamma_d + P_d) S_d \tag{7.15a}$$

$$Q_d S_d = C_d \Pi_d + (\mathcal{D}_d \Gamma_d + \Theta_{cd}) S_d \tag{7.15b}$$

where $P_d = \Xi_d F_d$*,* $\Theta_{cd} = \Upsilon_d F_d$*, and* $L_d = \Gamma_d - K_d \Pi_d S_d^{-1}$ *can be utilized to compute the state feedback control law* u_k *in Eq.(7.14).*

Proof. First, let us prove the sufficiency. Plugging Eq.(7.14) into the discrete system (7.4) leads to the closed-loop model as follows:

$$x_k = (\mathcal{A}_d + \mathcal{B}_d K_d)x_{k-1} + (\mathcal{B}_d L_d + P_d)q_k \tag{7.16}$$

To ensure [C1], the operator $\mathcal{A}_d + \mathcal{B}_d K_d$ needs to be strongly stable and the discrete-time solution takes the following form:

$$x_k = (\mathcal{A}_d + \mathcal{B}_d K_d)^k x_0 + \sum_{m=1}^{k} (\mathcal{A}_d + \mathcal{B}_d K_d)^{m-1}(\mathcal{B}_d L_d + P_d)q_{k+1-m} \tag{7.17}$$

By substituting Eq.(7.9) and Eq.(7.15) into Eq.(7.17), one gets:

$$\begin{aligned}
x_k &= (\mathcal{A}_d + \mathcal{B}_d K_d)^k x_0 + \sum_{m=1}^{k} (\mathcal{A}_d + \mathcal{B}_d K_d)^{m-1}[\mathcal{B}_d(\Gamma_d - K_d\Pi_d S_d^{-1}) + P_d]q_{k+1-m} \\
&= (\mathcal{A}_d + \mathcal{B}_d K_d)^k x_0 + \sum_{m=1}^{k} (\mathcal{A}_d + \mathcal{B}_d K_d)^{m-1}[(\mathcal{B}_d\Gamma_d + P_d)S_d - \mathcal{B}_d K_d\Pi_d)]q_{k-m} \\
&= (\mathcal{A}_d + \mathcal{B}_d K_d)^k x_0 + \sum_{m=1}^{k} (\mathcal{A}_d + \mathcal{B}_d K_d)^{m-1}[\Pi_d S_d - (\mathcal{A}_d + \mathcal{B}_d K_d)\Pi_d)]q_{k-m} \\
&= (\mathcal{A}_d + \mathcal{B}_d K_d)^k x_0 + \sum_{m=1}^{k} (\mathcal{A}_d + \mathcal{B}_d K_d)^{m-1}\Pi_d q_{k+1-m} \\
&\quad - \sum_{m=2}^{k+1} (\mathcal{A}_d + \mathcal{B}_d K_d)^{m-1}\Pi_d q_{k+1-m} \\
&= (\mathcal{A}_d + \mathcal{B}_d K_d)^k (x_0 - \Pi_d q_0) + \Pi_d q_k
\end{aligned} \tag{7.18}$$

Moreover, the discrete tracking error can be expressed as:

$$\begin{aligned}
e_k =& y_{ck} - y_{rk} \\
=& C_d x_{k-1} + \mathcal{D}_d u_k + \Theta_{cd} q_k - Q_d q_k \\
=& (C_d + \mathcal{D}_d K_d)x_{k-1} + (\mathcal{D}_d L_d + \Theta_{cd} - Q_d)q_k \\
=& (C_d + \mathcal{D}_d K_d)(\mathcal{A}_d + \mathcal{B}_d K_d)^{k-1}(x_0 - \Pi_d q_0) \\
&+ [(C_d + \mathcal{D}_d K_d)\Pi_d + (\mathcal{D}_d L_d + \Theta_{cd} - Q_d)S_d]q_{k-1}
\end{aligned} \tag{7.19}$$

Since $\mathcal{A}_d + \mathcal{B}_d K_d$ is a strongly stable operator, we have that $(\mathcal{A}_d + \mathcal{B}_d K_d)^k x \to 0$ as $k \to +\infty$ for all $x \in \mathcal{X}$. Therefore, x_k converges to $\Pi_d q_k$ in Eq.(7.18) and the discrete tracking error e_k goes to zero in Eq.(7.19) as $k \to +\infty$, which is guaranteed by the discrete Sylvester equations (7.15a)-(7.15b).

Now, we focus on the proof of the necessity and let us construct the following extended closed-loop system:

$$\begin{bmatrix} x_k \\ q_k \end{bmatrix} = \begin{bmatrix} \mathcal{A}_d + \mathcal{B}_d K_d & (\mathcal{B}_d L_d + P_d) S_d \\ 0 & S_d \end{bmatrix} \begin{bmatrix} x_{k-1} \\ q_{k-1} \end{bmatrix} \tag{7.20}$$

It is straightforward to conclude the solution of Eq.(7.20) by induction as follows:

$$\begin{bmatrix} x_k \\ \\ q_k \end{bmatrix} = \begin{bmatrix} (\mathcal{A}_d + \mathcal{B}_d K_d)^k x_0 + \sum\limits_{m=1}^{k} (\mathcal{A}_d + \mathcal{B}_d K_d)^{m-1} (\mathcal{B}_d L_d + P_d) q_{k+1-m} \\ S_d^k q_0 \end{bmatrix} \tag{7.21}$$

Given that $\mathcal{A}_d + \mathcal{B}_d K_d$ is strongly stable, $(\mathcal{A}_d + \mathcal{B}_d K_d)^k x_0 \to 0$ as $k \to +\infty$ and Eq.(7.21) indicates that $\begin{bmatrix} x_k \\ q_k \end{bmatrix} \to \begin{bmatrix} \Pi_d q_k \\ q_k \end{bmatrix}$ as $k \to +\infty$ and $\Pi_d \in \mathcal{L}(\mathbb{C}^{n_q}, \mathcal{X})$. To determine Π_d, we can construct the dynamical evolution of $\begin{bmatrix} x_k \\ q_k \end{bmatrix} - \begin{bmatrix} \Pi_d q_k \\ q_k \end{bmatrix}$ as the following homogeneous difference equation:

$$\begin{bmatrix} x_k \\ q_k \end{bmatrix} - \begin{bmatrix} \Pi_d q_k \\ q_k \end{bmatrix} = \begin{bmatrix} \mathcal{A}_d + \mathcal{B}_d K_d & (\mathcal{B}_d L_d + P_d) S_d \\ 0 & S_d \end{bmatrix} \times \left(\begin{bmatrix} x_{k-1} \\ q_{k-1} \end{bmatrix} - \begin{bmatrix} \Pi_d q_{k-1} \\ q_{k-1} \end{bmatrix} \right) \tag{7.22}$$

where the initial condition is defined as $\begin{bmatrix} x_0 \\ q_0 \end{bmatrix} - \begin{bmatrix} \Pi_d q_0 \\ q_0 \end{bmatrix} \in \Omega$ with $\Omega = \mathcal{X} \bigoplus \mathbb{C}^{n_q}$. The first component in Eq.(7.22) leads to $(\mathcal{A}_d + \mathcal{B}_d K_d)\Pi_d + (\mathcal{B}_d L_d + P_d) S_d = \Pi_d S_d$ which is identical to discrete-time Sylvester equation (7.15a). Furthermore, the discrete tracking error is described as:

$$\begin{aligned} e_k &= y_{ck} - y_{rk} \\ &= C_d x_{k-1} + \mathcal{D}_d u_k + \Theta_{cd} q_k - Q_d q_k \\ &= (C_d + \mathcal{D}_d K_d) x_{k-1} + (\mathcal{D}_d L_d + \Theta_{cd} - Q_d) q_k \\ &= \begin{bmatrix} C_d + \mathcal{D}_d K_d & (\mathcal{D}_d L_d + \Theta_{cd} - Q_d) S_d \end{bmatrix} \begin{bmatrix} x_{k-1} \\ q_{k-1} \end{bmatrix} \\ &\to \left[(C_d + \mathcal{D}_d K_d) \Pi_d + (\mathcal{D}_d L_d + \Theta_{cd} - Q_d) S_d \right] q_{k-1} \quad (\text{as } k \to +\infty) \end{aligned} \tag{7.23}$$

To realize perfect tracking, it is necessary to ensure that $(C_d + \mathcal{D}_d K_d)\Pi_d + (\mathcal{D}_d L_d + \Theta_{cd} - Q_d) S_d = 0$, which implies Eq.(7.15b) by substituting $L_d = \Gamma_d - K_d \Pi_d S_d^{-1}$. ■

7.3.3 Link between continuous and discrete regulator equations

The solutions of the proposed discrete-time regulator equations are linked to the associated continuous analogues, based on Theorem 14, Theorem 15 and Cayley-Tustin bilinear transform.

Before we proceed with the theorem showing the equivalent link, let us propose the following proposition:

Proposition 6. *Suppose that we have these "second-order" transfer functions (or equivalently the derivatives of transfer functions $\mathcal{G}_c(s)$, $\mathcal{G}_d(z)$, $\mathcal{T}_c(s)$, and $\mathcal{T}_d(z)$) for the continuous- and discrete-time system (7.1) and (7.4) accordingly as:*

$$\mathcal{G}_c^{(2)}(s) = C(sI - \mathcal{A})^{-2}\mathcal{B},\ s \in \rho(\mathcal{A})\backslash\{\delta\} \tag{7.24a}$$

$$\mathcal{G}_d^{(2)}(z) = C_d(zI - \mathcal{A}_d)^{-2}\mathcal{B}_d,\ z \in \rho(\mathcal{A}_d)\backslash\{-1\} \tag{7.24b}$$

$$\mathcal{T}_c^{(2)}(s) = C(sI - \mathcal{A})^{-2}\Xi,\ s \in \rho(\mathcal{A})\backslash\{\delta\} \tag{7.24c}$$

$$\mathcal{T}_d^{(2)}(z) = C_d(zI - \mathcal{A}_d)^{-2}\Xi_d,\ z \in \rho(\mathcal{A}_d)\backslash\{-1\} \tag{7.24d}$$

Using the Cayley-Tustin bilinear transform, it can be proved that the following relationships hold:

$$\mathcal{G}_d^{(2)}(z) = \frac{(\delta - s)^2}{2\delta} \times \mathcal{G}_c^{(2)}(s) \tag{7.25a}$$

$$\mathcal{T}_d^{(2)}(z) = \frac{(\delta - s)^2}{2\delta} \times \mathcal{T}_c^{(2)}(s) \tag{7.25b}$$

Proof. By substituting the discrete operators $(\mathcal{A}_d, \mathcal{B}_d, C_d, \Xi_d)$ given by Eq.(7.5) and $z = \frac{\delta+s}{\delta-s}$ into (7.24b), one can obtain

$$\begin{aligned}\mathcal{G}_d^{(2)}(z) &= \sqrt{2\delta}C(\delta I - \mathcal{A})^{-1}[2\delta(\delta - s)^{-1} - 2\delta(\delta I - \mathcal{A})^{-1}]^{-2}\sqrt{2\delta}(\delta I - \mathcal{A})^{-1}\mathcal{B} \\ &= \frac{(\delta - s)^2}{2\delta} \times C(sI - \mathcal{A})^{-2}\mathcal{B} \\ &= \frac{(\delta - s)^2}{2\delta} \times \mathcal{G}_c^{(2)}(s)\end{aligned} \tag{7.26}$$

In the similar manner, the proof of Eq.(7.25b) can be completed. ■

To reveal the relationship between (Γ, Π) and (Γ_d, Π_d), we provide the following theorem:

Theorem 16. *Let Assumptions 10-12 hold. By Cayley-Tustin transform (7.5) and (7.10), the solutions of continuous- and discrete-time Sylvester equations are linked by:*

(a). For diagonalizable $S = S_m,\ S_d = S_m^d$

$$\Gamma_d = \Gamma \tag{7.27a}$$

$$\Pi_d \phi_i^s = \frac{\delta + \lambda_i^s}{\sqrt{2\delta}} \Pi \phi_i^s = \frac{\sqrt{2\delta}\lambda_i^d}{\lambda_i^d + 1} \Pi \phi_i^s \tag{7.27b}$$

(b). For non-diagonalizable $S = S_n,\ S_d = S_n^d$

$$\Gamma_d = \Gamma \tag{7.27c}$$

$$\Pi_d \phi_1^s = \sqrt{\frac{\delta}{2}} \Pi \phi_1^s,\ \Pi_d \phi_2^s = \sqrt{\frac{\delta}{2}} \Pi \phi_2^s + \frac{1}{\delta}\sqrt{\frac{\delta}{2}} \Pi \phi_1^s \tag{7.27d}$$

Proof. Under the stated assumptions, we have that Theorem 14 and Theorem 15 hold. We first consider diagonalizable S and S_d, namely $S = S_m$, $S_d = S_m^d$. Based on simple manipulations of discrete Sylvester equations (7.15) on the eigenpair (λ_i^d, ϕ_i^d) of S_d, the discrete regulator gains (Γ_d, Π_d) can be found as:

$$\Pi_d \phi_i^d = \lambda_i^d (\lambda_i^d I - \mathcal{A}_d)^{-1} (\mathcal{B}_d \Gamma_d + P_d) \phi_i^d \tag{7.28a}$$

$$\Gamma_d \phi_i^d = [\mathcal{G}_d(\lambda_i^d)]^{-1} [Q_d - \mathcal{T}_d(\lambda_i^d) F_d] \phi_i^d \tag{7.28b}$$

where $\mathcal{G}_d(\lambda_i^d)$ and $\mathcal{T}_d(\lambda_i^d)$ are discrete-time transfer functions $\mathcal{G}_d(z)$ (from u_k to y_{ck}) and $\mathcal{T}_d(z)$ (from d_k to y_{ck}) with z evaluated at $z = \lambda_i^d$. Since $\lambda_i^d \in \rho(\mathcal{A}_d) \backslash \{-1\}$, $\mathcal{G}_d(\lambda_i^d)$ and $\mathcal{T}_d(\lambda_i^d)$ are always solvable.

Similarly, one can solve for the continuous regulator gains (Γ, Π) from continuous Sylvester equations (7.13) as below:

$$\Pi \phi_i^s = (\lambda_i^s I - \mathcal{A})^{-1} (\mathcal{B}\Gamma + P) \phi_i^s \tag{7.29a}$$

$$\Gamma \phi_i^s = [\mathcal{G}_c(\lambda_i^s)]^{-1} [Q - \mathcal{T}_c(\lambda_i^s) F] \phi_i^s \tag{7.29b}$$

where $\mathcal{G}_c(\lambda_i^s)$ and $\mathcal{T}_c(\lambda_i^s)$ are continuous-time transfer functions $\mathcal{G}_c(s)$ (from $u(t)$ to $y_c(t)$) and $\mathcal{T}_c(s)$ (from $d(t)$ to $y_c(t)$) with s evaluated at $s = \lambda_i^s$. Since $\lambda_i^s \in \rho(\mathcal{A}) \backslash \{\delta\}$, $\mathcal{G}_c(\lambda_i^s)$ and $\mathcal{T}_c(\lambda_i^s)$ are always solvable.

To proceed with the proof, one need to show the following relationships between the continuous- and discrete-time transfer functions:

$$\mathcal{G}_c(\lambda_i^s) = \mathcal{G}_d(\lambda_i^d),\ \mathcal{T}_c(\lambda_i^s) = \mathcal{T}_d(\lambda_i^d),\ \forall \lambda_i^s \in \sigma(S),\ \forall \lambda_i^d \in \sigma(S_d) \tag{7.30}$$

Under Assumptions 10-11 and Corollary 1, one can deduce that $\lambda_i^d = \frac{\delta + \lambda_i^s}{\delta - \lambda_i^s} \in \sigma(S_d) \subset \rho(\mathcal{A}_d)$ (and $\lambda_i^d \neq -1$ since $-1 \notin \sigma(S_d)$) which coincides with the bilinear mapping

$z = \frac{\delta+s}{\delta-s}$ with $z = \lambda_i^d$ and $s = \lambda_i^s$. Combining Eq.(7.7), we can infer Eq.(7.30). Note that $F = F_d$ and $Q = Q_d$, so one can finally conclude that $\Gamma = \Gamma_d$.

With the relationship between λ_i^s and λ_i^d shown in Corollary 1, we establish the correspondence between Π and Π_d in Eq.(7.29a) and Eq.(7.28a) as follows:

$$\begin{aligned}
\Pi_d \phi_i^d &= \lambda_i^d (\lambda_i^d I - \mathcal{A}_d)^{-1} (\mathcal{B}_d \Gamma_d + P_d) \phi_i^d \\
&= \lambda_i^d [2\delta(\delta - \lambda_i^s)^{-1} - 2\delta(\delta I - A)^{-1}]^{-1} \sqrt{2\delta} (\delta I - A)^{-1} (\mathcal{B}\Gamma + P) \phi_i^d \\
&= \lambda_i^d \sqrt{2\delta}^{-1} (\delta - \lambda_i^s)(\lambda_i^s I - \mathcal{A})^{-1} (\mathcal{B}\Gamma + P) \phi_i^s \quad (\text{with } \phi_i^d = \phi_i^s) \\
&= \lambda_i^d \sqrt{2\delta}^{-1} (\delta - \lambda_i^s) \Pi \phi_i^s \\
&= \frac{\sqrt{2\delta}\lambda_i^d}{\lambda_i^d + 1} \Pi \phi_i^s
\end{aligned} \tag{7.31}$$

Therefore, in the case that S and S_d are diagonalizable the solutions of Sylvester regulator equations in continuous- and discrete-time settings are related by Eq.(7.27a)-(7.27b).

Then we consider non-diagonalizable S and S_d, i.e., $S = S_n$ and $S_d = S_n^d$. By recalling Assumption 10, we can show that there are two eigenvectors associated with the eigenvalue 0 of S, and two eigenvectors associated with the eigenvalue 1 of S_d. Considering the multiplicity of eigenvalues 0 and 1, there are a standard eigenvector and a generalized eigenvector associated with S and S_d respectively.

From Corollary 1, we have the following relationship of the first (standard) eigenvector in S and S_d:

$$S\phi_1^s = \lambda^s \phi_1^s, \; S_d \phi_1^d = \lambda^d \phi_1^d \tag{7.32}$$

In this case, the solutions of continuous and discrete Sylvester regulator equations are related by Eq.(7.27a)-(7.27b). The proof is the same as the previous one so it is omitted. More specifically, due to $\lambda^s = 0$ and $\lambda^d = 1$, we have the following for the first (standard) eigenpair:

$$\Gamma_d = \Gamma \tag{7.33a}$$

$$\Pi_d \phi_1^s = \sqrt{\frac{\delta}{2}} \Pi \phi_1^s \tag{7.33b}$$

Now we need to fully consider the the action of Π and Π_d on the generalized eigenvectors of S and S_d, namely ϕ_2^s and ϕ_2^d as follows:

$$\Pi_d S_d \phi_2^d = \mathcal{A}_d \Pi_d \phi_2^d + (\mathcal{B}_d \Gamma_d + P_d) S_d \phi_2^d \tag{7.34a}$$

$$Q_d S_d \phi_2^d = C_d \Pi_d \phi_2^d + (\mathcal{D}_d \Gamma_d + \Theta_{cd}) S_d \phi_2^d \tag{7.34b}$$

By substituting $S_d \phi_2^d = \lambda^d \phi_2^d + \phi_1^d$ into Eq.(7.34), a directly algebraic manipulation leads to

$$\Pi_d \phi_2^d = (\lambda^d I - \mathcal{A}_d)^{-1} (\mathcal{B}_d \Gamma_d + P_d) \lambda^d \phi_2^d - (\lambda^d I - \mathcal{A}_d)^{-1} \lambda_d^{-1} \mathcal{A}_d \Pi_d \phi_1^d \tag{7.35a}$$

$$\Gamma_d \phi_2^d = [\mathcal{G}_d(\lambda^d)]^{-1} [Q_d - \mathcal{T}_d(\lambda^d) F_d] \phi_2^d + [\mathcal{G}_d(\lambda^d)]^{-1} [\mathcal{G}_d{}^{(2)}(\lambda^d) \Gamma_d + \mathcal{T}_d^{(2)}(\lambda^d) F_d] \phi_1^d \tag{7.35b}$$

where $\mathcal{G}_d^{(2)}(\lambda^d)$ and $\mathcal{T}_d^{(2)}(\lambda^d)$ represent the discrete-time "second-order" transfer functions $\mathcal{G}_d^{(2)}(z)$ and $\mathcal{T}_d^{(2)}(z)$ with z evaluated at $z = \lambda^d$. Similarly, by inserting the generalized eigenvector ϕ_2^s (with $S\phi_2^s = \lambda^s \phi_2^s + \phi_1^s$) in the continuous-time Sylvester regulator equations (7.13), it is straightforward to attain

$$\Pi \phi_2^s = (\lambda^s I - \mathcal{A})^{-1} (\mathcal{B}\Gamma + P) \phi_2^s - (\lambda^s I - \mathcal{A})^{-1} \Pi \phi_1^s \tag{7.36a}$$

$$\Gamma \phi_2^s = [\mathcal{G}_c(\lambda^s)]^{-1} [Q - \mathcal{T}_c(\lambda^s) F] \phi_2^s + [\mathcal{G}_c(\lambda^s)]^{-1} [\mathcal{G}_c{}^{(2)}(\lambda^s) \Gamma + \mathcal{T}_c^{(2)}(\lambda^s) F] \phi_1^s \tag{7.36b}$$

where $\mathcal{G}_c^{(2)}(\lambda^s)$ and $\mathcal{T}_c^{(2)}(\lambda^s)$ denote continuous-time "second-order" transfer functions $\mathcal{G}_c^{(2)}(s)$ and $\mathcal{T}_c^{(2)}(s)$ with s evaluated at $s = \lambda^s$. Applying the link (7.24) between discrete- and continuous-time "second-order" transfer functions, one can readily conclude that $\Gamma_d = \Gamma$ (i.e., Eq.(7.27c)) holds.

To show the relationship between Π_d and Π, we can substitute Eq.(7.28a) with $\phi_i^d = \phi_1^d$ into Eq.(7.35a) as follows:

$$\begin{aligned}
\Pi_d \phi_2^d &= (\lambda^d I - \mathcal{A}_d)^{-1} (\mathcal{B}_d \Gamma_d + P_d) \lambda^d \phi_2^d - (\lambda^d I - \mathcal{A}_d)^{-1} \mathcal{A}_d (\lambda^d I - \mathcal{A}_d)^{-1} (\mathcal{B}_d \Gamma_d + P_d) \phi_1^d \\
&= \lambda^d (\lambda^d I - \mathcal{A}_d)^{-1} (\mathcal{B}_d \Gamma_d + P_d) \phi_2^d + (\lambda^d I - \mathcal{A}_d)^{-1} (\mathcal{B}_d \Gamma_d + P_d) \phi_1^d \\
&\quad - (\lambda^d I - \mathcal{A}_d)^{-1} \Pi_d \phi_1^d
\end{aligned} \tag{7.37}$$

Based on the chain rule of the generalized eigenvectors ϕ_2^s and ϕ_2^d shown in Corollary 1, we have the following:

$$\phi_2^d = \frac{(\delta - \lambda^s)(\delta - S)}{2\delta} \phi_2^s = \frac{(\delta - \lambda^s)^2}{2\delta} \phi_2^s - \frac{\delta - \lambda^s}{2\delta} \phi_1^s \tag{7.38}$$

which can be further substituted in Eq.(7.37). Through simple algebraic manipulation, one can rewrite Eq.(7.37) as:

$$\begin{aligned}\Pi_d\phi_2^s =&\sqrt{\frac{\delta}{2}}(-\mathcal{A})^{-1}(\mathcal{B}\Gamma_d + P)\phi_2^s + \frac{2}{\delta}\sqrt{\frac{\delta}{2}}(-\mathcal{A})^{-1}(\mathcal{B}\Gamma_d + P)\phi_1^s \\ &- \frac{1}{\delta}(-\mathcal{A})^{-1}(\delta I - \mathcal{A})\Pi_d\phi_1^d \\ =&\sqrt{\frac{\delta}{2}}\Pi\phi_2^s + \sqrt{\frac{1}{2\delta}}\Pi\phi_1^s \end{aligned} \tag{7.39}$$

The last expression is induced by the use of Eq.(7.36a). This completes the whole proof. ■

Remark 21. *With the 1-1 correspondence (7.27), it can be seen that the solutions of discrete Sylvester equations (7.15) and its continuous counterparts (7.13) are linked via the Cayley-Tustin bilinear transform. Hence, one can solve for* (Γ_d, Π_d) *from discrete-time Sylvester equations to attain* (Γ, Π) *for continuous-time regulator design and vice versa.*

Regarding the solvability of the regulator equations, the non-resonance conditions of finite-dimensional systems have been generalized for the continuous-time linear infinite-dimensional systems in [10]. Following that, we establish the non-resonance conditions for the discrete state feedback regulator equations (7.15).

Lemma 4. *Let Assumptions 10-12 hold. The regulator equations (7.13) are solvable for every choice of* P *and* Q *if and only if no eigenvalue of* S *is a transmission zero of continuous-time plant (7.1), i.e.,* $\det \mathcal{G}_c(\lambda_i^s) \neq 0, \forall \lambda_i^s \in \sigma(S)$.

Proof. For the diagonalizable S having all eigenvalues on the imaginary axis (namely $S = S_m$), the proof is shown in [10, Corollary V.1]. For the case of non-diagonalizable S, i.e., $S = S_n$, we observe that Π is always solvable under stated assumptions, and the solvability of Γ depends on the invertibility of $\mathcal{G}_c(\lambda^s)$ (indeed $\lambda^s = 0$ and for simplicity we drop the subscript i in λ_i^s in this case) as shown in Eq.(7.36) that is same as the case of diagonalizable S but through more complicated manipulation. ■

In the similar manner, we can prove the non-resonance solvability criteria for discrete-time regulator equations (7.15) with a proper choice of the time discretization interval.

Corollary 2. *Let Assumptions 10-12 hold. The regulator equations (7.15) are solvable for every choice of P_d and Q_d if and only if no eigenvalue of S_d is a transmission zero of discrete-time plant (7.4), i.e.,* $\det \mathcal{G}_d(\lambda_i^d) \neq 0, \forall \lambda_i^d \in \sigma(S_d)$.

By combining Lemma 4 and Corollary 2, we show that the non-resonance conditions stay invariant under the Cayley-Tustin transformation.

Theorem 17. *Let Assumptions 10-12 hold. The non-resonance conditions in Lemma 4 and Corollary 2 are equivalent under the Cayley-Tustin bilinear transformation, and regulator equations (7.15) are solvable if and only if regulator equations (7.13) are solvable.*

Proof. Under Assumptions 10-11 and Corollary 1, we have $\lambda_i^d = \frac{\delta+\lambda_i^s}{\delta-\lambda_i^s} \in \sigma(S_d) \subset \rho(\mathcal{A}_d)$, $\lambda_i^s = \frac{\lambda_i^d-1}{\lambda_i^d+1}\delta \in \sigma(S) \subset \rho(\mathcal{A})$, where $\delta \notin \sigma(S)$ and $-1 \notin \sigma(S_d)$ with a proper choice of δ, which indicates $\mathcal{G}_c(\lambda_i^s) = \mathcal{G}_d(\lambda_i^d)$, for all $\lambda_i^s \in \sigma(S)$ and $\lambda_i^d \in \sigma(S_d)$ by using the 1-1 correspondence in continuous- and discrete-time transfer functions (7.7). Thus we have $\det \mathcal{G}_c(\lambda_i^s) \neq 0, \forall \lambda_i^s \in \sigma(S)$ if and only if $\det \mathcal{G}_d(\lambda_i^d) \neq 0, \forall \lambda_i^d \in \sigma(S_d)$. The proof is completed by combining Lemma 4 and Corollary 2. ■

7.3.4 Stabilizing feedback control gain

The 1-1 correspondence of exponential (strong) stability of the pairs $(\mathcal{A}, \mathcal{B})$ and $(\mathcal{A}_d, \mathcal{B}_d)$ has been addressed by [23]. In this chapter, we will provide a novel way to determine the discrete stabilizing feedback controller gain by finding its correspondence relationship with the associated continuous counterpart.

Lemma 5. *Given that $\mathcal{A}_c$ is an infinitesimal generator of the C_0-semigroup $\mathbb{T}_{\mathcal{A}_c}(t)$ on the Hilbert space, $\mathbb{T}_{\mathcal{A}_c}(t)$ is exponentially stable if and only if there exists a non-negative self-adjoint operator Q_c such that [4, The. 5.1.3]:*

$$\mathcal{A}_c^* Q_c + Q_c \mathcal{A}_c + M_c = 0, \ \text{on } \mathcal{D}(\mathcal{A}_c) \tag{7.40}$$

with $Q_c(\mathcal{D}(\mathcal{A}_c)) \subset \mathcal{D}(\mathcal{A}_c^*)$, where M_c is a positive definite design parameter.

With $\mathcal{A}_c = \mathcal{A} + \mathcal{B}K$, the Lemma 5 is linked to the following theorem.

Theorem 18. *Let Assumption 12 hold. If there exists a non-negative self-adjoint operator Q_c that solves the following algebraic Riccati equation [102, The. 1]:*

$$\mathcal{A}^* Q_c + Q_c \mathcal{A} + M_c - 2Q_c \mathcal{B}\mathcal{B}^* Q_c = 0, \text{ on } \mathcal{D}(\mathcal{A}) \tag{7.41}$$

where M_c is a positive definite design parameter such that $Q_c(\mathcal{D}(\mathcal{A})) \in \mathcal{D}(\mathcal{A}^*)$ and the stabilizing feedback control gain is $K = -\mathcal{B}^* Q_c$, then the closed-loop system is exponentially stable, i.e., $\mathcal{A} + \mathcal{B}K$ generates an exponentially stable C_0-semigroup. It can be shown that the Eq.(7.41) is equivalent to Eq.(7.40) by taking $K = -\mathcal{B}^* Q_c$ and $Q_c(\mathcal{D}(\mathcal{A})) \subset \mathcal{D}(\mathcal{A}^*)$, see [102]. Motivated by this, we aim at proposing a novel way to determine the discrete stabilizing control gain K_d.

By [4, Exe. 4.30], it can be shown that the operator Q_c that is the solution of the continuous-time Lyapunov equation (7.40) and Riccati equation (7.41) coincides with the solution Q_{cd} of the discrete-time Lyapunov equation:

$$\mathcal{A}_{cd}^* Q_{cd} \mathcal{A}_{cd} - Q_{cd} + M_{cd} = 0, \text{ on } \mathcal{X} \tag{7.42}$$

where M_{cd} is a positive definite design parameter such that $Q_{cd} \in \mathcal{L}(\mathcal{X})$.
To show that the discrete- and continuous-time Lyapunov equations share the same solution, we propose the following proposition.

Proposition 7. *Let Assumption 12 hold. Given that $M_c = C^* N_c C$, $M_{cd} = C_{cd}^* N_c C_{cd}$, $\mathcal{A}_{cd} = -I + 2\delta(\delta - \mathcal{A}_c)^{-1}$ and $C_{cd} = \sqrt{2\delta} C(\delta - \mathcal{A}_c)^{-1}$ by the Cayley-Tustin transform, the discrete Lyapunov equation (7.42) and its continuous version (7.40) share the same solution, i.e., $Q_c = Q_{cd}$, where N_c is a positive definite design parameter.*

Proof. Stemming from the continuous Lyapunov equation (7.40), we can demonstrate the following:

$$\begin{aligned}
& \mathcal{A}_c^* Q_c + Q_c \mathcal{A}_c + C^* N_c C = 0 \\
\Leftrightarrow \quad & -2\delta(\delta I - \mathcal{A}_c)^* Q_c - 2\delta Q_c(\delta I - \mathcal{A}_c) + 4\delta^2 Q_c + 2\delta C^* N_c C = 0 \\
\Leftrightarrow \quad & -2\delta Q_c(\delta I - \mathcal{A}_c)^{-1} - 2\delta[(\delta I - \mathcal{A}_c)^{-1}]^* Q_c + 4\delta^2[(\delta I - \mathcal{A}_c)^{-1}]^* Q_c(\delta I - \mathcal{A}_c)^{-1} \\
& + 2\delta[(\delta I - \mathcal{A}_c)^{-1}]^* C^* N_c C(\delta I - \mathcal{A}_c)^{-1} = 0 \\
\Leftrightarrow \quad & [-I + 2\delta(\delta I - \mathcal{A}_c)^{-1}]^* Q_c[-I + 2\delta(\delta I - \mathcal{A}_c)^{-1}] - Q_c + C_{cd}^* N_c C_{cd} = 0
\end{aligned}$$

$$\Leftrightarrow \qquad \mathcal{A}_{cd}^* Q_c \mathcal{A}_{cd} - Q_c + C_{cd}^* N_c C_{cd} = 0$$

The last expression implies that $Q_c = Q_{cd}$. ■

Then, we further investigate the link between solutions of the discrete-time Lyapunov equation and Riccati equation by the following corollary.

Corollary 3. *Let Assumption 12 hold. If there exist the non-negative operator Q_{cd} that solves the following discrete-time algebraic Riccati equation:*

$$\mathcal{A}_d^* Q_{cd} \mathcal{A}_d - Q_{cd} + C_d^* N_c C_d - K_d^* \left(2I + \mathcal{B}_d^* Q_{cd} \mathcal{B}_d + \mathcal{D}_d^* N_c \mathcal{D}_d\right) K_d = 0, \text{ on } \mathcal{X} \quad (7.43)$$

where N_c is a positive definite design parameter, then the strongly stabilizing feedback control gain is $K_d = -(I + \mathcal{B}_d^* Q_{cd} \mathcal{B}_d + \mathcal{D}_d^* N_c \mathcal{D}_d)^{-1} (\mathcal{B}_d^* Q_{cd} \mathcal{A}_d + \mathcal{D}_d^* N_c C_d)$.

As for the proof, one can take $\mathcal{A}_{cd} = \mathcal{A}_d + \mathcal{B}_d K_d$, $C_{cd} = C_d + \mathcal{D}_d K_d$ and $K_d = -(I + \mathcal{B}_d^* Q_{cd} \mathcal{B}_d + \mathcal{D}_d^* N_c \mathcal{D}_d)^{-1} (\mathcal{B}_d^* Q_{cd} \mathcal{A}_d + \mathcal{D}_d^* N_c C_d)$ in Eq.(7.42) which can be further simplified as Eq.(7.43).

Remark 22. *The continuous- and discrete-time stabilizing feedback control gains are given by: $K = -\mathcal{B}^* Q_c$ and $K_d = -(I + \mathcal{B}_d^* Q_{cd} \mathcal{B}_d + \mathcal{D}_d^* N_c \mathcal{D}_d)^{-1} (\mathcal{B}_d^* Q_{cd} \mathcal{A}_d + \mathcal{D}_d^* N_c C_d)$ where $Q_c = Q_{cd}$, so we can solve for discrete Q_{cd} and apply it for the construction of continuous K, and vice versa.*

7.4 Error feedback regulation

In this section, after a brief review of the continuous-time error feedback regulator design, the discrete-time error feedback regulator design is proposed. More specifically, the discrete error feedback output regulator equations are constructed and proved.

7.4.1 Continuous-time error feedback regulator

In comparison to the discrete error feedback regulator design, the continuous counterpart is reviewed shortly. A continuous-time error feedback regulator design is achieved by finding a regulator taking the following form:

$$\dot{r}(t) = \mathcal{G}_1 r(t) + \mathcal{G}_2 e(t), \; r(0) = r_0 \quad (7.44a)$$

$$u(t) = Hr(t) \tag{7.44b}$$

where $r(t) \in \Omega = \mathcal{X} \bigoplus \mathbb{C}^{n_q}$ for $t \in [0, +\infty)$, $\mathcal{G}_1 \in \mathcal{L}(\Omega)$, $\mathcal{G}_2 \in \mathcal{L}(Y, \Omega)$ and $H \in \mathcal{L}(\Omega, U)$, and only the error signal $e(t)$ is known in order to satisfy the following conditions:

(c3) The system

$$\dot{x}(t) = \mathcal{A}x(t) + \mathcal{B}Hr(t) \tag{7.45a}$$

$$\dot{r}(t) = \mathcal{G}_2 \mathcal{C}x(t) + \mathcal{G}_1 r(t) \tag{7.45b}$$

is exponentially stable when $q \equiv 0$, which implies $[\mathcal{A} \;\; \mathcal{B}H; \mathcal{G}_2\mathcal{C} \;\; \mathcal{G}_1]$ is an infinitesimal generator of an exponentially stable C_0-semigroup.

(c4) The tracking error $e(t) \to 0$ as $t \to +\infty$ for any given $x_0 \in \mathcal{X}$, $r_0 \in \Omega$ and $q^0 \in \mathbb{C}^{n_q}$.

To solve the continuous-time error feedback regulator design problem, the following theorem is often utilized.

Theorem 19. *Let Assumptions 11-13 hold. The continuous-time error feedback regulation problem is solvable if and only if there exist mappings $\Pi \in \mathcal{L}(\mathbb{C}^{n_q}, \mathcal{X})$ with $\Pi D(S) \subset D(\mathcal{A})$ and $\Gamma \in \mathcal{L}(\mathbb{C}^{n_q}, U)$ such that the following Sylvester equations hold [10, The. IV.2]:*

$$\Pi S = \mathcal{A}\Pi + \mathcal{B}\Gamma + P \tag{7.46a}$$

$$\mathcal{C}\Pi = Q \tag{7.46b}$$

where $P = \Xi F$, and $L = \Gamma - K\Pi$. With Π and Γ, the error feedback regulator is found by:

$$\dot{r}(t) = \mathcal{G}_1 r(t) + \mathcal{G}_2 e(t) \tag{7.47a}$$

$$u(t) = Hr(t) \tag{7.47b}$$

where $r(t) \in \Omega = \mathcal{X} \bigoplus \mathbb{C}^{n_q}$ and

$$\mathcal{G}_1 = \begin{bmatrix} \mathcal{A} + \mathcal{B}K - G_1\mathcal{C} & P + \mathcal{B}(\Gamma - K\Pi) + G_1 Q \\ -G_2\mathcal{C} & S + G_2 Q \end{bmatrix}$$

$$\mathcal{G}_2 = \begin{bmatrix} G_1 \\ G_2 \end{bmatrix}, H = \begin{bmatrix} K & \Gamma - K\Pi \end{bmatrix}$$

Here $K \in \mathcal{L}(\mathcal{X}, U)$, $G_1 \in \mathcal{L}(Y, \mathcal{X})$ *and* $G_2 \in \mathcal{L}(Y, \mathbb{C}^{n_q})$ *such that* $K \in \mathcal{L}(\mathcal{X}, U)$ *is an exponentially stabilizing feedback gain for the pair* $(\mathcal{A}, \mathcal{B})$ *and* $\mathcal{G}_2 = \begin{bmatrix} G_1 \\ G_2 \end{bmatrix}$ *is an exponentially stabilizing output injection gain for the pair* $\left(\begin{bmatrix} \mathcal{A} & P \\ 0 & S \end{bmatrix}, \begin{bmatrix} C & -Q \end{bmatrix} \right)$.

7.4.2 Discrete-time error feedback regulator

Discrete-time error feedback regulator design: Find a regulator having the following form:

$$r_k = \mathcal{G}_{1d} r_{k-1} + \mathcal{G}_{2d} e_{k-1}, \ k \geq 1 \tag{7.48a}$$

$$u_k = H_d r_k \tag{7.48b}$$

where $r_k \in \Omega = \mathcal{X} \bigoplus \mathbb{C}^{n_q}$, Ω is a Hilbert space, $\mathcal{G}_{1d} \in \mathcal{L}(\Omega)$, $\mathcal{G}_{2d} \in \mathcal{L}(Y, \Omega)$ and $H_d \in \mathcal{L}(\Omega, U)$, where only the error signal e_k is available, such that the following conditions hold:

(C3) The system

$$x_k = \mathcal{A}_d x_{k-1} + \mathcal{B}_d H_d r_k \tag{7.49a}$$

$$r_{k+1} = \mathcal{G}_{2d} \mathcal{C}_d x_{k-1} + (\mathcal{G}_{1d} + \mathcal{G}_{2d} \mathcal{D}_d H_d) r_k \tag{7.49b}$$

is strongly stable when $q_k \equiv 0$, which means $[\mathcal{A}_d \ \ \mathcal{B}_d H_d; \ \mathcal{G}_{2d} \mathcal{C}_d \ \ (\mathcal{G}_{1d} + \mathcal{G}_{2d} \mathcal{D}_d H_d)]$ is a strongly stable operator.

(C4) The tracking error $e_k \to 0$ as $k \to +\infty$ for any given $x_0 \in \mathcal{X}$, $r_0 \in \Omega$ and $q^0 \in \mathbb{C}^{n_q}$.

To address the discrete-time error feedback regulation problem, we propose the following theorem.

Theorem 20. *Under Assumptions 11-13, the discrete error feedback regulation problem is solvable if and only if there exist mappings* $\Pi_d \in \mathcal{L}(\mathbb{C}^{n_q}, \mathcal{X})$ *and* $\Gamma_d \in \mathcal{L}(\mathbb{C}^{n_q}, U)$ *such that the following discrete Sylvester equations hold:*

$$\Pi_d S_d = \mathcal{A}_d \Pi_d + (\mathcal{B}_d \Gamma_d + P_d) S_d \tag{7.50a}$$

$$Q_d S_d = C_d \Pi_d + (\mathcal{D}_d \Gamma_d + \Theta_{cd}) S_d \tag{7.50b}$$

where $P_d = \Xi_d F_d$, $\Theta_{cd} = \Upsilon_d F_d$, *and the discrete error feedback control law* u_k *can be computed as follows:*

$$r_k = \mathcal{G}_{1d} r_{k-1} + \mathcal{G}_{2d} e_{k-1},\ k \geq 1 \tag{7.51a}$$

$$u_k = H_d r_k \tag{7.51b}$$

where $r_k \in \Omega = \mathcal{X} \bigoplus \mathbb{C}^{n_q}$ *and*

$$\mathcal{G}_{1d} = \begin{bmatrix} \mathcal{A}_d + \mathcal{B}_d K_d - G_{1d}\iota & P_d + \mathcal{B}_d L_d - G_{1d}\nu \\ -G_{2d}\iota & S_d - G_{2d}\nu \end{bmatrix} \tag{7.52a}$$

$$\mathcal{G}_{2d} = \begin{bmatrix} G_{1d} \\ G_{2d} \end{bmatrix},\ H_d = [K_d \quad L_d] \tag{7.52b}$$

where $\iota = \mathcal{C}_d + \mathcal{D}_d K_d$, $\nu = \mathcal{D}_d L_d + \Theta_{cd} - Q_d$, $L_d = \Gamma_d - K_d \Pi_d S_d^{-1}$, $K_d \in \mathcal{L}(\mathcal{X}, U)$, $G_{1d} \in \mathcal{L}(Y, \mathcal{X})$ *and* $G_{2d} \in \mathcal{L}(Y, \mathbb{C}^{n_q})$, *such that* $\mathcal{A}_d + \mathcal{B}_d K_d$ *is a strongly stable operator and* $\mathcal{G}_{2d} = \begin{bmatrix} G_{1d} \\ G_{2d} \end{bmatrix}$ *is a strongly stabilizing output injection gain for the pair* $\left(\begin{bmatrix} \mathcal{A}_d & P_d \\ 0 & S_d \end{bmatrix}, \begin{bmatrix} \mathcal{C}_d & \Theta_{cd} - Q_d \end{bmatrix} \right)$.

Proof. Let us prove sufficiency first. In the regulator equation (7.51), one can take $r_k = \begin{bmatrix} \hat{x}_{k-1} \\ \hat{q}_k \end{bmatrix} \in \Omega$ as the estimated plant and exogenous states, which leads to

$$\begin{aligned}
\begin{bmatrix} \hat{x}_{k-1} \\ \hat{q}_k \end{bmatrix} &= \mathcal{G}_{1d} \begin{bmatrix} \hat{x}_{k-2} \\ \hat{q}_{k-1} \end{bmatrix} + \mathcal{G}_{2d} e_{k-1} \\
&= \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix} \begin{bmatrix} \hat{x}_{k-2} \\ \hat{q}_{k-1} \end{bmatrix} \\
&\quad + \begin{bmatrix} G_{1d} \\ G_{2d} \end{bmatrix} [\mathcal{C}_d x_{k-2} + \mathcal{D}_d u_{k-1} + \Theta_{cd} q_{k-1} - Q_d q_{k-1}] \\
&= \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix} \begin{bmatrix} \hat{x}_{k-2} \\ \hat{q}_{k-1} \end{bmatrix} + \begin{bmatrix} G_{1d}\mathcal{D}_d \\ G_{2d}\mathcal{D}_d \end{bmatrix} u_{k-1} \\
&\quad + \begin{bmatrix} G_{1d}\mathcal{C}_d & G_{1d}(\Theta_{cd} - Q_d) \\ G_{2d}\mathcal{C}_d & G_{2d}(\Theta_{cd} - Q_d) \end{bmatrix} \begin{bmatrix} x_{k-2} \\ q_{k-1} \end{bmatrix} \\
&= \begin{bmatrix} G_{11} + G_{1d}\mathcal{D}_d K_d & G_{12} + G_{1d}\mathcal{D}_d L_d \\ G_{21} + G_{2d}\mathcal{D}_d K_d & G_{22} + G_{2d}\mathcal{D}_d L_d \end{bmatrix} \begin{bmatrix} \hat{x}_{k-2} \\ \hat{q}_{k-1} \end{bmatrix} \\
&\quad + \begin{bmatrix} G_{1d}\mathcal{C}_d & G_{1d}(\Theta_{cd} - Q_d) \\ G_{2d}\mathcal{C}_d & G_{2d}(\Theta_{cd} - Q_d) \end{bmatrix} \begin{bmatrix} x_{k-2} \\ q_{k-1} \end{bmatrix}
\end{aligned} \tag{7.53}$$

$$u_k = [K_d \ \ L_d] \begin{bmatrix} \hat{x}_{k-1} \\ \hat{q}_k \end{bmatrix} \tag{7.54}$$

where $\mathcal{G}_{1d} = \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix}$. Then, we substitute u_k in Eq.(7.54) back into the plant (7.4a) and exo-system (7.9a) as follows:

$$\begin{bmatrix} x_{k-1} \\ q_k \end{bmatrix} = \begin{bmatrix} \mathcal{A}_d & P_d \\ 0 & S_d \end{bmatrix} \begin{bmatrix} x_{k-2} \\ q_{k-1} \end{bmatrix} + \begin{bmatrix} \mathcal{B}_d K_d & \mathcal{B}_d L_d \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \hat{x}_{k-2} \\ \hat{q}_{k-1} \end{bmatrix} \tag{7.55}$$

By combining Eq.(7.53) and Eq.(7.55) and introducing two estimation errors $e_{x_{k-1}} = x_{k-1} - \hat{x}_{k-1}$ and $e_{q_k} = q_k - \hat{q}_k$, one can attain:

$$\begin{aligned} \begin{bmatrix} e_{x_{k-1}} \\ e_{q_k} \end{bmatrix} &= \begin{bmatrix} \mathcal{A}_d - G_{1d}C_d & P_d - G_{1d}(\Theta_{cd} - Q_d) \\ -G_{2d}C_d & S_d - G_{2d}(\Theta_{cd} - Q_d) \end{bmatrix} \begin{bmatrix} x_{k-2} \\ q_{k-1} \end{bmatrix} \\ &+ \begin{bmatrix} (\mathcal{B}_d - G_{1d}\mathcal{D}_d)K_d - G_{11} & (\mathcal{B}_d - G_{1d}\mathcal{D}_d)L_d - G_{12} \\ -G_{2d}\mathcal{D}_d K_d - G_{21} & -G_{2d}\mathcal{D}_d L_d - G_{22} \end{bmatrix} \begin{bmatrix} \hat{x}_{k-2} \\ \hat{q}_{k-1} \end{bmatrix} \end{aligned} \tag{7.56}$$

Through direct calculation, one can have the following homogeneous difference equation for describing the error evolution dynamics:

$$\begin{bmatrix} e_{x_{k-1}} \\ e_{q_k} \end{bmatrix} = \begin{bmatrix} \mathcal{A}_d - G_{1d}C_d & P_d - G_{1d}(\Theta_{cd} - Q_d) \\ -G_{2d}C_d & S_d - G_{2d}(\Theta_{cd} - Q_d) \end{bmatrix} \begin{bmatrix} e_{x_{k-2}} \\ e_{q_{k-1}} \end{bmatrix} \tag{7.57}$$

With Theorem 13 and Assumption 13, one can readily conclude that $\begin{bmatrix} e_{x_{k-1}} \\ e_{q_k} \end{bmatrix}$ converges to zero with $k \to +\infty$, and obtain the following by combining Eq.(7.56) and Eq.(7.57):

$$\mathcal{G}_{1d} = \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix} = \begin{bmatrix} \mathcal{A}_d + \mathcal{B}_d K_d - G_{1d}\iota & P_d + \mathcal{B}_d L_d - G_{1d}\nu \\ -G_{2d}\iota & S_d - G_{2d}\nu \end{bmatrix} \tag{7.58}$$

where $\iota = C_d + \mathcal{D}_d K_d$, $\nu = \mathcal{D}_d L_d + \Theta_{cd} - Q_d$. Finally, from Eq.(7.54) we have $u_k \to K_d x_{k-1} + L_d q_k$ since $\hat{x}_{k-1} \to x_{k-1}$ and $\hat{q}_k \to q_k$ as $k \to +\infty$. Based on the proof of Theorem 15, we can conclude that the tracking error $e_k \to 0$ with $k \to +\infty$ and the error feedback regulation problem is solved given that Sylvester equations (7.50) hold.

Now we prove the necessity. Similarly, let us consider $r_k = \begin{bmatrix} \hat{x}_{k-1} \\ \hat{q}_k \end{bmatrix} \in \Omega$, $e_{x_{k-1}} = x_{k-1} - \hat{x}_{k-1}$ and $e_{q_k} = q_k - \hat{q}_k$. Substituting Eq.(7.51) and Eq.(7.52) into the extended system (7.55) leads to Eq.(7.57). Then one can plug the expression of u_k in Eq.(7.51b) into the plant system (7.4), and induce that

$$\begin{aligned} x_k &= \mathcal{A}_d x_{k-1} + \mathcal{B}_d(K_d \hat{x}_{k-1} + L_d \hat{q}_k) + P_d q_k \\ &= (\mathcal{A}_d + \mathcal{B}_d K_d)x_{k-1} - \mathcal{B}_d K_d e_{x_{k-1}} + (\mathcal{B}_d L_d + P_d)q_k - \mathcal{B}_d L_d e_{q_k} \end{aligned} \tag{7.59}$$

By combining Eq.(7.57) and Eq.(7.59), we denote $\Phi_{k-1} = [x_{k-1}; e_{x_{k-1}}; e_{q_k}]$ and obtain

$$\Phi_k = \mathcal{A}_1\Phi_{k-1} + P_1 q_k + \mathcal{B}_1 u_k \tag{7.60a}$$

$$q_k = S_d q_{k-1} \tag{7.60b}$$

$$e_k = \mathcal{C}_1\Phi_{k-1} + \mathcal{D}_d u_k + (\Theta_{cd} - Q_d)q_k \tag{7.60c}$$

where $\mathcal{B}_1 = [0; 0; 0]$, $\mathcal{C}_1 = [\mathcal{C}_d \;\; 0 \;\; 0]$, $P_1 = [\mathcal{B}_d L_d + P_d; 0; 0]$ and

$$\mathcal{A}_1 = \begin{bmatrix} \mathcal{A}_d + \mathcal{B}_d K_d & -\mathcal{B}_d K_d & -\mathcal{B}_d L_d \\ 0 & \mathcal{A}_d - G_{1d}\mathcal{C}_d & P_d - G_{1d}(\Theta_{cd} - Q_d) \\ 0 & -G_{2d}\mathcal{C}_d & S_d - G_{2d}(\Theta_{cd} - Q_d) \end{bmatrix}$$

Along this line, we can define an extended mapping $\Pi_1 = [\Pi_d; 0; 0] : \mathbb{C}^{n_q} \mapsto \mathcal{X} \bigoplus Y \bigoplus \mathbb{C}^{n_q}$ and then apply Theorem 15 to design a state feedback regulator for the system (7.60) as: $u_k = K_1\Phi_{k-1} + (\Gamma_1 - K_1\Pi_1 S_d^{-1})q_k$ where $K_1 = [K_d \;\; 0 \;\; 0]$ and $\Gamma_1 = \Gamma_d$. Then the following Sylvester equations are obtained for solving the corresponding operators Π_1 and $\Gamma_1 \in \mathcal{L}(\mathbb{C}^{n_q}, Y)$:

$$\Pi_1 S_d = \mathcal{A}_1\Pi_1 + (\mathcal{B}_1\Gamma_1 + P_1)S_d \tag{7.61a}$$

$$Q_d S_d = \mathcal{C}_1\Pi_1 + (\mathcal{D}_d\Gamma_1 + \Theta_{cd})S_d \tag{7.61b}$$

Under Assumptions 11-13 and Theorem 13, it is apparent that $\mathcal{A}_1$ is strongly stable, and hence $\mathcal{A}_1 + \mathcal{B}_1 K_1$ is strongly stable due to $\mathcal{B}_1 = 0$. Taking the first component of Eq.(7.61a) and Eq.(7.61b) into consideration, one can obtain:

$$\Pi_d S_d = (\mathcal{A}_d + \mathcal{B}_d K_d)\Pi_d + (\mathcal{B}_d L_d + P_d)S_d \tag{7.62a}$$

$$Q_d S_d = \mathcal{C}_d\Pi_d + (\mathcal{D}_d\Gamma_d + \Theta_{cd})S_d \tag{7.62b}$$

which further indicates Eq.(7.50) with $L_d = \Gamma_d - K_d\Pi_d S_d^{-1}$. ■

Corollary 4. *Let Assumptions 10-13 hold. With Cayley-Tustin transform (7.5) and (7.11), the corresponding relationships between* (Γ_d, Π_d) *and* (Γ, Π) *are established as below for the continuous- and discrete-time error feedback regulator designs.*

(a). For diagonalizable $S = S_m,\ S_d = S_m^d$

$$\Gamma_d = \frac{\sqrt{2\delta}}{\delta - \lambda_i^s}\Gamma \tag{7.63a}$$

$$\Pi_d \phi_i^s = \lambda_i^d \Pi \phi_i^s \tag{7.63b}$$

(b). For non-diagonalizable $S = S_n,\ S_d = S_n^d$

$$\Gamma_d \phi_1^s = \sqrt{\frac{2}{\delta}} \Gamma \phi_1^s,\ \Gamma_d \phi_2^s = \sqrt{\frac{2}{\delta}} \Gamma \phi_2^s + \frac{1}{\delta} \sqrt{\frac{2}{\delta}} \Gamma \phi_1^s \tag{7.63c}$$

$$\Pi_d \phi_i^s = \Pi \phi_i^s \tag{7.63d}$$

Proof. Under stated assumptions, we have that Theorem 19 and Theorem 20 hold. The next proof is similar to Theorem 16 so it is omitted. ■

We note that under Assumptions 10-13, the solvability of the state feedback regulator problem is equivalent to that of the error feedback regulator problem in a continuous- or discrete-time setting. Along this line, we can establish the same non-resonance solvability criteria for continuous-time error feedback regulator equations (7.46) and discrete-time error feedback regulator equations (7.50) as in Lemma 4 and Corollary 2 by including Assumption 13. Thus we can further show that the solvability of discrete-time error feedback regulator equations is equivalent to that of the continuous analogues.

Corollary 5. *Under Assumptions 10-13, we have the following assertions: a). regulator equations (7.46) are solvable if and only if no eigenvalue of* S *is a transmission zero of continuous-time plant (7.1), i.e.,* $\det \mathcal{G}_c(\lambda_i^s) \neq 0, \forall \lambda_i^s \in \sigma(S)$*; b). regulator equations (7.50) are solvable if and only if no eigenvalue of* S_d *is a transmission zero of discrete-time plant (7.4), i.e.,* $\det \mathcal{G}_d(\lambda_i^d) \neq 0, \forall \lambda_i^d \in \sigma(S_d)$*; c). the non-resonance conditions a) and b) are equivalent under the Cayley-Tustin bilinear transformation; and d) regulator equations (7.50) are solvable if and only if regulator equations (7.46) are solvable.*

Proof. The proof is similar to Lemma 4, Corollary 2 and Theorem 17 under Assumptions 10-13, so we omit it. ■

7.4.3 Discrete stabilizing output injection gain

In this section, we will provide a new way to solve for the discrete stabilizing output injection gain $\mathcal{G}_{2d}$. First let us revisit some main results on stabilization of continuous- and discrete-time systems from [4, 102].

Lemma 6. *Given that $\mathcal{A}_o$ is an infinitesimal generator of the C_0-semigroup $\mathbb{T}_{\mathcal{A}_o}(t)$ on the Hilbert space, $\mathbb{T}_{\mathcal{A}_o}(t)$ is exponentially stable if and only if there exists a non-negative self-adjoint operator Q_o such that [4]*

$$\mathcal{A}_o Q_o + Q_o \mathcal{A}_o^* + M_o = 0, \text{ on } \mathcal{D}(\mathcal{A}_o^*) \tag{7.64}$$

with $Q_o(\mathcal{D}(\mathcal{A}_o^*)) \subset \mathcal{D}(\mathcal{A}_o)$, where M_o is a positive definite design parameter.

Given that $\mathcal{A}_o = \begin{bmatrix} \mathcal{A} & P \\ 0 & S \end{bmatrix} - \mathcal{G}_2 \begin{bmatrix} C & -Q \end{bmatrix}$, we adopt the following theorem from [102] to solve for $\mathcal{G}_2$.

Theorem 21. *Let Assumption 13 hold. If there exists a non-negative self-adjoint operator Θ_1 and Θ_2 that solve the following algebraic Riccati equations [102, The. 5]:*

$$\mathcal{A}\Theta_1 + \Theta_1 \mathcal{A}^* - 2\Theta_1 C^* C \Theta_1 + M_{o1} = 0 \tag{7.65a}$$

$$P\Theta_2 + 2\Theta_1 C^* Q \Theta_2 + M_{o2} = 0 \tag{7.65b}$$

$$S\Theta_2 + \Theta_2 S^* - 2\Theta_2 Q^* Q \Theta_2 + M_{o3} = 0 \tag{7.65c}$$

where M_{o1} and M_{o3} are positive definite operator and matrix, respectively, and M_{o2} is determined based on solutions of Θ_1 and Θ_2 to ensure that $\begin{bmatrix} M_{o1} & M_{o2} \\ M_{o2}^* & M_{o3} \end{bmatrix}$ is positive definite and $\Theta_1(\mathcal{D}(\mathcal{A}^*)) \subset \mathcal{D}(\mathcal{A})$, then $\mathcal{G}_2 = \begin{bmatrix} G_1 \\ G_2 \end{bmatrix} = \begin{bmatrix} \Theta_1 C^* \\ -\Theta_2 Q^* \end{bmatrix}$ is an exponentially stabilizing output injection gain.

By [4, Exe. 4.30], it can be shown that the operator Q_o that solves the continuous Lyapunov equation (7.64) coincides with the solution Q_{od} (namely $[\Theta_1 \;\; 0; 0 \;\; \Theta_2] = [\Theta_{1d} \;\; 0; 0 \;\; \Theta_{2d}]$ as shown in Eq.(7.67)) of the following discrete-time Lyapunov equation:

$$\mathcal{A}_{od} Q_{od} \mathcal{A}_{od}^* - Q_{od} + M_{od} = 0, \text{ on } \mathcal{X} \tag{7.66}$$

where M_{od} is a positive definite design parameter and $Q_{od} \in \mathcal{L}(\mathcal{X})$ is a non-negative self-adjoint operator. To prove that, one can take $M_o = \mathcal{B} N_o \mathcal{B}^*$ in Eq.(7.64), $M_{od} = \mathcal{B}_{od} N_o \mathcal{B}_{od}^*$, $\mathcal{B}_{od} = \sqrt{2\delta}(\delta - \mathcal{A}_o)^{-1}\mathcal{B}$ and $\mathcal{A}_{od} = -I + 2\delta(\delta - \mathcal{A}_o)^{-1}$ by using the Cayley-Tustin transform, where N_o is a positive definite design parameter.

Given that $\mathcal{A}_{od} = \begin{bmatrix} \mathcal{A}_d & P_d \\ 0 & S_d \end{bmatrix} - \mathcal{G}_{2d} \begin{bmatrix} C_d & \Theta_{cd} - Q_d \end{bmatrix}$, we provide the following discrete Riccati equations for solving the discrete stabilizing output injection gain $\mathcal{G}_{2d}$.

Corollary 6. *Let Assumption 13 hold. If there exist the non-negative operators Θ_{1d} and Θ_{2d} that solve the following discrete-time algebraic Riccati equations*

$$\mathcal{A}_d\Theta_{1d}\mathcal{A}_d^* - \Theta_{1d} + P_d\Theta_{2d}P_d^* - G_{1d}(2R_1 + C_d\Theta_{1d}C_d^* + \Omega_d\Theta_{2d}\Omega_d^*)G_{1d}^* + M_{d1} = 0 \quad (7.67a)$$

$$(P_d - G_{1d}\Omega_d)\Theta_{2d}(S_d - G_{2d}\Omega_d)^* - (\mathcal{A}_d - G_{1d}C_d)\Theta_{1d}C_d^*G_{2d}^* + M_{d2} = 0 \quad (7.67b)$$

$$S_d\Theta_{2d}S_d^* - \Theta_{2d} - G_{2d}(2R_2 + C_d\Theta_{1d}C_d^* + \Omega_d\Theta_{2d}\Omega_d^*)G_{2d}^* + M_{d3} = 0 \quad (7.67c)$$

where R_1, R_2, M_{d3} and M_{d1} are positive definite matrices and operator, respectively, and M_{d2} is determined such that $\begin{bmatrix} M_{d1} & M_{d2} \\ M_{d2}^* & M_{d3} \end{bmatrix}$ is positive definite, then $\mathcal{G}_{2d} = \begin{bmatrix} G_{1d} \\ G_{2d} \end{bmatrix} = \begin{bmatrix} (\mathcal{A}_d\Theta_{1d}C_d^* + P_d\Theta_{2d}\Omega_d^*)(R_1 + C_d\Theta_{1d}C_d^* + \Omega_d\Theta_{2d}\Omega_d^*)^{-1} \\ S_d\Theta_{2d}\Omega_d^*(R_2 + C_d\Theta_{1d}C_d^* + \Omega_d\Theta_{2d}\Omega_d^*)^{-1} \end{bmatrix}$ is a strongly stabilizing output injection gain, where $\Omega_d = \Theta_{cd} - Q_d$.

Proof. Given $Q_{od} = \text{bdiag}(\Theta_{1d}, \Theta_{2d})$, $M_{od} = \begin{bmatrix} M_{d1} & M_{d2} \\ M_{d2}^* & M_{d3} \end{bmatrix}$ and $\mathcal{A}_{od} = \begin{bmatrix} \mathcal{A}_d & P_d \\ 0 & S_d \end{bmatrix} - \mathcal{G}_{2d} \begin{bmatrix} C_d & \Theta_{cd} - Q_d \end{bmatrix}$ with $\mathcal{G}_{2d} = \begin{bmatrix} G_{1d} \\ G_{2d} \end{bmatrix}$, $G_{1d} = (\mathcal{A}_d\Theta_{1d}C_d^* + P_d\Theta_{2d}\Omega_d^*)(R_1 + C_d\Theta_{1d}C_d^* + \Omega_d\Theta_{2d}\Omega_d^*)^{-1}$ and $G_{2d} = S_d\Theta_{2d}\Omega_d^*(R_2 + C_d\Theta_{1d}C_d^* + \Omega_d\Theta_{2d}\Omega_d^*)^{-1}$, simple algebraic manipulation of Eq.(7.66) leads to Eq.(7.67). ■

Remark 23. *To solve the algebraic Riccati equations (7.67), we provide the following steps:*

1. *Initialize R_1, R_2, M_{d1} and M_{d3};*
2. *Solve for Θ_{1d} and Θ_{2d} in Eq.(7.67a) and Eq.(7.67c) using numerical iteration methods (e.g. Newton-Kleinman iteration method [283, 284, 285]);*
3. *Find M_{d2} by substituting Θ_{1d} and Θ_{2d} into Eq.(7.67b), and check the positive definiteness of $M_{od} = \begin{bmatrix} M_{d1} & M_{d2} \\ M_{d2}^* & M_{d3} \end{bmatrix}$ and $Q_{od} = bdiag(\Theta_{1d}, \Theta_{2d})$, if not return to repeat steps (1)-(2).*

Remark 24. *Given that $M_o = \mathcal{B}N_o\mathcal{B}^*$ and $M_{od} = \mathcal{B}_{od}N_o\mathcal{B}_{od}^*$, the continuous- and discrete-time stabilizing output injection gains $\mathcal{G}_2$ and $\mathcal{G}_{2d}$ can be linked by Theorem 20 and Corollary 6. Thus one can solve for $\mathcal{G}_{2d}$ and then apply the result $(\Theta_{1d}, \Theta_{2d})$ for the construction of $\mathcal{G}_2$, and vice versa.*

7.5 Simulation

To verify the effectiveness and applicability of the proposed discrete-time regulator design methods, we provide three examples including two state feedback regulator designs for a first-order hyperbolic PDE (non-spectral system) with considerations of harmonic and polynomial exogenous signals respectively, and an error feedback regulator design for a 1-D heat equation (spectral system) to realize set-point reference control.

7.5.1 Example 1: state feedback regulator design for a first-order hyperbolic PDE (non-spectral system) with consideration of harmonic reference and disturbance

Let us consider a tubular reactor system described by a first-order hyperbolic partial differential equation model as follows:

$$z_t(\zeta,t) = -vz_\zeta(\zeta,t) + \psi(\zeta)z(\zeta,t) + b(\zeta)u(t) + f(\zeta)d(t) \tag{7.68a}$$

$$z(0,t) = 0,\ z(\zeta,0) = z_0(\zeta) \tag{7.68b}$$

$$y_c(t) = C_c z(\zeta,t) \tag{7.68c}$$

where $\zeta \in [0,1]$ and $t \in [0,+\infty)$ stand for spatial and temporal coordinates, respectively. We consider bounded input and disturbance operators as $b(\zeta) = 1$, $f(\zeta) = 0.5$ with model parameters $v = 2$ and $\psi(\zeta) = \sinh(\zeta)$. A bounded output operator is considered as $C_c := \int_0^1 \frac{1}{2\epsilon_c}\mathbf{1}_{[\zeta_c-\epsilon_c,\zeta_c+\epsilon_c]}(\cdot)d\zeta$, where $\mathbf{1}_{[a,b]}(\zeta)$ denotes the spatial shaping function: $\mathbf{1}_{[a,b]}(\zeta) = \begin{cases}1, \zeta \in [a,b]\\ 0, \text{otherwise}\end{cases}$. The system operator $\mathcal{A}$ is defined as: $\mathcal{A} := -v\frac{\partial}{\partial\zeta} + \sinh(\zeta)$ with the domain $\mathcal{D}(\mathcal{A}) = \{\phi(\zeta) \in \mathcal{X} | \phi(\zeta)$ is absolutely continuous, $\frac{d\phi(\zeta)}{d\zeta} \in \mathcal{X}$ and $\phi(0) = 0\}$. In addition, $\zeta_c = 0.5$, $\epsilon_c = 0.01$, and $z_0(\zeta) = 2\sin(2\pi\zeta)$.

In this example, we focus on harmonic reference and disturbance signals, so the discrete-time exo-system is designed with the diagonalizable $S_d = S_m^d = [0.9969, 0.0784; -0.0784, 0.9969]$ (with the continuous counterpart $S = S_m = [0, 0.05\pi; -0.05\pi, 0]$), $q^0 = [0;1]$, $Q_d = [2,0]$, and $F_d = [0,1]$. Along this line, the discrete disturbance and reference signals are generated as: $d_k = \cos(0.025k\pi)$ and $y_{rk} = 2\sin(0.025k\pi)$. Apparently, we have $\sigma(S) \subset \rho(\mathcal{A})$ ensuing that Assumption 11 holds. Then, the state feedback regulator is constructed using the procedures shown in Table 7.1.

Table 7.1: Construction of the discrete state feedback regulator.

Algorithm 1: Discrete state feedback regulator
Step 1: Solve discrete Sylvester equations (7.15) for Γ_d and Π_d
Step 2: Solve for continuous stabilizing feedback gain Q_c in Eq.(7.41), and then obtain K_d to stabilize the discrete operator $\mathcal{A}_d + \mathcal{B}_d K_d$
Step 3: Determine $L_d = \Gamma_d - K_d \Pi_d S_d^{-1}$ and simulate the exo-system and plant
Step 4: Construct the discrete state feedback regulation law (7.14) and apply to the discrete system (7.4)

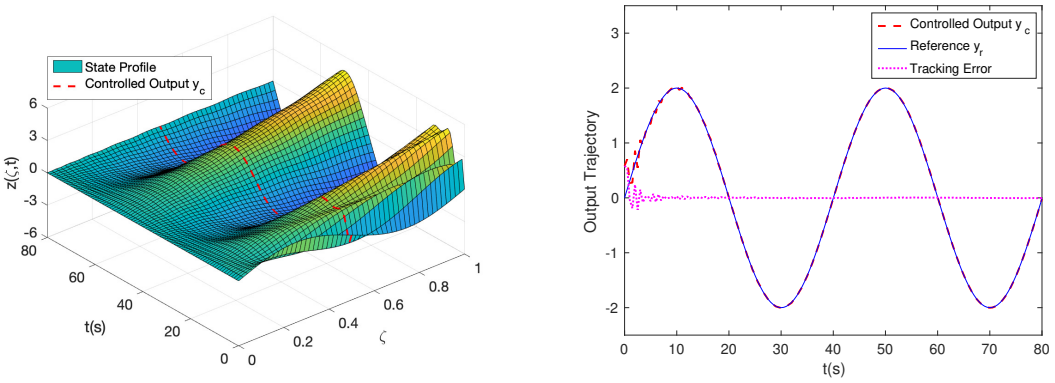
Through discrete Sylvester equations (7.15), the discrete feedforward gain can be solved as $\Gamma_d = [7.6629, -0.3479]$ and Π_d can be obtained as spatial functions correspondingly. To ensure the stability of the closed-loop system, the state feedback stabilizing gain is solved using Eq.(7.41) as bellow:

$$2\frac{dQ_c(\zeta)}{d\zeta} + \sinh(\zeta)Q_c(\zeta) + Q_c(\zeta)\sinh(\zeta) + M - 2Q_c(\zeta)^2 = 0,\ Q_c(1) = 0 \quad (7.69)$$

where M is a design parameter which is chosen as 0.001 in this example. It is straightforward to solve this ODE by the finite difference method, from which we obtain the feedback control law $u_k = K_d x_{k-1} + L_d q_k$ by performing Steps 3-4.

After 80 seconds of simulation, the regulation results are illustrated in Fig. 7.1. Specifically, the simulated time and spatial intervals are set as 0.5 second and 0.01. With the control action applied, the closed-loop state profile can follow the sinusoidal reference trend and reject undesired cosine disturbance as well. As shown in Fig. 7.1(b), it can be seen that controlled output y_c tracks the desired sinusoidal reference rapidly and the tracking error converges to zero under the closed-loop control.

In addition, the influence of the choice of the sampling time on the output regulation performance is investigated below. With the same regulation objective and overall simulation time of 80 seconds, different simulated time intervals (with ±5% and ±25% based on the chosen 0.50s) are implemented. As illustrated in Table 7.2, it is apparent that the tracking error increases approximately with the increase of discretization time interval and overall the relative error stays within a reasonable range (around 0.01).



(a) State regulation performance.

(b) Output regulation performance.

Figure 7.1: Output regulation of the transport equation with harmonic reference and disturbance signals.

Table 7.2: Comparison of regulation performance with different time discretization intervals.

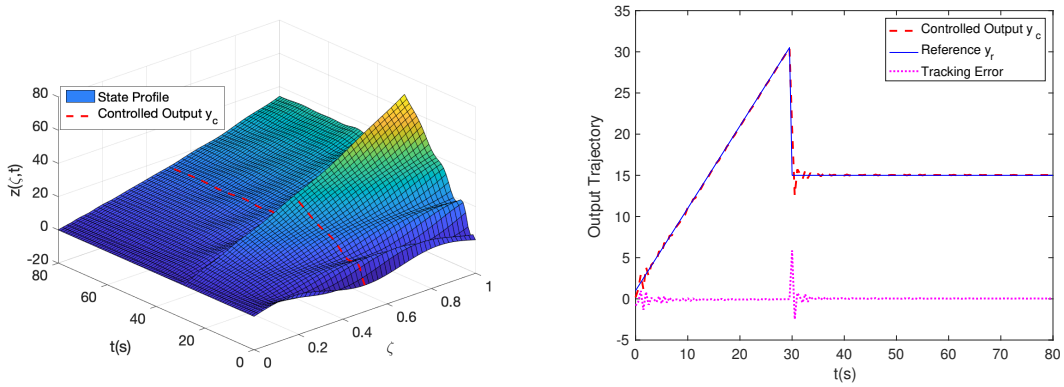
Δt (s)	0.375	0.475	0.50	0.525	0.625
Absolute Tracking Error $\lvert e_k \rvert$	0.0016	0.0009	0.0070	0.0022	0.0100

7.5.2 Example 2: state feedback regulator design for a first-order hyperbolic PDE (non-spectral system) with consideration of step-like and ramp-like reference and disturbance

In this example, we construct another state feedback regulator for the same first-order hyperbolic PDE model (7.68) with the same parameters considered in the first example. Differently to the Example 1, we aim at tracking ramp-like and step-like references by considering a non-diagonal exo-system in this example.

In this case, we consider the continuous exo-system with the non-diagonalizable $S = S_n = [0, 0; 1, 0]$, and the discrete counterpart $S_d = S_n^d = [1, 0; 0.5, 1]$ with $\Delta t = 0.5$ and $q^0 = [1; 0]$. By designing $Q_d = [1, 1]$ for $0 \leq k \leq 60$ and $Q_d = [15, 0]$ for $61 \leq k \leq 160$, the output reference signal is generated as $y_{rk} = \begin{cases} 1 + 0.5k, 0 \leq k \leq 60 \\ 15, \quad 61 \leq k \leq 160 \end{cases}$. In addition, a constant disturbance $d_k = 0.5$ is considered in this example with $F_d = [0.5, 0]$. By revisiting discrete Sylvester equations (7.15), the discrete feedforward gain can be solved as $\Gamma_d = [3.5819, 3.8319]$ for $0 \leq k \leq 60$, $\Gamma_d = [57.2292, 0]$ for $61 \leq k \leq 160$. Using the same stabilizing gain calculated in the first example, the

feedback control law $u_k = K_d x_{k-1} + L_d q_k$ can be computed by implementing Steps 2-4 given in Table 7.1. The initial condition in this case is taken as: $z_0(\zeta) = 6\sin(3\pi\zeta)$.



(a) State regulation performance.

(b) Output regulation performance.

Figure 7.2: Output regulation of the transport equation with polynomial reference and disturbance signals.

After 80 seconds of simulation, the output regulation performance are illustrated in Fig. 7.2. Using the constructed state feedback regulator, the output is steered to track the ramp-like and step-like reference signals and reject the undesired step disturbance simultaneously. In particular, the tracking error converges to zero rapidly as shown in Fig. 7.2 (b).

7.5.3 Example 3: error feedback regulator design for a 1-D heat equation (spectral system) with set-point reference control

In this case, we consider a 1-D heated bar model described by a parabolic PDE with Newman boundary conditions as follows:

$$x_t(\zeta,t) = x_{\zeta\zeta}(\zeta,t) + b(\zeta)u(t) + f(\zeta)d(t) \tag{7.70a}$$

$$x_\zeta(0,t) = 0 = x_\zeta(1,t) \tag{7.70b}$$

$$y_c(t) = C_c x(\zeta,t) \tag{7.70c}$$

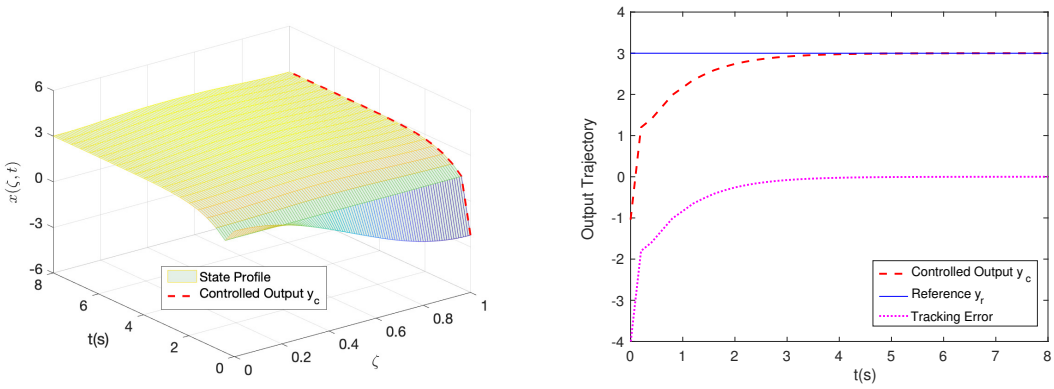
where $\zeta \in [0,1]$ and $t \in [0,+\infty)$ represent spatial and temporal coordinates, respectively. In addition, we consider spatially distributed actuation and disturbance that are characterized by: $b(\zeta) = \frac{1}{2\varepsilon_b}\mathbf{1}_{[\zeta_b-\varepsilon_b,\zeta_b+\varepsilon_b]}(\zeta)$ and $f(\zeta) = 1$. The goal is to regu-

late the output y_c with $C_c := \int_0^1 \frac{1}{2\varepsilon_c}\mathbf{1}_{[\zeta_c-\epsilon_c,\zeta_c+\epsilon_c]}(\cdot)d\zeta$. More specifically, we consider $\zeta_b = 0.5$, $\varepsilon_b = 0.3$, $\zeta_c = 0.99$ and $\epsilon_c = 0.01$.

It is apparent that the original state evolution operator $\mathcal{A} := \frac{\partial^2}{\partial\zeta^2}$ has the eigenvalues $\lambda_n = -(n\pi)^2$, $n \in \mathbb{N}$, which violates the Assumption 11 since $0 \in \sigma(\mathcal{A})$. To address this issue, we stabilize the system by introducing a stabilizing gain K as $K\Phi = -\beta\langle\Phi, \mathbf{1}\rangle$, with $\beta > 0$ [10]. For the stabilized system $\mathcal{A}_c := \frac{\partial^2}{\partial\zeta^2} + b(\zeta)K$ with $0 \notin \sigma(A)$, we have $\sigma(s) \subset \sigma(\mathcal{A}_c)$ so Assumption 11 is satisfied. The rest design of the discrete error feedback regulator follows the steps as shown in Table 7.3.

Table 7.3: Construction of the discrete error feedback regulator.

Algorithm 2: Discrete error feedback regulator
Step 1: Solve discrete Sylvester equations (7.50) for Γ_d and Π_d
Step 2: Solve for discrete stabilizing output injection gain Θ_{1d} and Θ_{2d} in Eq.(7.67), and then obtain $\mathcal{G}_{2d}$
Step 3: Determine $\mathcal{G}_{1d}$ and simulate r_k system (7.51a)
Step 4: Construct the discrete error feedback control law (7.51b) and apply it to the discrete system (7.4)



(a) State regulation performance.

(b) Output regulation performance.

Figure 7.3: Output regulation of the heat equation.

To track and reject step signals, we set $S = 0$ (hence $S_d = 1$), $F_d = 1$ and $Q_d = 3$. Based on the discrete output regulator equations (7.50), we obtain: $L_d = \Gamma_d = 8.7859$ and $H_d = [\mathbf{0} \quad 8.7859]$ due to $K_d = 0$ and $\mathbf{0}$ is a zero vector with proper dimension. By solving Riccati equations (7.67), we can determine the discrete stabilizing output injection gain $\mathcal{G}_{2d}$ and $\mathcal{G}_{1d}$, with $R_1 = R_2 = M_{3d} = 1$ and $M_{1d} = \mathbf{1}$ where $\mathbf{1}$ is an identity matrix with proper dimension. Then, one can simulate r_k system (7.51) to

generate control trajectory u_k, which is plugged into the original discrete plant (7.4). As shown in Fig. 7.3, the closed loop output y_c follows the set-point and the tracking error converges to zero rapidly. In addition, the discretization time interval is 0.2 second and the number of spatial nodes is 1001.

Remark 25. *As proposed in this chapter, the discrete-in-time regulator design provides a novel way to directly realize digital regulator design in a late-lumping manner. Most importantly, intrinsic properties of linear continuous systems are fully preserved in the discrete-time systems by the use of Cayley-Tustin transform. Therefore, the discrete-in-time regulator design method is beneficial to the corresponding design of continuous-time systems by using the equivalent relationships between continuous- and discrete-time systems established in this chapter.*

Remark 26. *Spatial discretization may have a potential effect on the time-discretized stability in the numerical realization stage. Particularly, dissipation and dispersion errors may be induced by an improper choice of spatial and temporal discretization intervals, which may eventually influence time-discretized stability. To reduce the dissipation and dispersion errors, the spatial and temporal discretization intervals need to be chosen small enough, see details in the reference [286, Cha. 7.2, 7.3].*

7.6 Conclusion

In this chapter, discrete-time state and error feedback output regulators are designed for a class of linear distributed parameter systems with bounded input and output operators. The Cayley-Tustin bilinear transform is used for model discretization without spatial approximation or model order reduction. Based on the discretized plant and exogenous systems, discrete-time Sylvester equations are formulated and solved for discrete state and error feedback regulator designs. The solvability of discrete-time regulator equations is proved and linked to the associated continuous version. Given the solutions of discrete-time Sylvester regulator equations, one can attain the corresponding continuous solutions and vice versa. To stabilize the closed-loop systems, a novel way to determine the stabilizing output injection gain (and its dual problem) is provided. Finally, three output regulators are designed and simulated, including

harmonic, step-like and ramp-like reference control for a first-order hyperbolic PDE system and set-point tracking for a 1-D heat equation, which verify the feasibility of the proposed method. The discrete-in-time design and continuous-time design can be beneficial to each other in determining the stabilizing controller gain, stabilizing output injection gain and regulator gain. Hence, the proposed discrete-in-time design has the potential to be utilized in digital control applications, such as sampled-data control of distributed parameter systems.

Chapter 8

Discrete-time Model-Based Output Regulation of Fluid Flow Systems

8.1 Introduction

Flow control and manipulation play an important role in the realm of aerodynamics and hydrodynamics especially when it comes to drag reduction, lift enhancement and turbulence suppression. Generally speaking, there are mainly three approaches to cope with general fluid dynamics problems, namely, theoretical analysis, numerical simulation, and experimental study. Considering that experimental study often requires a prohibitive amount of time and cost, while numerical simulation heavily relies on advanced computational technology, computing capacity and most importantly on the availability of accurate, robust and flexible dynamic models, this chapter seeks to propose an efficient, computational implementable and scalable modelling method for fluid flow output regulation and manipulation.

Differently from the dynamics of lumped parameter systems, fluid dynamics often take place in both time and space domains and their state evolves on infinite-dimensional Hilbert spaces, which requires relatively complex spatiotemporal modelling techniques. Mathematically stated, most fluid dynamics models are governed by partial differential equations (PDEs) and/or delay equations, leading to general distributed parameter systems (DPSs). For instance, the Ginzburg-Landau equation (GLE) involves a first-order temporal derivative, and first- and second-order spatial derivatives with complex model coefficients, which dramatically increases the difficulty of accurate modelling and corresponding control designs. In addition, one may need

Figure 8.1: Vortex shedding in the 2D flow behind a cylinder [2]

to address multiple spatial variables (e.g. three spatial components in Navier-Stokes equation) and even higher-order derivatives and nonlinear terms (e.g. fourth-order spatial derivatives and nonlinear multiplication term in Kuramoto-Sivashinsky equation). Hence, these considerations stated above provide a strong motivation to seek advanced modelling and control techniques for effective and implementable flow control of fluid dynamics systems.

Among the aforementioned fluid dynamics processes, vortex shedding has attracted increasing attention, due to their wide existence manifested by vortex formation when flows pass submerged obstacles with Reynolds numbers larger than the critical values. More specifically, a schematic diagram illustrating the vortex shedding phenomenon in a 2D flow behind a cylinder is given in Fig. 8.1, where it clearly shows an unstable vortex shedding and its evolution [2]. Additionally, there is also strong interest in falling thin film phenomena described by the Kuramoto-Sivashinsky equation (KSE), which as a representative PDE flow model accounts for a wide range of complex phenomena, such as falling film fronts dynamics, unstable flame fronts evolution, phase turbulence in Belousov-Zhabotinsky reaction-diffusion systems and interfacial instabilities between multiple viscous phases [287, 288, 289]. As shown in Fig. 8.2, a two-phase annular falling flow in a vertical tube is illustrated using a schematic [3]. For the sake of brevity, this chapter considers vortex shedding phenomena and falling thin film processes as two representative examples, and review some existing work on modelling and control of CGLE and KSE sequentially.

Regarding vortex shedding analysis and suppression, plenty of studies have been carried out experimentally and using numerical simulation. From an experimental perspective, it has been revealed that the laminar Kármán vortex can be suppressed within a certain range of Reynolds number by several distinct approaches, includ-

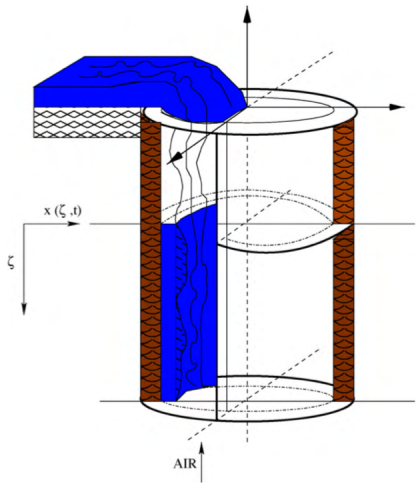


Figure 8.2: Schematic framework of a two-phase annular flow in vertical pipe modelled by Kuramoto-Sivashinsky Equation [3].

ing: external oscillating the cylinder normal to the mean flow [290], feedback control through suction and blowing treatment on the surface [291, 292, 293], and acoustic feedback of signals collected from hot-wires in the wake of the cylinder [294]. In addition, the Navier-Stokes equation (NSE) has been explored to model the dynamics of the cylinder wake theoretically [295, 296, 297]. However, owing to the inherent complexity of the Navier-Stokes equation, most of the related work has been conducted numerically. On the other hand, the complex Ginzburg-Landau equation (CGLE) with appropriate coefficients was suggested as a simplified model to describe vortex shedding processes in [298]. Along the line of controller designs, a proportional feedback controller was proposed for Kármán vortex shedding suppression with Reynolds numbers close to the critical value ($\mathrm{Re_c} \approx 47$ based on the cylinder diameter) [299]. In addition, a non-linear one-dimensional Ginzburg-Landau wake model at 20% above the critical Reynolds number was controlled using a conventional proportional-integral-derivative (PID) controller and a non-linear fuzzy controller [300]. For feedback boundary control, the backstepping approach has been extensively utilized for stabilization of 1D and 2D CGLE systems [301, 2, 58, 16]. The developed controllers were validated using computational fluid dynamics (CFD) simulations [302, 303] and extended to 3D scenarios [304]. A hybrid method and evolution strategies were deployed to study 2D and 3D vortex evolution of cylinder wakes in [305, 306]. From an optimal control perspective, a model predictive controller (MPC) was proposed

to solve the problem of CGLE stabilization with consideration of input and state constraints [1]. These studies on CGLE are oriented on stabilizing control while work related to the output regulation of CGLE is limited, which motivates this contribution.

When it comes to flow control of falling thin film processes modelled by the Kuramoto-Sivashinsky equation, many control methods have been developed using recent advances in the area of control of distributed parameter systems. Among these, one important contribution lies in the stabilization of the KSE model, including: global stabilization of KSE by in-domain output feedback control [307, 308] and through boundary control [309, 15]. A single-input-single-output (SISO) and multiple-input-single-output (MISO) boundary model predictive controllers were presented for KSE stabilization in the presence of input and state constraints using a truncated modal decomposition [310, 3, 311]. A zero dynamics inverse design method was proposed for tracking regulation of a nonlinear KSE with two boundary actuators [312]. For sampled-data control of KSE, a spatially distributed controller is constructed for local stabilization of KSE by using a time-delay approach and a descriptor method [313]. Recently, a delayed boundary controller was designed for global stabilization of a linear KSE by means of the spectral decomposition and the Artstein transform [314]. Although these contributions have provided elegant solutions and controller designs to guarantee the exponential stability of the closed-loop system, most of them are conducted in a continuous-time setting, which at the implementation level needs to be realized in a sampled-data computationally feasible setting and in addition brings another layer of complexity and questions to be addressed. Furthermore, a realizable sampled-data servo-control design and not just stabilization is needed for manipulation of various fluid dynamics systems represented by the Kuramoto-Sivashinsky equation. Moreover, the realization of continuous-in-time designs in the sampled-data setting with finite computational resources has not been fully addressed since temporal and spatial approximations (and/or model reduction) need to be performed for control algorithm realization with the hope that approximate controllers account for the infinite-dimensional nature of underlying distributed parameter models. Hence, the motivation behind this chapter is not to take the path of continuous designs first, then approximate at the realization level (early lumping),

but to discretize the model in time by application of structure-preserving Cayley-Tustin discretization (linear system properties including stability, controllability and observability are preserved [97, 119]) and conduct discrete-in-time regulator design for complex fluid flow systems without any model lumping or spatial approximation. In this way, both regulator design and computing control law realization can be accomplished in the natural discrete-time sampled-data computational setting of modern microchips.

Therefore, this chapter is devoted to the development of effective, computational realizable and scalable servo controllers in the discrete-time setting for fluid flow output regulation of linear CGLE and KSE systems. In particular, the Cayley-Tustin transform is deployed for infinite-dimensional model discretization in the time domain and it does not induce any spectral decomposition or spatial approximation. Then, the internal model principle [247] is revisited and extended to the output regulation of infinite-dimensional discrete-time systems. By further establishing a finite-dimensional discrete-time exogenous system, discrete-time regulator equations are formulated and utilized for solving a state feedback regulation problem. As for model stabilization, a pole-shifting approach is employed for continuous PDE models and linked to their discrete counterparts. Additionally, an output feedback regulator design is completed by the development of finite-dimensional (exo-system) and infinite-dimensional (fluid flow-plant) observers. Furthermore, we emphasize that proposed design is applicable for a general class of Riesz-spectral distributed parameter systems, and can be extended to DPS models other than CGLE and KSE systems.

The remainder of this chapter is organized as follows. Section 8.2 presents a necessary preliminary, including: general continuous-time PDE model description, and time discretization with the aid of the Cayley-Tustin approach. In Section 8.3, a finite-dimensional discrete-time exogenous system is provided and a discrete output regulator design framework is given. To demonstrate the feasibility and effectiveness of the proposed method, two representative PDE flow models, i.e., CGLE and KSE systems, are analyzed in detail. The model formulation, spectrum analysis, analytic resolvent determination, and simulation studies for two models are shown in Sections 8.4 and 8.5, respectively. Finally, a conclusion is drawn in Section 8.6.

8.2 Preliminary

8.2.1 Notations

In this chapter, the following notation is used. Suppose that X and V are two Hilbert spaces and $A : X \mapsto V$ is a linear operator. $\mathcal{L}(X, V)$ denotes the set of linear bounded operators from X to V. If $X = V$, we simply write $\mathcal{L}(X)$. The domain, spectrum, resolvent set and resolvent operator of a linear operator A are denoted by: $\mathcal{D}(A)$, $\sigma(A)$, $\rho(A)$, and $\mathcal{R}(s, A) = (sI - A)^{-1}$ with $s \in \rho(A)$, respectively. We denote the space X_1 as the space $\mathcal{D}(A)$ with the norm $\|x\|_1 = \|(\beta I - A)x\|$, and the space X_{-1} as the completion of X with the norm $\|z\|_{-1} = \|(\beta I - A)^{-1}z\|$, where $\forall x \in \mathcal{D}(A)$, $\forall z \in X$, and $\beta \in \rho(A)$. The constructed spaces are related as follows: $X_1 \subset X \subset X_{-1}$, with each inclusion being dense and continuous embedding [42]. The extension of A to X_{-1} is still denoted as A, and C_Λ represents the Λ-extension of C, i.e., $C_\Lambda x = \lim_{\lambda \to +\infty} \lambda C \mathcal{R}(\lambda, A)x$, where the domain of C_Λ consists of those elements $x \in X$ for which the limit exists. Additionally, the inner product is denoted by $\langle \cdot, \cdot \rangle$, and $L^2(0, l)^m$ with a positive integer m denotes a Hilbert space of a m-dimensional vector of the real functions that are square integrable over $[0, l]$ with a spatial length l.

8.2.2 Model description

As fluid flow dynamics often takes place in both temporal and spatial domains, the mathematical models are described by partial differential equations (PDE). In general, we consider a continuous-time infinite-dimensional system having the following abstract form:

$$\frac{\partial x}{\partial t}(\xi, t) = Ax(\xi, t) + Bu(t) + Ed(t),\ x(0) = x_0 \tag{8.1a}$$

$$y_c(t) = C_c x(\xi, t) + D_c u(t) + F_c d(t),\ t \geq 0 \tag{8.1b}$$

$$y_m(t) = C_m x(\xi, t) + D_m u(t) + F_m d(t) \tag{8.1c}$$

on a complex Hilbert space $X = L^2((0, l), \mathbb{C})$, i.e., the spatial state $x(\cdot, t) \in X$, where $\xi \in [0, l]$ and $t \in [0, \infty)$ represent spatial and temporal variables. We denote the input $u(t) \in L^2_{loc}([0, \infty), U)$, disturbance $d(t) \in L^2_{loc}([0, \infty), D)$, and the controlled and measured outputs $y_c(t), y_m(t) \in L^2_{loc}([0, \infty), Y)$, where U, D and Y are assumed to be finite-dimensional Hilbert spaces with $\dim U = \dim D = \dim Y = 1$. Additionally,

$A : \mathcal{D}(A) \subset X \mapsto X$ is an infinitesimal generator of a C_0-semigroup $\mathbb{T}_A(t)$ on X. For simplicity, we consider bounded control operator $B \in \mathcal{L}(U, X)$, bounded disturbance operators $E \in \mathcal{L}(D, X)$, $F_c, F_m \in \mathcal{L}(D, Y)$, and bounded measured output operator $C_m \in \mathcal{L}(X, Y)$. The controlled output operator $C_c \in \mathcal{L}(X_1, Y)$ is considered to be unbounded and assumed to be admissible for $\mathbb{T}_A(t)$. To account for well-posedness, one needs to replace C_c by $C_\Lambda x = \lim_{\lambda \to +\infty} \lambda C_c \mathcal{R}(\lambda, A)x$, where $x \in X$ and $\lambda \in \rho(A)$. Thus this framework leads to a well-posed regular system [21, 42, 111]. For ease of notation, we will use C_c to denote C_Λ in what follows. The transfer functions can be expressed as:

$$\mathcal{G}_c(s) = C_c \mathcal{R}(s, A)B + D_c \tag{8.2a}$$

$$\mathcal{T}_c(s) = C_c \mathcal{R}(s, A)E + F_c \tag{8.2b}$$

$$\mathcal{G}_m(s) = C_m \mathcal{R}(s, A)B + D_m \tag{8.2c}$$

$$\mathcal{T}_m(s) = C_m \mathcal{R}(s, A)E + F_m \tag{8.2d}$$

with $s \in \rho(A)$ and $\mathcal{R}(s, A) = (sI - A)^{-1}$ is the resolvent operator, and $\mathcal{G}_c(s)$ and $\mathcal{T}_c(s)$ stand for continuous-time transfer functions from $u(t)$ to $y_c(t)$ and from $d(t)$ to $y_c(t)$, respectively. Likewise, $\mathcal{G}_m(s)$ and $\mathcal{T}_m(s)$ are transfer functions from $u(t)$ to $y_m(t)$ and from $d(t)$ to $y_m(t)$.

8.2.3 Model time-discretization

In order to preserve system properties (such as stability, controllability, and observability) in a discretization process, the Cayley-Tustin time-discretization approach is facilitated to transform a continuous-time model into its discrete analogue [97, 119]. For the considered linear infinite-dimensional continuous-time invariant system (8.1), one deploys the Crank-Nicolson discretization for a given time discretization interval h as follows:

$$\frac{x(\xi, kh) - x(\xi, (k-1)h)}{h} \approx A \frac{x(\xi, kh) + x(\xi, (k-1)h)}{2} + Bu(kh) + Ed(kh) \tag{8.3a}$$

$$y_c(kh) \approx C_c \frac{x(\xi, kh) + x(\xi, (k-1)h)}{2} + D_c u(kh) + F_c d(kh) \tag{8.3b}$$

$$y_m(kh) \approx C_m \frac{x(\xi, kh) + x(\xi, (k-1)h)}{2} + D_m u(kh) + F_m d(kh) \tag{8.3c}$$

with $x(0) = x_0 \in X,\ k \geq 1$. Above time discretization admits an implicit mid-point integration rule and is reversible in time (namely: symmetric) due to the fact that Eq.(8.3a) stays invariant when one interchanges $x(\xi, kh) \leftrightarrow x(\xi, (k-1)h)$ and $h \leftrightarrow -h$, see details in [119]. Additionally, this time discretization scheme is symplectic in the sense that it preserves Hamiltonian properties of the system [119]. Simple algebraic manipulations lead to the following infinite-dimensional discrete-time state space model:

$$x_k = A_d x_{k-1} + B_d u_k + E_d d_k \tag{8.4a}$$

$$y_{ck} = C_{cd} x_{k-1} + D_{cd} u_k + \Upsilon_{cd} d_k \tag{8.4b}$$

$$y_{mk} = C_{md} x_{k-1} + D_{md} u_k + \Upsilon_{md} d_k \tag{8.4c}$$

with $x(0) = x_0 \in X,\ k \geq 1$, where x_k, u_k, d_k, y_{ck} and y_{mk} denote the discrete-time state, input, disturbance, controlled output and measured output. It is noted that the discrete input is given by the integration as $\frac{u_k}{\sqrt{h}} = \frac{1}{h}\int_{(k-1)h}^{kh} u(t)dt$ on the interval $t \in [(k-1)h, kh]$, and it has been shown that the Cayley-Tustin discretization is a convergent time discretization scheme for input-output stable system nodes (with $\dim U = \dim D = \dim Y = 1$) in the sense that $\|\frac{u_k}{\sqrt{h}} - u(t)\| \to 0$ as $h \to 0^+$, see [97]. Similar expressions hold for the approximation of d_k, y_{ck} and y_{mk}. The associated discrete-time operators are given by:

$$\begin{aligned}\begin{bmatrix} A & B & E \\ C_c & D_c & F_c \\ C_m & D_m & F_m \end{bmatrix} \to \begin{bmatrix} A_d & B_d & E_d \\ C_{cd} & D_{cd} & \Upsilon_{cd} \\ C_{md} & D_{md} & \Upsilon_{md} \end{bmatrix} \\ = \begin{bmatrix} -I + 2\delta\mathcal{R}(\delta, A) & \sqrt{2\delta}\mathcal{R}(\delta, A)B & \sqrt{2\delta}\mathcal{R}(\delta, A)E \\ \sqrt{2\delta}C_c\mathcal{R}(\delta, A) & \mathcal{G}_c(\delta) + D_c & \mathcal{T}_c(\delta) + F_c \\ \sqrt{2\delta}C_m\mathcal{R}(\delta, A) & \mathcal{G}_m(\delta) + D_m & \mathcal{T}_m(\delta) + F_m \end{bmatrix}\end{aligned} \tag{8.5}$$

where $\mathcal{G}_c(\delta)$, $\mathcal{T}_c(\delta)$, $\mathcal{G}_m(\delta)$, and $\mathcal{T}_m(\delta)$ are the transfer functions $\mathcal{G}_c(s)$, $\mathcal{T}_c(s)$, $\mathcal{G}_m(s)$, and $\mathcal{T}_m(s)$ respectively with s evaluated at $\delta = 2/h \in \rho(A)$. Note that there are feedforward operators D_{cd}, Υ_{cd}, D_{cd} and Υ_{md} introduced in the discrete-time setting (8.4) after performing the Cayley-Tustin discretization, which are not necessarily present in the continuous model setting (8.1) (i.e., when $D_c = F_c = D_m = F_m = 0$). Furthermore, we denote the discrete-time transfer functions from u_k to y_{ck} and from d_k to y_{ck} by $\mathcal{G}_{cd}(z) = C_{cd}(zI - A_d)^{-1}B_d + D_{cd}$ and $\mathcal{T}_{cd}(z) = C_{cd}(zI - A_d)^{-1}E_d + \Upsilon_{cd}$

respectively, where $z \in \rho(A_d)\backslash\{-1\}$. Similar definitions hold for transfer functions $\mathcal{G}_{md}(z)$ and $\mathcal{T}_{md}(z)$.

Remark 27. *The Cayley-Tustin discretization scheme brings an obvious technical advantage that enables one to avoid direct treatment of unbounded operators in the continuous-time setting by instead dealing with bounded operators in the discrete-time setting.*

Remark 28. *It is shown that the continuous- and discrete-time transfer functions satisfy the following relationship [97]:*

$$\mathcal{G}_{cd}(z) = \mathcal{G}_c(\frac{z-1}{z+1}\delta),\ \mathcal{T}_{cd}(z) = \mathcal{T}_c(\frac{z-1}{z+1}\delta),\ z \in \rho(A_d)\backslash\{-1\} \tag{8.6a}$$

$$\mathcal{G}_c(s) = \mathcal{G}_{cd}(\frac{\delta+s}{\delta-s}),\ \mathcal{T}_c(s) = \mathcal{T}_{cd}(\frac{\delta+s}{\delta-s}),\ s \in \rho(A)\backslash\{\delta\} \tag{8.6b}$$

which provides a way for finding $\mathcal{G}_{cd}(z)$ and $\mathcal{T}_{cd}(z)$ from their continuous counterparts, and vice versa. Thus, a stable, continuous-time transfer function is holomorphic and bounded on $\mathbb{C}^+$ if and only if the corresponding discrete-time transfer function is holomorphic and bounded on $\mathbb{D}^+$, see details in [24, 23].

Remark 29. *For a conservative infinite-dimensional linear system with a scalar inner transfer function and a zero initial state, the discrete-time state trajectory converges to the corresponding continuous-time one in some Hilbert norm sense as the time discretization interval goes to zero [239]. It was also shown that the discrete state converges to the continuous one as time increases for the first-order evolution differential equations (with zero input) in Hilbert space with bounded and unbounded A operators depending on the smoothness of the initial conditions [315, 316, 317, 318, 319], which was further extended to the cases of initial values problems and boundary values problems of second-order evolution differential equations [320, 321, 322]. Nevertheless, the state approximation induced by the Cayley-Tustin transformation does not hinder its use in the output regulation problem since transfer functions from the input to the output are preserved from the continuous-time model to the discrete-time one, which plays a central role in the output regulation context.*

In order to proceed with the regulator design, the following concepts are introduced.

Definition 10. *A C_0-semigroup $\mathbb{T}_A(t)$ on the Hilbert space X is exponentially stable if there exist positive constants M and α such that:*

$$\|\mathbb{T}_A(t)\| \leq Me^{-\alpha t},\ \forall t \in [0,\infty) \tag{8.7}$$

and it is strongly stable if $\|\mathbb{T}_A(t)x\| \to 0$ as $t \to +\infty$ for all $x \in X$. $\mathbb{T}_A(t)$ is β-exponentially stable if (8.7) holds for $-\alpha < \beta$, i.e., its stability margin is at least $-\beta$. A_d is power stable if there exist positive constants M and $\gamma < 1$ such that:

$$\|A_d^k\| \leq M\gamma^k,\ \forall k \in \mathbb{N} \tag{8.8}$$

and A_d is strongly stable if $A_d^k x \to 0$ as $k \to +\infty$ for all $x \in X$ [4].

Definition 11. *Assume that A is the infinitesimal generator of the C_0-semigroup $\mathbb{T}_A(t)$ on the Hilbert space X, $B \in \mathcal{L}(U,X)$, $C_m \in \mathcal{L}(X,Y)$, where U and Y are finite-dimensional Hilbert spaces. Let $\Sigma(A,B,C_m)$ denote the state linear system $\dot{z}(t) = Az(t) + Bu(t)$, $y(t) = C_m z(t)$, $t \geq 0$, $z(0) = z_0 \in X$, where generating operators A, B, and C_m are defined as above, and state, input and output spaces are X, U, and Y. If there exist $K \in \mathcal{L}(X,U)$ and $L \in \mathcal{L}(Y,X)$ such that $A + BK$ and $A + LC_m$ generate exponentially stable C_0-semigroup $\mathbb{T}_{BK}(t)$ and $\mathbb{T}_{LC}(t)$, then the system $\Sigma(A,B,C_m)$ is exponentially stabilizable and detectable. If $\mathbb{T}_{BK}(t)$ and $\mathbb{T}_{LC}(t)$ are β-exponentially stable, then the system $\Sigma(A,B,C_m)$ is β-exponentially stabilizable and detectable [4].*

8.3 Output regulation

In this section, an output regulator design method is developed for output regulation of a class of infinite-dimensional systems in a discrete-time setting. In particular, a discrete-time output regulator design is proposed for self-adjoint Riesz-spectral PDE systems, based on the discretized distributed parameter flow plant and discrete exogenous system.

8.3.1 Exogenous system

To construct the disturbance (to be rejected) and the reference signals (to be tracked), a discrete-time finite-dimensional exogenous system (i.e., exo-system) is introduced

as follows:

$$q_k = S_d q_{k-1},\ q_0 = q^0 \in \mathbb{C}^n,\ k \geq 1 \tag{8.9a}$$

$$d_k = F_d q_k \tag{8.9b}$$

$$y_{rk} = Q_d q_k \tag{8.9c}$$

where q_k, d_k, and y_{rk} represent the exogenous state, disturbance, and reference signals in the discrete-time setting. Moreover, S_d denotes a discrete-time evolution matrix of state q_k and is of $n \times n$ dimension. More specifically, it is assumed that S_d has distinct eigenvalues placed on the boundary of the unit disc, i.e., $\lambda_i^d = \lambda_{Re} + \lambda_{Im} j$ where $i = 1, ..., n$, $j^2 = -1$, $\lambda_{Re} \in [0, 1]$, $\lambda_{Im} \in [0, 1]$ and $\lambda_{Re}^2 + \lambda_{Im}^2 = 1$. Thus, S_d accounts for step-like and sinusoid-like signals. In order to reconstruct the full state information from the reference single y_{rk}, it is assumed that (S_d, Q_d) is observable. Additionally, we suppose that F_d and Q_d have proper dimensions to generate disturbance and reference signals of interest.

8.3.2 State feedback regulator design

The main purpose of output regulation is to realize system stabilization, disturbance rejection and reference tracking. Normally, it can be mathematically stated as constructing a discrete-time state feedback regulator of the following form:

$$u_k = K_d x_{k-1} + L_d q_k,\ k \geq 1 \tag{8.10}$$

where $K_d \in \mathcal{L}(X, U)$, $L_d \in \mathcal{L}(\mathbb{C}^n, U)$, such that the following conditions are ensured:

(1) The discrete-time closed-loop system operator $A_d + B_d K_d$ is strongly stable;

(2) The discrete-time tracking error $e_k = y_{ck} - y_{rk} \to 0$ as $k \to +\infty$ for any given $x_0 \in X$ and $q^0 \in \mathbb{C}^n$.

For this design, all state information of the plant and the exo-system is assumed to be known in (8.10), which literally interprets the definition of "state feedback regulator". Combining the discrete-time plant and exogenous models, the following theorem states the necessary and sufficient condition for the discrete-time state feedback regulator design:

Theorem 22. *Suppose that* (A, B) *is exponentially stabilizable and* $\sigma(S_d) \subset \rho(A_d)$. *The discrete state feedback regulation problem is solvable if and only if there exist mappings* $\Pi_d \in \mathcal{L}(\mathbb{C}^n, X)$ *and* $\Gamma_d \in \mathcal{L}(\mathbb{C}^n, U)$ *such that the following discrete Sylvester equations hold:*

$$\Pi_d S_d = A_d \Pi_d + (B_d \Gamma_d + P_d) S_d \tag{8.11a}$$

$$Q_d S_d = C_{cd} \Pi_d + (D_{cd} \Gamma_d + \Theta_{cd}) S_d \tag{8.11b}$$

where $P_d = E_d F_d$, $\Theta_{cd} = \Upsilon_{cd} F_d$, *and* $L_d = \Gamma_d - K_d \Pi_d S_d^{-1}$ *are utilized to compute the state feedback control law* u_k *in Eq.(8.10). Here,* E_d *and* Υ_{cd} *are defined in Eqs.(8.4)-(8.5).*

Proof. First, we prove the sufficiency. Substituting Eq.(8.10) into the discrete system (8.4) leads to the closed-loop model as follows:

$$x_k = (A_d + B_d K_d) x_{k-1} + (B_d L_d + P_d) q_k \tag{8.12}$$

By induction, the discrete-time state solution can be found as:

$$\begin{aligned} x_k = {} & (A_d + B_d K_d)^k x_0 \\ & + \sum_{m=1}^{k} (A_d + B_d K_d)^{m-1} (B_d L_d + P_d) q_{k+1-m} \end{aligned} \tag{8.13}$$

By substituting Eq.(8.9) and Eq.(8.11) into Eq.(8.13), one obtains:

$$\begin{aligned} & x_k - (A_d + B_d K_d)^k x_0 \\ & = \sum_{m=1}^{k} (A_d + B_d K_d)^{m-1} [B_d (\Gamma_d - K_d \Pi_d S_d^{-1}) + P_d] q_{k+1-m} \\ & = \sum_{m=1}^{k} (A_d + B_d K_d)^{m-1} [(B_d \Gamma_d + P_d) S_d - B_d K_d \Pi_d)] q_{k-m} \\ & = \sum_{m=1}^{k} (A_d + B_d K_d)^{m-1} [\Pi_d S_d - (A_d + B_d K_d) \Pi_d)] q_{k-m} \\ & = \sum_{m=1}^{k} (A_d + B_d K_d)^{m-1} \Pi_d q_{k+1-m} - \sum_{m=2}^{k+1} (A_d + B_d K_d)^{m-1} \Pi_d q_{k+1-m} \end{aligned} \tag{8.14}$$

Then, the last expression further induces:

$$x_k = (A_d + B_d K_d)^k (x_0 - \Pi_d q_0) + \Pi_d q_k \tag{8.15}$$

Moreover, the discrete tracking error can be expressed as:

$$\begin{aligned} e_k =& y_{ck} - y_{rk} \\ =& C_{cd}x_{k-1} + D_{cd}u_k + \Theta_{cd}q_k - Q_d q_k \\ =& (C_{cd} + D_{cd}K_d)x_{k-1} + (D_{cd}L_d + \Theta_{cd} - Q_d)q_k \\ =& (C_{cd} + D_{cd}K_d)(A_d + B_d K_d)^{k-1}(x_0 - \Pi_d q_0) \\ &+ [(C_{cd} + D_{cd}K_d)\Pi_d + (D_{cd}L_d + \Theta_{cd} - Q_d)S_d]q_{k-1} \end{aligned} \tag{8.16}$$

Under the assumption that (A, B) is exponentially stabilizable, it is shown in [23] that (A_d, B_d) is strongly stabilizable with a proper choice of $\delta \in \rho(A)$. Thus we can find $K_d \in \mathcal{L}(X, U)$ such that $A_d + B_d K_d$ is a strongly stable operator, which indicates $(A_d + B_d K_d)^k x \to 0$ as $k \to +\infty$ for all $x \in X$. Therefore, x_k converges to $\Pi_d q_k$ in Eq.(8.15) and the discrete tracking error e_k goes to zero as $k \to +\infty$ in Eq.(8.16), which is ensured by the discrete Sylvester equations (8.11a)-(8.11b).

Now, we show the proof of the necessity by constructing the following extended closed-loop system:

$$\begin{bmatrix} x_k \\ q_k \end{bmatrix} = \begin{bmatrix} A_d + B_d K_d & (B_d L_d + P_d)S_d \\ 0 & S_d \end{bmatrix} \begin{bmatrix} x_{k-1} \\ q_{k-1} \end{bmatrix} \tag{8.17}$$

It is straightforward to obtain the solution of Eq.(8.17) by induction as follows:

$$\begin{bmatrix} x_k \\ \\ q_k \end{bmatrix} = \begin{bmatrix} (A_d + B_d K_d)^k x_0 + \sum_{m=1}^{k} (A_d + B_d K_d)^{m-1}(B_d L_d + P_d)q_{k+1-m} \\ \\ S_d^k q_0 \end{bmatrix} \tag{8.18}$$

Given that $A_d + B_d K_d$ is strongly stable, $(A_d + B_d K_d)^k x_0 \to 0$ as $k \to +\infty$ and Eq.(8.18) indicates that $[x_k; q_k] \to [\Pi_d q_k; q_k]$ with $k \to +\infty$ and $\Pi_d \in \mathcal{L}(\mathbb{C}^n, X)$. To determine Π_d, we can construct the dynamical evolution of $w_k = [x_k; q_k] - [\Pi_d q_k; q_k]$ as the following homogeneous difference equation:

$$w_k = \begin{bmatrix} A_d + B_d K_d & (B_d L_d + P_d)S_d \\ 0 & S_d \end{bmatrix} w_{k-1} \tag{8.19}$$

where the initial condition is defined as $w_0 = [x_0; q_0] - [\Pi_d q_0; q_0]$ with $\mathcal{H}_e = X \bigoplus \mathbb{C}^n$. The first component in Eq.(8.19) leads to $(A_d + B_d K_d)\Pi_d + (B_d L_d + P_d)S_d = \Pi_d S_d$ which is identical to discrete-time Sylvester equation (8.11a). Furthermore, the discrete tracking error is described as:

$$e_k = y_{ck} - y_{rk} \tag{8.20}$$

$$
\begin{aligned}
&= C_{cd}x_{k-1} + D_{cd}u_k + \Theta_{cd}q_k - Q_d q_k \\
&= (C_{cd} + D_{cd}K_d)x_{k-1} + (D_{cd}L_d + \Theta_{cd} - Q_d)q_k \\
&= \begin{bmatrix} C_{cd} + D_{cd}K_d & (D_{cd}L_d + \Theta_{cd} - Q_d)S_d \end{bmatrix} \begin{bmatrix} x_{k-1} \\ q_{k-1} \end{bmatrix} \\
&\rightarrow [(C_{cd} + D_{cd}K_d)\Pi_d + (D_{cd}L_d + \Theta_{cd} - Q_d)S_d]\, q_{k-1} \\
&\qquad (\text{as } k \rightarrow +\infty)
\end{aligned}
$$

To realize perfect tracking, one needs to ensure that $(C_{cd} + D_{cd}K_d)\Pi_d + (D_{cd}L_d + \Theta_{cd} - Q_d)S_d = 0$, which can be simplified as Eq.(8.11b) with $L_d = \Gamma_d - K_d\Pi_d S_d^{-1}$. ■

By projecting the eigenvalue pair (λ_i^d, ϕ_i^d) of S_d on the discrete Sylvester equations (8.11), it is straightforward to determine the discrete regulator gains (L_d, Π_d) as:

$$\Pi_d \phi_i^d = \lambda_i^d (\lambda_i^d I - A_d)^{-1}(B_d L_d + P_d)\phi_i^d \tag{8.21a}$$

$$L_d \phi_i^d = [\mathcal{G}_{cd}(\lambda_i^d)]^{-1}[Q_d - \mathcal{T}_{cd}(\lambda_i^d)F_d]\phi_i^d \tag{8.21b}$$

where $\mathcal{G}_{cd}(\lambda_i^d)$ is the discrete transfer function from u_k to y_{ck} with z evaluated at $z = \lambda_i^d$. Similarly, $\mathcal{T}_{cd}(\lambda_i^d)$ is the discrete transfer function from d_k to y_{ck} with z evaluated at $z = \lambda_i^d$.

Remark 30. *In this design the assumption that $\lambda_i^d \in \sigma(S_d) \subset \rho(A_d)$ can be ensured by adjusting the time discretization interval h. To ensure the solvability of the discrete Sylvester equations (8.11), we need to further assume that $\mathcal{G}_{cd}(\lambda_i^d) \neq 0$, $\forall \lambda_i^d \in \sigma(S_d)$, such that $\mathcal{G}_{cd}(\lambda_i^d)$ is invertible in Eq.(8.21).*

Remark 31. *The 1-1 correspondence between discrete- and continuous-time transfer functions shown in Eq. (8.6) provides a constructive way in evaluating the discrete transfer functions $(\mathcal{G}_{cd}(\lambda_i^d), \mathcal{T}_{cd}(\lambda_i^d))$ by using their continuous counterparts.*

For the control law (8.10), what remains is to provide a convenient way to solve for the stabilizing controller gain K_d. In order to address this issue, the following theorem is proposed:

Theorem 23. *Suppose that a self-adjoint Riesz-spectral operator A is an infinitesimal generator of the C_0-semigroup $\mathbb{T}_A(t)$ on the Hilbert space X, $B \in \mathcal{L}(\mathbb{C}, X)$. Assume that A has simple eigenvalues and the spectrum $\{\lambda_n, \phi_n\}$ of A can be decomposed into an unstable part $\{\lambda_{n_u}^u, \phi_{n_u}^u\}$ (with $\lambda_{n_u}^u \geq 0$) and a stable part $\{\lambda_{n_s}^s, \phi_{n_s}^s\}$*

(with $\lambda^s_{n_s} < 0$). Then the discrete stabilizing controller gain $K_d \in \mathcal{L}(X, \mathbb{C})$ can be obtained as $K_d\phi = -\sum_{n_u=1}^{N} \beta^d_{n_u} \langle \phi, \phi^u_{n_u} \rangle$, where $\beta^d_{n_u}$ satisfies $|\frac{\delta+\lambda^u_{n_u}-\sqrt{2\delta}\beta^d_{n_u} b_{n_u}}{\delta-\lambda^u_{n_u}}| < 1$, with $b_{n_u} = \left\langle B, \phi^u_{n_u} \right\rangle \neq 0$, such that the Cayley-Tustin discretized system $A_d + B_d K_d$ is strongly stable, where $\delta = \frac{2}{h} \in \rho(A)$ and h denotes the discretization time interval.

Proof. Given that a self-adjoint Riesz-spectral operator A is an infinitesimal generator of the C_0-semigroup $\mathbb{T}_A(t)$ on the Hilbert space X, it is straightforward to show that A has the following decomposition:

$$Ax = \sum_{n=1}^{+\infty} \lambda_n \langle x, \phi_n \rangle \phi_n \tag{8.22}$$

where λ_n and ϕ_n, with $n \in \mathbb{N}$, are eigenvalues and eigenfunctions of A, see [4].
For $\Sigma(A_d, B_d, C_{md})$ obtained by Cayley-Tustin transform (8.5) of $\Sigma(A, B, C_m)$, it is straightforward to obtain operator decompositions as follows:

$$\mathcal{R}x = \sum_{n=1}^{+\infty} \frac{1}{\delta - \lambda_n} \langle x, \phi_n \rangle \phi_n \tag{8.23a}$$

$$A_d x = \sum_{n=1}^{+\infty} \frac{\delta + \lambda_n}{\delta - \lambda_n} \langle x, \phi_n \rangle \phi_n \tag{8.23b}$$

$$B_d x = \sum_{n=1}^{+\infty} \frac{\sqrt{2\delta}}{\delta - \lambda_n} \langle Bx, \phi_n \rangle \phi_n \tag{8.23c}$$

$$C_{md} x = \sum_{n=1}^{+\infty} \frac{\sqrt{2\delta}}{\delta - \lambda_n} \langle x, \phi_n \rangle C_m \phi_n \tag{8.23d}$$

Now, it is apparent that by the Cayley-Tustin transform one can map $\lambda_n \in \mathbb{C}^-$ of the S-plane into the interior section of the unit disc on the Z-plane (except -1), and vice versa.

Suppose that the spectrum $\{\lambda_n, \phi_n\}$ of A can be decomposed into an unstable part $\{\lambda^u_{n_u}, \phi^u_{n_u}\}$ (with $\lambda^u_{n_u} \geq 0$) and a stable part $\{\lambda^s_{n_s}, \phi^s_{n_s}\}$ (with $\lambda^s_{n_s} < 0$). By Eq. (III.6) in the reference [10], a bounded stabilizing controller gain K can be chosen as $K\phi = -\sum_{n_u=1}^{N} \beta_{n_u} \langle \phi, \phi^u_{n_u} \rangle$ with some positive β_{n_u}, so that the finite set of countable unstable eigenvalues of A can be shifted to the left side of the complex plane as follows:

$$(A + BK)x = \sum_{n_u=1}^{N} \lambda^u_{n_u} \left\langle x, \phi^u_{n_u} \right\rangle \phi^u_{n_u} - \beta_{n_u} \langle x, \phi^u_{n_u} \rangle B$$

$$+ \sum_{n_s=N+1}^{+\infty} \lambda_{n_s}^s \left\langle x, \phi_{n_s}^s \right\rangle \phi_{n_s}^s \tag{8.24}$$

where N denotes the number of possibly unstable eigenvalues and hence it is the rank of K. Thus, the original continuous-time system is stabilized, see [4].

Furthermore, one can determine the discrete-time stabilizing operator $K_d\phi = -\sum_{n=1}^{N} \beta_{n_u}^d \langle \phi, \phi_{n_u}^u \rangle$ with some design parameter $\beta_{n_u}^d$ as:

$$\begin{aligned}(A_d + B_d K_d)x = \sum_{n_u=1}^{N} [&\frac{\delta + \lambda_{n_u}^u}{\delta - \lambda_{n_u}^u} \left\langle x, \phi_{n_u}^u \right\rangle \phi_{n_u}^u \\ &- \frac{\sqrt{2\delta}\beta_{n_u}^d}{\delta - \lambda_{n_u}^u} \left\langle B\langle x, \phi_{n_u}^u \rangle, \phi_{n_u}^u \right\rangle \phi_{n_u}^u] \\ &+ \sum_{n_s=N+1}^{+\infty} \frac{\delta + \lambda_{n_s}^s}{\delta - \lambda_{n_s}^s} \left\langle x, \phi_{n_s}^s \right\rangle \phi_{n_s}^s \end{aligned} \tag{8.25}$$

as it can be ensured that $\langle B, \phi_{n_u}^u \rangle = b_{n_u} \neq 0$ for $n_u = 1, 2, ..., N$, it is straightforward to solve the discrete-time stabilizing gain satisfying $|\frac{\delta + \lambda_{n_u}^u - \sqrt{2\delta}\beta_{n_u}^d b_{n_u}}{\delta - \lambda_{n_u}^u}| < 1$. ■

Corollary 7. *Under the assumptions in Theorem 22 and $C_m \in \mathcal{L}(X, \mathbb{C})$, the discrete stabilizing output injection gain $L_{1d} \in \mathcal{L}(\mathbb{C}, X)$ can be obtained as $L_{1d}C_{md}\phi = -\sum_{n_u=1}^{N} \gamma_{n_u}^d \langle C_{md}, \phi \rangle \phi_{n_u}^u$, where $\gamma_{n_u}^d$ satisfies $|\frac{\delta + \lambda_{n_u}^u - \sqrt{2\delta}c_{n_u}\gamma_{n_u}^d}{\delta - \lambda_{n_u}^u}| < 1$, with $c_{n_u} = \left\langle C_m, \phi_{n_u}^u \right\rangle \neq 0$, such that the discretized system using Cayley-Tustin method $A_d + L_{1d}C_{md}$ is strongly stable, where $\delta = \frac{2}{h} \in \rho(A)$ and h denotes the discretization time interval.*

Proof. It is straightforward to show that this is a dual problem of Theorem 23. Hence, the proof can be completed if one takes $C_{md} = B_d^*$ and $L_{1d} = K_d^*$. ■

Remark 32. *In general, Theorem 23 and Corollary 7 can be extended to the case with A being a general Riesz-spectral operator by finding corresponding eigenfunctions of A^* operator and the case with $U = \mathbb{C}^p$ and $Y = \mathbb{C}^q$ through complex manipulation.*

Based on the series expressions in Eq.(8.23a) and spectral decomposition, one can design a stabilizing controller (and/or observer) gain. For the resolvent and the corresponding discrete operators (A_d, B_d, C_{md}) given by closed-form expressions, we provide the following theorem to guarantee the equivalence of the stability property of the discretized stabilized system using continuous- and discrete-time stabilizing gains by following [23].

Theorem 24. *Given an infinitesimal generator A of the C_0-semigroup $\mathbb{T}_A(t)$ on the Hilbert space X, $B \in \mathcal{L}(\mathbb{C}, X)$, $K \in \mathcal{L}(X, \mathbb{C})$ ensuring that $A + BK$ is exponentially stable, the continuous- and discrete-time stabilizing controller gain are linked by the following expression:*

$$K_d = \sqrt{2\delta}K(\delta - A - BK)^{-1} \tag{8.26}$$

where $\delta \in \rho(A) \cap \rho(A + BK)$, and $K_d \in \mathcal{L}(X, \mathbb{C})$ is the discrete-time stabilizing gain in the sense that $A_d + B_d K_d$ is strongly stable, where A_d and B_d are the corresponding discrete-time operators of A and B using the Cayley-Tustin transformation.

Proof. The stability of the closed-loop system in the continuous-time ($A_c = A + BK$) and discrete-time settings ($A_d + B_d K_d$) are related as follows:

$$\begin{aligned} A_{cd} &= -I + 2\delta(\delta - A_c)^{-1} \\ &= -I + 2\delta(\delta - A - BK)^{-1} \end{aligned} \tag{8.27}$$

$$\begin{aligned} A_{dc} &= A_d + B_d K_d \\ &= -I + 2\delta(\delta - A)^{-1} \cdot (I + \frac{1}{\sqrt{2\delta}} BK_d) \end{aligned} \tag{8.28}$$

where A_{cd} represents the closed-loop system that is first stabilized in the continuous setting by designing a continuous-time stabilizing gain K and then subsequently discretized, while A_{dc} denotes the system that is first discretized in time and then stabilized in the discrete setting by finding a discrete-time stabilizing gain K_d. Hence, by holding the equality of A_{cd} and A_{dc} expressions, one can directly calculate the solution of K_d in terms of K as follow:

$$K_d = \sqrt{2\delta}K(\delta - A - BK)^{-1} \tag{8.29}$$

Hence the proof is completed and it demonstrates the invariance of the Cayley-Tustin discretization with respect to closed-loop stabilization. ■

Corollary 8. *For an infinitesimal generator A of the C_0-semigroup $\mathbb{T}_A(t)$ on the Hilbert space X, $C_m \in \mathcal{L}(X, \mathbb{C})$, the continuous- and discrete-time stabilizing output injection gains can be linked by $L_{1d} = \sqrt{2\delta}(\delta - A - LC_m)^{-1}L$, such that the discretized observer error systems $A_{do} = A_d + L_{1d}C_{md}$ and $A_{od} = -I + 2\delta(\delta - A - LC_m)^{-1}$ share the same stability property by using continuous- and discrete-time output injection gains*

$L \in \mathcal{L}(\mathbb{C}, X)$ and $L_{1d} \in \mathcal{L}(\mathbb{C}, X)$ based on the Cayley-Tustin transformation, where $\delta \in \rho(A) \cap \rho(A + LC_m)$.

Proof. The proof is similar to Theorem 24 and hence can be completed by taking $C_{md} = B_d^*$ and $L_{1d} = K_d^*$. ■

8.3.3 Output feedback regulator design

Considering the unavailability and/or potentially prohibitive costs of installing spatially distributed sensing devices, the implementation of a state feedback compensator is not realistic. Hence, the output feedback regulator is more preferred in practical use. Along this line, the state feedback regulator is extended to an output feedback regulator, which is mathematically described as:

$$u_k = K_d \hat{x}_{k-1} + L_d \hat{q}_k, \ k \geq 1 \tag{8.30}$$

where $K_d \in \mathcal{L}(X, U)$, $L_d \in \mathcal{L}(\mathbb{C}^n, U)$, where $\hat{x}_{k-1}$ and $\hat{q}_k$ denote the estimated states of the plant and exogenous systems.

Along this line, one can construct Luenberger-type observers for plant and exo-system respectively as follows:

$$\hat{x}_k = A_d \hat{x}_{k-1} + B_d u_k + E_d \hat{d}_k + L_{1d}(y_{mk} - \hat{y}_{mk}) \tag{8.31a}$$

$$\hat{y}_{mk} = C_{md} \hat{x}_{k-1} + D_{md} u_k + \Upsilon_{md} \hat{d}_k \tag{8.31b}$$

$$\hat{q}_{k+1} = S_d \hat{q}_k + L_{2d}(y_{rk} - \hat{y}_{rk}) \tag{8.31c}$$

$$\hat{d}_k = F_d \hat{q}_k \tag{8.31d}$$

$$\hat{y}_{rk} = Q_d \hat{q}_k \tag{8.31e}$$

where $\hat{x}_0 \in X$, $\hat{q}_0 = \hat{q}^0 \in \mathbb{C}^n$, $k \geq 1$.

By direct manipulation of Eqs.(8.30)-(8.31) the following form is obtained:

$$\hat{x}_k^e = \begin{bmatrix} O_1 & O_2 \\ 0 & O_3 \end{bmatrix} \hat{x}_{k-1}^e + \begin{bmatrix} L_{1d} & 0 \\ 0 & L_{2d} \end{bmatrix} \begin{bmatrix} y_{mk} \\ y_{rk} \end{bmatrix} \tag{8.32a}$$

$$u_k = \begin{bmatrix} K_d & L_d \end{bmatrix} \hat{x}_{k-1}^e \tag{8.32b}$$

where $\hat{x}_k^e = [\hat{x}_k; \hat{q}_{k+1}]$, with $k \geq 1$, and

$$O_1 = A_d - L_{1d} C_{md} + (B_d - L_{1d} D_{md}) K_d$$

$$O_2 = (E_d - L_{1d}\Upsilon_{md})F_d + (B_d - L_{1d}D_{md})L_d$$
$$O_3 = S_d - L_{2d}Q_d$$

Finally, one can deploy Corollary 7 or Corollary 8 to determine the stabilizing output injection gain L_{1d} of the plant observer and apply pole-placement to find a stabilizing observer gain L_{2d} for the exo-system.

In the ensuing sections, two representative examples (CGLE and KSE) of fluid flow systems are given to illustrate the feasibility and applicability of the proposed regulator design method.

Remark 33. *In the output regulation problem, there are two components in the control laws (8.10) and (8.30), including a feedback control part accounting for (closed-loop) model stabilization and a feedforward control part ensuring the tracking error converging to zero as time goes to infinity. In general, one can adopt some existing feedback techniques to realize model stabilization of nonlinear distributed parameter systems, e.g., by using interpolants and projections methods [323, 324]. However, the feedforward control is much more difficult to find since it is determined by a set of nonlinear partial differential and algebraic equations, which is a nonlinear counterpart of the regulator equations encountered in the linear output regulation problems. The difficulties lie in the solvability of the so-called nonlinear regulator equations. To address that, a zero dynamics design method can be applied as in [312, 325], also see [326, Chapter 3] for nonlinear lumped parameter systems. Since this chapter is focused on the discrete-time linear output regulator design of fluid flow systems, the nonlinear regulator design will not be detailed in this chapter.*

Remark 34. *The presented Cayley-Tustin transformation is capable of transforming a continuous-time infinite-dimensional systems with possible unbounded operators (boundary control and observation) into a discrete-time infinite-dimensional model with all bounded operators. However, for the unstable system (e.g. Ginzburg-Landau equation) considered in this chapter, the corresponding model stabilization via boundary control could not be achieved by using traditional methods, and the most common practice to solve boundary control problem is to use backstepping methods, which is beyond the scope of this chapter. Another possible approach is to utilize model predictive*

controller design proposed in [65]. Since this chapter is focused on the discrete-time linear output regulator design of fluid flow systems with bounded control and observation operators, the boundary control and observation will not be addressed in this chapter.

8.4 Complex Ginzburg-Landau flow model

In this section, a linearized complex Ginzburg-Landau equation (CGLE), which takes form of a complex parabolic partial differential equation (PDE), is considered. More specifically, a discrete-time CGLE model is generated without spatial approximation or model reduction using the Cayley-Tustin discretization method. Additionally, a resolvent operator is found in a closed analytic form and deployed in the realization of the discrete CGLE model.

8.4.1 CGLE model description

Based on the model developments from [1, 58], a linearized complex Ginzburg-Landau equation is given as follows:

$$\frac{\partial x}{\partial t}(\bar{\xi},t) = a_1\frac{\partial^2 x(\bar{\xi},t)}{\partial \bar{\xi}^2} + a_2(\bar{\xi})\frac{\partial x(\bar{\xi},t)}{\partial \bar{\xi}} + a_3(\bar{\xi})x(\bar{\xi},t) \tag{8.33a}$$

$$x_{\bar{\xi}}(0,t) = \bar{u}(t) \tag{8.33b}$$

$$x_{\bar{\xi}}(\bar{\xi}_d,t) = 0 \tag{8.33c}$$

where $x(\bar{\xi},t) \in \bar{\mathcal{X}}$ is a complex-valued function with spatial variable $\bar{\xi} \in [0,\bar{\xi}_d] \subset \mathbb{R}$, and temporal variable $t \in [0,\infty)$. $\bar{\mathcal{X}} = L^2((0,\bar{\xi}_d),\mathbb{C})$ denotes a complex Hilbert space. In addition, a_1 is a positive constant, and $a_2(\bar{\xi})$ and $a_3(\bar{\xi})$ are two complex spatial functions. By applying an invertible state transformation $w(\bar{\xi},t) = x(\bar{\xi},t)g(\bar{\xi})$, $g(\bar{\xi}) = \exp(\frac{1}{2a_1}\int_0^{\bar{\xi}} a_2(\eta)d\eta)$ and spatial scaling $\xi = \frac{\bar{\xi}_d-\bar{\xi}}{\bar{\xi}_d}$, the convective term is eliminated as:

$$\frac{\partial w}{\partial t}(\xi,t) = b_1\frac{\partial^2 w(\xi,t)}{\partial \xi^2} + b_2(\xi)w(\xi,t) \tag{8.34a}$$

$$w_\xi(1,t) = u(t) \tag{8.34b}$$

$$w_\xi(0,t) = 0 \tag{8.34c}$$

with notations:

$$b_1 = \frac{a_1}{\bar{\xi}_d^2} \tag{8.35a}$$

$$b_2(\xi) = -\frac{1}{2}a_2'(\xi) - \frac{1}{4a_1}a_2^2(\xi) + a_3(\xi) \tag{8.35b}$$

where $w(\xi,t) \in \mathcal{X} = L^2((0,1),\mathbb{C})$, $\xi \in [0,1]$, and $a_2'(\xi)$ denotes the spatial derivative of $a_2(\xi)$. Moreover, $u(t)$ is the corresponding input of the scaled system (8.34).

By [42, Rem. 10.1.6], the boundary actuation can be transformed into an abstract in-domain control described by a spatial distribution function $\mathcal{B}(\xi)$ by solving an inner product formula as below:

$$\langle \mathcal{L}_o\phi, \psi\rangle = \langle \phi, \mathcal{A}^*\psi\rangle + \langle G\phi, \mathcal{B}^*\psi\rangle \tag{8.36}$$

where $\forall\phi \in \mathcal{D}(\mathcal{L}_o)$, $\forall\psi \in \mathcal{D}(\mathcal{A}^*)$, and $\mathcal{L}_o := b_1\frac{\partial^2}{\partial\xi^2}+b_2(\xi)$ with $\mathcal{D}(\mathcal{L}_o) = \mathcal{H}^1((0,1),\mathbb{C})$. The boundary control is denoted by G, namely, $G\phi := \phi_\xi(1)$. By introducing $\mathcal{X}_1 = \mathrm{Ker}\,(G)$, we obtain $\mathcal{A} = \mathcal{L}_o|_{\mathcal{X}_1}$ with the same definition as $\mathcal{L}_o$, but a different domain as $\mathcal{D}(\mathcal{A}) = \{\phi \in \mathcal{X} | \phi \in \mathcal{H}^1((0,1),\mathbb{C}) \cap \mathrm{Ker}\,(G)\}$. It can be found that $\mathcal{A}^* = b_1\frac{\partial^2}{\partial\xi^2} + \overline{b_2(\xi)}$ and $\mathcal{D}(\mathcal{A}^*) = \mathcal{D}(\mathcal{A})$. It is straightforward to obtain

$$\langle G\phi, \mathcal{B}^*\psi\rangle = \phi_\xi(1)\psi^*(1) \tag{8.37}$$

where $\forall\phi \in \mathcal{D}(\mathcal{L}_o)$, and $\forall\psi \in \mathcal{D}(\mathcal{A}^*)$. Comparing this with the fact that $G\phi = \phi_\xi(l)$, it follows that $\mathcal{B}(\xi) = \delta(\xi-1)$. For the sake of simplicity, we deploy the approximation $\mathcal{B}(\xi) \approx \frac{1}{2\varepsilon}\mathbf{1}_{[\xi_b-\varepsilon,\xi_b+\varepsilon]}(\xi)$ in what follows, where $\mathbf{1}_{[a,b]}(\xi)$ denotes the spatial shaping function: $\mathbf{1}_{[a,b]}(\xi) = \begin{cases} 1, \xi \in [a,b] \\ 0, \text{otherwise} \end{cases}$.
Therefore, a standard infinite-dimensional state-space model is formulated for the considered CGLE model as:

$$\frac{\partial w}{\partial t}(\xi,t) = \mathcal{A}w(\xi,t) + \mathcal{B}u(t) + \mathcal{E}d(t) \tag{8.38a}$$

$$y_c(t) = \mathcal{C}_c w(\xi,t) \tag{8.38b}$$

$$y_m(t) = \mathcal{C}_m w(\xi,t) \tag{8.38c}$$

where $u(t) \in L^2_{loc}([0,\infty),U)$, $d(t) \in L^2_{loc}([0,\infty),U_d)$, and $y(t) \in L^2_{loc}([0,\infty),Y)$, with U, U_d and Y being finite-dimensional spaces. More specifically, we consider $\dim U = \dim U_d = \dim Y = 1$. In addition, we consider $\mathcal{A} : \mathcal{D}(\mathcal{A}) \subset \mathcal{X} \mapsto \mathcal{X}$ being an

infinitesimal generator of a C_0-semigroup $\mathbb{T}(t)$ on $\mathcal{X}$, a bounded control operator $\mathcal{B} \in \mathcal{L}(U, \mathcal{X})$, a point observation operator $C_c \in \mathcal{L}(\mathcal{X}_1, Y)$, and a bounded disturbance operator $\mathcal{E} \in \mathcal{L}(U_d, \mathcal{X})$. More specifically, we aim to steer a flow at point ξ_c, so the output of interest is given as: $C_c := \int_0^1 \delta(\xi - \xi_c)(\cdot) d\xi$, where $\delta(\xi - \xi_c)$ denotes the Dirac delta function. For model well-posedness, we employ $C_\Lambda x = \lim_{\lambda \to +\infty} C_c \lambda (\lambda I - \mathcal{A})^{-1} x$ to replace C_c, where I is an identity operator, $x \in \mathcal{X}$, $\lambda \in \rho(\mathcal{A})$ (see details in [111]). As for observation, we introduce a bounded operator C_m as: $C_m(\xi) = \frac{1}{2\upsilon} \mathbf{1}_{[\xi_m - \upsilon, \xi_m + \upsilon]}(\xi)$. The design objective is to realize output reference tracking, disturbance rejection, and model stabilization. In order to find the stabilizing controller and observer gains, we provide a spectrum analysis utilized in pole-shifting design.

8.4.2 Spectrum analysis

It can be shown in several ways that for a constant spatial function $b_2(\xi)$ the spectrum of $\mathcal{A}$ can be found analytically as follows:

$$\lambda_n = b_2 - b_1 n^2 \pi^2 \tag{8.39a}$$

$$\phi_n = \sqrt{2} \cos(n\pi\xi) \tag{8.39b}$$

with $n \in \mathbb{N}$. As for $n = 0$, one has $(\lambda_0, \phi_0) = (\bar{b}_2, \mathbf{1}(\xi))$. However, for an arbitrary complex spatial function $b_2(\xi)$, it is not simple to find the spectrum characteristic of $\mathcal{A}$. For simplicity, we take maximum value (w.r.t. real part) of the spatial function $b_2(\xi)$ as $\bar{b}_2$ to approximate the original function $b_2(\xi)$, resulting in $\mathcal{A} := b_1 \frac{\partial^2}{\partial \xi^2} + \bar{b}_2$. Then, the spectrum of $\mathcal{A}$ naturally follows Eq.(8.39) with b_2 replaced by $\bar{b}_2$, which is given as follows:

$$\lambda_n = \bar{b}_2 - b_1 n^2 \pi^2 \tag{8.40a}$$

$$\phi_n = \sqrt{2} \cos(n\pi\xi) \tag{8.40b}$$

with $n = 1, 2, ...$ For $n = 0$, one has $(\lambda_0, \phi_0) = (\bar{b}_2, \mathbf{1}(\xi))$, which will be further exploited for the output regulator design in the ensuing sections.

8.4.3 CGLE resolvent operator

One of the most important steps in discrete regulator design is to find the resolvent operator from the continuous model so as to realize the corresponding discrete-

time model. To achieve this, Laplace transformation is usually performed on the continuous-time model (8.38). Considering that the resolvent operator purely depends on $\mathcal{A}$, one can drop $\mathcal{B}$ and $\mathcal{E}$ and directly apply Laplace transformation leading to:

$$\frac{\partial}{\partial \xi}\begin{bmatrix} w(\xi,s) \\ w_\xi(\xi,s) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ \frac{s-b_2}{b_1} & 0 \end{bmatrix}\begin{bmatrix} w(\xi,s) \\ w_\xi(\xi,s) \end{bmatrix} - \begin{bmatrix} 0 \\ \frac{1}{b_1} \end{bmatrix} w(\xi,0)$$

By defining $w_e(\xi,s) = [w(\xi,s); w_\xi(\xi,s)]$, it is straightforward to obtain the solution of $w_e(\xi,s)$ as follow:

$$w_e(\xi,s) = e^{M\xi} w_e(0,s) + \int_0^\xi e^{M(\xi-\eta)} A_0 w(\eta,0) d\eta \tag{8.41}$$

where

$$A_0 = \begin{bmatrix} 0 \\ -\frac{1}{b_1} \end{bmatrix},\ M = \begin{bmatrix} 0 & 1 \\ \frac{s-b_2}{b_1} & 0 \end{bmatrix}$$

After some simple algebraic manipulations, one can obtain $e^{M\xi} = [M_{ij}(\xi,s)]_{2\times 2}$, with $i,j = 1,2$, as below

$$e^{M\xi} = \begin{bmatrix} \cosh(\sqrt{\frac{s-b_2}{b_1}}\xi) & \sqrt{\frac{b_1}{s-b_2}}\sinh(\sqrt{\frac{s-b_2}{b_1}}\xi) \\ \sqrt{\frac{s-b_2}{b_1}}\sinh(\sqrt{\frac{s-b_2}{b_1}}\xi) & \cosh(\sqrt{\frac{s-b_2}{b_1}}\xi) \end{bmatrix}$$

Substituting boundary conditions $w_\xi(1,s) = 0 = w_\xi(0,s)$ into Eq.(8.41), one can solve for $w(0,s)$ so that the resolvent operator is determined in the closed analytic form:

$$\begin{aligned} \mathcal{R}(s,\mathcal{A})(\cdot) &= \frac{M_{11}(\xi,s)}{b_1 M_{21}(l,s)} \int_0^1 M_{22}(1-\eta,s)(\cdot)d\eta - \frac{1}{b_1}\int_0^\xi M_{12}(\xi-\eta,s)(\cdot)d\eta \\ &= \frac{1}{\sqrt{b_1(s-b_2)}} \times \Big[-\int_0^\xi \sinh(w_s(\xi-\eta))(\cdot)d\eta \\ &\quad + \frac{\cosh(w_s\xi)}{\sinh(w_s)} \int_0^1 \cosh(w_s(1-\eta))(\cdot)d\eta \Big] \end{aligned} \tag{8.42}$$

where $w_s = \sqrt{\frac{s-b_2}{b_1}}$. Then, a direct calculation leads to the expressions of $(\mathcal{A}_d, \mathcal{B}_d, \mathcal{C}_d, \mathcal{D}_d)$ as:

$$\begin{aligned} \mathcal{A}_d(\cdot) &= -(\cdot) - \frac{2\delta}{\sqrt{b_1(\delta-b_2)}} \int_0^\xi \sinh(w_\delta(\xi-\eta))(\cdot)d\eta \\ &\quad + \frac{2\delta\cosh(w_\delta\xi)}{\sqrt{b_1(\delta-b_2)}\sinh(w_\delta)} \int_0^1 \cosh(w_\delta(1-\eta))(\cdot)d\eta \end{aligned} \tag{8.43a}$$

$$\mathcal{B}_d = \begin{cases} \frac{\sqrt{2\delta}}{2\varepsilon(\delta-b_2)}\Big[1-\cosh\left(w_\delta\left(\xi-\xi_b+\varepsilon\right)\right)+ \\ \qquad \frac{\cosh(w_\delta\xi)\cosh(w_\delta(\xi-\xi_b+\varepsilon))}{\sinh(w_\delta)}\Big], \xi\in[\xi_b-\varepsilon,\xi_b+\varepsilon] \\ \frac{\sqrt{2\delta}\cosh(w_\delta\xi)\cosh(w_\delta(\xi-\xi_b+\varepsilon))}{2\varepsilon(\delta-b_2)\sinh(w_\delta)}, \text{otherwise} \end{cases} \tag{8.43b}$$

$$\begin{aligned} \mathcal{C}_{cd}(\cdot) = &-\frac{\sqrt{2\delta}}{\sqrt{b_1(\delta-b_2)}}\int_0^{\xi_c}\sinh(w_\delta(\xi_c-\eta))(\cdot)d\eta \\ &+\frac{\sqrt{2\delta}\cosh(w_\delta\xi_c)}{\sqrt{b_1(\delta-b_2)}\sinh(w_\delta)}\int_0^{1}\cosh(w_\delta(1-\eta))(\cdot)d\eta \end{aligned} \tag{8.43c}$$

$$\mathcal{D}_{cd} = \frac{\cosh\left(w_\delta\xi_c\right)\cosh\left(w_\delta\left(\xi_c-\xi_b+\varepsilon\right)\right)}{2\varepsilon\left(\delta-b_2\right)\sinh\left(w_\delta\right)},\ \xi_c<\xi_b-\varepsilon \tag{8.43d}$$

where $w_\delta=\sqrt{\frac{\delta-b_2}{b_1}}$. With $\xi_c<\xi_b-\varepsilon<1$, we note that

$$\lim_{s\to+\infty}\mathcal{G}_c(s)=\lim_{\delta\to+\infty}\mathcal{D}_{cd}(\delta)=0$$

In a similar fashion, we can obtain analytic expressions of $\mathcal{D}_{md}$ and other discrete operators. With $\xi_m+\upsilon<\xi_b-\varepsilon$, one can further infer that

$$\lim_{s\to+\infty}\mathcal{T}_c(s)=\lim_{\delta\to+\infty}\mathcal{D}_{md}(\delta)=0$$

which implies that the system (8.38) is a well-posed regular system [121].

8.4.4 Simulation study

In the simulation section, the designed discrete-time regulator is implemented to regulate the real part of the controlled output of the linearized CGLE model (8.38). More specifically, sinusoidal signals generated by the discrete-time exo-system are deployed as disturbance and reference signals. In addition, the model parameters of the considered CGLE model in this chapter are adopted from [303, 1], and are given in Table 8.1.

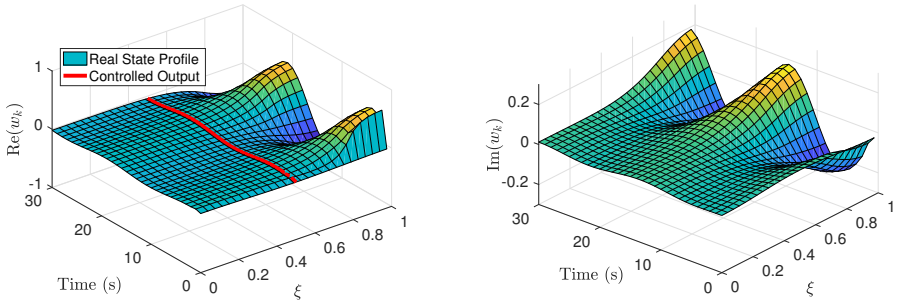
In this case, the disturbance distribution is described by $\mathcal{E}(\xi)=\mathbf{1}_{[0,0.5]}(\xi)$, and other numerical parameters are taken as $\Delta\xi=0.00125$, $h=0.5$, $\xi_d=1.5$, $\xi_c=0.5$, $\xi_b=0.9$, $\varepsilon=0.1$, $\xi_m=0.1$ and $\upsilon=0.1$. Additionally, the initial condition utilized here is given as: $\hat{w}_0=(\sqrt{2}-\sqrt{2}i)\times(0.01\times\cos(2\pi\xi)-0.0003\times\cos(4\pi\xi))$. By applying Theorem 23 and Corollary 7, K_d and L_{1d} can be consequently determined. By performing pole placement, the real part of the eigenvalues of $(S_d-L_{2d}Q_d)$ are placed at -0.1. In order to generate sinusoidal signals, we take $S_d=[0.9824,0.1868;-0.1868,0.9824]$,

Table 8.1: Parameters considered for the CGLE model (where $j^2 = -1$) [1]

Parameter	Numerical Value
a_1	0.01667
$a_2(\bar{\xi})$	$(0.1697 + 0.04939j)\bar{\xi}^2 - (0.1748 + 0.06535j)\bar{\xi}$ $-0.09061 + 0.001485j$
$a_3(\bar{\xi})$	$(0.1563 - 0.001352j)\bar{\xi}^4 + (-1590 + 0.6278j)\bar{\xi}^3$ $+(0.3958 - 1.8577j)\bar{\xi}^2 + (-1.6852 + 1.6759j)\bar{\xi}$ $+1.2645 - 0.2489j$

$\hat{q}^0 = [-0.4; 1.4]$, $F_d = [0, 0.010]$, and $Q_d = [0.030, 0]$, leading to periodic reference and disturbance signals as: $y_{rk} = 0.03 \times \sin(0.06k\pi)$ and $d_k = 0.01 \times \cos(0.06k\pi)$. Revisiting Eq.(8.21b) and Eq.(8.6), the discrete feedforward gain can be solved as $L_d = [0.0224, 0.0340]$, which completes the control action u_k.

After simulation of 30 seconds, the closed-loop state and output evolution profiles are depicted in Fig. 8.3 and Fig. 8.4. It is apparent that the designed output regulator can stabilize the originally unstable system, and the spatiotemporal profile shows the expected periodic behaviour. From the perspective of output tracking performance, one can clearly see that the real part of the controlled output follows the desired reference signal and the tracking error converges to zero quickly, which demonstrates the effectiveness of the proposed output regulator design.



(a) Real part (b) Imaginary part

Figure 8.3: State evolution of closed-loop GLE system in the case of regulation of the real part of the controlled output.

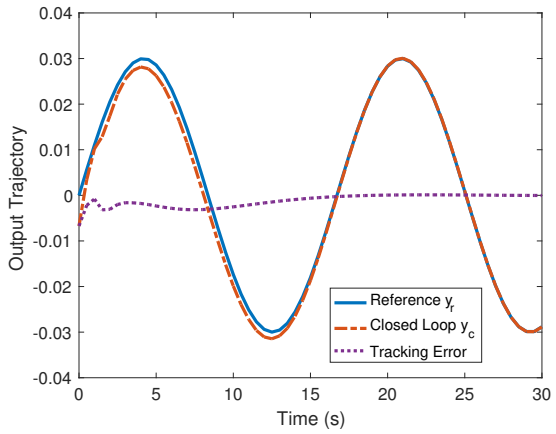


Figure 8.4: Output trajectory of closed-loop GLE system in the case of regulation of the real part of the controlled output.

8.5 Kuramoto-Sivashinsky equation

In this section, a continuous-time nonlinear Kuramoto-Sivashinsky equation (KSE) is introduced to describe the falling thin film dynamics. For the sake of simplicity, a linear KSE is achieved by performing linearization. In the same manner, the Cayley-Tustin transformation is utilized for time discretization for the linear KSE model. Moreover, an explicit closed-form solution is obtained for the corresponding resolvent operator, which is exploited for the discrete-time regulator design.

8.5.1 KSE model description

In this section, a general nonlinear Kuramoto-Sivashinsky equation with an in-domain actuation is considered as follows [327]:

$$\begin{aligned}\frac{\partial x}{\partial t}(\zeta,t) &+ \nu x_{\zeta\zeta\zeta\zeta}(\zeta,t) + x_{\zeta\zeta}(\zeta,t) \\ &+ x_{\zeta}(\zeta,t)\, x(\zeta,t) + B(\zeta)u(t) + E(\zeta)d(t) = 0\end{aligned} \tag{8.44}$$

with boundary conditions and initial condition:

$$x(0,t) = 0,\ x(l,t) = 0,\ x_{\zeta}(0,t) = 0,\ x_{\zeta}(l,t) = 0 \tag{8.45a}$$

$$x(\zeta,0) = x_0(\zeta) \tag{8.45b}$$

where x describes the thickness of a thin falling flow, as illustrated in Fig. 8.2. $t \in [0,\infty)$ and $\zeta \in [0,l]$ denote temporal and spatial variables, respectively. The total

length of the vertical pipe is denoted by l. Additionally, x_t represents the first-order temporal derivative of state x, while x_ζ, $x_{\zeta\zeta}$ and $x_{\zeta\zeta\zeta\zeta}$ stand for first-order, second-order and fourth-order derivatives of state x with respect to space. $u(t)$ denotes a manipulated variable representing the angle at which the air front is acting on the film-annulus separation points, and $B(\zeta)$ is a spatial function describing the control action along the vertical pipe. In particular, we consider $B(\zeta) = \frac{1}{2\varepsilon}\mathbf{1}_{[\zeta_b-\varepsilon,\zeta_b+\varepsilon]}(\xi)$. $d(t)$ represents a distributed disturbance characterized by a bounded spatial function $E(\zeta)$. After linearizing the KSE (8.44) around the spatially uniform steady state $x_{ss}(\zeta) = 0$, a linear KSE is attained as follow [311]:

$$x_t(\zeta,t) + \nu x_{\zeta\zeta\zeta\zeta}(\zeta,t) + x_{\zeta\zeta}(\zeta,t) + B(\zeta)u(t) + E(\zeta)d(t) = 0 \tag{8.46}$$

with:

$$x(0,t) = 0,\ x(l,t) = 0,\ x_\zeta(0,t) = 0,\ x_\zeta(l,t) = 0 \tag{8.47a}$$

$$x(\zeta,0) = x_0(\zeta) \tag{8.47b}$$

To complete the KSE system, we define an controlled output $y_c(t)$ and an measured output $y_m(t)$ as below:

$$y_c(t) = C_c x(\zeta,t) \tag{8.48a}$$

$$y_m(t) = C_m x(\zeta,t) \tag{8.48b}$$

where C_c represents a point observation described by $C_c(\cdot) = \int_0^l \delta(\zeta-\zeta_c)(\cdot)\,d\zeta$, with $\delta(\zeta-\zeta_c)$ denoting the Dirac delta function. In other words, the controlled output extracts state information at a specific spatial point ζ_c of interest, i.e., $y_c(t) = x(\zeta_c,t)$. In addition, a bounded operator C_m is introduced to describe $y_m(t)$ as: $C_m(\zeta) = \frac{1}{2\upsilon}\mathbf{1}_{[\zeta_m-\upsilon,\zeta_m+\upsilon]}(\zeta)$. In doing so, a continuous-time KSE system is constructed in the following abstract state-space form:

$$\frac{\partial x}{\partial t}(\zeta,t) = \mathcal{A}x(\zeta,t) + \mathcal{B}u(t) + \mathcal{E}d(t) \tag{8.49a}$$

$$y_c(t) = C_c x(\zeta,t) \tag{8.49b}$$

$$y_m(t) = C_m x(\zeta,t) \tag{8.49c}$$

where $\mathcal{A} := -\nu\frac{\partial^4}{\partial\zeta^4} - \frac{\partial^2}{\partial\zeta^2}$ with domain $\mathcal{D}(\mathcal{A}) = \{\phi(\zeta) \in L_2(0,l) | \phi, \phi_\zeta, \phi_{\zeta\zeta}, \phi_{\zeta\zeta\zeta}$ are abs. con., $\phi_{\zeta\zeta\zeta} \in L_2(0,l), \phi(0) = 0, \phi(l) = 0, \phi_\zeta(0) = 0, \phi_\zeta(l) = 0\}$. In addition, we have

$\mathcal{B} := B(\zeta)$ and $\mathcal{E} := E(\zeta)$. This standard model structure taking the same form as Eq. (8.1) is suitable for model time discretization through the Cayley-Tustin approach as described previously.

8.5.2 KSE resolvent operator

In a similar manner as we solved the resolvent operator of CGLE model, we aim to determine the resolvent operator of KSE model in this section. Differently, it needs more complex manipulation to solve for the KSE resolvent operator $\mathcal{R}$ owing to the higher order derivatives in the state evolution operator $\mathcal{A} := -\nu\frac{\partial^4}{\partial\zeta^4} - \frac{\partial^2}{\partial\zeta^2}$. By directly applying Laplace transform to the linearized Kuramoto-Sivashinsky equation (8.46) and ignoring the input and disturbance, one achieves the following:

$$x_{\zeta\zeta\zeta\zeta}(\zeta, s) = -\frac{s}{\nu}x(\zeta, s) - \frac{1}{\nu}x_{\zeta\zeta}(\zeta, s) + \frac{1}{\nu}x_0(\zeta) \tag{8.50}$$

To utilize boundary conditions for solving the state solution, we introduce new state derivatives to Eq.(8.50) as follows:

$$\frac{\partial \mathbf{x}}{\partial \zeta} = A\mathbf{x} + \mathbf{x}_0 \tag{8.51}$$

where $\mathbf{x} = [x; x_\zeta; x_{\zeta\zeta}; x_{\zeta\zeta\zeta}]$, $\mathbf{x}_0 = [0; 0; 0; \frac{1}{\nu}x_0(\zeta)]$, and

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -\frac{s}{\nu} & 0 & -\frac{1}{\nu} & 0 \end{bmatrix}$$

Direct integration of Eq.(8.51) leads to:

$$\mathbf{x}(\zeta, s) = e^{A\zeta}\mathbf{x}(0, s) + \int_0^\zeta e^{A(\zeta-\eta)}\mathbf{x}_0(\eta)d\eta \tag{8.52}$$

where $e^{A\zeta} = [a_{ij}(\zeta, s)]_{4\times4}$, with $i, j = 1, 2, 3, 4$. Thus, substituting the boundary conditions $x(0, s) = 0$ and $x_\zeta(0, s) = 0$ into Eq.(8.52) leads to the following simplification:

$$\begin{bmatrix} x(\zeta, s) \\ x_\zeta(\zeta, s) \end{bmatrix} = \begin{bmatrix} a_{13}(\zeta, s) & a_{14}(\zeta, s) \\ a_{23}(\zeta, s) & a_{24}(\zeta, s) \end{bmatrix} \begin{bmatrix} x_{\zeta\zeta}(0, s) \\ x_{\zeta\zeta\zeta}(0, s) \end{bmatrix} + \int_0^\zeta \begin{bmatrix} a_{14}(\zeta - \eta, s) \\ a_{24}(\zeta - \eta, s) \end{bmatrix} x_0(\eta)\, d\eta \tag{8.53}$$

Then, the remaining two boundary conditions $x(l,s)=0$ and $x_\zeta(l,s)=0$ in Eq.(8.47a) are deployed in Eq.(8.53) such that $x_{\zeta\zeta}(0,s)$ and $x_{\zeta\zeta\zeta}(0,s)$ can be determined as follows:

$$\begin{bmatrix} x_{\zeta\zeta}(0,s) \\ x_{\zeta\zeta\zeta}(0,s) \end{bmatrix} = -\frac{1}{v}M^{-1}\int_0^l \begin{bmatrix} a_{14}(l-\eta,s) \\ a_{24}(l-\eta,s) \end{bmatrix} x_0(\eta)\,d\eta \tag{8.54}$$

where $M = [a_{13}(l,s), a_{14}(l,s); a_{23}(l,s), a_{24}(l,s)]$. The invertibility of M must be checked to ensure that Eq.(8.54) holds.

Finally, one can obtain the solution of $x(\zeta,s)$ by plugging Eq.(8.54) into Eq.(8.53), which leads to the associated resolvent operator of the KSE system as follows:

$$\begin{aligned}\mathcal{R}(\zeta,\mathcal{A})(\cdot) =& \frac{1}{v}\int_0^\zeta a_{14}(\zeta-\eta,s)(\cdot)d\eta \\ &+ \frac{a_{13}(\zeta,s)}{v\left[a_{23}(l,s)\,a_{14}(l,s)-a_{13}(l,s)\,a_{24}(l,s)\right]} \\ &\times \int_0^l \left[a_{14}(l-\eta,s)\,a_{24}(l,s)-a_{24}(l-\eta,s)\,a_{14}(l,s)\right](\cdot)d\eta \\ &- \frac{a_{14}(\zeta,s)}{v\left[a_{23}(l,s)\,a_{14}(l,s)-a_{13}(l,s)\,a_{24}(l,s)\right]} \\ &\times \int_0^l \left[a_{14}(l-\eta,s)\,a_{23}(l,s)-a_{24}(l-\eta,s)\,a_{13}(l,s)\right](\cdot)d\eta \end{aligned} \tag{8.55}$$

where due to the specific structure of A, an analytic expression of resolvent operator is found by:

$$a_{23}(\zeta,s) = \frac{v\left(\varpi\sinh(\varpi\zeta)-\omega\sinh(\omega\zeta)\right)}{\sqrt{1-4sv}} \tag{8.56a}$$

$$a_{13}(\zeta,s) = \frac{v\left(\cosh(\varpi\zeta)-\cosh(\omega\zeta)\right)}{\sqrt{1-4sv}} \tag{8.56b}$$

$$a_{24}(\zeta,s) = \frac{v\left(\cosh(\varpi\zeta)-\cosh(\omega\zeta)\right)}{\sqrt{1-4sv}} \tag{8.56c}$$

$$\begin{aligned} a_{14}(\zeta,s) =& \frac{\left(1-\sqrt{1-4sv}\right)}{4s^2\sqrt{(2-8sv)}} \times \left[\left(2sv-1-\sqrt{1-4sv}\right)\right. \\ &\left.\times\sqrt{2}\varpi\sinh(\varpi\zeta)+2sv\sqrt{2}\omega\sinh(\omega\zeta)\right] \end{aligned} \tag{8.56d}$$

where $\varpi = \sqrt{\frac{-1+\sqrt{1-4sv}}{2v}}$ and $\omega = \sqrt{\frac{-1-\sqrt{1-4sv}}{2v}}$. It is observed that, for a given positive real s, ϖ stays as a complex number while ω can be a complex or real number with different choices of v. Due to the existence and multiplication with hyperbolic functions ("sinh" and "cosh"), it is straightforward to check that $a_{ij}(\zeta,s)\in L^2((0,l),\mathbb{R})$, $\forall s\in\mathbb{R}$, with $i,j=1,2$.

The closed-form expressions of discrete-time operators in Eq.(8.4) corresponding to the discrete KSE system can be determined by simply substituting Eq.(8.55)-(8.56) back into Eq.(8.5). By comparing the order of s in $\mathcal{R}(\zeta, \mathcal{A})$, it can be shown that for any given $\xi_c = 0.5$ and $\xi_m + \upsilon \leq \xi_b - \varepsilon$

$$\lim_{s\to+\infty} \mathcal{G}_c(s) = \lim_{\delta\to+\infty} \mathcal{D}_{cd}(\delta) = 0$$

$$\lim_{s\to+\infty} \mathcal{T}_c(s) = \lim_{\delta\to+\infty} \mathcal{D}_{md}(\delta) = 0$$

which implies that the system (8.49) is a well-posed regular system [121].

8.5.3 Simulation study

In this section, the proposed discrete-time output regulator is implemented to the KSE system and the results are discussed. To demonstrate the effectiveness of the developed regulator, we consider an unstable KSE with $\nu = -3$. The characteristic equation is found by $\nu s^4 + s^2 = \lambda$ (or equivalently written as $\left(s^2 + \frac{1}{2\nu}\right)^2 - \frac{\lambda}{\nu} - \frac{1}{4\nu^2} = 0$), so the eigenvalue spectrum is roughly given by the range $(-\infty, \frac{1}{4\nu}]$ [311]. By revisiting Theorem 23 and Corollary 7, K_d and L_{1d} can be found.

To achieve asymptotic tracking and disturbance rejection of periodic signals, we take $S_d = [0.9995, 0.0314; -0.0314, 0.9995]$, $\hat{q}^0 = [-0.2; 1.2]$, $F_d = [0, 0.1]$, $Q_d = [0.01, 0]$, $\mathcal{E}(\zeta) = \mathbf{1}_{[0,1]}(\zeta)$, $\zeta_b = 0.98$, $\varepsilon = 0.01$, $\zeta_m = 0.1$, $\upsilon = 0.1$, and $\zeta_c = 0.5$. The reference and disturbance signals are generated as: $d_k = 0.1 \times \cos(0.009k\pi)$ and $y_{rk} = 0.01 \times \sin(0.009k\pi)$. Using Eq.(8.21b) and Eq.(8.6), the discrete feedforward gain is found as $L_d = [2640.6559, -23.2574]$, which leads to the control law u_k. Additionally, pole placement is applied to move the real part of eigenvalues of $(S_d - L_{2d}Q_d)$ to -0.5.

The simulated pipe length is taken as $l = 1$m with a spatial interval $\Delta l = 0.005$m. Moreover, a time discretization interval $h = 0.1$s is chosen with total simulation of 40 seconds. As shown in Fig. 8.5, the state is steered to reject a cosine disturbance and follow a sinusoid wave using the closed-loop control. As for the output tracking performance, it is apparent that the controlled output rapidly converges to the desired reference and the tracking error goes to near zero around 30s, as illustrated in Fig. 8.6, which further verifies the feasibility of the proposed regulator design.

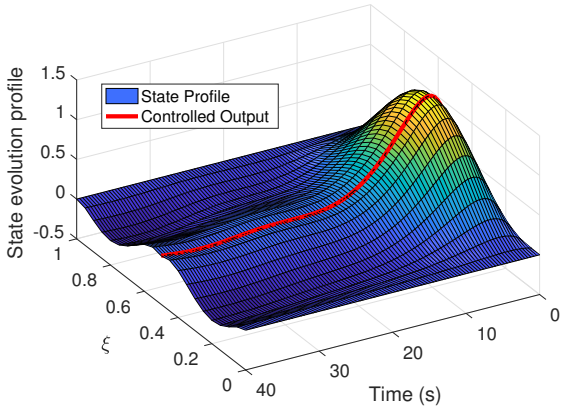


Figure 8.5: State evolution of the closed-loop KSE system.

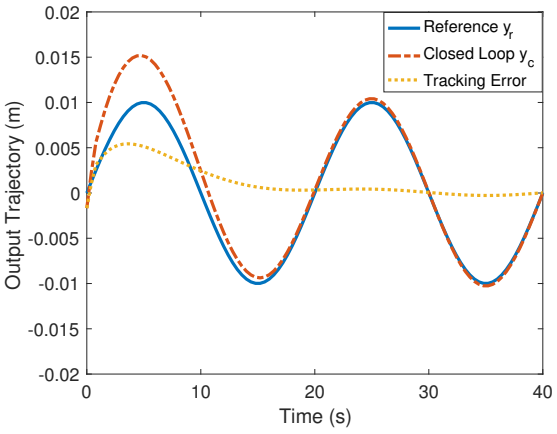


Figure 8.6: Output regulation performance of the closed-loop KSE system.

8.6 Conclusion

In this chapter, discrete-time output regulators are designed for PDE model-based fluid flow manipulation and regulation. To model the vortex shedding process and falling thin film dynamics, the linearized complex Ginzburg-Landau and Kuramoto-Sivashinsky models are utilized. For realistic implementation of regulators in digital computer systems, the Cayley-Tustin time discretization method is utilized for discrete-in-time analysis with system properties preserved and no spatial approximation. Standard finite-dimensional continuous-time regulator design framework is extended to an infinite-dimensional discrete-time setting with application to CGLE

and KSE systems. As key components for the formulation of a discrete-time model (A_d, B_d, C_d, D_d), the resolvent operators corresponding to CGLE and KSE are found in analytic forms, which are used to show the well-posedness of CGLE and KSE models. Simulation results show that the developed design method is capable of stabilizing the system, tracking the periodic output references in the CGLE model and higher-order dynamics in the KSE case. Undesired periodic disturbance signals are rejected for both systems. In the future work, the implementation and realization of the proposed design will be explored experimentally on vortex shedding and falling thin film control problems.

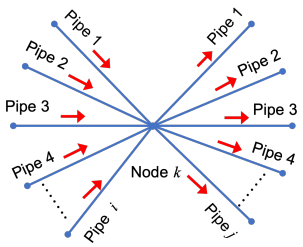
Chapter 9

Discrete-time Modeling and Output Regulation of Gas Pipeline Network

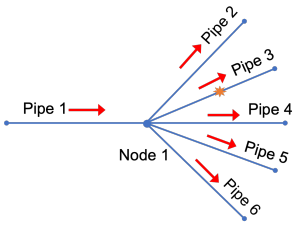
9.1 Introduction

Natural gas transportation through pipeline networks is of significance for energy planning, scheduling, and operations management. With the increasing need for energy storage and transportation, more and more sophisticated and large-scale pipe networks have been established to gather and distribute oil/gas from spatially distributed reservoirs to local consumers. However, in reality, pipeline operations often involve complex interactions between various supplies and demands and might encounter pressure surges and oscillations [328]. On the other hand, the full state information of pipeline systems is not often available due to either physical constraints on sensor implementations or prohibitive costs of installation of spatially-distributed sensors, leading to limited measurements, which poses another challenge for pipeline condition monitoring and regulation. Hence, it is of practical importance to design output feedback regulators to meet various operating and scheduling requirements on distributed gas pipeline networks.

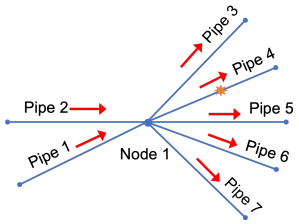
To model flow dynamics in pipeline systems, the real-time transient model (RTTM), as a representative of the first principle modelling approach, has gained a lot of attention from academia and industry [329]. Essentially, the real-time transient model is formulated based on mass, momentum, and energy balance laws, making it capable of capturing the physical nature of pipeline transport flow that is difficult to ensure by using data-driven methods. When it comes to pipeline engineering practices, the



(a) A gas pipe network of $N = i + j$ pipes.



(b) A gas pipe network of $N = 6$ pipes.



(c) A gas pipe network of $N = 7$ pipes.

Figure 9.1: Different architectures of star-shaped gas pipe networks.

RTTM method has been utilized for mathematical modelling of liquid and gas flows [329, 330], single straight pipeline and pipe network [331, 332], and pipeline systems with associated accessories (such as pumps, valves, reducers, etc.) [333, 235]. Among these, a special interest has been given to star-shaped gas pipeline networks due to their wide existence in realistic pipeline network topologies as shown in Fig. 9.1(a). More specifically, the coupling conditions (also referred to as Rankine-Hugoniot condition) at the vertex of the gas pipe networks and related modelling analysis have been intensively investigated [334]. Among these, relevant studies on classical solutions, feedback stabilization, and time-delay boundary stabilization of such systems were reported in [335, 332, 336]. However, inadequate efforts have been made toward output regulation of pipeline network systems although servo control is significant for energy planning and flow regulation of pipeline network systems, which motivates this chapter.

Output regulation has been an active control topic during the past decades. The main idea of output regulation is to track desired output references and reject undesired disturbances, while ensuring the stability of the closed-loop system. As stated, there are two distinct design problems in the realm of output regulation, i.e., state feedback regulator design and output/error feedback regulator design. In general,

full state information is assumed to be available for a state feedback regulator design while only output or tracking error is known in output/error feedback regulator design problems. Following the pioneering work by Francis on the internal model control [247], the output regulation theory has been developed and extended to various linear and nonlinear lumped parameter systems (LPSs) [248, 251]. Furthermore, the internal model control theory has been applied for output regulation of distributed parameter systems (DPSs) governed by partial differential equations (PDEs), including parabolic PDE systems [10, 337] and hyperbolic PDE systems [338, 13].

On the other hand, the backstepping approach has been developed for output regulation of DPSs, based on Krstic's work [17, 15]. As a general model structure that is similar to the single pipeline model, a class of 2×2 linear bidirectional coupled hyperbolic PDE systems has been studied in output regulator design problems using the backstepping approach. Specifically, 2×2 linear hyperbolic systems with boundary and in-domain disturbances as well as co-located sensing and actuation were considered, and backstepping transforms were performed to acquire the control action to annihilate the disturbance effect in [263, 264]. Recently, this work was further extended to a class of general linear heterodirectional hyperbolic systems with spatially-varying coefficients [267] and ODE-PDE-ODE cascade systems [339].

However, the existing contributions have not fully addressed the output regulation of gas pipeline network systems. Furthermore, most of the work has been done in a continuous-time setting and the question arises in how to ensure the convergence of late-lumping controllers in numerical realizations [18]. Considering that system theoretic developments often go in parallel for continuous-time and discrete-time systems, it is common practice to derive results for one class of systems and then map these over to the other by using a bilinear transformation in finite-dimensional system theory [19]. Since it is often possible to use some bilinear transform to avoid repeating tedious derivations if results have been obtained for continuous systems and similar results are needed for the discrete ones and vice versa [19]. When it comes to infinite-dimensional systems, Cayley bilinear transform has been used in establishing 1-1 correspondences of continuous- and discrete-time systems in terms of conservativity (energy preserving) [19, 20], reachability and observability [19], stability [21, 22], stabilizability, controllability and observability [4, 23], and optimality [24].

These facts motivate us to propose a novel discrete-time regulator design method for pipeline networks by using the Cayley-Tustin transform, and the main contributions are the following: 1). The linear continuous-time model with unbounded operators is transformed to a linear discrete-time infinite-dimensional model with all bounded operators, and the essential continuous-time properties (including stability, controllability and observability) stay invariant under the Cayley-Tustin transformation; 2). The discrete output regulator is designed without the usual technical difficulties present in the continuous-time infinite-dimensional case and its solvability conditions are provided in the form of discrete output regulator equations. The proposed design method is validated with applications in output regulation of a single gas pipeline and a stared-shaped gas pipeline network under consideration of coupling conditions at junction joints.

The rest of this chapter is presented in sections. Following the model description, steady states analysis, model time discretization, and resolvent operator determination, the discrete-time single gas pipeline model are provided in Section 9.2. Section 9.3 presents a star-shaped gas pipe network model and its discretization, with consideration of coupling conditions at junction joints. Regarding output regulation, details on state and output feedback regulator design and realization are provided in Section 9.4. An internal stability analysis is provided in Section 9.5 for the considered pipeline system. To verify the feasibility and applicability of the developed output regulator design, two numerical examples are illustrated in Section 9.6. Finally, conclusions are drawn in Section 9.7.

9.2 Single pipe model for gas transportation

In this section, the isothermal Euler equations in the form of 1st-order coupled nonlinear hyperbolic PDEs are introduced for pipeline gas flow modelling. In particular, a linearized model is obtained by linearizing the original nonlinear model around some given operating mode of interest. By the use of Cayley-Tustin transform, a discrete-time state-space model for describing the single pipeline system is established and realized by determining the resolvent operator, which is amenable to the discrete output regulator design.

9.2.1 Model description

In this chapter, we consider that: 1) Each pipe segment is a rigid buried pipe implying that the cross-sectional area change along the stream of gas can be negligible, and the parameters are given in Table 9.1; 2) The single-phase gas flow is assumed to be a 1-D flow indicating that mass flux, density and pressure are only functions of time and position along the pipe axis; 3) All states are subsonic, i.e., $\frac{|q|}{\rho} \ll a$, where ρ, q, and a represent the gas density, mass flux and speed of sound in the considered gas flow ($q = \rho v$, v denotes mass flow velocity); 4) The flow is viscous (viscosity causes shear stresses in the moving flow); 5) There is no vacuum present; 6) The flowing process is isothermal which implies $p/\rho = zRT/(M_g) = a^2$, where p is the pressure, z represents the natural gas compressibility factor, R stands for the universal gas constant, T is the absolute gas temperature, and M_g is the gas molecular. The last assumption implies that temperature changes due to pressure changes and friction effects can be neglected in the flow system. This is motivated by the fact that most of the pipeline infrastructures are buried underground, and for example, in the case of a buried gasoline pipeline, the heat flux between the flow and surrounding soil is negligible [241].

Table 9.1: Pipeline parameters

Item	Notation	Numerical Value
Length	l	$10,000$ m
Speed of Sound	a	370.7010 m/s
Friction Coefficient	λ	0.011
Gravity Acceleration	g	9.81 m/s^2
Diameter	D	0.5 m
Inclination Angle	α	-0.00256

Based on the continuity and momentum equations, the 1-D transient gas flow model is constructed for a single pipeline system as follows [340, 330]:

$$\rho_t + q_x = 0 \tag{9.1a}$$

$$q_t + \left(q^2/\rho + a^2\rho\right)_x = -\lambda \frac{q|q|}{2D\rho} - g\rho \sin\alpha \tag{9.1b}$$

Using the equation of state of ideal gas, one obtains:

$$p/\rho = zRT/(M_g) = a^2 \tag{9.2}$$

with boundary conditions:

$$q(0,t) = q_0 + u(t),\ p(l,t) = p_0,\ \rho(l,t) = \rho_0 \tag{9.3}$$

where $x \in [0, l]$ and $t \in [0, \infty)$ are spatial and temporal coordinates, and subscripts t and x refer to derivatives in time and x-direction respectively. The length of the straight pipe is l, and the quantity of $g \sin \alpha$ is the x-component of the gravity acceleration g. λ is a dimensionless friction factor described by the following Chen's equation. Additionally, the boundary input is denoted by $u(t)$. For high-pressure gas pipes, typical values are taken as: $q/\rho \approx 10$ m/s, $a \approx 300$ m/s, and $a^2\rho \approx 7$ MPa. In particular, the momentum term q^2/ρ is of order 10^{-3} while the friction term is of order 10^{-2} for a pipe with length $l = 100$ km and operational time of 1 hour [341, 340]. Hence, the momentum term q^2/ρ can be dropped for the sake of simplicity, leading to the Weymouth equation as follows:

$$\rho_t + q_x = 0 \tag{9.4a}$$

$$q_t + a^2\rho_x = -\lambda \frac{q|q|}{2D\rho} - g\rho \sin \alpha \tag{9.4b}$$

Inserting Eq.(9.2) into the system (9.4) yields

$$p_t + a^2 q_x = 0 \tag{9.5a}$$

$$q_t + p_x = -\lambda \frac{q|q|a^2}{2Dp} - \frac{gp \sin \alpha}{a^2} \tag{9.5b}$$

$$q(0,t) = q_0 + u(t),\ p(l,t) = p_0 \tag{9.5c}$$

which are often referred to as the *isothermal Euler equations* [334, 336, 332, 330] and can also be expressed by pressure p and mass flow rate M (with $M = \rho v A$). As for the characterization of solutions, spaces and regularity of the nonlinear isothermal Euler equations, we refer to the references [342, 343, 344, 345, 346, 347]. In this chapter, we assume steady state friction, which can be explicitly determined by the following Chen's equation [348]:

$$\frac{1}{\sqrt{\lambda}} = -2 \log \left(\frac{\varepsilon}{3.7065D} - \frac{5.0452}{\mathrm{Re}} \log \left(\frac{1}{2.8257} \left(\frac{\varepsilon}{D} \right)^{1.1098} + \frac{5.8506}{0.8981\mathrm{Re}} \right) \right)$$

where Re represents the Reynolds number with Re= $\rho v D/\nu$, and ν is the dynamic viscosity. ε denotes the pipe roughness, and we suppose that the considered straight pipe have the same value of friction factor.

9.2.2 Steady states analysis

By setting the temporal derivative terms equal to zero, we have the following:

$$a^2\tilde{q}_x = 0,\ \tilde{p}_x = -\lambda\frac{\tilde{q}|\tilde{q}|a^2}{2D\tilde{p}} - \frac{g\tilde{p}\sin\alpha}{a^2} \tag{9.6a}$$

$$\tilde{q}(0) = q_0,\ \tilde{p}(l) = p_0 \tag{9.6b}$$

where $\tilde{q}$ and $\tilde{p}$ denote steady states of mass flux and pressure. It can be seen that $\tilde{q}$ and $\tilde{p}$ are spatial functions and do not depend on time. For simplicity, we adopt some physical values of incoming gas flow as: $p_0 = 5$ MPa, $q_0 = 278.7455$ kg/(m^{2}s), $u_0 = 7.661$ m/s, $\rho_0 = 36.385$ kg/m^3 from [330]. Based on that, the steady states are numerically calculated and depicted in Fig. 9.2.

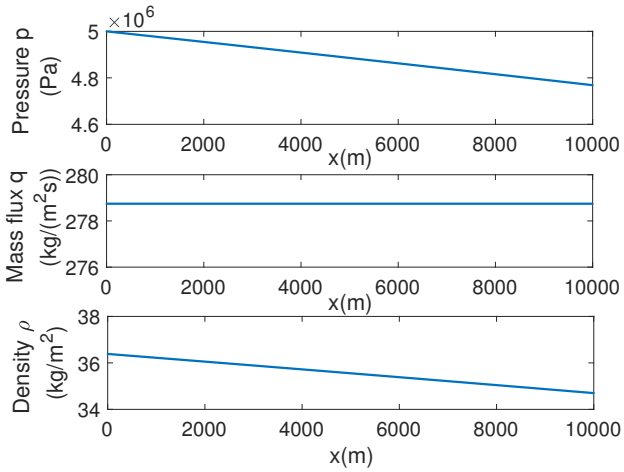


Figure 9.2: Steady states of pressure, mass flux and density.

Along this line, the transient pressure and mass flux are introduced as $\bar{p}(x,t) = p(x,t) - \tilde{p}(x,t)$ and $\bar{q}(x,t) = q(x,t) - \tilde{q}(x,t)$. In the ensuing sections, we assume that the mass flux is always positive, i.e., $q > 0$. Thus, one can derive the transient dynamics with respect to the new states $\bar{p}(x,t)$ and $\bar{q}(x,t)$ by linearizing the system (9.5a)-(9.5c) around the determined steady states as follows:

$$\frac{\partial}{\partial t}\begin{bmatrix}\bar{p}(x,t)\\ \bar{q}(x,t)\end{bmatrix} + A\frac{\partial}{\partial x}\begin{bmatrix}\bar{p}(x,t)\\ \bar{q}(x,t)\end{bmatrix} + B\begin{bmatrix}\bar{p}(x,t)\\ \bar{q}(x,t)\end{bmatrix} = 0 \tag{9.7}$$

where $A = [0, a^2; 1, 0]$, $B = [0, 0; Q, F]$, $Q = 2g\sin\alpha/a^2 + \tilde{p}_x/\tilde{p}$ and $F = \lambda a^2\tilde{q}/(D\tilde{p})$. Combining boundary conditions (9.5c) and (9.6b) leads to the new boundary condi-

tions as:

$$\bar{q}(0,t) = u(t),\ \bar{p}(l,t) = 0 \tag{9.8}$$

After simple algebraic manipulation, a boundary control system (BCS) of the gas flow pipe model is constructed as follows:

$$\frac{\partial}{\partial t} w(x,t) = \mathcal{A}_1 w(x,t),\ w(x,0) = w_0(x) \tag{9.9a}$$

$$\mathcal{B}_1 w(x,t) = u(t), y_c(t) = C_c w(x,t), y_m(t) = C_m w(x,t) \tag{9.9b}$$

where $w(x,t) = [\bar{p}(x,t); \bar{q}(x,t)]$ and $w(\cdot,t) \in \mathcal{X}$, with $\mathcal{X} = L^2((0,l)^2, \mathbb{R})$ defined as a separable Hilbert space. The input is denoted by: $u(t) \in L^2_{loc}([0,\infty), U)$, and the controlled and measured outputs are $y_c, y_m \in L^2_{loc}([0,\infty), Y)$, where U and Y are assumed to be finite-dimensional spaces. Specifically, the continuous-time operators are denoted as follows:

$$\mathcal{A}_1 = -\begin{bmatrix} 0 & a^2 \frac{\partial}{\partial x} \\ \frac{\partial}{\partial x} + Q & F \end{bmatrix}, \mathcal{B}_1 = \left[0,\ \int_0^l \delta(x)(\cdot)\, dx\right] \tag{9.10a}$$

$$C_c = \left[\int_0^l \delta(x - x_c)(\cdot)\, dx, 0\right] \tag{9.10b}$$

$$C_m = \text{bdiag}\left(\int_0^l \delta(x - x_{m1})(\cdot)\, dx, \int_0^l \delta(x - x_{m2})(\cdot)\, dx\right) \tag{9.10c}$$

where δ denotes the Dirac delta function, and "bdiag" represents a block diagonal matrix. We note that the domain of operator $\mathcal{A}_1$ is $\mathcal{D}(\mathcal{A}_1) = \{w(x) = [w_1(x); w_2(x)] \in \mathcal{X} | w_1(l) = 0, w_1(x) \text{ and } w_2(x) \text{ are abs. cont.}\}$. In this chapter, there is no constraint posed on the definition of the controlled and measured outputs, which indicates that the controlled and measured outputs are not necessarily to be co-located. To proceed with the regulator design in the ensuing sections, the following abstract linear state-space model is developed:

$$\frac{\partial}{\partial t} w(x,t) = \mathcal{A} w(x,t) + \mathcal{B} u(t),\ w(x,0) = w_0(x) \tag{9.11a}$$

$$y_c(t) = C_c w(x,t),\ y_m(t) = C_m w(x,t) \tag{9.11b}$$

where $\mathcal{A}$ takes the same expression as $\mathcal{A}_1$, but a different domain as: $\mathcal{D}(\mathcal{A}) = \{[\phi_1(x); \phi_2(x)] \in \mathcal{X} \ | \phi_1(l) = 0, \phi_2(0) = 0, \phi_1(x) \text{ and } \phi_2(x) \text{ are abs. cont.}\}$. By [42, Rem. 10.1.6], $\mathcal{B}$ can be obtained through solving the following inner product formula:

$$\langle \mathcal{A}_1 w, \psi \rangle = \langle w, \mathcal{A}^* \psi \rangle + \langle \mathcal{B}_1 w, \mathcal{B}^* \psi \rangle \tag{9.12}$$

where $\psi \in \mathcal{D}(\mathcal{A}^*)$, $w \in \mathcal{D}(\mathcal{A}_1)$. By solving the inner product formula $\langle \mathcal{A}\phi, \psi \rangle = \langle \phi, \mathcal{A}^*\psi \rangle$, the adjoint operator $\mathcal{A}^*$ can be found as: $\mathcal{A}^* = [0, -Q + \frac{\partial}{\partial x}; a^2 \frac{\partial}{\partial x}, -F]$, with the associated domain defined as $\mathcal{D}(\mathcal{A}^*) = \{[\psi_1(x); \psi_2(x)] \in \mathcal{X} \,|\, \psi_1(l) = 0, \psi_2(0) = 0,$ $\psi_1(x)$ and $\psi_2(x)$ are abs. cont.$\}$. Thus, a direct computation of Eq.(9.12) leads to:

$$\left\langle \mathcal{B}_1 \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}, \mathcal{B}^* \begin{bmatrix} \psi_1 \\ \psi_2 \end{bmatrix} \right\rangle = a^2 u(t) \psi_1(0) \tag{9.13}$$

which implies that $\mathcal{B} = [a^2 \delta(x); 0]$, with δ defined as the Dirac delta function.

In the abstract linear system (9.11), we assume $\mathcal{A} : \mathcal{D}(\mathcal{A}) \subset \mathcal{X} \mapsto \mathcal{X}$ is an infinitesimal generator of a C_0-semigroup $\mathbb{T}(t)$ on $\mathcal{X}$. We introduce the space $\mathcal{X}_1$ as the space $\mathcal{D}(\mathcal{A})$ with the norm $\|x\|_1 = \|(\beta I - \mathcal{A})x\|$, and the space $\mathcal{X}_{-1}$ as the completion of $\mathcal{X}$ with the norm $\|z\|_{-1} = \|(\beta I - \mathcal{A})^{-1} z\|$, where $\forall x \in \mathcal{D}(\mathcal{A})$, $\forall z \in \mathcal{X}$, and $\beta \in \rho(\mathcal{A})$. The constructed spaces are linked by $\mathcal{X}_1 \subset \mathcal{X} \subset \mathcal{X}_{-1}$, with each inclusion being dense and continuous embedding. The restriction of $\mathcal{A}$ to $\mathcal{X}$ that is the generator of $\mathbb{T}(t)$ on $\mathcal{X}_1$ is also denoted by $\mathcal{A}$, leading to $\mathcal{A} \in \mathcal{L}(\mathcal{X}_1, \mathcal{X})$ [42]. The operators $\mathcal{B} \in \mathcal{L}(U, \mathcal{X}_{-1})$, and $C_c, C_m \in \mathcal{L}(\mathcal{X}_1, Y)$ are assumed to be admissible to account for boundary control and observation. For $w_0 \in \mathcal{X}$, $u(t) \in L^2_{loc}([0, \infty), U)$, the functions $w : [0, \infty) \to \mathcal{X}$ and $y_c, y_m \in L^2_{loc}([0, \infty), Y)$ have the weak solutions as:

$$w(t) = \mathbb{T}(t) w_0 + \int_0^t \mathbb{T}(t - \tau) B u(\tau) d\tau \tag{9.14a}$$

$$y_c(t) = C_{c\Lambda} \Big[\mathbb{T}(t) w_0 + \int_0^t \mathbb{T}(t - \tau) B u(\tau) d\tau \Big] \tag{9.14b}$$

where $C_{c\Lambda}$ denotes the Λ-extension of C_c, i.e., $C_{c\Lambda} w = \lim_{\lambda \to +\infty} C_c \lambda (\lambda I - \mathcal{A})^{-1} w$, and the domain of $C_{c\Lambda}$ contains all $w \in \mathcal{X}$ for which the above limit exists. In addition, the solution of y_m has the same expression as y_c with $C_{c\Lambda}$ being replace by $C_{m\Lambda}$ that is the Λ-extension of C_m. To let a well-posed system have the usual state-space presentation (9.11), C_c and C_m have to be replaced by $C_{c\Lambda}$ and $C_{m\Lambda}$, respectively.

As in [121], we define the transfer function from the input $u(t)$ to the controlled output $y_c(t)$ in the following form:

$$\mathcal{G}_c(s) = C_{c\Lambda}(sI - \mathcal{A})^{-1} \mathcal{B}, \text{ for } s \in \mathbb{C}^+_\sigma \cap \rho(\mathcal{A}) \tag{9.15}$$

where σ is the maximum of growth index of a well-posed linear system and the growth bound of the semigroup $\mathbb{T}(t)$, and $\mathbb{C}^+_\sigma$ denotes the right open half-plane in $\mathbb{C}$

with $\mathbb{C}_\sigma^+ = \{s \in \mathbb{C} | \operatorname{Re} s > \sigma\}$ delimited by $\sigma \in \mathbb{R}$. Similarly, $\mathcal{G}_m(s)$ can be defined by replacing $\mathcal{C}_{c\Lambda}$ with $\mathcal{C}_{m\Lambda}$ in (9.15). Moreover, it can be shown that the model (9.11) is well-posed following the structure of the 2×2 port-Hamiltonian system [170, Lem. 4.1], and it is indeed regular and will be shown later on. For ease of notation, we keep using $\mathcal{C}_c$ and $\mathcal{C}_m$ to denote their Λ-extensions $\mathcal{C}_{c\Lambda}$ and $\mathcal{C}_{m\Lambda}$ in the ensuing sections.

9.2.3 Model time-discretization

To address the boundary (or point) control and observation, we deploy the Cayley-Tustin bilinear transform to convert the continuous-time model with unbounded operators into a discrete-time infinite-dimensional model with all bounded operators. Meanwhile, the essential properties of the continuous-time system stay invariant under this transformation, including conservativity [19, 20], reachability and observability [19], stability [21, 22], stabilizability, controllability and observability [4, 23], and optimality [24]. Furthermore, the 1-1 correspondences between the continuous model with its discrete counterpart can be utilized in digital control (or sampled-data control). To start with, we apply the Cayley-Tustin transform to the linear continuous model (9.11) with a given time discretization interval Δt as follows [97]:

$$\frac{w(k\Delta t) - w((k-1)\Delta t)}{\Delta t} \approx \mathcal{A}\frac{w(k\Delta t) + w((k-1)\Delta t)}{2} + \mathcal{B}u(k\Delta t),\ k \geq 1 \tag{9.16a}$$

$$y_c(k\Delta t) \approx \mathcal{C}_c\frac{w(k\Delta t) + w((k-1)\Delta t)}{2} \tag{9.16b}$$

$$y_m(k\Delta t) \approx \mathcal{C}_m\frac{w(k\Delta t) + w((k-1)\Delta t)}{2} \tag{9.16c}$$

As shown in Eq.(9.16), this discretization framework follows an implicit midpoint integration rule. Furthermore, this framework is a symmetric and symplectic integration scheme leading to a structure- and energy-preserving time discretization [119]. The discrete input is given by the mean value sampling as $u_k/\sqrt{\Delta t} = 1/\Delta t \int_{(k-1)\Delta t}^{k\Delta t} u(t)dt$. It can be shown that $\|u_k/\sqrt{\Delta t} - u(t)\| \to 0$ on the interval $[(k-1)\Delta t, k\Delta t]$ as $\Delta t \to 0^+$, and similar expressions hold for y_{ck} and y_{mk} [97]. Through simple algebraic manipulations, one can obtain the following discrete-time infinite-dimensional state-space model:

$$w_k = \mathcal{A}_d w_{k-1} + \mathcal{B}_d u_k,\ k \geq 1 \tag{9.17a}$$

$$y_{ck} = \mathcal{C}_{cd} w_{k-1} + \mathcal{D}_{cd} u_k \tag{9.17b}$$

$$y_{mk} = \mathcal{C}_{md} w_{k-1} + \mathcal{D}_{md} u_k \tag{9.17c}$$

where one denotes w_k, u_k, y_{ck} and y_{mk} as the discrete-time state, input, controlled and measured outputs, with the associated discrete-time operators given as follows:

$$\begin{bmatrix} \mathcal{A}_d & \mathcal{B}_d \\ \mathcal{C}_{cd} & \mathcal{D}_{cd} \\ \mathcal{C}_{md} & \mathcal{D}_{md} \end{bmatrix} = \begin{bmatrix} -I + 2\delta \mathcal{R}(\delta, \mathcal{A}) & \sqrt{2\delta}\mathcal{R}(\delta, \mathcal{A})\mathcal{B} \\ \sqrt{2\delta}\mathcal{C}_c \mathcal{R}(\delta, \mathcal{A}) & \mathcal{G}_c(\delta) \\ \sqrt{2\delta}\mathcal{C}_m \mathcal{R}(\delta, \mathcal{A}) & \mathcal{G}_m(\delta) \end{bmatrix} \tag{9.18}$$

where $\delta = 2/\Delta t \in \mathbb{R}^+$ with $\mathbb{R}^+$ denoting the set of all positive real numbers, and $\delta \in \mathbb{C}_\sigma^+ \cap \rho(\mathcal{A})$. $\mathcal{R}(\delta, \mathcal{A})$ represents the resolvent operator $\mathcal{R}(s, \mathcal{A}) = (sI - \mathcal{A})^{-1}$ with s evaluated at δ. $\mathcal{G}_c(\delta)$ and $\mathcal{G}_m(\delta)$ denote the transfer functions $\mathcal{G}_c(s)$ and $\mathcal{G}_m(s)$ with s evaluated at δ, respectively.

Remark 35. *With Cayley-Tustin discretization being applied to system (9.11), the continuous-time unbounded operators are all converted to discrete-time bounded operators in Eq.(9.17), where the operators are defined as:* $\mathcal{A}_d \in \mathcal{L}(\mathcal{X})$, $\mathcal{B}_d \in \mathcal{L}(U, \mathcal{X})$, $\mathcal{C}_{cd}, \mathcal{C}_{md} \in \mathcal{L}(\mathcal{X}, Y)$, $\mathcal{D}_{cd}, \mathcal{D}_{md} \in \mathcal{L}(U, Y)$, *see details in the reference [279, Sec. IV.B].*

Remark 35 provides a technical advantage for addressing the continuous model (9.11) with unbounded operators. Furthermore, it can be recognized that there are feedthrough operators $\mathcal{D}_{cd}$ and $\mathcal{D}_{md}$ existing in the discrete-time setting (9.17) after applying Cayley-Tustin transform, which are not necessarily present in the continuous model (9.11).

As in [24], we define the transfer function (also called characteristic function) from u_k to y_{ck} for the discrete-time system (9.17) as follows:

$$\mathcal{G}_d(z) = \mathcal{C}_{cd}(zI - \mathcal{A}_d)^{-1}\mathcal{B}_d + \mathcal{D}_{cd} \tag{9.19a}$$

for $z \in \rho(\mathcal{A}_d)$. Then, we establish a 1-1 equivalent relationship between continuous- and discrete-time transfer functions via the Cayley-Tustin transform.

Lemma 7. *With the well-posed continuous system (9.11), and its discrete counterpart (9.17) induced by Cayley-Tustin transform, the following relationship holds*

$$\mathcal{G}_c(s) = \mathcal{G}_d(z) \tag{9.20}$$

where $z = \frac{\delta + s}{\delta - s} \in \rho(\mathcal{A}_d) \backslash \{-1\}$ *and* $s = \frac{z-1}{z+1}\delta \in \mathbb{C}_\sigma^+ \cap \rho(\mathcal{A}) \backslash \{\delta\}$.

Proof. The proof is similar to [24, Lem. 8] by replacing z with $1/z$ under the consideration of $\delta \in \mathbb{R}^+$. ■

The same 1-1 relationship between continuous- and discrete-time transfer functions with respect to bounded control and observation operators has been shown in [23] (for $\delta = 1$). Apparently, for $\delta > r \geq 0$, Cayley-Tustin transform maps the right half-plane $\mathbb{C}_r^+$ bijectively the exterior disc $\mathbb{D}_r^+$ with center $r/(\delta - r)$ and radius $\delta/(\delta - r)$. For more results on Cayley transform, see references [19, 23, 24, 21].

9.2.4 Resolvent operator

In order to realize the discrete-time model (9.17), the resolvent operator needs to be determined from the continuous-time counterpart (9.11). Considering that the resolvent operator only depends on operator $\mathcal{A}$, one can apply Laplace transformation to Eq.(9.7) or (9.11a) by neglecting $\mathcal{B}$ as follows:

$$w(x,s) = e^{Mx}w(0,s) + \int_0^x e^{M(x-\eta)}A_0{}^{-1}w(\eta,0)d\eta \tag{9.21}$$

where $A_0 = [0, a^2; 1, 0]$ and $M = -[Q, F+s; s/a^2, 0]$. For simplicity, one can denote $e^{Mx} = [M_{ij}(x,s)]_{2\times 2}$ with $i, j = 1, 2$. By checking the nonzero elements off the diagonal of matrix A_0, it is straightforward to show that A_0 is invertible, which guarantees the solvability of Eq.(9.21). More specifically, the analytical expression of e^{Mx} is given by:

$$\begin{aligned} e^{Mx} &= \begin{bmatrix} M_{11}(x,s) & M_{12}(x,s) \\ M_{21}(x,s) & M_{22}(x,s) \end{bmatrix} \\ &= e^{-\frac{Qx}{2}} \times \begin{bmatrix} \cosh(\frac{\mathrm{H}}{2a}x) - \frac{Qa}{\mathrm{H}}\sinh(\frac{\mathrm{H}}{2a}x) & -\frac{2(F+s)a}{\mathrm{H}}\sinh(\frac{\mathrm{H}}{2a}x) \\ -\frac{2s}{\mathrm{H}a}\sinh(\frac{\mathrm{H}}{2a}x) & \cosh(\frac{\mathrm{H}}{2a}x) + \frac{Qa}{\mathrm{H}}\sinh(\frac{\mathrm{H}}{2a}x) \end{bmatrix} \end{aligned} \tag{9.22}$$

where $H = \sqrt{Q^2a^2 + 4s^2 + 4Fs}$. Substituting boundary conditions $\bar{p}(l,t) = 0$ and $\bar{q}(0,t) = 0$ into Eq.(9.21), one can solve for $\bar{p}(0,s)$ and then determine the resolvent operator as follows:

$$w(x,s) = \mathcal{R}(s,\mathcal{A})w(x,0) \tag{9.23}$$

where $\mathcal{R}(s,\mathcal{A}) = [\mathcal{R}_{ij}(s,\mathcal{A})]_{2\times 2}$, with $i, j = 1, 2$. The explicit resolvent expressions are provided are provided in the Appendix. From the limit of the transfer function as $s \to +\infty$, we can show that the system (9.11) is regular as:

$$\lim_{s\to+\infty} \mathcal{G}_c(s) = \lim_{s\to+\infty} C_{c\Lambda}(sI - \mathcal{A})^{-1}\mathcal{B}$$

$$= \lim_{s\to+\infty} \frac{4(F+s)ae^{-\frac{Qx_c}{2}}\left[\exp(\frac{\mathrm{H}}{2a}(l-x_c)) - \exp(\frac{\mathrm{H}}{2a}(x_c-l))\right]}{(H-Qa)\exp(\frac{\mathrm{H}}{2a}l) + (H+Qa)\exp(-\frac{\mathrm{H}}{2a}l)}$$
$$= 0 \ \ (\text{with}\, H = \sqrt{Q^2a^2 + 4s^2 + 4Fs}) \tag{9.24}$$

Thus, the networked pipeline system is also regular due to the fact that the cascade connection of two regular systems is regular [121]. Substituting the resolvent operator $\mathcal{R}(s,\mathcal{A})$ back to Eq.(9.18) with s evaluated at $2/\Delta t$, one can obtain the discrete-time model (9.17) of the single gas pipeline.

9.3 Gas pipe network model

In this section, we consider a general star-shaped gas pipeline network as shown in Fig. 9.1(a). Based on the continuity and momentum balance laws and the equation of state of ideal gas, a 1-D transient gas flow pipe network is established as follows [340]:

$$\rho_t^{(i)} + q_x^{(i)} = 0 \tag{9.25a}$$
$$q_t^{(i)} + \Big(\frac{q^{(i)2}}{\rho^{(i)}} + a^2\rho^{(i)}\Big)_x = -\lambda^{(i)}\frac{q^{(i)}|q^{(i)}|}{2D^{(i)}\rho^{(i)}} - g\rho^{(i)}\sin(\alpha^{(i)}) \tag{9.25b}$$
$$p^{(i)}/\rho^{(i)} = zRT/(M_g) = a^2 \tag{9.25c}$$

with boundary conditions given below:

$$q^{(i)}(0,t) = q_0^{(i)} + u^{(i)}(t),\ p^{(i)}(l,t) = p_0^{(i)}, \rho^{(i)}(l,t) = \rho_0^{(i)} \tag{9.26}$$

By performing model linearization (around the same steady states for brevity) and model time discretization as in the previous sections, one can obtain the discrete-time gas pipe network model as:

$$\frac{\partial}{\partial t}w^{(i)}(x,t) = \mathcal{A}^{(i)}w^{(i)}(x,t) + \mathcal{B}^{(i)}u^{(i)}(t) \tag{9.27a}$$
$$y_c^{(i)}(t) = C_c^{(i)}w^{(i)}(x,t),\ y_m^{(i)}(t) = C_m^{(i)}w^{(i)}(x,t) \tag{9.27b}$$

where $w^{(i)}(x,0) = w_0^{(i)}(x)$, $w^{(i)}(x,t) = [\bar{p}^{(i)}(x,t); \bar{q}^{(i)}(x,t)]$, and the corresponding continuous-time operators in Eq.(9.10) are: $\mathcal{A}^{(i)} := -[0, a^2\frac{\partial}{\partial x}; \frac{\partial}{\partial x} + Q^{(i)}, F^{(i)}]$, $\mathcal{B}^{(i)} = \mathcal{B}$, $C_c^{(i)} = C_c$, $C_m^{(i)} = C_m$, where $\mathcal{D}(\mathcal{A}^{(i)}) = \{[\phi_1^i(x); \phi_2^i(x)] \in \mathcal{X} |\, \phi_1^i(l) = 0,\ \phi_2^i(0) = 0,$

$\phi_1^i(x)$ and $\phi_2^i(x)$ are abs. cont.}. In addition, the superscript i corresponds to the quantities of the i-th pipe segment.

The coupling conditions at pipe junctions exhibit a significant difference between the single gas pipeline model and the gas pipeline network. To be precise, we model the coupling conditions by neglecting the pressure and mass loss at the vertex joint k as follows:

$$\bar{p}_k^{(i)}(l,t) = \bar{p}_k^{(j)}(0,t),\ (i \in I^k,\ j \in J^k,\ k \in K) \tag{9.28a}$$

$$\sum_{i \in I^k} \bar{q}_k^{(i)}(l,t) + \sum_{i \in J^k} \bar{q}_k^{(j)}(0,t) = 0 \tag{9.28b}$$

where I^k denotes all pipes with incoming flow toward the junction node k, and J^k denotes all pipes with outgoing flow from the junction node k. In addition, we denote the set of all junction nodes as K. The coupling conditions (9.28) have been widely utilized for gas pipe network modelling and analysis, see references [334, 336, 332].

9.3.1 SIMO gas pipe network

The gas pipe network model (9.27) represents a general architecture as boundary control and observation are considered for each pipe segment, while it is complex and may not be the case when it comes to engineering practice. In order to address this issue, we focus on gas pipe network systems with boundary actuations applied at upstream pipe ends. Firstly, we consider a simple star-shaped pipe net with a single inlet and multiple outlets (SIMO) as shown in Fig. 9.1(b). For this architecture, boundary input is applied only at upstream of pipe 1, i.e., $\bar{q}^{(1)}(0,t) = u(t)$. Based on the coupling conditions (9.28), we have the relationship: $\bar{p}^{(1)}(l,t) = \bar{p}^{(3)}(0,t)$. Furthermore, we assume the incoming mass flux to pipe 3 is proportional to the outgoing mass flux from pipe 1, i.e., $q^{(3)}(0,t) = \gamma q^{(1)}(l,t)$, where the flow directionality is neglected, which allows two PDE systems of pipe 1 and pipe 3 to be connected as a cascade PDE system with a goal of controlling the pressure at some position in pipe 3 (marked as an orange star in Fig. 9.1(b)). Moreover, we measure the upstream pressure and downstream mass flux of all considered pipes. In this case, we need to measure pipe 1 and pipe 3 as: $q^{(1)}(l,t)$, $p^{(1)}(0,t)$, and $q^{(3)}(l,t)$ (since $p^{(3)}(0,t)$ is given by $p^{(1)}(l,t)$). Mathematically, the dynamics of the cascade system (pipe 1 and

pipe 3) are described in the continuous-time setting as below:

$$\frac{\partial}{\partial t}w^e(x,t) = \mathcal{A}^e w^e(x,t) + \mathcal{B}^e u^e(t) \tag{9.29a}$$

$$y_c^e(t) = C_c^e w^e(x,t),\ y_m^e(t) = C_m^e w^e(x,t) \tag{9.29b}$$

where $w^e(x,0) = w_0^e(x)$, $w^e(x,t) = [w^{(1)}(x,t); w^{(3)}(x,t)]$ and $\mathcal{D}(\mathcal{A}^e) = \{[\phi_1(x); \phi_2(x); \phi_3(x); \phi_4(x)] \in L^2(0,l)^4 |\ \phi_2(0) = 0,\ \phi_1(l) = \phi_3(0),\ \phi_4(0) = \gamma\phi_2(l),\ \phi_3(l) = 0,\ \phi_1(x),\ \phi_2(x),\ \phi_3(x)$ and $\phi_4(x)$ are abs. cont.$\}$. The corresponding extended operators are given as follows:

$$\mathcal{A}^e := \text{bdiag}(\mathcal{A}^{(1)}, \mathcal{A}^{(3)}),\ \mathcal{B}^e := [\mathcal{B}^{(1)};\ 0] \tag{9.30a}$$

$$C_c^e := \begin{bmatrix} 0 & C_c^{(3)} \end{bmatrix},\ C_m^e := \text{bdiag}(C_m^{(1)}, C_m^{(3)}) \tag{9.30b}$$

By performing model discretization as in Eq.(9.16), a discrete pipe network model is established as:

$$w_k^e = \mathcal{A}_d^e w_{k-1}^e + \mathcal{B}_d^e u_k^e,\ k \geq 1 \tag{9.31a}$$

$$y_{ck}^e = C_{cd}^e w_{k-1}^e + \mathcal{D}_{cd}^e u_k^e \tag{9.31b}$$

$$y_{mk}^e = C_{md}^e w_{k-1}^e + \mathcal{D}_{md}^e u_k^e \tag{9.31c}$$

where the associated discrete operators take same expressions as in Eq.(9.18) with all operators being replaced by their extended counterparts. Consequently, the resolvent operator corresponding to the pipe network model is also extended as:

$$w^e(x,s) = \mathfrak{R}(s, \mathcal{A}^e) w^e(x,0) \tag{9.32}$$

where $\mathfrak{R}(s, \mathcal{A}^e) = [\mathfrak{R}_{ij}(s, \mathcal{A}^e)]_{4\times 4}$, with $i, j = 1, 2, 3, 4$. By applying boundary conditions associated with $\mathcal{A}^e$, one can attain the analytic expressions of the extended resolvent operator as shown in the Appendix. Hence, we convert the single-inlet-multiple-outlet pipe network into a two-pipe cascade system (9.31).

9.3.2 MIMO gas pipe network

In this section, we consider a gas pipe framework with two inlets and multiple outlets (MIMO) as shown in Fig. 9.1(c). Similarly, we can formulate the associated continuous-time network model as Eq.(9.29), where $w^E(x,t) = [w^{(1)}(x,t); w^{(2)}(x,t);$

$w^{(3)}(x,t)]$, $u^E(t) = [u^{(1)}(t), u^{(2)}(t)]$ with $\mathcal{D}(\mathcal{A}^E) = \{[\phi_1(x); \phi_2(x); \phi_3(x); \phi_4(x); \phi_5(x); \phi_6(x)] \in L^2(0,l)^6 | \ \phi_2(0) = 0,\ \phi_5(l) = 0,\ \phi_1(l) = \phi_3(l) = \phi_5(0), \phi_4(0) = \gamma[\phi_2(l)+\phi_4(l)]$, $\phi_i(x)$ (with i=1,...,6) is abs. cont.$\}$. The corresponding extended operators are given as follows:

$$\mathcal{A}^E := \text{bdiag}(\mathcal{A}^{(1)}, \mathcal{A}^{(2)}, \mathcal{A}^{(3)}) \tag{9.33a}$$

$$\mathcal{B}^E := [\mathcal{B}^{(1)} \ \ 0; 0 \ \ \mathcal{B}^{(2)}; 0 \ \ 0] \tag{9.33b}$$

$$C_c^E := \begin{bmatrix} 0 & 0 & C_c^{(3)} \end{bmatrix}, \ C_m^E := \text{bdiag}(C_m^{(1)}, C_m^{(2)}, C_m^{(3)}) \tag{9.33c}$$

By applying the Cayley-Tustin transform, the corresponding discrete-time model of the multi-inlet-multi-outlet pipe network can be derived as Eq.(9.31), which can be completed by determining the resolvent operator corresponding to the extended state evolution operator $\mathcal{A}^E$. Considering the space limitation, the process of solving the resolvent operator is omitted here.

Remark 36. *Generally, a star-shaped gas pipe network with multiple inlets and multiple outlets, as shown in Fig. 9.1(a), can be modelled as a cascade PDE system with multiple inputs and multiple outputs as shown in Eq.(9.29) and Eq.(9.31).*

Remark 37. *To account for input constraints (e.g. same or different amplitude constraints for different inputs), the semiglobal approach developed in [349] for output regulation of discrete-time finite-dimensional systems subject to input saturation can be further extended.*

9.4 Discrete output regulator design

In this section, discrete-time output regulator designs for infinite-dimensional discrete-time gas pipe network systems are presented. Based on the discrete exogenous and plant models, discrete-time regulator equations are constructed for a state feedback regulator design. Considering that in reality the full state information of the pipeline system and exogenous system (i.e., exo-system) may not be available, an observer-based output feedback regulator is proposed.

9.4.1 Exo-system description

In order to generate reference signals, a finite-dimensional discrete-time exo-system is considered as follows:

$$q_k = S_d q_{k-1},\ q_0 = q^0 \in \mathbb{C}^n \tag{9.34a}$$

$$y_{rk} = Q_d q_k,\ k \geq 1 \tag{9.34b}$$

where q_k and y_{rk} are the exogenous state and output reference in the discrete-time language. In particular, S_d is the discrete-time state evolution matrix of $n \times n$ dimension, and we assume that S_d has distinct eigenvalues placed on the boundary of the unit disc, i.e., $\lambda_i^d = \nu_i + \iota_i j$ where $j^2 = -1$ and $\nu_i^2 + \iota_i^2 = 1$ with $\nu_i, \iota_i \in \mathbb{R}$, and $i = 1, 2, ..., n$. Hence, S_d is capable of generating step-like and harmonic signals, and S_d is invertible. In order to estimate the exogenous state from the reference y_{rk}, we further assume that (S_d, Q_d) is observable. In addition, Q_d has proper dimension for generating reference signals.

For clarification, we introduce the following stability concepts.

Definition 12. *The C_0-semigroup $\mathbb{T}(t)$ on $\mathcal{X}$ is exponentially stable if there exist positive constants M and α such that:*

$$\|\mathbb{T}(t)\| \leq Me^{-\alpha t},\ \forall t \in \mathbb{R}^+$$

and it is strongly stable if $\|\mathbb{T}(t)x\| \to 0$ as $t \to +\infty$ for all $x \in \mathcal{X}$. $\mathcal{A}_d$ is power stable if there exist positive constants $M \geq 1$ and $\gamma < 1$ such that:

$$\|\mathcal{A}_d^k\| \leq M\gamma^k,\ \forall k \in \mathbb{N}$$

and $\mathcal{A}_d$ is strongly stable if $\mathcal{A}_d^k x \to 0$ as $k \to +\infty$ for all $x \in \mathcal{X}$ [4, 23].

9.4.2 Discrete state feedback regulator design

In this section, we aim to design a discrete state feedback regulator by assuming all state information is available. The discrete-time state feedback regulator design problem is addressed by finding a discrete regulator in the following form:

$$u_k = K_d w_{k-1} + L_d q_k \tag{9.35}$$

where $K_d \in \mathcal{L}(\mathcal{X}, U)$, $L_d \in \mathcal{L}(\mathbb{C}^n, U)$, and all state information of the plant and the exo-system is assumed to be known, such that: (1) $\mathcal{A}_d + \mathcal{B}_d K_d$ generates a strongly stable operator; (2) the discrete-time tracking error $e_k = y_{ck} - y_{rk} \to 0$ as $k \to +\infty$ for any given $w_0 \in \mathcal{X}$ and $q^0 \in \mathbb{C}^n$.

In order to determine the feedback gain K_d to guarantee the stability or to adjust the convergence rate of the closed-loop system, we introduce the following lemma:

Lemma 8. *An infinite-dimensional discrete-time system $w_k = \mathcal{A}_{cd} w_{k-1}$ is power stable if and only if there exists a positive self-adjoint operator Q_{cd} such that*

$$\mathcal{A}_{cd}^* Q_{cd} \mathcal{A}_{cd} - Q_{cd} + M_{cd} = 0 \tag{9.36}$$

on $\mathcal{X}$, where M_{cd} is a positive definite design parameter.

Proof. The proof is similar to [4, Exe. 4.29c] and [350] if we let the observability gramian L_C be Q_{cd}, $A = \mathcal{A}_{cd}$ and $C^*C = M_{cd}$, respectively. ■

By substituting $\mathcal{A}_{cd} = \mathcal{A}_d + \mathcal{B}_d K_d$ and $M_{cd} = (\mathcal{C}_{cd} + \mathcal{D}_{cd} K_d)^* N_{cd} (\mathcal{C}_{cd} + \mathcal{D}_{cd} K_d)$ into Eq.(9.36), one can construct the following discrete-time Riccati equation:

$$\mathcal{A}_d^* Q_{cd} \mathcal{A}_d - Q_{cd} + \mathcal{C}_{cd}^* N_{cd} \mathcal{C}_{cd} - K_d^* \left(2\mu I + \mathcal{B}_d^* Q_{cd} \mathcal{B}_d + \mathcal{D}_{cd}^* N_{cd} \mathcal{D}_{cd}\right) K_d = 0 \tag{9.37}$$

on $\mathcal{X}$. μ and N_{cd} are positive definite design parameters such that $Q_{cd} \in \mathcal{L}(\mathcal{X})$ and the feedback gain is $K_d = -(\mu I + \mathcal{B}_d^* Q_{cd} \mathcal{B}_d + \mathcal{D}_{cd}^* N_{cd} \mathcal{D}_{cd})^{-1} (\mathcal{B}_d^* Q_{cd} \mathcal{A}_d + \mathcal{D}_{cd}^* N_{cd} \mathcal{C}_{cd})$. Similar discrete-time Riccati equations can be found in [4, Exe. 6.35] and [24, Lem. 6] where the connection between the discrete-time Riccati equation and its continuous-time counterpart is established by the use of Cayley transform for the optimal control problem. Clearly, we note that the solution of Q_{cd} is not unique, since we can have different choices of the design parameters μ and N_{cd}. Since the power stability induces the strong stability, one can obtain a strongly stabilizing gain K_d by solving the discrete-time Riccati equation (9.37).

To determine the feedforward gain L_d in Eq.(9.35), a common assumption is made that the spectrum of S_d is included in the resolvent set of $\mathcal{A}_d$ (i.e., $\sigma(S_d) \subset \rho(\mathcal{A}_d)$), with some proper time discretization interval in the Cayley-Tustin transform. The pair $(\mathcal{A}_d, \mathcal{B}_d)$ is assumed to be strongly stabilizable. Thus, we provide the following discrete-time regulator equations motivated by the similar design of continuous-time output regulator equations [10, The. IV.1]:

Theorem 25. *Under the assumptions that $\sigma(S_d) \subset \rho(\mathcal{A}_d)$ and $(\mathcal{A}_d, \mathcal{B}_d)$ is strongly stabilizable, the discrete state feedback regulation problem is solvable if and only if there exist mappings $\Pi_d \in \mathcal{L}(\mathbb{C}^n, \mathcal{X})$ and $\Gamma_d \in \mathcal{L}(\mathbb{C}^n, U)$ such that the following discrete regulator equations hold:*

$$\Pi_d S_d = \mathcal{A}_d \Pi_d + \mathcal{B}_d \Gamma_d S_d \tag{9.38a}$$

$$Q_d S_d = \mathcal{C}_{cd} \Pi_d + \mathcal{D}_{cd} \Gamma_d S_d \tag{9.38b}$$

where $L_d = \Gamma_d - K_d \Pi_d S_d^{-1}$ can be utilized to compute the state feedback control law u_k in Eq.(9.35).

Proof. First, we prove the sufficiency. For brevity, we denote $\mathcal{A}_{cd} = \mathcal{A}_d + \mathcal{B}_d K_d$. By substituting Eq.(9.35) into the discrete system (9.17), the closed-loop model is obtained as follows:

$$w_k = \mathcal{A}_{cd} w_{k-1} + \mathcal{B}_d L_d q_k \tag{9.39}$$

By induction, the solution of w_k takes the following form:

$$w_k = (\mathcal{A}_{cd})^k w_0 + \sum_{m=1}^{k} (\mathcal{A}_{cd})^{m-1} \mathcal{B}_d L_d q_{k+1-m} \tag{9.40}$$

Plugging exo-system model (9.34) and Eq.(9.38) into Eq.(9.40) leads to:

$$\begin{aligned}
w_k &= (\mathcal{A}_{cd})^k w_0 + \sum_{m=1}^{k} (\mathcal{A}_{cd})^{m-1} \mathcal{B}_d (\Gamma_d - K_d \Pi_d S_d^{-1}) q_{k+1-m} \\
&= (\mathcal{A}_{cd})^k w_0 + \sum_{m=1}^{k} (\mathcal{A}_{cd})^{m-1} [\Pi_d S_d - \mathcal{A}_{cd} \Pi_d] q_{k-m} \\
&= (\mathcal{A}_{cd})^k w_0 + \sum_{m=1}^{k} (\mathcal{A}_{cd})^{m-1} \Pi_d q_{k+1-m} - \sum_{m=2}^{k+1} (\mathcal{A}_{cd})^{m-1} \Pi_d q_{k+1-m} \\
&\Rightarrow w_k = (\mathcal{A}_{cd})^k (w_0 - \Pi_d q_0) + \Pi_d q_k
\end{aligned} \tag{9.41}$$

Then, the discrete tracking error can be expressed as:

$$\begin{aligned}
e_k &= y_{ck} - y_{rk} \\
&= \mathcal{C}_{cd} w_{k-1} + \mathcal{D}_{cd} u_k - Q_d q_k \\
&= (\mathcal{C}_{cd} + \mathcal{D}_{cd} K_d) w_{k-1} + (\mathcal{D}_{cd} L_d - Q_d) q_k
\end{aligned}$$

$$= (C_{cd} + \mathcal{D}_{cd}K_d)(\mathcal{A}_{cd})^{k-1}(w_0 - \Pi_d q_0)$$
$$+ [(C_{cd} + \mathcal{D}_{cd}K_d)\Pi_d + (\mathcal{D}_{cd}L_d - Q_d)S_d]q_{k-1} \tag{9.42}$$

Under the assumption that $\mathcal{A}_{cd}$ is a strongly stable operator, we have that $(\mathcal{A}_{cd})^{k-1}(w_0 - \Pi_d q_0) \to 0$ as $k \to +\infty$ for all on $(w_0 - \Pi_d q_0) \in \mathcal{X}$. Therefore, $w_k - \Pi_d q_k$ converges to zero in Eq.(9.41) and the discrete tracking error e_k goes to zero as $k \to +\infty$ in Eq.(9.42), which is guaranteed by the discrete regulator equations (9.38a)-(9.38b).

Now, we show the proof of the necessity by constructing the following extended closed-loop system:

$$\begin{bmatrix} w_k \\ q_k \end{bmatrix} = \begin{bmatrix} \mathcal{A}_{cd} & \mathcal{B}_d L_d S_d \\ 0 & S_d \end{bmatrix} \begin{bmatrix} w_{k-1} \\ q_{k-1} \end{bmatrix} \tag{9.43}$$

It is straightforward to find the solution of Eq.(9.43) by induction as follows:

$$\begin{bmatrix} w_k \\ q_k \end{bmatrix} = \begin{bmatrix} (\mathcal{A}_{cd})^k w_0 + \sum_{m=1}^{k} (\mathcal{A}_{cd})^{m-1}\mathcal{B}_d L_d q_{k+1-m} \\ S_d^k q_0 \end{bmatrix} \tag{9.44}$$

By assuming that $\mathcal{A}_{cd}$ is strongly stable, $(\mathcal{A}_{cd})^k w_0 \to 0$ as $k \to +\infty$ and Eq.(9.44) indicates that $[w_k;\ q_k] \to [\Pi_d q_k;\ q_k]$ with $k \to +\infty$ and $\Pi_d \in \mathcal{L}(\mathbb{C}^n, \mathcal{X})$. To determine Π_d, we construct the dynamics of $z_k = [w_k;\ q_k] - [\Pi_d q_k;\ q_k]$ by the following homogeneous difference equation:

$$z_k = \begin{bmatrix} \mathcal{A}_{cd} & \mathcal{B}_d L_d S_d \\ 0 & S_d \end{bmatrix} z_{k-1} \tag{9.45}$$

where the initial condition is defined as $[w_0; q_0] - [\Pi_d q_0; q_0] \in \mathcal{H}_e$ with $\mathcal{H}_e = \mathcal{X} \bigoplus \mathbb{C}^n$. The first component in Eq.(9.45) leads to the discrete-time regulator equation (9.38a). Furthermore, the discrete tracking error is described as:

$$\begin{aligned} e_k &= y_{ck} - y_{rk} \\ &= C_{cd} w_{k-1} + \mathcal{D}_{cd} u_k - Q_d q_k \\ &= \begin{bmatrix} C_{cd} + \mathcal{D}_{cd}K_d & (\mathcal{D}_{cd}L_d - Q_d)S_d \end{bmatrix} \begin{bmatrix} w_{k-1} \\ q_{k-1} \end{bmatrix} \\ &\to [(C_{cd} + \mathcal{D}_{cd}K_d)\Pi_d + (\mathcal{D}_{cd}L_d - Q_d)S_d]\, q_{k-1} \\ &\qquad (\text{as } k \to +\infty) \end{aligned} \tag{9.46}$$

To realize a perfect tracking, it is necessary to ensure that $(C_{cd}+\mathcal{D}_{cd}K_d)\Pi_d+(\mathcal{D}_{cd}L_d - Q_d)S_d = 0$, which indicates the regulator equation (9.38b) with $L_d = \Gamma_d - K_d\Pi_d S_d^{-1}$.

■

Through simple manipulations of discrete regulator equations (9.38a)-(9.38b) with the eigenpair (λ_i^d, ϕ_i^d) of S_d, the discrete regulator gains (Π_d, Γ_d) can be solved by:

$$\Pi_d \phi_i^d = \lambda_i^d (\lambda_i^d I - \mathcal{A}_d)^{-1} \mathcal{B}_d \Gamma_d \phi_i^d \tag{9.47a}$$

$$\Gamma_d \phi_i^d = [\mathcal{G}_d(\lambda_i^d)]^{-1} Q_d \phi_i^d \tag{9.47b}$$

where $\mathcal{G}_d(\lambda_i^d)$ is the discrete transfer functions $\mathcal{G}_d(z)$ with z evaluated at $\lambda_i^d \in \sigma(\mathcal{S}_d)$. To ensure the solvability of the regulator equations (9.38a)-(9.38b), we assume that $\mathcal{G}_d(\lambda_i^d) \neq 0$ for all $\lambda_i^d \in \rho(S_d)$.

9.4.3 Discrete output feedback regulator design

Considering that the state information of the exo-system and pipeline model may not be available, we formulate the following observer-based output feedback regulator:

$$u_k = K_d \hat{w}_{k-1} + L_d \hat{q}_k \tag{9.48}$$

where $K_d \in \mathcal{L}(\mathcal{X}, U)$, $L_d \in \mathcal{L}(\mathbb{C}^n, U)$. For state estimation, we propose two Luenberger observers for the infinite- and finite-dimensional systems. First, let us construct an observer for the discrete-time infinite-dimensional single pipe system (9.17) as follows:

$$\hat{w}_k = \mathcal{A}_d \hat{w}_{k-1} + \mathcal{B}_d u_k + H(y_{mk} - \hat{y}_{mk}) \tag{9.49a}$$

$$\hat{y}_{mk} = C_{md} \hat{w}_{k-1} + \mathcal{D}_{md} u_k \tag{9.49b}$$

By introducing an estimation error $e_k^w = w_k - \hat{w}_k$, it is straightforward to derive the error dynamics through combining the observer model (9.49) with the plant model (9.17) as follows:

$$e_k^w = (\mathcal{A}_d - H C_{md}) e_{k-1}^w \tag{9.50}$$

To determine the discrete output injection gain H, we introduce the following lemma that is a dual version of Lemma 2.

Lemma 9. *An infinite-dimensional discrete-time system $w_k = \mathcal{A}_{od} w_{k-1}$ is power stable if and only if there exists a positive self-adjoint operator Q_{od} satisfying*

$$\mathcal{A}_{od} Q_{od} \mathcal{A}_{od}^* - Q_{od} + M_{od} = 0 \tag{9.51}$$

on $\mathcal{X}$, where M_{od} is a positive definite design parameter.

Proof. The proof is similar to [4, Exe. 4.29b] and [350] if we replace the controllability gramian L_B by Q_{od}, $A = \mathcal{A}_{od}$ and $BB^* = M_{od}$, respectively. ■

By plugging $\mathcal{A}_{od} = \mathcal{A}_d - HC_{md}$ and $M_{od} = (\mathcal{B}_d - H\mathcal{D}_{md})N_{od}(\mathcal{B}_d - H\mathcal{D}_{md})^*$ in Eq.(9.51), we obtain the following discrete-time Riccati equation:

$$\mathcal{A}_d Q_{od} \mathcal{A}_d^* - Q_{od} + \mathcal{B}_d N_{od} \mathcal{B}_d^* - H\left(2\vartheta I + C_{md} Q_{od} C_{md}^* + \mathcal{D}_{md} N_{od} \mathcal{D}_{md}^*\right) H^* = 0 \quad (9.52)$$

on $\mathcal{X}$. ϑ and N_{od} are positive definite design parameters so that $Q_{od} \in \mathcal{L}(\mathcal{X})$ and the discrete stabilizing output injection gain is given by $H = (\mathcal{A}_d Q_{od} C_{md}^* + \mathcal{B}_d N_{od} \mathcal{D}_{md}^*)(\vartheta I + C_{md} Q_{od} C_{md}^* + \mathcal{D}_{md} N_{od} \mathcal{D}_{md}^*)^{-1}$. Similar discrete-time Riccati equations can be found in [4, Exe. 6.35] and [24, Lem. 6]. Particularly, the output injection gain is not unique given that one can choose different design parameters ϑ and N_{od}.

Based on the modern control theory of finite-dimensional systems, one can design an observer for the exo-system (9.34) as below:

$$\hat{q}_k = S_d \hat{q}_{k-1} + G(y_{r(k-1)} - \hat{y}_{r(k-1)}),\ \hat{q}_0 = \hat{q}^0 \in \mathbb{C}^n \quad (9.53a)$$

$$\hat{y}_{rk} = Q_d \hat{q}_k,\ k \geq 1 \quad (9.53b)$$

Along this line, the estimation error is defined as $e_k^q = q_k - \hat{q}_k$ and its dynamics is described as:

$$e_k^q = (S_d - GQ_d) e_{k-1}^q,\ e_0^q = e^{q_0} \in \mathbb{C}^n,\ k \geq 1 \quad (9.54)$$

Therein, G can be determined such that $S_d - GQ_d$ is a Hurwitz matrix by performing pole placement. The estimated states $\hat{w}_k$ and $\hat{q}_k$ can be eventually substituted into Eq.(9.48) to attain the control law u_k. Hence, the discrete-time output regulator design is completed and will be verified by two simulation examples in Section 9.6.

9.5 Stability analysis

In this section, we provide insight into the internal stability of the continuous-time system (9.11) by using a Lyapunov functional analysis. Based on that, the strong stability of the discrete-time model (9.17) can be ensured.

Theorem 26. *The system (9.11) is exponentially stable with* $u = 0$.

Proof. A symmetric hyperbolic system is obtained by applying a linear transformation $T = [0.5, -0.5a; 0.5, 0.5a]$ to the system (9.11) as below:

$$\bar{w}(x,t) = Tw(x,t) \tag{9.55}$$

Along this line, the transformed state equations are given as:

$$\frac{\partial}{\partial t}\bar{w}(x,t) = \mathcal{A}_{\bar{w}}\bar{w}(x,t),\ \bar{w}(x,0) = Tw_0(x) \tag{9.56}$$

with $\mathcal{A}_{\bar{w}} = T\mathcal{A}T^{-1} = \Lambda_o\frac{\partial}{\partial x} + \Sigma_o$, where $\Delta = \frac{aQ-F}{2}$, $\Pi = \frac{aQ+F}{2}$, $\Lambda_o = [a, 0; 0, -a]$, and $\Sigma_o = [\Delta, \Pi; -\Delta, -\Pi]$.

In order to further convert Σ_o into a symmetric matrix form while Λ_o stays invariant, the following similarity transform is deployed as $\hat{w}(x,t) = \hat{T}\bar{w}(x,t) = [\sqrt{-\Delta}, 0; 0, \sqrt{\Pi}]w(x,t)$. Thus, one can attain new transformed state equations as:

$$\frac{\partial}{\partial t}\hat{w}(x,t) = \mathcal{A}_{\hat{w}}\hat{w}(x,t),\ \hat{w}(x,0) = \hat{T}Tw_0(x) \tag{9.57}$$

where $\mathcal{A}_{\hat{w}} = \hat{T}\mathcal{A}_{\bar{w}}\hat{T}^{-1} = \Lambda\frac{\partial}{\partial x} + \Sigma$, with $\Lambda = [a, 0; 0, -a]$, and $\Sigma = [\Delta, \sqrt{-\Delta\Pi}; \sqrt{-\Delta\Pi}, -\Pi]$. Apparently, $\mathcal{A}_{\hat{w}} := \Lambda\frac{\partial}{\partial x} + \Sigma$ with the domain $\mathcal{D}(\mathcal{A}_{\hat{w}}) = \{\phi = [\phi_1(x), \phi_2(x)]^T \in L^2(0,l)^2, \sqrt{\Pi}\phi_1(l) + \sqrt{-\Delta}\phi_2(l) = 0, \sqrt{\Pi}\phi_1(0) - \sqrt{-\Delta}\phi_2(0) = 0.\ \phi_1(x)$ and $\phi_2(x)$ are abs. cont.$\}$ generates a C_0-semigroup on $\mathcal{X}$. In this case, we can simply check that $\Delta < 0$ and $\Pi > 0$ leading to a real value of $(-\Delta\Pi)$.

Motivated by [169], we construct a Lyapunov functional $V_\theta : \mathcal{X} \to \mathbb{R}^+$ as follows:

$$V_\theta(\hat{w}(x,t)) = \int_0^l \hat{w}(x,t)^* e^{x\theta\Lambda}\hat{w}(x,t)dx \tag{9.58}$$

For a given smooth initial condition, one can calculate the temporal derivative of the considered Lyapunov functional $\dot{V}_\theta(\hat{w}(x,t))$ as below:

$$\begin{aligned}
&\int_0^l \frac{\partial}{\partial t}\hat{w}(x,t)^* e^{x\theta\Lambda}\hat{w}(x,t)dx + \int_0^l \hat{w}(x,t)^* e^{x\theta\Lambda}\frac{\partial}{\partial t}\hat{w}(x,t)dx \\
=&\int_0^l [\Lambda\hat{w}_x(x,t) + \Sigma\hat{w}(x,t)]^* e^{x\theta\Lambda}\hat{w}(x,t)dx \\
&+ \int_0^l \hat{w}(x,t)^* e^{x\theta\Lambda}[\Lambda\hat{w}_x(x,t) + \Sigma\hat{w}(x,t)]dx \\
=&\hat{w}(x,t)^*\Lambda e^{x\theta\Lambda}\hat{w}(x,t)|_0^l - \int_0^l \hat{w}(x,t)^*\theta\Lambda^2 e^{x\theta\Lambda}\hat{w}(x,t)dx
\end{aligned}$$

$$
\begin{aligned}
&+\int_0^l \hat{w}(x,t)^*[\Sigma^* e^{x\theta\Lambda}+e^{x\theta\Lambda}\Sigma]\hat{w}(x,t)dx\\
=&\hat{w}(l,t)^*\Lambda e^{l\theta\Lambda}\hat{w}(l,t)-\hat{w}(0,t)^*\Lambda\hat{w}(0,t)\\
&-\int_0^l \hat{w}(x,t)^* e^{\frac{1}{2}x\theta\Lambda}[\theta\Lambda^2-2\Sigma]e^{\frac{1}{2}x\theta\Lambda}\hat{w}(x,t)dx\\
&+\int_0^l \hat{w}(x,t)^* e^{\frac{1}{2}x\theta\Lambda}\Omega(x,t)e^{\frac{1}{2}x\theta\Lambda}\hat{w}(x,t)dx\\
\leq&\hat{w}(l,t)^*\Lambda_{po} e^{l\theta\Lambda}\hat{w}(l,t)+\hat{w}(0,t)^*\Lambda_{po}\hat{w}(0,t)\\
&-\int_0^l \hat{w}(x,t)^*\Lambda_{po} e^{x\theta\Lambda}\hat{w}(x,t)dx\\
&-\int_0^l \hat{w}(x,t)^* e^{\frac{1}{2}x\theta\Lambda}[\theta\Lambda^2-\Lambda_{po}-2\Sigma]e^{\frac{1}{2}x\theta\Lambda}\hat{w}(x,t)dx\\
&+\int_0^l \hat{w}(x,t)^* e^{\frac{1}{2}x\theta\Lambda}\Omega(x,t)e^{\frac{1}{2}x\theta\Lambda}\hat{w}(x,t)dx
\end{aligned}
\tag{9.59}
$$

where $\Omega(x,t)=e^{\frac{1}{2}x\theta\Lambda}\Sigma e^{-\frac{1}{2}x\theta\Lambda}+e^{-\frac{1}{2}x\theta\Lambda}\Sigma e^{\frac{1}{2}x\theta\Lambda}-2\Sigma$, and $\Lambda_{po}=aI$ with the identity matrix I.

It is straightforward to show that the summation of first three terms are negative in the last expression in Eq.(9.59), and there exists some positive constant M leading to:

$$
z^*(\theta\Lambda^2-\Lambda_{po}-2\Sigma^*)z\geq\frac{\theta}{M}z^*z,\ \forall z\in\mathbb{R}^{2\times 1} \tag{9.60}
$$

Furthermore, one can consider some large constant $\varkappa$ to ensure the following:

$$
\int_0^l \hat{w}(x,t)^* e^{\frac{1}{2}x\theta\Lambda}\Omega(x,t)e^{\frac{1}{2}x\theta\Lambda}\hat{w}(x,t)dx\leq\varkappa\theta^2 V_\theta(\hat{w}(x,t)) \tag{9.61}
$$

Therefore, for some small values of θ (eg: $\theta=\frac{1}{2M\varkappa}$), one can show that

$$
\dot{V}_\theta(\hat{w}(x,t))\leq-\frac{\theta}{2M}V_\theta(\hat{w}(x,t)) \tag{9.62}
$$

The last expression implies that the original continuous-time system (9.11) is exponentially stable. By the [23, Lem. 2.2], one can finally conclude that the associated discrete-time system (9.17) is internally strongly stable. ■

Remark 38. *The proposed proof of internal stability extends the previous results in [169] to the case with nontrivial boundary conditions, so that the internally strong stability of the associated discrete-time system is subsequently guaranteed.*

Remark 39. *In a similar manner, one can verify the internally strong stability of the star-shaped pipe network system (9.31).*

9.6 Numerical simulation

In this section, two realistic gas pipeline systems are simulated to verify the effectiveness of the proposed discrete-time output regulator design, including step signal tracking for a straight pipeline and harmonic signal tracking for a star-shaped network. In both simulation cases, we adopt pipeline parameters given in Table 9.1 and the equilibrium profiles shown in Fig. 9.2 for each pipe segment. In addition, we take $\Delta t = 5$ s, $\Delta x = 10$ m and $x_c = l/2$ in both cases.

9.6.1 Example 1: step signal tracking for a straight pipeline

In this example, we consider step-like reference signal tracking for a single straight gas pipeline. Based on the target reference signals, the discrete-time exo-system is designed as $S_d = 1$, $q_0 = 1$, $\hat{q}_0 = 0.8$, and $Q_d = \begin{cases} 20000, t \in [0, 500) \\ 40000, t \in [500, 1000] \end{cases}$. Thus, the discrete-time reference signal is generated as: $y_{rk} = \begin{cases} 20000, t \in [0, 500) \\ 40000, t \in [500, 1000] \end{cases}$. Revisiting Eq.(9.47a)-(9.47b), the discrete regulator gains Π_d and Γ_d can be solved as $\Gamma_d = \begin{cases} 23.4810, t \in [0, 500) \\ 46.9619, t \in [500, 1000] \end{cases}$. The linearized model derived from the original nonlinear model around different equilibrium points will have different model parameters. Hence, for implementing regulators to better suit the original nonlinear cases and the linearized models with different model parameters, in general we need to solve the regulator equations for each tracking task.

In this case, we calculate the discrete stabilizing gain K_d and output injection gain H via Eq.(9.37) and Eq.(9.52) numerically, where $x_{m1} = 0$ and $x_{m2} = l$ are chosen as the sensor locations. As for the observer design of the exogenous system, we apply the pole placement and the desired eigenvalue of error system (9.54) is set as -0.5. After 1000 seconds of simulation, the results are depicted in Fig. 9.3 and Fig. 9.4. It is apparent that the intrinsic dynamics of the pipeline model is stable since the open-loop output y_c converges to zero. With the control law applied to the upstream mass flux, the controlled pressure at $x_c = l/2$ can track the pre-designed step reference signal as illustrated in Fig. 9.3. From the output regulation perspective, it can be clearly seen that controlled output y_c can realize perfect tracking and the tracking error converges to near zero rapidly under the closed-loop control, as shown in Fig. 9.4. By

implementing the designed discrete regulator to the nonlinear continuous model (9.5) and the linearized continuous model (9.9), it can be seen that the designed discrete regulator is capable of perfectly regulating the linearized continuous (LC) model and steering nonlinear continuous (NC) model to track small reference (2×10^4 Pa) while the tracking performance of the NC model starts to degrade as the tracking reference increases to 4×10^4 Pa.

Remark 40. *The proposed discrete regulator design can be utilized to regulate the linearized continuous model (9.9) and approximately steer the nonlinear continuous model (9.5) for tracking small references close to the nominal equilibrium points.*

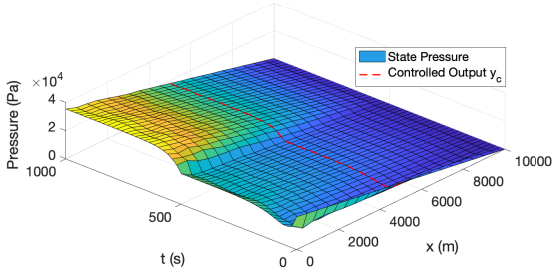


Figure 9.3: Closed-loop pressure profile of the single pipe system.

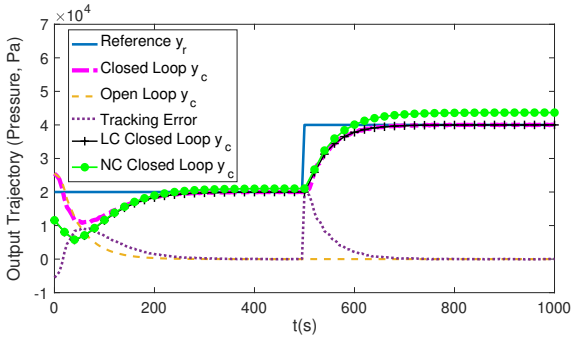


Figure 9.4: Reference tracking performance of the single pipe system.

9.6.2 Example 2: periodic signal tracking for a star-shaped network

In this case, we design an output regulator to track a periodic signal for the single-inlet-two-outlet gas pipeline network (9.31), and γ is taken as 0.5. To generate the target reference signal, the discrete-time exo-system is designed as $S_d = [0.9518, 0.3066;$

$-0.3066, 0.9518]$, $q_0 = [0; 1]$, $\hat{q}_0 = [-0.2; 1.2]$, and $Q_d = [10000, 0]$. Thus, the discrete reference signal is generated as: $y_{rk} = 10000 \times \sin(\pi \Delta t k/50)$. Recalling Eq.(9.47a)-(9.47b), we can calculate discrete regulator gains Π_d and Γ_d as $\Gamma_d = [-145.1998, -194.0138]$. In this example, we rely on the model internal stability and do not apply feedback control. By using pole placement, we design an observer for the finite-dimensional exo-system.

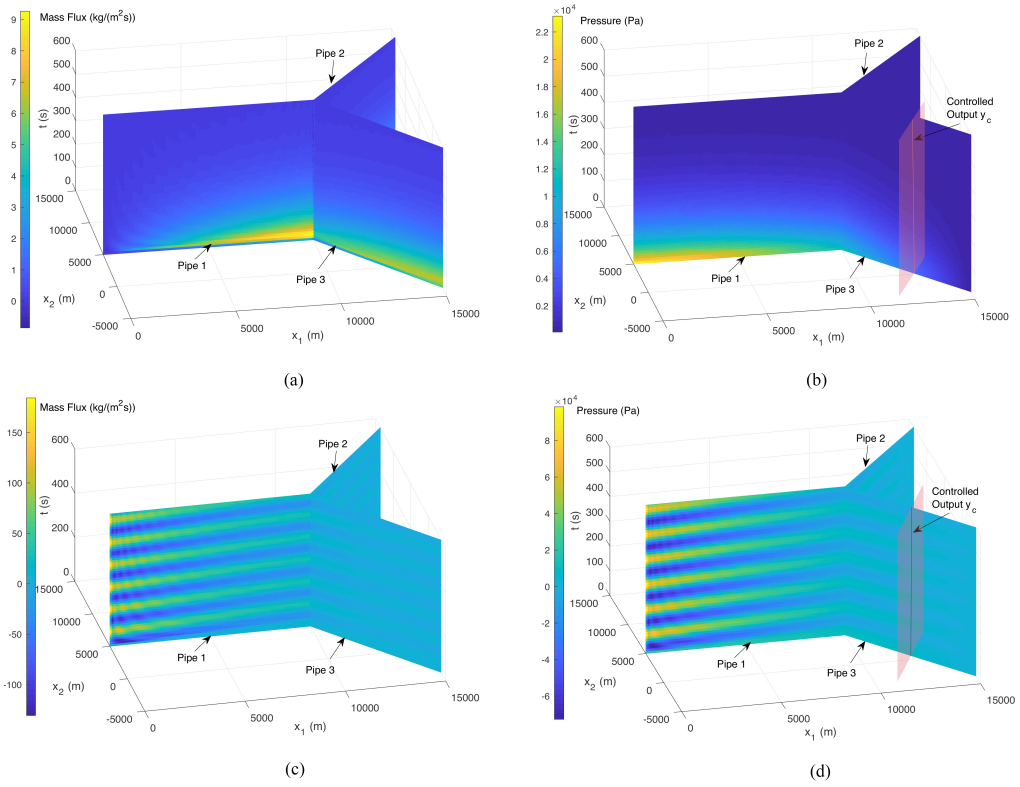


Figure 9.5: State evolution of the pipe network: (a) open-loop mass flow flux evolution; (b) open-loop pressure evolution; (c) closed-loop mass flow flux evolution; (d) closed-loop pressure evolution.

After simulation of 600 seconds, the state and output evolution profiles are illustrated in Fig. 9.5 and Fig. 9.6. It can be seen that intrinsic dynamics of the pipeline network is stable because the open-loop profiles of mass flow flux and pressure converges to zero, as shown in Fig. 9.5 (a)-(b). By implementing the control action at the upstream end, the controlled pressure at the middle point ($x_c = l/2$) of pipe 3 is capable of tracking the sinusoid reference signal as shown in Fig. 9.5 (c)-(d). More

specifically, it can be observed that the closed-loop controlled output at first steps is not smooth due to fact that there are still non-ignorable estimation errors. With time increasing, the controlled output y_c converges to the desired reference signal and the tracking error decays to near zero as illustrated in Fig. 9.6.

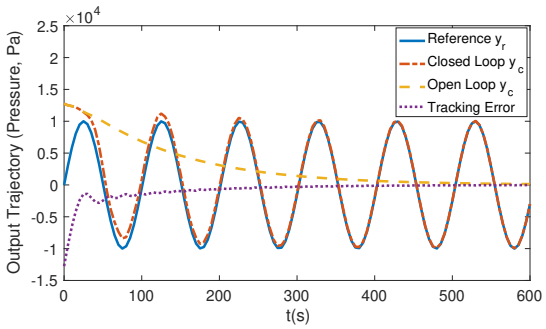


Figure 9.6: Reference tracking performance of the pipe network.

9.7 Conclusions

In this chapter, a discrete-time output regulator for gas flow pipeline network systems described by linearized first-order coupled hyperbolic equations was developed. The major conclusions are drawn as follows: 1) by using Cayley-Tustin bilinear transform, the linearized continuous-time infinite-dimensional model with unbounded operators was transformed into a discrete-time infinite-dimensional model with all bounded operators; 2) closed-form analytic expressions of discrete model operators of the pipeline gas network were obtained under consideration of the coupling conditions at junctions; 3) based on the internal model control theory, novel discrete-time output regulator equations were formulated and the corresponding solvability conditions were provided; 4) Discrete-time Riccati equations were proposed for calculating the controller and observer gains. Finally, two numerical examples have demonstrated that the presented discrete regulator design is effective and has a good servo-control performance. The further benefit of this design methodology can be integrated into supervisory control and data acquisition (SCADA) systems of large-scale pipe network systems for energy scheduling and pressure surge suppression by considering a disturbance observer design in a future work.

Chapter 10

Conclusions and Future Work

10.1 Conclusions

In this thesis, we systematically studied the state/output, parameter and model estimation and soft sensing of linear distributed and lumped parameter systems under the consideration of plant and measurement noises/disturbances, physical constraints and varying operating conditions, and extended the continuous-time internal model control theory into discrete-time regulation of linear distributed parameter systems. Fruitful applications of the proposed techniques have been provided.

In Chapter 2, this thesis addressed state estimation problems of linear deterministic and stochastic transport-reaction distributed parameter systems. For deterministic systems, continuous- and discrete-time infinite-dimensional Luenberger observers were designed and their performances were numerically compared. For stochastic systems, discrete-time two-step and one-step Kalman filters were proposed by extending the standard finite-dimensional Kalman filter theory. The prominent feature of the observer and filter designs was the application of the Cayley-Tustin transformation, leading to a discrete-time infinite-dimensional model, while preserving essential model properties. The design results were demonstrated through illustrative examples of chemical transport-reaction processes.

To complement the works in Chapter 2, Chapter 3 developed a linear moving horizon estimator for constrained output estimation of general regular linear distributed parameter systems with unbounded disturbance and observation operators. The Cayley-Tustin transformation was utilized to approximate the continuous DPS model with a discrete distributed parameter model. Important results on the discrete-

time Riccati equation and exact observability were provided. Based on that, the optimality and stability analysis for the proposed full information estimation and moving horizon estimator was given. Two representative examples including Schrödinger equation and wave equation demonstrated the effectiveness of the proposed design.

Chapter 4 is an extension of the proposed estimator design from Chapter 3. In Chapter 4, we proposed moving horizon estimator designs for switching regular linear infinite-dimensional systems with bounded disturbances and unknown (and unpredictable) switching modes. Similarly to Chapters 2-3, we have applied Cayley-Tustin transformation to transform the continuous model with unbounded operators to the discrete-time model with all bounded operators. By introducing *ad-hoc* observability concepts (i.e., simultaneous exact (and approximate) observability), we have proposed suitable observability conditions for finding a solution of the estimation problem. This was accomplished by extending the observability concepts of switching linear and nonlinear discrete-time LPSs [161, 157] and non-switching linear continuous-time DPSs [4, 42]. Based on that, we have provided the stability analysis of the proposed MHE algorithm. A heat exchanger and a damped Rayleigh beam equation were taken as two representative examples to verify the proposed algorithm.

In Chapter 5, this thesis proposed an online soft sensor algorithm (i.e., transfer slow feature analysis (TSFA)) for dynamically predicting the quality variables of complex industrial systems. In a transfer learning manner, the proposed algorithm can dynamically transfer the model parameters of multiple source systems to the target system that has limited output labels. As a prerequisite of TSFA, the probabilistic slow feature analysis (PSFA) was applied to learn source domain models (with a preference for slow features) based on the input and output measurements from each source domain. More specifically, a linear discrete-time lumped parameter model with Gaussian noises was considered in both algorithms. Variational Bayesian inference was deployed for dynamical model learning in PSFA and dynamical learning of weighting functions in TSFA. Finally, we have demonstrated the applicability of the proposed algorithm through a simulation example, a public dataset, and an industrial SAGD process.

Chapter 6 is an extension of the proposed estimator design from Chapter 2. In Chapter 6, we developed a hybrid estimator for pipeline condition monitoring and

leak diagnosis. To be more precise, the model-based discrete Luenberger observer and the data-driven support vector machine model were incorporated into the hybrid estimator design. Stemming from the governing equations (i.e., two coupled nonlinear first-order hyperbolic PDEs), a linear discrete-time infinite-dimensional pipeline model was obtained using linearization and Cayley-Tustin transformation, based on which, the discrete Luenberger observer design was carried out. By using the observer, various normal working situations and different leakage cases were simulated. Two support vector machine models (i.e., OCSVM and SVR) were trained and tested for leak detection, localization, and size estimation. Through simulations on a large scale pipeline system, the effectiveness of the proposed design was validated.

In Chapter 7, this thesis addressed discrete-time state and error feedback output regulator designs for a class of linear distributed parameter systems with bounded control and observation operators. Likewise, the Cayley-Tustin transformation was deployed for model time discretization while no spatial approximation or model order reduction was required. We formulated and proved the discrete-time regulator equations for solving the discrete state and error feedback regulator designs by using the PDE plant model and the exogenous system in discrete-time setting. Along this line, the discrete- and continuous-time state and error feedback regulator designs were shown to be equivalent under the Cayley-Tustin transformation. Moreover, we further proved that the solvability of discrete- and continuous-time regulator equations are equivalent under standard assumptions. Discrete-time Riccati equations were proposed for stabilizing output injection gain determination (and its dual problem). Finally, three examples on step-like, ramp-like and harmonic reference control of a first-order transport equation and a heat equation demonstrated the proposed results.

Chapters 8-9 are extensions and applications of the proposed regulator design from Chapter 7. In Chapter 8, we proposed the discrete-time output regulator designs for fluid flow systems (i.e., spectral PDE models) with unbounded controlled output operator. In particular, we considered the complex Ginzburg-Landau equation (CGLE) and Kuramoto-Sivashinsky equation (KSE) as two representative fluid flow PDE models. Similarly, the Cayley-Tustin time discretization method was utilized to transform the continuous-time PDE model into the discrete-time model. Based on that, we provided a link between discrete- and continuous-time closed-loop system

stabilizing gains (and its dual problem). The analytic solutions of resolvent operators and the discrete operators of CGLE and KSE were provided, which indicated that both systems are well-posed regular systems. Two simulation examples validated the developed regulator design.

In Chapter 9, we further extended the discrete-time output regulator designs to networked hyperbolic PDE systems. More specifically, we have investigated the continuous- and discrete-time modelling of gas flow pipeline network systems in terms of linearized first-order coupled hyperbolic PDEs. In particular, we considered unbounded control and observation operators in the continuous-time model, and converted them all into bounded operators in the discrete-time setting by using Cayley-Tustin transformation. The regularity and stability of the pipeline network system were proved. Similarly to Chapter 7, the discrete-time regulator equations were proved and the discrete-time Riccati equation was used to solve for the discrete-time output injection gain. Finally, the applicability of the proposed design was verified through two numerical examples.

10.2 Future work

The observer, filter, estimator and soft sensor designs for state/output, parameter and model estimation and soft sensing (output prediction) of linear distributed and lumped parameter systems have been systematically investigated in this thesis. Nevertheless, there are still some open questions in this important topic. For instance, soft sensor designs for linear distributed parameter systems using variational approaches have not been touched. In the cases of tubular reactors, online measurements of species concentrations are usually not available, and hence how to address state estimation of tubular reactor processes using limited, delayed or slow-rated measurements can be interesting and challenging.

Considering that many practical systems are modelled by PDE-ODE coupled systems, state/output and parameter estimation of PDE-ODE coupled systems is of practical and theoretical significance. For example, the vehicular traffic flow models and flexible spacecraft models are generally modelled by PDE-ODE coupled systems. The switching of dynamic parameters can happen in the ODE part and/or the PDE

part that might involve temporal and/or spatial varying parameters, which deserves more research in the future. Moreover, estimation of switching systems that are composed of continuous-time models and discrete-time models could be investigated, especially from the perspective of sampled-data systems with multi-rate measurements.

By using the Cayley-Tustin transformation, the continuous-time internal model control theory has been extended to the discrete-time infinite-dimensional systems. There are some open questions that need to be further answered. First of all, the Cayley-Tustin transformation follows an implicit midpoint rule, which has second-order accuracy. Along this line, exploring advanced discretization schemes with higher order accuracy is interesting and can be possibly done [119]. Moreover, research on different types of controller designs of discrete-time infinite-dimensional systems could be considered in future study, e.g., passivity/dissipativity based control, robust control and etc.

Bibliography

[1] Mojtaba Izadi, Charles R Koch, and Stevan S Dubljevic. Model predictive control of Ginzburg-Landau equation. In *Active Flow and Combustion Control 2018*, pages 75–90. Springer, 2019.

[2] Ole Morten Aamo and Miroslav Krstic. Global stabilization of a nonlinear Ginzburg-Landau model of vortex shedding. *Eur. J. Control*, 10(2):105–116, 2004.

[3] Stevan Dubljevic. Boundary model predictive control of Kuramoto–Sivashinsky equation with input and state constraints. *Comput. Chem. Eng.*, 34(10):1655–1661, 2010.

[4] Ruth F Curtain and Hans Zwart. *An introduction to infinite-dimensional linear systems theory*, volume 21. New York, NY: Springer, 1995.

[5] Martin Weiss and George Weiss. Optimal control of stable weakly regular linear systems. *Math. Control Signals Syst.*, 10(4):287–330, 1997.

[6] Olof Staffans. Quadratic optimal control of stable well-posed linear systems. *Trans. Am. Math. Soc.*, 349(9):3679–3715, 1997.

[7] Ruth F Curtain. Infinite-dimensional filtering. *SIAM Journal on Control*, 13(1):89–104, 1975.

[8] Ruth F Curtain. A survey of infinite-dimensional filtering. *SIAM Review*, 17(3):395–411, 1975.

[9] Minxin Zhang and Kirsten Morris. Sensor choice for minimum error variance estimation. *IEEE Trans. Autom. Control*, 63(2):315–330, 2017.

[10] Christopher I Byrnes, István G Laukó, David S Gilliam, and Victor I Shubov. Output regulation for linear distributed parameter systems. *IEEE Trans. Autom. Control*, 45(12):2236–2252, 2000.

[11] Richard Rebarber and George Weiss. Internal model based tracking and disturbance rejection for stable well-posed systems. *Automatica*, 39(9):1555–1569, 2003.

[12] Eero Immonen. Practical output regulation for bounded linear infinite-dimensional state space systems. *Automatica*, 43(5):786–794, 2007.

[13] Vivek Natarajan, David S Gilliam, and George Weiss. The state feedback regulator problem for regular linear systems. *IEEE Trans. Autom. Control*, 59(10):2708–2723, 2014.

[14] Joachim Deutscher. Backstepping design of robust state feedback regulators for linear 2× 2 hyperbolic systems. *IEEE Trans. Autom. Control*, 62(10):5240–5247, 2017.

[15] Miroslav Krstic and Andrey Smyshlyaev. *Boundary control of PDEs: A course on backstepping designs*, volume 16. Siam, 2008.

[16] Ole Morten Aamo, Andrey Smyshlyaev, Miroslav Krstic, and Bjarne A Foss. Output feedback boundary control of a Ginzburg–Landau model of vortex shedding. *IEEE Trans. Autom. Control*, 52(4):742–748, 2007.

[17] Rafael Vazquez, Miroslav Krstic, and Jean-Michel Coron. Backstepping boundary stabilization and state estimation of a 2× 2 linear hyperbolic system. In *Decision and Control and European Control Conference (CDC-ECC), 2011 50th IEEE Conference on*, pages 4937–4942. IEEE, 2011.

[18] Jean Auriol, Kirsten A Morris, and Florent Di Meglio. Late-lumping backstepping control of partial differential equations. *Automatica*, 100:247–259, 2019.

[19] Raimund Ober and Stephen Montgomery-Smith. Bilinear transformation of infinite-dimensional state-space systems and balanced realizations of nonrational transfer functions. *SIAM J. Control Opt.*, 28(2):438–465, 1990.

[20] Jarmo Malinen, Olof Staffans, and George Weiss. When is a linear system conservative? *Quarterly of Applied Mathematics*, 64(1):61–91, 2006.

[21] Olof Staffans. *Well-posed linear systems*, volume 103. Cambridge University Press, 2005.

[22] Bao-Zhu Guo and Hans Zwart. On the relation between stability of continuous- and discrete-time evolution equations via the Cayley transform. *Integr. Equat Oper. Th.*, 54(3):349–383, 2006.

[23] Ruth F Curtain and Job Oostveen. Bilinear transformations between discrete- and continuous-time infinite-dimensional linear systems. In *Proceedings of the International Symposium MMAR'97, Miedzyzdroje, Poland*, pages 861–870, 1997.

[24] Mark R Opmeer and Ruth F Curtain. New Riccati equations for well-posed linear systems. *Syst. Control. Lett.*, 52(5):339–347, 2004.

[25] Hartmut Logemann and Stuart Townley. Discrete-time low-gain control of uncertain infinite-dimensional systems. *IEEE Trans. Autom. Control*, 42(1):22–37, 1997.

[26] Richard Rebarber and Stuart Townley. Generalized sampled data feedback control of distributed parameter systems. *Syst. Control. Lett.*, 34(5):229–240, 1998.

[27] Richard Rebarber and Stuart Townley. Nonrobustness of closed-loop stability for infinite-dimensional systems under sample and hold. *IEEE Trans. Autom. Control*, 47(8):1381–1385, 2002.

[28] Hartmut Logemann, Richard Rebarber, and Stuart Townley. Generalized sampled-data stabilization of well-posed linear infinite-dimensional systems. *SIAM J. Control Opt.*, 44(4):1345–1369, 2005.

[29] Michael P Polis and Raymond E Goodson. Parameter identification in distributed systems: A synthesizing overview. *Proceedings of the IEEE*, 64(1):45–61, 1976.

[30] Harmon Ray. *Advanced Process Control*. McGraw-Hill, New York, 1981.

[31] Denis Dochain. State and parameter estimation in chemical and biochemical processes: a tutorial. *J. Process Control*, 13(8):801–818, 2003.

[32] Hartmut Logemann, Richard Rebarber, and Stuart Townley. Stability of infinite-dimensional sampled-data systems. *Trans. Am. Math. Soc.*, 355(8):3301–3328, 2003.

[33] Thomas Banks and Karl Kunisch. *Estimation techniques for distributed parameter systems*. Springer Science & Business Media, 2012.

[34] Jarinah Mohd Ali, N. Ha Hoang, M.A. Hussain, and Denis Dochain. Review and classification of recent observers applied in chemical process systems. *Comput. Chem. Eng.*, 76:27 – 41, 2015.

[35] Dan Simon. *Optimal state estimation: Kalman, H_∞ and nonlinear approaches*. John Wiley & Sons, 2006.

[36] James Blake Rawlings and David Q Mayne. *Model predictive control: Theory and design*. Nob Hill Pub., 2009.

[37] Kenneth R Muske, James B Rawlings, and Jay H Lee. Receding horizon recursive state estimation. In *1993 American Control Conference*, pages 900–904. IEEE, 1993.

[38] Christopher V Rao, James B Rawlings, and Jay H Lee. Constrained linear state estimation—a moving horizon approach. *Automatica*, 37(10):1619–1628, 2001.

[39] Angelo Alessandri, Marco Baglietto, and Giorgio Battistelli. Receding-horizon estimation for discrete-time linear systems. *IEEE Trans. Autom. Control*, 48(3):473–478, 2003.

[40] Kirsten Morris. *Controller Design for Distributed Parameter Systems*. Springer, 2020.

[41] Birgit Jacob and Hans J Zwart. *Linear port-Hamiltonian systems on infinite-dimensional spaces*, volume 223. Springer Science & Business Media, 2012.

[42] Marius Tucsnak and George Weiss. *Observation and control for operator semigroups*. Springer Science & Business Media, 2009.

[43] Zheng-Hua Luo, Bao-Zhu Guo, and Ömer Morgül. *Stability and stabilization of infinite dimensional systems with applications*. Springer Science & Business Media, 2012.

[44] Kirsten Morris. *Control of systems governed by partial differential equations*. in ed. W.S. Levine, The Control Handbook, Second Edition, Control System Advanced Methods. CRC Press, 2010.

[45] Orest V Iftime and Michael A Demetriou. Optimal control of switched distributed parameter systems with spatially scheduled actuators. *Automatica*, 45(2):312–323, 2009.

[46] Kalle Mikkola et al. *Infinite-dimensional linear systems, optimal control and algebraic Riccati equations*. Helsinki University of Technology, 2002.

[47] George Weiss and Hans Zwart. An example in linear quadratic optimal control. *Syst. Control. Lett.*, 33(5):339–349, 1998.

[48] Chlng-Tsan Lo. Optimal control of counter-current distributed-parameter systems. *Int. J. Contr.*, 18(2):273–288, 1973.

[49] Kazufumi Ito and Karl Kunisch. Receding horizon optimal control for infinite dimensional systems. *ESAIM: Control Optim. Calc. Var.*, 8:741–760, 2002.

[50] Ilyasse Aksikas, Adrian Fuxman, J Fraser Forbes, and Joseph J Winkin. Lq control design of a class of hyperbolic pde systems: Application to fixed-bed reactor. *Automatica*, 45(6):1542–1548, 2009.

[51] Ilyasse Aksikas, Joseph J Winkin, and Denis Dochain. Optimal lq-feedback regulation of a nonisothermal plug flow reactor model by spectral factorization. *IEEE Trans. Autom. Control*, 52(7):1179–1193, 2007.

[52] Eero Immonen and Seppo Pohjolainen. Output regulation of periodic signals for dps: an infinite-dimensional signal generator. *IEEE Trans. Autom. Control*, 50(11):1799–1804, 2005.

[53] Timo Hämäläinen and Seppo Pohjolainen. Robust regulation of distributed parameter systems with infinite-dimensional exosystems. *SIAM J. Control Opt.*, 48(8):4846–4873, 2010.

[54] Joachim Deutscher. A backstepping approach to the output regulation of boundary controlled parabolic PDEs. *Automatica*, 57:56–64, 2015.

[55] Joachim Deutscher. Finite-time output regulation for linear 2× 2 hyperbolic systems using backstepping. *Automatica*, 75:54–62, 2017.

[56] Lassi Paunonen. Controller design for robust output regulation of regular linear systems. *IEEE Trans. Autom. Control*, 61(10):2974–2986, 2016.

[57] Eugenio Aulisa and David Gilliam. *A practical guide to geometric regulation for distributed parameter systems*. Chapman and Hall/CRC, 2015.

[58] Ole Morten Aamo, Andrey Smyshlyaev, and Miroslav Krstic. Boundary control of the linearized Ginzburg–Landau model of vortex shedding. *SIAM J. Control Opt.*, 43(6):1953–1971, 2005.

[59] Iasson Karafyllis, Miroslav Krstic, and Katerina Chrysafi. Adaptive boundary control of constant-parameter reaction–diffusion PDEs using regulation-triggered finite-time identification. *Automatica*, 103:166–179, 2019.

[60] Federico Bribiesca-Argomedo and Miroslav Krstic. Backstepping-forwarding control and observation for hyperbolic PDEs with Fredholm integrals. *IEEE Trans. Autom. Control*, 60(8):2145–2160, 2015.

[61] Nael H El-Farra, Antonios Armaou, and Panagiotis D Christofides. Analysis and control of parabolic pde systems with input constraints. *Automatica*, 39(4):715–725, 2003.

[62] Qingqing Xu and Stevan Dubljevic. Linear model predictive control for transport-reaction processes. *AIChE J.*, 63(7):2644–2659, 2017.

[63] Stevan Dubljevic, Nael H El-Farra, Prashant Mhaskar, and Panagiotis D Christofides. Predictive control of parabolic PDEs with state and control constraints. *Int. J. Robust Nonlinear Control*, 16(16):749–772, 2006.

[64] Thang Van Pham, Didier Georges, and Gildas Besançon. Predictive control with guaranteed stability for water hammer equations. *IEEE Trans. Autom. Control*, 59(2):465–470, 2013.

[65] Stevan Dubljevic and Jukka-Pekka Humaloja. Model predictive control for regular linear systems. *Automatica*, 119:109066, 2020.

[66] Panagiotis D Christofides and J Chow. Nonlinear and robust control of pde systems: Methods and applications to transport-reaction processes. *Appl. Mech. Rev.*, 55(2):B29–B30, 2002.

[67] Antonios Armaou and Panagiotis D Christofides. Robust control of parabolic pde systems with time-dependent spatial domains. *Automatica*, 37(1):61–69, 2001.

[68] Yury V Orlov. *Discontinuous systems: Lyapunov analysis and robust synthesis under uncertainty conditions*. Springer Science & Business Media, 2008.

[69] Andrey Smyshlyaev and Miroslav Krstic. *Adaptive control of parabolic PDEs*. Princeton University Press, 2010.

[70] Henrik Anfinsen and Ole Morten Aamo. *Adaptive control of hyperbolic PDEs*. Springer, 2019.

[71] Delio Mugnolo. *Semigroup methods for evolution equations on networks*, volume 20. Springer, 2014.

[72] Miroslav Krstic. *Delay compensation for nonlinear, adaptive, and PDE systems*. Springer, 2009.

[73] Héctor Ramírez, Yann Le Gorrec, Alessandro Macchelli, and Hans Zwart. Exponential stabilization of boundary controlled port-Hamiltonian systems with dynamic feedback. *IEEE Trans. Autom. Control*, 59(10):2849–2855, 2014.

[74] Jukka-Pekka Humaloja and Lassi Paunonen. Robust regulation of infinite-dimensional port-Hamiltonian systems. *IEEE Trans. Autom. Control*, 63(5):1480–1486, 2017.

[75] Peter L Falb. Infinite-dimensional filtering: The Kalman-Bucy filter in Hilbert space. *Inf. Control*, 11:102—137, 1967.

[76] Ruth F Curtain and AJ Pritchard. The infinite-dimensional Riccati equation. *J. Math. Anal. Appl*, 47(1):43–57, 1974.

[77] Ruth F Curtain and A J Pritchard. The infinite-dimensional Riccati equation for systems defined by evolution operators. *SIAM J. Control Optim.*, 14(5):951–983, 1976.

[78] Ruth F Curtain and Anthony J Pritchard. Infinite dimensional linear systems theory. Technical report, Springer-Verlag, 1978.

[79] Jarinah Mohd Ali, N Ha Hoang, Mohamed Azlan Hussain, and Denis Dochain. Review and classification of recent observers applied in chemical process systems. *Comput. Chem. Eng.*, 76:27–41, 2015.

[80] Rudolph E Kalman and Richard S Bucy. New results in linear filtering and prediction theory. *J. Basic Eng.*, 83(3):95–108, 1961.

[81] David Luenberger. An introduction to observers. *IEEE Trans. Autom. Control*, 16(6):596–602, 1971.

[82] Karl Johan Åström and Peter Eykhoff. System identification—a survey. *Automatica*, 7(2):123–162, 1971.

[83] Thomas K Yu and John H Seinfeld. Observability and optimal measurement location in linear distributed parameter systems. *Int. J. Control*, 18(4):785–799, 1973.

[84] David G Luenberger. Observing the state of a linear system. *IEEE transactions on military electronics*, 8(2):74–80, 1964.

[85] Rudolph Emil Kalman. A new approach to linear filtering and prediction problems. *J. Basic Eng.*, 82(1):35–45, 1960.

[86] Sudarshan Kumar and John Seinfeld. Optimal location of measurements in tubular reactors. *Chem. Eng. Sci.*, 33(11):1507–1516, 1978.

[87] Thomas Yu, John Seinfeld, and Harmon Ray. Filtering in nonlinear time delay systems. *IEEE Trans. Autom. Control*, 19(4):324–333, 1974.

[88] Vande Wouwer and Michael Zeitz. State estimation in distributed parameter systems. *Control Systems, Robotics and Automation–Volume XIV: Nonlinear, Distributed, and Time Delay Systems-III*, page 92, 2009.

[89] Zulkifli Hidayat, Robert Babuska, Bart De Schutter, and Alfredo Nunez. Observers for linear distributed-parameter systems: A survey. In *2011 IEEE international symposium on robotic and sensors environments (ROSE)*, pages 166–171. IEEE, 2011.

[90] Denis Dochain. State observers for tubular reactors with unknown kinetics. *J. Process Control*, 10(2-3):259–268, 2000.

[91] Denis Dochain. State observation and adaptive linearizing control for distributed parameter (bio) chemical reactors. *Int. J. Adapt. Control Signal Process.*, 15(6):633–653, 2001.

[92] Efrén Aguilar-Garnica, Juan Paulo García-Sandoval, and Carlos González-Figueredo. A robust monitoring tool for distributed parameter plug flow reactors. *Comput. Chem. Eng.*, 35(3):510–518, 2011.

[93] Matthias Bitzer and Michael Zeitz. Design of a nonlinear distributed parameter observer for a pressure swing adsorption plant. *J. Process Control*, 12(4):533–543, 2002.

[94] Antonio A Alonso, Ioannis G Kevrekidis, Julio R Banga, and Christos E Frouzakis. Optimal sensor location and reduced order observer design for distributed process systems. *Comput. Chem. Eng.*, 28(1-2):27–35, 2004.

[95] Giuseppe Da Prato and Jerzy Zabczyk. *Stochastic equations in infinite dimensions*. Cambridge university press, 2014.

[96] Karl J. Astrom and Bjorn Wittenmark. *Computer-controlled Systems: Theory and Design (2Nd Ed.)*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1990.

[97] Ville Havu and Jarmo Malinen. The Cayley transform as a time discretization scheme. *Numer. Funct. Anal. Optim.*, 28(7-8):825–851, 2007.

[98] Gene F Franklin, J David Powell, and Michael L Workman. *Digital control of dynamic systems*, volume 3. Addison-wesley Menlo Park, CA, 1998.

[99] John von Neumann. Allgemeine eigenwerttheorie hermitescher funktionaloperatoren. *Mathematische Annalen*, 102(1):49–131, 1930.

[100] Nicolaas Cornelis Besseling. *Stability analysis in continuous and discrete time*. University of Twente, 2012.

[101] Ernst Haier, Christian Lubich, and Gerhard Wanner. *Geometric Numerical integration: structure-preserving algorithms for ordinary differential equations*. Springer, 2006.

[102] Xiaodong Xu and Stevan Dubljevic. Output regulation problem for a class of regular hyperbolic systems. *Int. J. Control*, 89(1):113–127, 2016.

[103] Joseph Winkin, Denis Dochain, and Phillipe Ligarius. Dynamical analysis of distributed parameter tubular reactors. *Automatica*, 36:349–361, 2000.

[104] Delattre Cedric, Denis Dochain, and Joseph Winkin. Sturm-Liouville systems are Riesz-spectral operators. *Int. J. Appl. Math. Comput. Sci.*, 13:481–484, 2003.

[105] Dirk Vries, Karel J Keesman, and Hans Zwart. Luenberger boundary observer synthesis for Sturm–Liouville systems. *Int. J. Control*, 83(7):1504–1514, 2010.

[106] Alfredo Germani, Leopoldo Jetto, and M Piccioni. Galerkin approximation for optimal linear filtering of infinite-dimensional linear systems. *SIAM J. Control Optim.*, 26(6):1287–1305, 1988.

[107] Frank Allgöwer, Thomas A Badgwell, Joe S Qin, James B Rawlings, and Steven J Wright. Nonlinear predictive control and moving horizon estimation—an introductory overview. In *Advances in control*, pages 391–449. Springer, 1999.

[108] Angelo Alessandri, Marco Baglietto, and Giorgio Battistelli. Moving-horizon state estimation for nonlinear discrete-time systems: New stability results and approximation schemes. *Automatica*, 44(7):1753–1765, 2008.

[109] Dietmar Salamon. Realization theory in Hilbert space. *Mathematical Systems Theory*, 21(1):147–164, 1988.

[110] George Weiss. The representation of regular linear systems on Hilbert spaces. *Control and estimation of distributed parameter systems*, 91:401–416, 1989.

[111] Marius Tucsnak and George Weiss. Well-posed systems—the LTI case and beyond. *Automatica*, 50(7):1757–1779, 2014.

[112] Olof J Staffans. On the discrete and continuous time infinite-dimensional algebraic Riccati equations. *Syst. Control. Lett.*, 29(3):131–138, 1996.

[113] Ruth F Curtain. Regular linear systems and their reciprocals: applications to Riccati equations. *Syst. Control. Lett.*, 49(2):81–89, 2003.

[114] Michael A Demetriou. Design of consensus and adaptive consensus filters for distributed parameter systems. *Automatica*, 46(2):300–311, 2010.

[115] Bao-Zhu Guo and Huaiqiang Yu. Optimal state estimation for non-time invertible evolutionary systems. *SIAM J. Control Optim.*, 54(5):2754–2786, 2016.

[116] Xiaosong Hu, Dongpu Cao, and Bo Egardt. Condition monitoring in advanced battery management systems: Moving horizon estimation using a reduced electrochemical model. *IEEE/ASME Trans. Mechatron.*, 23(1):167–178, 2017.

[117] Sönke Rhein, Tilman Utz, and Knut Graichen. Model predictive control and moving horizon estimation of a large-scale chemical reactor model. *IFAC Proceedings Volumes*, 46(26):121–126, 2013.

[118] Agus Hasan and Lars Imsland. Moving horizon estimation in managed pressure drilling using distributed models. In *2014 IEEE Conference on Control Applications (CCA)*, pages 605–610. IEEE, 2014.

[119] Ernst Hairer, Christian Lubich, and Gerhard Wanner. *Geometric numerical integration: structure-preserving algorithms for ordinary differential equations*, volume 31. Springer Science & Business Media, 2006.

[120] Mattia Bruschetta, Giorgio Picci, and Alessandro Saccon. A variational integrators approach to second order modeling and identification of linear mechanical systems. *Automatica*, 50(3):727–736, 2014.

[121] George Weiss. Transfer functions of regular linear systems. I. characterizations of regularity. *Trans. Am. Math. Soc.*, 342(2):827–854, 1994.

[122] George Weiss and Richard Rebarber. Optimizability and estimatability for infinite-dimensional linear systems. *SIAM J. Control Opt.*, 39(4):1204–1232, 2000.

[123] Mark R Opmeer and Olof J Staffans. Optimal input-output stabilization of infinite-dimensional discrete time-invariant linear systems by output injection. *SIAM J. Control Optim.*, 48(8):5084–5107, 2010.

[124] Mark R Opmeer and Olof J Staffans. Optimal control on the doubly infinite continuous time axis and coprime factorizations. *SIAM J. Control Optim.*, 52(3):1958–2007, 2014.

[125] Ruth Curtain and Hans Zwart. *Introduction to Infinite-Dimensional Systems Theory: A State-Space Approach*. Springer, 2020.

[126] Christopher V Rao. *Moving horizon strategies for the constrained monitoring and control of nonlinear discrete-time systems*. PhD thesis, University of Wisconsin–Madison, 2000.

[127] Ruth Frances Curtain and George Herbert Weiss. Well posedness of triples of operators (in the sense of linear systems theory). *Control and Estimation of Distributed Parameter Systems (F. Kappel, K. Kunisch, and W. Schappacher, eds.), Birkhäuser, Basel,*, pages 41–59, 1989.

[128] Bao-Zhu Guo and Zhi-Chao Shao. Regularity of a Schrödinger equation with dirichlet control and colocated observation. *Syst. Control. Lett.*, 54(11):1135–1142, 2005.

[129] Bao-Zhu Guo and Kun-Yi Yang. Output feedback stabilization of a one-dimensional Schrödinger equation by boundary observation with time delay. *IEEE Trans. Autom. Control*, 55(5):1226–1232, 2010.

[130] Ruth F Curtain. The Salamon—Weiss class of well-posed infinite-dimensional linear systems: a survey. *IMA J. Math. Control Inf.*, 14(2):207–223, 1997.

[131] Irena Lasiecka and Roberto Triggiani. L_2 (Σ)-regularity of the boundary to boundary operator $B*L$ for hyperbolic and petrowski PDEs. In *Abstr. Appl. Anal.*, volume 2003. Hindawi, 2003.

[132] Bao-Zhu Guo and Cheng-Zhong Xu. The stabilization of a one-dimensional wave equation by boundary feedback with noncollocated observation. *IEEE Trans. Autom. Control*, 52(2):371–377, 2007.

[133] Claudio Canuto, M Yousuff Hussaini, Alfio Quarteroni, and Thomas A Zang. *Spectral methods: fundamentals in single domains*. Springer Science & Business Media, 2007.

[134] David A Kopriva. *Implementing spectral methods for partial differential equations: Algorithms for scientists and engineers*. Springer Science & Business Media, 2009.

[135] Andrey Smyshlyaev, Bao-Zhu Guo, and Miroslav Krstic. Arbitrary decay rate for euler-bernoulli beam by backstepping boundary feedback. *IEEE Trans. Autom. Control*, 54(5):1134–1140, 2009.

[136] Yuri Orlov and Denis Dochain. Discontinuous feedback stabilization of minimum-phase semilinear infinite-dimensional systems with application to chemical tubular reactor. *IEEE Trans. Autom. Control*, 47(8):1293–1304, 2002.

[137] Amol Sasane. Stability of switching infinite-dimensional systems. *Automatica*, 41(1):75–78, 2005.

[138] Liguo Zhang and Christophe Prieur. Stochastic stability of markov jump hyperbolic systems with application to traffic flow control. *Automatica*, 86:29–37, 2017.

[139] Sebastian Peitz and Stefan Klus. Koopman operator-based model reduction for switched-system control of PDEs. *Automatica*, 106:184–191, 2019.

[140] Pierre-Olivier Lamare. Supervisory switching control for linear hyperbolic systems. *Automatica*, 105:64–70, 2019.

[141] Christian Clason, Armin Rund, Karl Kunisch, and Richard C Barnard. A convex penalty for switching control of partial differential equations. *Syst. Control. Lett.*, 89:66–73, 2016.

[142] Christophe Prieur, Antoine Girard, and Emmanuel Witrant. Stability of switched linear hyperbolic systems by Lyapunov techniques. *IEEE Trans. Autom. Control*, 59(8):2196–2202, 2014.

[143] Yan Zhao, Jianbin Qiu, Shengyuan Xu, Wenguo Li, and Junli Wu. Boundary observer-based control for hyperbolic PDE–ODE cascade systems with stochastic jumps. *Automatica*, 119:109089, 2020.

[144] Guojie Zheng, Jiandong Xiong, Xin Yu, and Chao Xu. Stabilization for infinite dimensional switched linear systems. *IEEE Trans. Autom. Control*, 2020.

[145] Yury Orlov, Yiming Lou, and Panagiotis D Christofides. Robust stabilization of infinite-dimensional systems using sliding-mode output feedback control. *Int. J. Control*, 77(12):1115–1136, 2004.

[146] Christian Clason, Kazufumi Ito, and Karl Kunisch. A convex analysis approach to optimal controls with switching structure for partial differential equations. *ESAIM: COCV*, 22(2):581–609, 2016.

[147] Pierre-Olivier Lamare, Antoine Girard, and Christophe Prieur. Switching rules for stabilization of linear systems of conservation laws. *SIAM J. Control Opt.*, 53(3):1599–1624, 2015.

[148] Saurabh Amin, Falk M Hante, and Alexandre M Bayen. Exponential stability of switched linear hyperbolic initial-boundary value problems. *IEEE Trans. Autom. Control*, 57(2):291–301, 2011.

[149] Hua-Cheng Zhou, Jian-Hua Chen, and George Weiss. Fast switching between infinite-dimensional linear systems. In *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*, pages 53–58. IEEE, 2017.

[150] Falk M Hante and Mario Sigalotti. Converse Lyapunov theorems for switched systems in banach and hilbert spaces. *SIAM J. Control Opt.*, 49(2):752–770, 2011.

[151] Saurabh Amin, Falk M Hante, and Alexandre M Bayen. Stability analysis of linear hyperbolic systems with switching parameters and boundary conditions. In *2008 47th IEEE Conference on Decision and Control*, pages 2081–2086. IEEE, 2008.

[152] Oswaldo LV Costa and Marcelo D Fragoso. Discrete-time LQ-optimal control problems for infinite markov jump parameter systems. *IEEE Trans. Autom. Control*, 40(12):2076–2088, 1995.

[153] Michael S Branicky. Multiple Lyapunov functions and other analysis tools for switched and hybrid systems. *IEEE Trans. Autom. Control*, 43(4):475–482, 1998.

[154] Daniel Liberzon and A Stephen Morse. Basic problems in stability and design of switched systems. *IEEE Control Syst.*, 19(5):59–70, 1999.

[155] Georg Böker and Jan Lunze. Stability and performance of switching Kalman filters. *Int. J. Control*, 75(16-17):1269–1281, 2002.

[156] Prashant Mhaskar, Nael H El-Farra, and Panagiotis D Christofides. Predictive control of switched nonlinear systems with scheduled mode transitions. *IEEE Trans. Autom. Control*, 50(11):1670–1680, 2005.

[157] Yafeng Guo and Biao Huang. Moving horizon estimation for switching nonlinear systems. *Automatica*, 49(11):3270–3281, 2013.

[158] Van Tri Nguyen, Didier Georges, and Gildas Besançon. Calculus of variations approach for state and parameter estimation in switched 1D hyperbolic PDEs. *Optim. Control Appl. Methods*, 39(3):1182–1201, 2018.

[159] Leobardo Camacho-Solorio and Miroslav Krstic. Boundary observers for the expected value of a randomly switching reaction-diffusion PDE. In *2018 IEEE Conference on Decision and Control (CDC)*, pages 1335–1340. IEEE, 2018.

[160] Qing Sun, Cheng-Chew Lim, Peng Shi, and Fei Liu. Design and stability of moving horizon estimator for markov jump linear systems. *IEEE Trans. Autom. Control*, 64(3):1109–1124, 2018.

[161] Angelo Alessandri, Marco Baglietto, and Giorgio Battistelli. Receding-horizon estimation for switching discrete-time linear systems. *IEEE Trans. Autom. Control*, 50(11):1736–1748, 2005.

[162] Christopher V Rao, James B Rawlings, and David Q Mayne. Constrained state estimation for nonlinear discrete-time systems: stability and moving horizon approximations. *IEEE Trans. Autom. Control*, 48(2):246–258, 2003.

[163] Javier Andres Villegas, Hans Zwart, Yann Le Gorrec, and Bernhard Maschke. Exponential stability of a class of boundary control systems. *IEEE Trans. Autom. Control*, 54(1):142–147, 2009.

[164] George Weiss and Ruth F Curtain. Exponential stabilization of a Rayleigh beam using collocated control. *IEEE Trans. Autom. Control*, 53(3):643–654, 2008.

[165] Marius Tucsnak and George Weiss. Simultaneous exact controllability and some applications. *SIAM J. Control Opt.*, 38(5):1408–1427, 2000.

[166] Einar Hille and Ralph Saul Phillips. *Functional analysis and semi-groups*, volume 31. American Mathematical Soc., 1996.

[167] Erwin Kreyszig. *Introductory functional analysis with applications*, volume 1. wiley New York, 1978.

[168] Thibaut Besson, Abdoua Tchousso, and Cheng-Zhong Xu. Exponential stability of a class of hyperbolic PDE models from chemical engineering. In *Proceedings of the 45th IEEE Conference on Decision and Control*, pages 3974–3978. IEEE, 2006.

[169] Cheng-Zhong Xu and Gauthier Sallet. Exponential stability and transfer functions of processes governed by symmetric hyperbolic systems. *ESAIM: COCV*, 7:421–442, 2002.

[170] Hans Zwart, Yann Le Gorrec, Bernhard Maschke, and Javier Villegas. Well-posedness and regularity of hyperbolic boundary control systems on a one-dimensional spatial domain. *ESAIM: COCV*, 16(4):1077–1093, 2010.

[171] Cheng-Zhong Xu. Exact observability and exponential stability of infinite-dimensional bilinear systems. *Math. Control Signals Systems*, 9(1):73–93, 1996.

[172] Bopeng Rao. Optimal energy decay rate in a damped Rayleigh beam. *Contemporary Mathematics*, 209:211–230, 1997.

[173] Jun-Min Wang and Siu-Pang Yung. Stability of a nonuniform Rayleigh beam with indefinite damping. *Syst. Control. Lett.*, 55(10):863–870, 2006.

[174] Bao-Zhu Guo, Jun-Min Wang, and Cui-Lian Zhou. On the dynamic behavior and stability of controlled connected Rayleigh beams under pointwise output feedback. *ESAIM: COCV*, 14(3):632–656, 2008.

[175] Said Hadd. Exact controllability of infinite dimensional systems persists under small perturbations. *J. Evol. Equ.*, 5(4):545–555, 2005.

[176] Klaus-Jochen Engel and Rainer Nagel. One-parameter semigroups for linear evolution equations. In *Semigroup forum*, volume 63, pages 278–280. Springer, 2001.

[177] Adilson José De Assis and Rubens Maciel Filho. Soft sensors development for on-line bioreactor state estimation. *Comput. Chem. Eng.*, 24(2-7):1099–1103, 2000.

[178] Bao Lin, Bodil Recke, Jørgen KH Knudsen, and Sten Bay Jørgensen. A systematic approach for soft sensor development. *Comput. Chem. Eng.*, 31(5-6):419–425, 2007.

[179] Petr Kadlec, Bogdan Gabrys, and Sibylle Strandt. Data-driven soft sensors in the process industry. *Comput. & Chem. Eng.*, 33(4):795–814, 2009.

[180] Andrew I Schein, Alexandrin Popescul, Lyle H Ungar, and David M Pennock. Methods and metrics for cold-start recommendations. In *Proc. 25th Annual Int. ACM SIGIR Conf. Res. and Dev. in Inf. Retrv*, pages 253–260, 2002.

[181] Jie Yu and S Joe Qin. Multimode process monitoring with Bayesian inference-based finite Gaussian mixture models. *AIChE J.*, 54(7):1811–1829, 2008.

[182] Sinno Jialin Pan and Qiang Yang. A survey on transfer learning. *IEEE Trans. Knowl. Data Eng.*, 22(10):1345–1359, 2009.

[183] Karl Weiss, Taghi M Khoshgoftaar, and DingDing Wang. A survey of transfer learning. *J. Big data*, 3(1):9, 2016.

[184] Junyao Xie, Laibin Zhang, Lixiang Duan, and Jinjiang Wang. On cross-domain feature fusion in gearbox fault diagnosis under various operating conditions based on transfer component analysis. In *2016 IEEE Int. Conf. on Prognostics and Health Manag. (ICPHM)*, pages 1–6. IEEE, 2016.

[185] Liang Guo, Yaguo Lei, Saibo Xing, Tao Yan, and Naipeng Li. Deep convolutional transfer learning network: A new method for intelligent fault diagnosis of machines with unlabeled data. *IEEE Trans. Ind. Electron.*, 66(9):7316–7325, 2018.

[186] Xiang Li, Wei Zhang, Nan-Xi Xu, and Qian Ding. Deep learning-based machinery fault diagnostics with domain adaptation across sensors at different places. *IEEE Trans. Ind. Electron.*, 2019.

[187] Ruqiang Yan, Fei Shen, Chuang Sun, and Xuefeng Chen. Knowledge transfer for rotary machine fault diagnosis. *IEEE Sensors J.*, 2019.

[188] Yi Liu, Chao Yang, Kaixin Liu, Bocheng Chen, and Yuan Yao. Domain adaptation transfer learning soft sensor for product quality prediction. *Chemometr. Intell. Lab.*, 192:103813, 2019.

[189] Liangjun Feng and Chunhui Zhao. Fault description based attribute transfer for zero-sample industrial fault diagnosis. *IEEE Trans. Ind. Informat.*, 2020.

[190] Yoshua Bengio, Aaron Courville, and Pascal Vincent. Representation learning: A review and new perspectives. *IEEE Trans. Pattern Anal. Mach. Intell.*, 35(8):1798–1828, 2013.

[191] Svante Wold, Kim Esbensen, and Paul Geladi. Principal component analysis. *Chemometr. Intell. Lab.*, 2(1-3):37–52, 1987.

[192] Paul Geladi and Bruce R Kowalski. Partial least-squares regression: a tutorial. *Analytica chimica acta*, 185:1–17, 1986.

[193] Wouter Favoreel, Bart De Moor, and Peter Van Overschee. Subspace state space system identification for industrial processes. *J. Process Control*, 10(2-3):149–155, 2000.

[194] Laurenz Wiskott and Terrence J Sejnowski. Slow feature analysis: Unsupervised learning of invariances. *Neural computation*, 14(4):715–770, 2002.

[195] Richard Turner and Maneesh Sahani. A maximum-likelihood interpretation for slow feature analysis. *Neural computation*, 19(4):1022–1038, 2007.

[196] Chao Shang, Biao Huang, Fan Yang, and Dexian Huang. Probabilistic slow feature analysis-based representation learning from massive process data for soft sensor modeling. *AIChE J.*, 61(12):4126–4139, 2015.

[197] Lei Fan, Hariprasad Kodamana, and Biao Huang. Semi-supervised dynamic latent variable modeling: I/O probabilistic slow feature analysis approach. *AIChE J.*, 65(3):964–979, 2019.

[198] Yanjun Ma and Biao Huang. Bayesian learning for dynamic feature extraction with application in soft sensing. *IEEE Trans. Ind. Electron.*, 64(9):7171–7180, 2017.

[199] Liang Ge, Jing Gao, and Aidong Zhang. OMS-TL: A framework of online multiple source transfer learning. In *Proc. 22nd ACM Int. Conf. Inf. & Knowl. Manag.*, pages 2423–2428, 2013.

[200] Qingyao Wu, Hanrui Wu, Xiaoming Zhou, Mingkui Tan, Yonghui Xu, Yuguang Yan, and Tianyong Hao. Online transfer learning with multiple homogeneous or heterogeneous sources. *IEEE Trans. Knowl. Data Eng.*, 29(7):1494–1507, 2017.

[201] Priyank Jaini, Zhitang Chen, Pablo Carbajal, Edith Law, Laura Middleton, Kayla Regan, Mike Schaekermann, George Trimponias, James Tung, and Pascal Poupart. Online Bayesian transfer learning for sequential data modeling. 2016.

[202] Alireza Karbalayghareh, Xiaoning Qian, and Edward R Dougherty. Optimal Bayesian transfer regression. *IEEE Signal Process. Lett.*, 25(11):1655–1659, 2018.

[203] Christopher M Bishop. *Pattern recognition and machine learning*. Springer, New York, USA, 2006.

[204] Qinliang Su, Xuejun Liao, Changyou Chen, and Lawrence Carin. Nonlinear statistical learning with truncated Gaussian graphical models. In *Int. Conf. on Mach. Learn.*, pages 1948–1957, 2016.

[205] Dirk Ostwald, Evgeniya Kirilina, Ludger Starke, and Felix Blankenburg. A tutorial on variational bayes for latent linear stochastic time-series models. *J. Math. Psychol.*, 60:1–19, 2014.

[206] Matthew James Beal et al. *Variational algorithms for approximate Bayesian inference*. University of London Press, London, UK, 2003.

[207] Yanjun Ma, Shunyi Zhao, and Biao Huang. Multiple-model state estimation based on variational Bayesian inference. *IEEE Trans. Autom. Control*, 64(4):1679–1685, 2018.

[208] Michael E Tipping and Christopher M Bishop. Probabilistic principal component analysis. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 61(3):611–622, 1999.

[209] Sinno Jialin Pan, Ivor W Tsang, James T Kwok, and Qiang Yang. Domain adaptation via transfer component analysis. *IEEE Trans. Neural Netw.*, 22(2):199–210, 2010.

[210] Yonghui Xu, Sinno Jialin Pan, Hui Xiong, Qingyao Wu, Ronghua Luo, Huaqing Min, and Hengjie Song. A unified framework for metric transfer learning. *IEEE Trans. Knowl. Data Eng.*, 29(6):1158–1171, 2017.

[211] Changan Yi, Yonghui Xu, Han Yu, Yuguang Yan, Yang Liu, et al. Multi-component transfer metric learning for handling unrelated source domain samples. *Knowledge-Based Systems*, pages 106–132, 2020.

[212] Yaroslav Ganin and Victor Lempitsky. Unsupervised domain adaptation by backpropagation. In *International conference on machine learning*, pages 1180–1189. PMLR, 2015.

[213] Yaroslav Ganin, Evgeniya Ustinova, Hana Ajakan, Pascal Germain, Hugo Larochelle, François Laviolette, Mario Marchand, and Victor Lempitsky. Domain-adversarial training of neural networks. *J. Mach. Learn. Res.*, 17(1):2096–2030, 2016.

[214] James J Downs and Ernest F Vogel. A plant-wide industrial process control problem. *Comput. Chem. Eng.*, 17(3):245–255, 1993.

[215] Shen Yin, Steven X Ding, Adel Haghani, Haiyang Hao, and Ping Zhang. A comparison study of basic data-driven fault diagnosis and process monitoring methods on the benchmark Tennessee Eastman process. *J. Process Control*, 22(9):1567–1581, 2012.

[216] Li Wang, Huaiping Jin, Xiangguang Chen, Jiayu Dai, Kai Yang, and Dongxiang Zhang. Soft sensor development based on the hierarchical ensemble of Gaussian process regression models for nonlinear and non-Gaussian chemical processes. *Ind. Eng. Chem. Res.*, 55(28):7704–7719, 2016.

[217] Lawrence Ricker. Decentralized control of the Tennessee Eastman challenge process. *J. Process Control*, 6(4):205–221, 1996.

[218] Andreas Bathelt, N Lawrence Ricker, and Mohieddine Jelali. Revision of the Tennessee Eastman process model. *IFAC-PapersOnLine*, 48(8):309–314, 2015.

[219] Ben C Juricek, Dale E Seborg, and Wallace E Larimore. Identification of the Tennessee Eastman challenge process with subspace methods. *Control Eng. Pract.*, 9(12):1337–1351, 2001.

[220] L Billmann and Rolf Isermann. Leak detection methods for pipelines. *Automatica*, 23(3):381–385, 1987.

[221] Pal-Stefan Murvay and Ioan Silea. A survey on gas leak detection and localization techniques. *J. Loss Prev. Process Ind.*, 25(6):966–973, 2012.

[222] Wei Liang, Laibin Zhang, Qingqing Xu, and Chunying Yan. Gas pipeline leakage detection based on acoustic technology. *Eng. Fail. Anal.*, 31:1–7, 2013.

[223] Xiaoben Liu, Hong Zhang, Yinshan Han, Mengying Xia, and Wei Zheng. A semi-empirical model for peak strain prediction of buried x80 steel pipelines under compression and bending at strike-slip fault crossings. *J. Nat. Gas Sci. Eng.*, 32:465–475, 2016.

[224] Raido Puust, Zoran Kapelan, Dragan Savic, and Tiit Koppel. A review of methods for leakage management in pipe networks. *Urban Water J.*, 7(1):25–45, 2010.

[225] Andrew F Colombo, Pedro Lee, and Bryan W Karney. A selective literature review of transient-based leak detection methods. *J. Hydro-Environ. Res.*, 2(4):212–227, 2009.

[226] Peter Black. A review of pipeline leak detection technology. In *Pipeline systems*, pages 287–298. Springer, 1992.

[227] Jun Zhang. Designing a cost-effective and reliable pipeline leak-detection system. *Pipes and Pipelines Int.*, 42(1):20–26, 1997.

[228] Chet Sandberg, Jim Holmes, Ken McCoy, and HEINRICH Koppitsch. The application of a continuous leak detection system to pipelines and associated equipment. *IEEE Trans. Ind Appl.*, 25(5):906–909, 1989.

[229] Dae Shik Kim, Sungho Shin, Go Bong Choi, Kwang Ho Jang, Jung Chul Suh, and Jong Min Lee. Diagnosis of partial blockage in water pipeline using support vector machine with fault-characteristic peaks in frequency domain. *Can. J. Civ. Eng.*, 44(9):707–714, 2017.

[230] Jiheon Kang, Youn-Jong Park, Jaeho Lee, Soo-Hyun Wang, and Doo-Seop Eom. Novel leakage detection by ensemble cnn-svm and graph-based localization in water distribution systems. *IEEE Trans. Ind. Electron.*, 65(5):4279–4289, 2018.

[231] Jinjiang Wang, Junyao Xie, Rui Zhao, Kezhi Mao, and Laibin Zhang. A new probabilistic kernel factor analysis for multisensory data fusion: Application to tool condition monitoring. *IEEE Trans. Instrum. Meas.*, 65(11):2527–2537, 2016.

[232] Salvatore Belsito, Paolo Lombardi, Paolo Andreussi, and Sanjoy Banerjee. Leak detection in liquefied gas pipelines by artificial neural networks. *AIChE J.*, 44(12):2675–2688, 1998.

[233] Marco Ferrante, Bruno Brunone, and Silvia Meniconi. Wavelets for the analysis of transient pressure signals for leak detection. *J. Hydraul. Eng.*, 133(11):1274–1282, 2007.

[234] Prashanth Reddy, Shankar Narasimhan, Murty Bhallamudi, and Saikot Bairagi. Leak detection in gas pipeline networks using an efficient state estimator. part-I: Theory and simulations. *Comput. Chem. Eng.*, 35(4):651–661, 2011a.

[235] Marcos Guillen, Jean F Dulhoste, Gildas Besancon, Ignacio Rubio Scola, Rafael Santos, and Didier Georges. Leak detection and location based on improved pipe model and nonlinear observer. In *Control Conference (ECC), 2014 European*, pages 958–963. IEEE, 2014.

[236] Ole Morten Aamo. Leak detection, size estimation and localization in pipe flows. *IEEE Trans. Autom. Control*, 61(1):246–251, 2016.

[237] Shouxi Wang, John Joseph Carroll, et al. Leak detection for gas and liquid pipelines by transient modeling. In *International Oil & Gas Conference and Exhibition in China*, pages 1–9. Society of Petroleum Engineers, 2006.

[238] Shin Je Lee, Jong Min Lee, Jingbo Wu, Frank Allgöwer, Jung Chul Suh, and Gibaek Lee. Consensus algorithm-based approach to fundamental modeling of water pipe networks. *AIChE J.*, 63(9):3860–3870, 2017.

[239] Jarmo Malinen. Tustin's method for final state approximation of conservative dynamical systems. *IFAC Proceedings Volumes*, 44(1):4564–4569, 2011.

[240] Victor L Streeter, E Benjamin Wylie, and Keith W Bedford. *Fluid mechanics. WCB.* McGraw-Hill, 1998.

[241] Pál Szilas. *Production and transport of oil and gas*, volume 3. Elsevier, 2010.

[242] Sašo Blažič, Drago Matko, and Gerhard Geiger. Simple model of a multi-batch driven pipeline. *Math. Comput. Simul.*, 64(6):617–630, 2004.

[243] Johan AK Suykens and Joos Vandewalle. Least squares support vector machine classifiers. *Neural Process. Lett.*, 9(3):293–300, 1999.

[244] Junshui Ma and Simon Perkins. Time-series novelty detection using one-class support vector machines. In *Neural Networks, 2003. Proceedings of the International Joint Conference on*, volume 3, pages 1741–1745. IEEE, 2003.

[245] Chih-Chung Chang and Chih-Jen Lin. Libsvm: a library for support vector machines. *ACM Trans. Intell. Syst. Technol.*, 2(3):27, 2011.

[246] Guido F Smits and Elizabeth M Jordaan. Improved svm regression using mixtures of kernels. In *Neural Networks, 2002. IJCNN'02. Proceedings of the 2002 International Joint Conference on*, volume 3, pages 2785–2790. IEEE, 2002.

[247] Bruce A Francis and Walter Murray Wonham. The internal model principle of control theory. *Automatica*, 12(5):457–465, 1976.

[248] Edward Davison. The robust control of a servomechanism problem for linear time-invariant multivariable systems. *IEEE Trans. Autom. Control*, 21(1):25–34, 1976.

[249] Cn Johnson. Accomodation of external disturbances in linear regulator and servomechanism problems. *IEEE Trans. Autom. Control*, 16(6):635–644, 1971.

[250] Gunnar Bengtsson. Output regulation and internal models–a frequency domain approach. *Automatica*, 13(4):333–345, 1977.

[251] Alberto Isidori and Christopher I Byrnes. Output regulation of nonlinear systems. *IEEE Trans. Autom. Control*, 35(2):131–140, 1990.

[252] Christopher I Byrnes, F Delli Priscoli, Alberto Isidori, and W Kang. Structurally stable output regulation of nonlinear systems. *Automatica*, 33(3):369–385, 1997.

[253] Andrea Serrani, Alberto Isidori, and Lorenzo Marconi. Semi-global nonlinear output regulation with adaptive internal model. *IEEE Trans. Autom. Control*, 46(8):1178–1194, 2001.

[254] Emilia Fridman. Output regulation of nonlinear systems with delay. *Syst. Control. Lett.*, 50(2):81–93, 2003.

[255] Pradeep Kumar Gillella, Xingyong Song, and Zongxuan Sun. Time-varying internal model-based control of a camless engine valve actuation system. *IEEE Trans. Control Syst. Technol.*, 22(4):1498–1510, 2014.

[256] Zeng Qiu, Mario Santillo, Mrdjan Jankovic, and Jing Sun. Composite adaptive internal model control and its application to boost pressure control of a turbocharged gasoline engine. *IEEE Trans. Control Syst. Technol.*, 23(6):2306–2315, 2015.

[257] Abraham Castellanos Silva, Ioan Doré Landau, and Petros Ioannou. Robust direct adaptive regulation of unknown disturbances in the vicinity of low-damped complex zeros-application to AVC. *IEEE Trans. Control Syst. Technol.*, 24(2):733–740, 2016.

[258] Simone De Marco, Lorenzo Marconi, Robert Mahony, and Tarek Hamel. Output regulation for systems on matrix lie-groups. *Automatica*, 87:8–16, 2018.

[259] Seppo Pohjolainen. Robust multivariable PI-controller for infinite dimensional systems. *IEEE Trans. Autom. Control*, 27(1):17–30, 1982.

[260] Toshihiro Kobayashi. Regulator problem for infinite-dimensional systems. *Syst. Control. Lett.*, 3(1):31–39, 1983.

[261] Osvaldo Maria Grasselli, Sauro Longhi, Antonio Tornambè, and Paolo Valigi. Robust output regulation and tracking for linear periodic systems under structured uncertainties. *Automatica*, 32(7):1015–1019, 1996.

[262] Timo Hamalainen and Seppo Pohjolainen. A finite-dimensional robust controller for systems in the CD-algebra. *IEEE Trans. Autom. Control*, 45(3):421–431, 2000.

[263] Ole Morten Aamo. Disturbance rejection in 2 x 2 linear hyperbolic systems. *IEEE Trans. Autom. Control*, 58(5):1095–1106, 2013.

[264] Henrik Anfinsen and Ole Morten Aamo. Disturbance rejection in the interior domain of linear 2 x 2 hyperbolic systems. *IEEE Trans. Autom. Control*, 60(1):186–191, 2015.

[265] Agus Hasan. Disturbance attenuation of n+ 1 coupled hyperbolic PDEs. In *53rd IEEE Conference on Decision and Control*, pages 2058–2064. IEEE, 2014.

[266] Henrik Anfinsen and Ole Morten Aamo. Disturbance rejection in general heterodirectional 1-d linear hyperbolic systems using collocated sensing and control. *Automatica*, 76:230–242, 2017.

[267] Joachim Deutscher. Output regulation for general linear heterodirectional hyperbolic systems with spatially-varying coefficients. *Automatica*, 85:34–42, 2017.

[268] Xiaodong Xu and Stevan Dubljevic. Output regulation for a class of linear boundary controlled first-order hyperbolic PIDE systems. *Automatica*, 85:43–52, 2017.

[269] Hua-Cheng Zhou and George Weiss. Solving the regulator problem for a 1-D Schrödinger equation via backstepping. *IFAC-PapersOnLine*, 50(1):4516–4521, 2017.

[270] Joachim Deutscher and Simon Kerschbaum. Output regulation for coupled linear parabolic PIDEs. *Automatica*, 100:360–370, 2019.

[271] Wen Kang and Bao-Zhu Guo. Stabilisation of unstable cascaded heat partial differential equation system subject to boundary disturbance. *IET Control. Theory Appl.*, 10(9):1027–1039, 2016.

[272] Jian-Jun Gu and Jun-Min Wang. Backstepping state feedback regulator design for an unstable reaction-diffusion PDE with long time delay. *J. Dyn. Control Syst.*, 24(4):563–576, 2018.

[273] Joachim Deutscher and Jakob Gabriel. Robust state feedback regulator design for general linear heterodirectional hyperbolic systems. *IEEE Trans. Autom. Control*, 2018.

[274] Ravi Mantri, Ali Saberi, Zongli Lin, and Anton A Stoorvogel. Output regulation for linear discrete-time systems subject to input saturation. *Int. J. Robust Nonlinear Control: IFAC-Affiliated Journal*, 7(11):1003–1021, 1997.

[275] Gang Feng and Tiejun Zhang. Output regulation of discrete-time piecewise-linear systems with application to controlling chaos. *IEEE Trans. Circuits Syst., II, Exp. Briefs*, 53(4):249–253, 2006.

[276] Jie Huang. The cooperative output regulation problem of discrete-time linear multi-agent systems by the adaptive distributed observer. *IEEE Trans. Autom. Control*, 62(4):1979–1984, 2017.

[277] Zhenqing Ke, Hartmut Logemann, and Richard Rebarber. A sampled-data servomechanism for stable well-posed systems. *IEEE Trans. Autom. Control*, 54(5):1123–1128, 2009.

[278] Masashi Wakaiki and Hideki Sano. Sampled-data output regulation of well-posed infinite-dimensional systems. *arXiv preprint arXiv:1901.07254*, 2019.

[279] Lassi Paunonen. Robust output regulation for continuous-time periodic systems. *IEEE Trans. Autom. Control*, 62(9):4363–4375, 2017.

[280] George Weiss. Regular linear systems with feedback. *Math. Control Signals Systems*, 7(1):23–57, 1994.

[281] Dietmar Salamon. *Control and observation of neutral systems*. Number 91. Pitman Advanced Publishing Program, 1984.

[282] Ruth F Curtain and Alejandro Rodriguez. Necessary and sufficient conditions for J-spectral factorizations with a J-lossless property for infinite-dimensional systems in continuous and discrete time. *Linear algebra and its applications*, 203:327–358, 1994.

[283] G Hewer. An iterative technique for the computation of the steady state gains for the discrete optimal regulator. *IEEE Trans. Autom. Control*, 16(4):382–384, 1971.

[284] David Kleinman. Stabilizing a discrete, constant, linear system with application to iterative methods for solving the Riccati equation. *IEEE Trans. Autom. Control*, 19(3):252–254, 1974.

[285] John Sevier Gibson and I Gary Rosen. Numerical approximation for the infinite-dimensional discrete-time optimal linear-quadratic regulator problem. *SIAM J. Control Optim.*, 26(2):428–451, 1988.

[286] James William Thomas. *Numerical partial differential equations: finite difference methods*, volume 22. Springer Science & Business Media, 2013.

[287] Yoshiki Kuramoto and Toshio Tsuzuki. On the formation of dissipative structures in reaction-diffusion systems: reductive perturbation approach. *Prog. Theor. Phys.*, 54(3):687–699, 1975.

[288] Daniel Michelson and Gregory Sivashinsky. Nonlinear analysis of hydrodynamic instability in laminar flames: II. numerical experiments. *Acta Astronautica*, 4(11-12):1207–1221, 1977.

[289] Gregory Sivashinsky. Nonlinear analysis of hydrodynamic instability in laminar flames: I. derivation of basic equations. In *Dynamics of Curved Fronts*, pages 459–488. Elsevier, 1988.

[290] Eberhard Berger. Suppression of vortex shedding and turbulence behind oscillating cylinders. *Phys. Fluids*, 10(9):S191–S193, 1967.

[291] Kimon Roussopoulos. Feedback control of vortex shedding at low reynolds numbers. *J. Fluid Mech.*, 248:267–296, 1993.

[292] XY Huang. Feedback control of vortex shedding from a circular cylinder. *Exp. Fluids*, 20(3):218–224, 1996.

[293] Max D Gunzburger and Hyung-Chun Lee. Feedback control of karman vortex shedding. *J. Appl. Mech.*, 63(3):828–835, 1996.

[294] JE Ffowcs Williams and BC Zhao. The active control of vortex shedding. *J. Fluids Struct.*, 3(2):115–122, 1989.

[295] Ching-Jen Chen and Hamn-Ching Chen. Finite analytic numerical method for unsteady two-dimensional navier-stokes equations. *J. Comput. Phys.*, 53(2):209–226, 1984.

[296] Ronald D Henderson. Details of the drag curve near the onset of vortex shedding. *Phys. Fluids*, 7(9):2102–2104, 1995.

[297] Milan Milovanovic, Marek Grayer, Joris Michielsen, and Ole Morten Aamo. Model-based stabilization of vortex shedding with cfd verification. In *Proceedings of the 48h IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference*, pages 8252–8257. IEEE, 2009.

[298] Patrick Huerre and Peter A Monkewitz. Local and global instabilities in spatially developing flows. *Annu. Rev. Fluid Mech.*, 22(1):473–537, 1990.

[299] Kimon Roussopoulos and Peter A Monkewitz. Nonlinear modelling of vortex shedding control in cylinder wakes. *Phys. D: Nonlinear Phenom.*, 97(1-3):264–273, 1996.

[300] Kelly Cohen, Stefan Siegel, Thomas McLaughlin, Eric Gillies, and James Myatt. Closed-loop approaches to control of a wake flow modeled by the Ginzburg–Landau equation. *Computers & Fluids*, 34(8):927–949, 2005.

[301] Ole Morten Aamo, Andrey Smyshlyaev, Miroslav Krstic, and Bjarne Anton Foss. Stabilization of a Ginzburg-Landau model of vortex shedding by output feedback boundary control. In *2004 43rd IEEE Conference on Decision and Control (CDC)(IEEE Cat. No. 04CH37601)*, volume 3, pages 2409–2416. IEEE, 2004.

[302] Milan Milovanovic and Ole Morten Aamo. Attenuation of vortex shedding in cfd simulations by model-based output feedback control. In *49th IEEE Conference on Decision and Control (CDC)*, pages 2979–2984. IEEE, 2010.

[303] Milan Milovanovic and Ole Morten Aamo. Attenuation of vortex shedding by model-based output feedback control. *IEEE Trans. Control Syst. Technol.*, 21(3):617–625, 2012.

[304] Milan Milovanovic and Ole Morten Aamo. Stabilization of 3d Ginzburg-Landau equation by model-based output feedback control. *IFAC Proceedings Volumes*, 44(1):14435–14439, 2011.

[305] Philippe Poncet. Topological aspects of three-dimensional wakes behind rotary oscillating cylinders. *J. Fluid Mech.*, 517:27–53, 2004.

[306] Philippe Poncet, Georges-Henri Cottet, and Petros Koumoutsakos. Control of three-dimensional wakes using evolution strategies. *Comptes Rendus Mecanique*, 333(1):65–77, 2005.

[307] Panagiotis D Christofides and Antonios Armaou. Global stabilization of the Kuramoto–Sivashinsky equation via distributed output feedback control. *Syst. Control. Lett.*, 39(4):283–294, 2000.

[308] Antonios Armaou and Panagiotis D Christofides. Feedback control of the Kuramoto–Sivashinsky equation. *Phys. D: Nonlinear Phenom.*, 137(1-2):49–61, 2000.

[309] Wei-Jiu Liu and Miroslav Krstić. Stability enhancement by boundary control in the Kuramoto–Sivashinsky equation. *Nonlinear Anal. Theory Methods Appl.*, 43(4):485–507, 2001.

[310] Stevan Dubljevic. Model predictive control of Kuramoto–Sivashinsky equation with state and input constraints. *Chem. Eng. Sci.*, 65(15):4388–4396, 2010.

[311] Yu Yang and Stevan Dubljevic. Boundary model predictive control of thin film thickness modelled by the Kuramoto–Sivashinsky equation with input and state constraints. *J. Process Control*, 23(9):1362–1379, 2013.

[312] Christopher I Byrnes, David S Gilliam, C Hu, and Victor I Shubov. Zero dynamics boundary control for regulation of the Kuramoto–Sivashinsky equation. *Math. Comput. Model.*, 52(5-6):875–891, 2010.

[313] Wen Kang and Emilia Fridman. Distributed sampled-data control of Kuramoto–Sivashinsky equation. *Automatica*, 95:514–524, 2018.

[314] Patricio Guzmán, Swann Marx, and Eduardo Cerpa. Stabilization of the linear Kuramoto-Sivashinsky equation with a delayed boundary control. In *in Proc. 3rd IFAC CPDE/CDPS Joint Workshop*. IFAC, 2019.

[315] Damir Z Arov and Ivan P Gavrilyuk. A method for solving initial value problems for linear differential equations in Hilbert space based on the Cayley transform. *Numer. Funct. Anal. Optim.*, 14(5-6):459–473, 1993.

[316] Ivan P Gavrilyuk and Vladimir L Makarov. The Cayley transform and the solution of an initial value problem for a first order differential equation with an unbounded operator coefficient in Hilbert space. *Numer. Funct. Anal. Optim.*, 15(5-6):583–598, 1994.

[317] Damir Zyamovich Arov, Ivan P Gavrilyuk, and Vladimir L Makarov. Representation and approximation of solutions of initial value problems for differential equations in Hilbert space based on the Cayley transform. *Pitman Research Notes in Mathematics Series*, pages 40–40, 1995.

[318] Ivan P Gavrilyuk and Vladimir L Makarov. Representation and approximation of the solution of an initial value problem for a first order differential equation in banach spaces. *Zeitschrift für Analysis und ihre Anwendungen*, 15(2):495–527, 1996.

[319] Ivan P Gavrilyuk and Roderick Melnik. Constructive approximations of the convection-diffusion-absorption equation based on the Cayley transform technique. In *Computational Mechanics: New Trends and Applications (Proceed-*

ings of the Fourth World Congress on Computational Mechanics), pages 1–14. CIMNE/IACM, 1998.

[320] Ivan P Gavrilyuk and Vladimir L Makarov. Explicit and approximate solutions of second-order evolution differential equations in Hilbert space. *Numer. Methods Partial Differ. Equ.*, 15(1):111–131, 1999.

[321] Ivan P Gavrilyuk, Volodymyr L Makarov, and Nataliya V Mayko. Weighted estimates of the Cayley transform method for abstract differential equations. *Comput. Methods Appl. Math.*, 1(ahead-of-print), 2020.

[322] Nataliya Mayko. Super-exponential rate of convergence of the Cayley transform method for an abstract differential equation. *Cybern. Syst. Anal.*, pages 1–12, 2020.

[323] Abderrahim Azouani, Eric Olson, and Edriss S Titi. Continuous data assimilation using general interpolant observables. *J. Nonlinear Sci.*, 24(2):277–304, 2014.

[324] Evelyn Lunasin and Edriss S Titi. Finite determining parameters feedback control for distributed nonlinear dissipative systems-a computational study. *Evol. Equ. Control Theory*, 6(4):535–557, 2017.

[325] Kaijun Yang, Jun Zheng, and Guchuan Zhu. Asymptotic output tracking for a class of semilinear parabolic equations: A semianalytical approach. *Int. J. Robust Nonlinear Control*, 29(8):2471–2493, 2019.

[326] Jie Huang. *Nonlinear output regulation: theory and applications*, volume 8. SIAM, 2004.

[327] Liang-Heng Chen and Hsueh-Chia Chang. Nonlinear waves on liquid film surfaces-II. Bifurcation analyses of the long-wave equation. *Chem. Eng. Sci.*, 41(10):2477–2486, 1986.

[328] D Marchesin and PJ Paes-Leme. Shocks in gas pipelines. *SIAM J. Sci. Statist. Comput.*, 4(1):105–116, 1983.

[329] Rolf Isermann. Process fault detection based on modeling and estimation methods: A survey. *Automatica*, 20(4):387–404, 1984.

[330] Andrzej Osiadacz. *Simulation and analysis of gas networks*. Gulf Publishing Company, Houston, TX, 1987.

[331] Thang Van Pham, Didier Georges, and Gildas Besancon. Predictive control with guaranteed stability for water hammer equations. *IEEE Trans. Autom. Control*, 59(2):465–470, 2014.

[332] Martin Gugat and Markus Dick. Time-delayed boundary feedback stabilization of the isothermal Euler equations with friction. *Math. Control Relat. Fields*, 1(4):469–491, 2011.

[333] Cristina Verde. Multi-leak detection and isolation in fluid pipelines. *Control Eng. Pract.*, 9(6):673–682, 2001.

[334] Mapundi K Banda, Michael Herty, and Axel Klar. Coupling conditions for gas networks governed by the isothermal Euler equations. *NHM*, 1(2):295–314, 2006.

[335] Markus Dick, Martin Gugat, and Günter Leugering. Classical solutions and feedback stabilization for the gas flow in a sequence of pipes. *Netw. Heterog. Media.*, 5(4):691–709, 2010.

[336] Martin Gugat and Michaël Herty. Existence of classical solutions and feedback stabilization for the flow in gas networks. *ESAIM: Control Optim. Calc. Var.*, 17(1):28–51, 2011.

[337] Huai-Ning Wu, Hong-Du Wang, and Lei Guo. Finite dimensional disturbance observer based control for nonlinear parabolic PDE systems via output feedback. *J. Process Control*, 48:25–40, 2016.

[338] Wei Wu and Ching-Tien Liou. Output regulation of two-time-scale hyperbolic PDE systems. *J. Process Control*, 11(6):637–647, 2001.

[339] Joachim Deutscher, Nicole Gehring, and Richard Kern. Output feedback control of general linear heterodirectional hyperbolic ODE–PDE–ODE systems. *Automatica*, 95:472–480, 2018.

[340] Michael Herty, Jan Mohring, and V Sachers. A new model for gas flow in pipe networks. *Math. Methods Appl. Sci.*, 33(7):845–855, 2010.

[341] Andrzej J Osiadacz, Maciej Chaczykowski, et al. Comparison of isothermal and and non-isothermal transient models. In *PSIG Annual Meeting*. Pipeline Simulation Interest Group, 1998.

[342] Alberto Bressan, Sunčica Čanić, Mauro Garavello, Michael Herty, and Benedetto Piccoli. Flows on networks: recent results and perspectives. *EMS Surv. Math. Sci.*, 1(1):47–111, 2014.

[343] Kun Zhao et al. On the isothermal compressible Euler equations with frictional damping. *Commun. Math. Anal.*, 9(2):77–97, 2010.

[344] Gui-Qiang Chen and James Glimm. Global solutions to the compressible Euler equations with geometrical structure. *Commun. Math. Phys.*, 180(1):153–193, 1996.

[345] Myoungjean Bae, Gui-Qiang Chen, and Mikhail Feldman. Regularity of solutions to regular shock reflection for potential flow. *Invent. Math.*, 175(3):505–543, 2009.

[346] Eitan Tadmor and Dongming Wei. On the global regularity of sub-critical Euler-poisson equations with pressure. *arXiv preprint math/0609250*, 2006.

[347] Martin Gugat, Michael Herty, Axel Klar, Günther Leugering, and Veronika Schleper. Well-posedness of networked hyperbolic systems of balance laws. In *Constrained optimization and optimal control for partial differential equations*, pages 123–146. Springer, 2012.

[348] Ning Hsing Chen. An explicit equation for friction factor in pipe. *Ind. Eng. Chem. Fundam.*, 18(3):296–297, 1979.

[349] Zongli Lin, Anton A Stoorvogel, and Ali Saberi. Output regulation for linear systems subject to input saturation. *Automatica*, 32(1):29–47, 1996.

[350] K Maciej Przyluski. The lyapunov equation and the problem of stability for linear bounded discrete-time systems in hilbert space. *Applied Mathematics and Optimization*, 6(1):97–112, 1980.

Appendix

The resolvent operator in Eq. (9.23) is shown as follows:

$$
\begin{aligned}
\mathcal{R}_{11}(s,\mathcal{A})(\cdot) &= \frac{-M_{11}(x,s)}{a^2 M_{11}(l,s)}\int_0^l M_{12}(l-\eta,s)(\cdot)d\eta + \int_0^x \frac{1}{a^2}M_{12}(x-\eta,s)(\cdot)d\eta \\
&= \frac{e^{-\frac{Qx}{2}}\left(\cosh(\frac{\mathrm{H}}{2a}x) - \frac{Qa}{\mathrm{H}}\sinh(\frac{\mathrm{H}}{2a}x)\right)}{\cosh(\frac{\mathrm{H}}{2a}l) - \frac{Qa}{\mathrm{H}}\sinh(\frac{\mathrm{H}}{2a}l)}\int_0^l \frac{2(F+s)}{\mathrm{H}a}e^{\frac{Q\eta}{2}}\sinh(\frac{\mathrm{H}}{2a}(l-\eta))(\cdot)d\eta \\
&\quad - \int_0^x \frac{2(F+s)}{\mathrm{H}a}e^{-\frac{Q(x-\eta)}{2}}\sinh(\frac{\mathrm{H}}{2a}(x-\eta))(\cdot)d\eta \\
\mathcal{R}_{12}(s,\mathcal{A})(\cdot) &= -\frac{M_{11}(x,s)}{M_{11}(l,s)}\int_0^l M_{11}(l-\eta,s)(\cdot)d\eta + \int_0^x M_{11}(x-\eta,s)(\cdot)d\eta \\
&= -\frac{e^{-\frac{Qx}{2}}\left(\cosh(\frac{\mathrm{H}}{2a}x) - \frac{Qa}{\mathrm{H}}\sinh(\frac{\mathrm{H}}{2a}x)\right)}{\cosh(\frac{\mathrm{H}}{2a}l) - \frac{Qa}{\mathrm{H}}\sinh(\frac{\mathrm{H}}{2a}l)} \\
&\quad \times \int_0^l e^{\frac{Q\eta}{2}}\left(\cosh(\frac{\mathrm{H}}{2a}(l-\eta)) - \frac{Qa}{\mathrm{H}}\sinh(\frac{\mathrm{H}}{2a}(l-\eta))\right)(\cdot)d\eta \\
&\quad + \int_0^x e^{-\frac{Q(x-\eta)}{2}}\left(\cosh(\frac{\mathrm{H}}{2a}(x-\eta)) - \frac{Qa}{\mathrm{H}}\sinh(\frac{\mathrm{H}}{2a}(x-\eta))\right)(\cdot)d\eta \\
\mathcal{R}_{21}(s,\mathcal{A})(\cdot) &= \frac{-M_{21}(x,s)}{a^2 M_{11}(l,s)}\int_0^l M_{12}(l-\eta,s)(\cdot)d\eta + \int_0^x \frac{1}{a^2}M_{22}(x-\eta,s)(\cdot)d\eta \\
&= \frac{-2se^{-\frac{Qx}{2}}\sinh(\frac{\mathrm{H}}{2a}x)}{\cosh(\frac{\mathrm{H}}{2a}l) - \frac{Qa}{\mathrm{H}}\sinh(\frac{\mathrm{H}}{2a}l)}\int_0^l \frac{2(F+s)}{\mathrm{H}^2a^2}e^{\frac{Q\eta}{2}}\sinh(\frac{\mathrm{H}}{2a}(l-\eta))(\cdot)d\eta \\
&\quad + \int_0^x \frac{1}{a^2}e^{-\frac{Q(x-\eta)}{2}}\left(\cosh(\frac{\mathrm{H}}{2a}(x-\eta)) + \frac{Qa}{\mathrm{H}}\sinh(\frac{\mathrm{H}}{2a}(x-\eta))\right)(\cdot)d\eta \\
\mathcal{R}_{22}(s,\mathcal{A})(\cdot) &= -\frac{M_{21}(x,s)}{M_{11}(l,s)}\int_0^l M_{11}(l-\eta,s)(\cdot)d\eta + \int_0^x M_{21}(x-\eta,s)(\cdot)d\eta \\
&= \frac{\frac{2s}{Ha}e^{-\frac{Qx}{2}}\sinh(\frac{\mathrm{H}}{2a}x)}{\cosh(\frac{\mathrm{H}}{2a}l) - \frac{Qa}{\mathrm{H}}\sinh(\frac{\mathrm{H}}{2a}l)} \\
&\quad \times \int_0^l e^{\frac{Q\eta}{2}}\left(\cosh(\frac{\mathrm{H}}{2a}(l-\eta)) - \frac{Qa}{\mathrm{H}}\sinh(\frac{\mathrm{H}}{2a}(l-\eta))\right)(\cdot)d\eta \\
&\quad - \int_0^x \frac{2s}{\mathrm{H}a}e^{-\frac{Q(x-\eta)}{2}}\sinh(\frac{\mathrm{H}}{2a}(x-\eta))(\cdot)d\eta
\end{aligned}
$$

where $H = \sqrt{Q^2a^2 + 4s^2 + 4Fs}$.

The resolvent operator in Eq. (9.32) is given as follows:

$$
\begin{aligned}
\mathfrak{R}_{11}(s,\mathcal{A}^e) :&= -\frac{M_{11}(x,s)}{M_{11}^2(l,s)+\gamma M_{12}(l,s)M_{21}(l,s)} \\
&\quad\times\int_0^l\left(\frac{1}{a^2}M_{11}(l,s)M_{12}(l-\eta,s)+\frac{\gamma}{a^2}M_{12}(l,s)M_{22}(l-\eta,s)\right)(\cdot)d\eta \\
&\quad+\int_0^x\frac{1}{a^2}M_{12}(x-\eta,s)(\cdot)d\eta \\
\mathfrak{R}_{12}(s,\mathcal{A}^e) :&= -\frac{M_{11}(x,s)}{M_{11}^2(l,s)+\gamma M_{12}(l,s)M_{21}(l,s)} \\
&\quad\times\int_0^l\left(M_{11}(l,s)M_{11}(l-\eta,s)+\gamma M_{12}(l,s)M_{21}(l-\eta,s)\right)(\cdot)d\eta \\
&\quad+\int_0^x M_{11}(x-\eta,s)(\cdot)d\eta \\
\mathfrak{R}_{13}(s,\mathcal{A}^e) :&= -\frac{M_{11}(x,s)\int_0^l\frac{1}{a^2}M_{12}(l-\eta,s)(\cdot)d\eta}{M_{11}^2(l,s)+\gamma M_{12}(l,s)M_{21}(l,s)} \\
\mathfrak{R}_{14}(s,\mathcal{A}^e) :&= -\frac{M_{11}(x,s)\int_0^l M_{11}(l-\eta,s)(\cdot)d\eta}{M_{11}^2(l,s)+\gamma M_{12}(l,s)M_{21}(l,s)} \\
\mathfrak{R}_{21}(s,\mathcal{A}^e) :&= -\frac{M_{21}(x,s)}{M_{11}^2(l,s)+\gamma M_{12}(l,s)M_{21}(l,s)} \\
&\quad\times\int_0^l\left(\frac{1}{a^2}M_{11}(l,s)M_{12}(l-\eta,s)+\frac{\gamma}{a^2}M_{12}(l,s)M_{22}(l-\eta,s)\right)(\cdot)d\eta \\
&\quad+\int_0^x\frac{1}{a^2}M_{22}(x-\eta,s)(\cdot)d\eta \\
\mathfrak{R}_{22}(s,\mathcal{A}^e) :&= -\frac{M_{21}(x,s)}{M_{11}^2(l,s)+\gamma M_{12}(l,s)M_{21}(l,s)} \\
&\quad\times\int_0^l\left(M_{11}(l,s)M_{11}(l-\eta,s)+\gamma M_{12}(l,s)M_{21}(l-\eta,s)\right)(\cdot)d\eta \\
&\quad+\int_0^x M_{21}(x-\eta,s)(\cdot)d\eta \\
\mathfrak{R}_{23}(s,\mathcal{A}^e) :&= -\frac{M_{21}(x,s)\int_0^l\frac{1}{a^2}M_{12}(l-\eta,s)(\cdot)d\eta}{M_{11}^2(l,s)+\gamma M_{12}(l,s)M_{21}(l,s)} \\
\mathfrak{R}_{24}(s,\mathcal{A}^e) :&= -\frac{M_{21}(x,s)\int_0^l M_{11}(l-\eta,s)(\cdot)d\eta}{M_{11}^2(l,s)+\gamma M_{12}(l,s)M_{21}(l,s)} \\
\mathfrak{R}_{31}(s,\mathcal{A}^e) :&= -\frac{M_{11}(x,s)M_{11}(l,s)+\gamma M_{12}(x,s)M_{21}(l,s)}{M_{11}^2(l,s)+\gamma M_{12}(l,s)M_{21}(l,s)} \\
&\quad\times\int_0^l\left[\frac{1}{a^2}M_{11}(l,s)M_{12}(l-\eta,s)+\frac{\gamma}{a^2}M_{12}(l,s)M_{22}(l-\eta,s)\right](\cdot)d\eta
\end{aligned}
$$

$$
\begin{aligned}
&+\int_0^l [\frac{1}{a^2} M_{11}(x,s) M_{12}(l-\eta,s) + \frac{\gamma}{a^2} M_{12}(x,s) M_{22}(l-\eta,s)](\cdot) d\eta \\
\mathfrak{R}_{32}(s,\mathcal{A}^e) :&= -\frac{M_{11}(x,s)M_{11}(l,s) + \gamma M_{12}(x,s)M_{21}(l,s)}{M_{11}^2(l,s) + \gamma M_{12}(l,s)M_{21}(l,s)} \\
&\times \int_0^l [M_{11}(l,s)M_{11}(l-\eta,s) + \gamma M_{12}(l,s)M_{21}(l-\eta,s)](\cdot) d\eta \\
&+\int_0^l [M_{11}(x,s)M_{11}(l-\eta,s) + \gamma M_{12}(x,s)M_{21}(l-\eta,s)](\cdot) d\eta \\
\mathfrak{R}_{33}(s,\mathcal{A}^e) :&= -\frac{M_{11}(x,s)M_{11}(l,s) + \gamma M_{12}(x,s)M_{21}(l,s)}{M_{11}^2(l,s) + \gamma M_{12}(l,s)M_{21}(l,s)} \\
&\times \int_0^l \frac{1}{a^2} M_{12}(l-\eta,s) d\eta + \int_0^x \frac{1}{a^2} M_{12}(x-\eta,s) d\eta \\
\mathfrak{R}_{34}(s,\mathcal{A}^e) :&= -\frac{M_{11}(x,s)M_{11}(l,s) + \gamma M_{12}(x,s)M_{21}(l,s)}{M_{11}^2(l,s) + \gamma M_{12}(l,s)M_{21}(l,s)} \\
&\times \int_0^l M_{11}(l-\eta,s) d\eta + \int_0^x M_{11}(x-\eta,s) d\eta \\
\mathfrak{R}_{41}(s,\mathcal{A}^e) :&= -\frac{M_{21}(x,s)M_{11}(l,s) + \gamma M_{22}(x,s)M_{21}(l,s)}{M_{11}^2(l,s) + \gamma M_{12}(l,s)M_{21}(l,s)} \\
&\times \int_0^l [\frac{1}{a^2} M_{11}(l,s) M_{12}(l-\eta,s) + \frac{\gamma}{a^2} M_{12}(l,s) M_{22}(l-\eta,s)](\cdot) d\eta \\
&+\int_0^l [\frac{1}{a^2} M_{21}(x,s) M_{12}(l-\eta,s) + \frac{\gamma}{a^2} M_{22}(x,s) M_{22}(l-\eta,s)](\cdot) d\eta \\
\mathfrak{R}_{42}(s,\mathcal{A}^e) :&= -\frac{M_{21}(x,s)M_{11}(l,s) + \gamma M_{22}(x,s)M_{21}(l,s)}{M_{11}^2(l,s) + \gamma M_{12}(l,s)M_{21}(l,s)} \\
&\times \int_0^l [M_{11}(l,s)M_{11}(l-\eta,s) + \gamma M_{12}(l,s)M_{21}(l-\eta,s)](\cdot) d\eta \\
&+\int_0^l [M_{21}(x,s)M_{11}(l-\eta,s) + \gamma M_{22}(x,s)M_{21}(l-\eta,s)](\cdot) d\eta \\
\mathfrak{R}_{43}(s,\mathcal{A}^e) :&= -\frac{M_{21}(x,s)M_{11}(l,s) + \gamma M_{22}(x,s)M_{21}(l,s)}{M_{11}^2(l,s) + \gamma M_{12}(l,s)M_{21}(l,s)} \int_0^l \frac{1}{a^2} M_{12}(l-\eta,s) d\eta \\
&+\int_0^x \frac{1}{a^2} M_{22}(x-\eta,s) d\eta \\
\mathfrak{R}_{44}(s,\mathcal{A}^e) :&= -\frac{M_{21}(x,s)M_{11}(l,s) + \gamma M_{22}(x,s)M_{21}(l,s)}{M_{11}^2(l,s) + \gamma M_{12}(l,s)M_{21}(l,s)} \int_0^l M_{11}(l-\eta,s) d\eta \\
&+\int_0^x M_{21}(x-\eta,s) d\eta
\end{aligned}
$$

where we adopt the notation of $M_{ij}(x,s)$, where $i = j = 1, 2$, from the Eq. (9.22).