Estimation and Prediction of Convection–Diffusion–Reaction Systems from Point Measurements

Dirk Vries

Promotor:

Prof. dr. ir. G. van Straten
Meet-, Regel- en Systeemtechniek
Wageningen Universiteit

Copromotoren:

Dr. ir. K.J. Keesman
Universitair hoofddocent
Leerstoelgroep Meet-, Regel- en Systeemtechniek
Wageningen Universiteit

Dr. H.J. Zwart
Universitair hoofddocent
Leerstoelgroep Stochastische Systeem-, Signaal- en Besturingstheorie
Universiteit Twente

Leden van de promotiecommissie:

Prof. dr. A. Bagchi
Universiteit Twente

Prof. dr. D. Dochain
Université Catholique de Louvain, Belgique

Prof. dr. J. Molenaar
Wageningen Universiteit

Prof. dr. M. Verhaegen
Technische Universiteit Delft

Dit onderzoek is uitgevoerd binnen de onderzoekschool Dutch Institute of Systems and Control (disc).

Estimation and Prediction of Convection–Diffusion–Reaction Systems from Point Measurements

Dirk Vries

Proefschrift

ter verkrijging van de graad van doctor
op gezag van de rector magnificus
van Wageningen Universiteit,
Prof. dr. M.J. Kropff,
in het openbaar te verdedigen
op vrijdag 6 juni 2008
des morgens te elf uur in de Aula.

Estimation and Prediction of Convection–Diffusion–Reaction Systems from Point Measurements, 2008

Dirk Vries

Ph.D. thesis Wageningen Universiteit, Wageningen, The Netherlands — with summary in Dutch.

Keywords: system identification, distributed parameter systems, estimation, prediction, observation, convection–diffusion–reaction models, UV disinfection, climate room control

ISBN: 987–90–8504–857–2

Voorwoord

Dit proefschrift is het resultaat geworden van vier jaar promotieonderzoek, begonnen in maart 2004. De achtergrond van dit werk valt wellicht beter te begrijpen met een korte voorgeschiedenis. Dit onderzoekswerk maakte deel uit van het STW-project 'Modeling and Control of Flows'*. In dit project zijn de hoofdrolspelers de systeem- en regeltechniekgroepen van Wageningen Universiteit (WU) en Universiteit Twente (UT). Bij beide universiteiten is een promovendus aangesteld: Simon van Mourik (UT) en ikzelf (WU). Ook zijn verscheidene industriële partners van waterverwerking, agrotechnologie, en kastechnologie betrokken geweest. Gedurende deze periode hebben deze 'gebruikers' hun praktijkervaringen gedeeld met de onderzoekers.

'Modeling and Control of Flows' is een brede omschrijving van een onderwerp, en zonder nadere specificatie zou voor de bestudering hiervan decennia moeten worden uitgetrokken. Gelukkig was die er wel. Enerzijds van het doel: het in kaart brengen van ruimtelijke veranderingen in een tijdsvariërend model voor stromingsverschijnselen om uiteindelijk (energie)kostenbesparingen te kunnen realiseren door 'slimme' besturing. Anderzijds een toespitsing op het gebied van de toepassingen: (1) waterzuivering en (2) koeling van bioproducten. Bovendien is er een onderverdeling (maar ook samenwerking) geweest op de werkgebieden: regeling en modelreductietechnieken in Twente en systeemidentificatietechnieken in Wageningen.

De afgelopen vier jaar zijn een aftasting en leerperiode geweest met betrekking tot de praktische en wetenschappelijke invulling van het project. De onderzoeksvrijheid die ik daarbij gekregen heb, waardeerde ik zeer. Uiteindelijk hebben een aantal vragen en antwoorden geleid tot wetenschappelijke publicaties en dit proefschrift. Meer specifiek gesproken: verscheidene systeemidentificatietechnieken met betrekking tot schatting en predictie zijn voorgesteld en uitgewerkt. Deze technieken kunnen in de praktijk worden toegepast om efficiënte, modelgebaseerde besturing mogelijk te maken.

Bij de totstandkoming van dit werk ben ik enkele woorden van dank verschuldigd. Allereerst wil ik mijn dagelijkse begeleiders Karel en Hans (iets minder dagelijks) van harte bedanken. Jullie ideeën, kritische blik, capaciteiten om technische details begrijpbaar te maken, geduld en bovenal, tijd, hebben geleid tot een uitstekende arbeidsomstandigheid! Ik waardeer dat enorm. Simon, jij ook bedankt voor de samenwerking en de droge humor die daarmee gepaard ging. Over (droge) humor en precisie gesproken, velen zullen het vermoedelijk met mij eens zijn dat in dat opzicht Johan Ploegaert uniek is. Met andere woorden, Johan, bedankt voor al je voorbereidende werk en prettige samenwerking! Verder dank ik Gerrit, niet alleen

*Vrij vertaald: modelleren en besturen van stromingen

voor je kritische noten in de laatste fase maar ook voor je bereidheid om van dienst te zijn wanneer nodig. Ook noem ik graag collega-promovendi. Als eerste Timo, we hebben leuke discussies gehad over het werk, software en maatschappelijke kwesties. Bas, Tijmen, Zita, Martijn, Stefan, Hadiyanto en Djaeni: hopelijk kunnen we contact blijven houden. Andere collegae: Marja, je bent een echte sfeermaker met een sociale antenne die (in ieder geval hoorbaar) boven andere (MRS[†])-antenne's uitsteekt (dit mag je als compliment beschouwen). Verder ben ik alle leerstoelgroepmedewerkers dank verschuldigd voor de prettige werksfeer en interessante discussies tijdens de koffiepauzes.

Last, but not least: alle vrienden, kennissen en familieleden die voor de nodige afleiding hebben kunnen zorgen. In het bijzonder Liesbeth, Pim en mijn ouders. Ger en Marina, jullie begrip en steun tijdens mijn studie aan de universiteit en mijn promotieonderzoek (alhoewel de inhoud van het proefschrift jullie misschien niet helemaal duidelijk is) zijn van onschatbare waarde geweest. Lieve Judith, je praktische instelling, zorg en toeverlaat hebben uiteindelijk indirect een bijdrage gehad in dit resultaat. Ik ben je heel dankbaar en ik hoop nog een lange tijd met je te kunnen delen.

[†]Meet-, Regel- en Systeemtechniek

“The most erroneous stories are those we think we know best—and therefore never scrutinize or question.”

Stephen Jay Gould, 1941–2002.

Abstract

DIFFERENT PROCEDURES with respect to estimation and prediction of systems characterized by convection, diffusion and reactions on the basis of point measurement data, have been studied. Two applications of these convection-diffusion-reaction (CDR) systems have been used as a case study of the proposed estimation and prediction methods. One is a climate room for bulk storage of agricultural produce (Case A) and the other is a UV disinfection process used in water treatment, food industry and greenhouse cultivation (Case B).

An essential step in the implementation of estimation and prediction for these types of systems is model reduction. The proposed procedures not only differ by the nature of the estimation and prediction method, but also with respect to early or late model reduction. In the context of this thesis, early model reduction encompasses approximation of the infinite-dimensional system to finite-dimensional form *before* estimation and prediction is worked out, whereas in late model reduction, the approximation step is applied *after* synthesis of an infinite-dimensional estimator (observer) or predictor.

The first contribution of this thesis is an identification approach with output-error (OE) modelling techniques that links important physical parameters in a reduced order model to the OE parameters. This technique is illustrated by Case A, using real experimental data. Local parametric sensitivity analysis shows how physical parameters affect the dominant time constant in an identified, first order output-error model.

The second contribution is a realization approach from a discrete-time linear finite-dimensional system affine in parameters to linear regressive form. The resulting linear regression form allows the formulation of a convex parameter estimation and prediction problem. Such an approach is attractive for reduced order, discretized CDR models with specific boundary conditions. For such models, it turns out that the response and regressor functions can be formulated explicitly as functions of the number of compartments, sensor and actuator location. Once available, they can further be used for a priori identifiability checks, parameter and input sensitivity analysis. Results are illustrated by two diffusion examples with different boundary conditions.

Finally, the last contributions are a static and a dynamic boundary observer for CDR systems. Detectability and observability results aid in the design of a static gain boundary observer of an infinite-dimensional system where only boundary measurements are available. The dynamic observer is synthesized by formulating an H_∞–filtering problem in a linear fractional transformation framework in order to cope with disturbances on the input and output of the system. Both observer synthesis approaches are illustrated by a CDR model of Case B.

Contents

1

Introduction

CONVECTION AND DIFFUSION are two different mechanisms in energy and mass transfer phenomena which have their application in many physical, chemical, biological, environmental and even sociological processes. At the one hand there is diffusion which may be best described by the movement of a random particle, also known as Brownian motion. At the other hand there is convection which describes the forced or free movement of particles or air/fluid flows. Interactions between particles, species or even humans may be described by reaction terms.

Since convection-diffusion-reaction (CDR) processes play an important role in many applications, there is a need for understanding, prediction [79] and control. For these applications, it is natural to study CDR processes from a system theoretical and control engineering viewpoint.

1.1 Motivation

In this work, we focus on estimation and prediction methods *which preserve, as much as possible, the physical interpretation* of a model. Estimation and prediction is considered to belong to the research area of system identification. Roughly speaking, system identification is the study of model determination from measurement data of a target system with a specific modeling objective, starting from a model set.

Two engineering applications are used as show cases in this thesis:

A. Climate control of post-harvest food storage facilities. The controlled variables are generally the air temperature and the moisture content of the air. The goal is to effectively cool or heat the climate room such that a balance between spatial uniformity of product temperature and minimal energy costs is obtained. Within this context, climate control on the basis of a CDR model seems to be a natural choice, where the dynamic model describes the spatial distribution of agricultural produce temperature. Control designs on the basis of (identified) physical models have been reported in [27, 70, 71, 98].

B. Deactivation of micro-organisms by ultra-violet (UV) light. In current process control strategies the UV lamp strength is manipulated on the basis of on-line turbidity measurements of the fluid and minimal required UV dosage for worst case scenarios, see e.g. chapter 14 in [64]. For tighter process control, the following candidate variable may be controlled: the concentration of living micro-organisms at the outlet. In both cases, flow velocity in the disinfection reactor and dispersion of the micro-organisms determine the residence time distribution of biomass. UV disinfection technology is nowadays used in water treatment, food industry and agricultural applications, see e.g. [32, 42, 56, 61, 66].

Although these systems may differ at first sight, the underlying processes may be described by CDR equations which belong to the larger class of *D*istributed *P*arameter

*S*ystems* (DPS). Furthermore, other similarities are:

- The geometry of the application has fixed boundaries;
- The input/output of the system under study is characterized by boundary actuation (ventilator, inlet concentration) and the output by point or boundary observations.
- The ultimate modeling objective is process control.

The applications will be described and studied in Chapter 2 in more detail.

1.2 Modeling approaches

The use of physical modeling is not necessarily needed for control of CDR processes. So-called black box models may be used for this purpose as well. Some advantages for this model type are: (i) flexibility in model characterization, including input-output models; (ii) broad availability of powerful methods for system identification and (iii) no physical knowledge is required. However, there are some disadvantages. These disadvantages show up as advantages in the 'white box' approach and are mentioned below.

In contrast to a black box approach, an attractive property of a *white box* modeling approach is that (i) it provides a physical or deterministic interpretation to the process and (ii) it is an elegant way to use past or a priori knowledge. The advantage of a physical interpretation is obvious: once we have a physical property of the system like the product temperature or dimensions of the environment, only an experimental validation of the changed (state) variable is needed (under the presumption that no other mass or heat transport laws come into play). This makes the extra effort in system identification, experiment design (including sensor/actuator placement) and (optimal) control solutions less laborious compared to black-box approaches.

1.3 Model reduction

Depending on the exact formulation of a particular CDR model, say Σ, explicit solutions to the partial differential equations (pde's) are only known for a very limited class of DPS under specific input signals (see e.g. [18]). So, in the majority of monitoring, prediction and control problems, numerical simulation is needed. This requires that the infinite-dimensional system (or DPS) (Σ) is reduced to a finite dimensional form (Σ_N).

This is illustrated schematically in Figure 1.1.

*Notably, the word 'system' is used, where strictly speaking a *model of the system* is meant. Outside the context of system identification, both terms will be used for a 'model of a system' in order to keep in line with commonly used terminology.

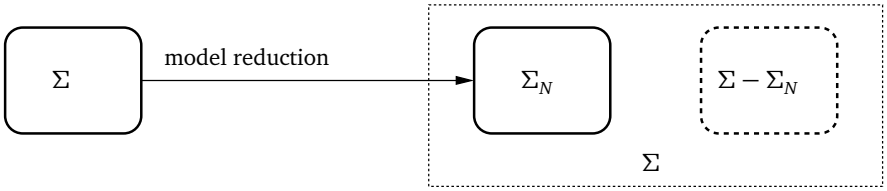


Figure 1.1: Schematic representation of model reduction

A dynamic, distributed parameter system (DPS) derived from physical knowledge can be written in state space form. In general, physical parameters appear in this set of (partial) differential equations. In most physical processes, the parameters can only be obtained by calibrating and validating these model parameters to available input-output data (observations), representing the external behavior (outputs) of the dynamic system. Roughly speaking, this calibration and validation process of the model candidate is called parameter estimation. Apart from some exclusive cases, model reduction is also needed for parameter estimation.

Because of the importance of model reduction, we categorize some reduction tools for distributed parameter systems in 'early stage' and 'late stage' techniques. In chemical engineering, this is also referred to as 'early lumping' and 'late lumping'[†].

'Early stage' reduction. By 'early stage' model reduction, it is meant that the DPS is approximated by a finite dimensional system.

'Late stage' reduction. By 'late stage' model reduction, it is meant that an observer or predictor of a DPS is formulated as an infinite dimensional system and approximated to finite dimensions before implementation in practice.

The following short overview covers a limited selection of reduction methods for DPS. Since the aim is to provide an estimation/prediction procedure which preserves the physical model structure or parameters in a transparent way, finite element or finite volume methods are not discussed. These methods are largely used in fluid dynamics applications, see for implementations in convection–diffusion systems e.g. [14].

Discretization by Finite Differences. In chemical engineering, discretization of a DPS by finite differences is probably still the most popular method [29]. The main reasons for this are: (i) its implementation is one of the simplest ways to approximate a differential operator and (ii) the resulting model is easy to interpret from a 'physical' modeling viewpoint. For instance, the approximation of a plug flow reactor can be characterized by a cascade of stirred tank reactors. With experimental tracer methods, it is shown in [38] that the hydrodynamics of a fixed bed reactor can be approximated reasonably well. For

[†]a lumped (parameter) model is a finite-dimensional model

the convection–diffusion equation, numerous accurate schemes on the basis of finite differences are proposed, see e.g. [1, 35, 116] or the textbook [96].

Weighted residual methods. This family of model reduction methods [37, 79] can be described as follows. Suppose we have an abstract linear evolution equation[‡]:

$$\dot{z} - (P)(z) = 0$$

with P a differential operator, $z \in Z$ the state variables in state space (e.g. a Hilbert space) on a spatial domain $\eta \in \Omega$ and $\dot{z}$ a first order derivative of z with respect to time t. For notational convenience, the arguments t and η of z are dropped. The solutions z are to be approximated by $\tilde{z}$ which is written as a linear combination of basis functions:

$$z \approxeq \tilde{z} = \sum_{i=1}^{N} a_i \varphi_i.$$

In general, substitution of $\tilde{z}$ into the evolution equation results in the following residuals:

$$\varepsilon = (P)(z) - (P)(\tilde{z}) \neq 0$$

The idea in the weighted residual methods is to force the residual to zero in some average sense over the domain Ω by variation of β, i.e.,

$$J = \int_{\Omega} \beta_i \varepsilon \mathrm{d}\eta = 0, \quad i = 1, 2, \ldots N$$

where β_i define the weights on each residual and is further specified below. Without diving deeper into the subject, we mention some frequently used sub-methods:

- Collocation method, where residuals ε are forced to be zero at discrete points (collocation points) $\eta_i \in \Omega$, by choosing β_i as a displaced delta function: $\beta_i := \delta(\eta - \eta_i)$. The sub-domain method can be considered as a modification of the collocation method.
- Least squares method, where $\beta_i := \frac{\partial}{\partial a_i}\varepsilon$ which boils down to the minimization of $\int_{\Omega} \varepsilon^2(\eta) d\eta$.
- Galerkin method, where $\beta_i := \frac{\partial}{\partial a_i}\tilde{z} = \varphi_i$. In the case that the evolution equation is perturbed by a non-linear term, so-called non-linear Galerkin methods [49] improve the approximation accuracy.

[‡]A more complete description is given in Appendix B, Section B.1.3

- Orthogonality methods, where the weight vector is chosen such that it is orthogonal to the residual vector.
- Method of moments, where for one-dimensional space domains: $\beta_i := \eta^{i-1}$, $i = 1, 2, \ldots, N$.

Singular perturbation approximation. Perturbation theory can be used to reduce the order of dynamical systems [54]. A singular perturbation problem is an asymptotic expansion problem containing a small parameter δ that cannot be uniformly approximated by setting δ to zero. If the approximation is exact, the above problem is called a regular perturbation problem. In a large class of singular perturbation problems (e.g. reaction-diffusion systems[65] or predator-prey models), the (spatial or time) domain may be divided into two sub-domains. For each domain an asymptotic series may be obtained. By suitably choosing a matching condition at some point in the domain, it is possible (sometimes by trial and error) to obtain an accurate approximation in a large part of the whole domain. Singular perturbation approximation is also reported for CDR processes, see e.g. [77].

***P*roper *O*rthogonal *D*ecomposition.** POD is also referred to as Principal Component Analysis or Karhunen-Loève expansion [34, 74, 75]. The idea is to let the solution of a pde $z(\eta, t)$ be represented by an (empirical) infinite linear combination of orthogonal functions, analogous to a Fourier series representation: $z(\eta, t) \approx \sum_{i=1}^{N} a_i(t)\varphi_i$ with the expectation that the approximation becomes exact as $N \to \infty$. The POD of z can be obtained by seeking a sequence of functions φ_i, with i up to some value of N such that an optimal approximation in some norm $\|\cdot\|$, e.g. the least-squares norm, is obtained. The discrete version of POD is in fact the well-known singular value decomposition of matrices.

Moment matching. A complex-valued function $g : \Omega \mapsto \mathbb{C}$ defined on an open subset Ω of the complex plane is called *analytical* if it can be represented by $g(s) = \sum_{i=0}^{\infty} c_i(s - s_0)^i$ exponentially converging in a neighborhood of every point $s_0 \in \Omega$. The number c_i is called the i-th moment of $g(s_0)$. The moment matching problem ([2]) is formulated as follows:

Given a positive integer $n > 0$, a sequence $\{s_k\}_{k=1,\ldots,N}$, a sequence of positive integers $\{m_k\}_{k=1,\ldots,N}$ and a function g which is analytical in a neighborhood of points $\{s_k\}_{k=1,\ldots,N}$, then find a strictly proper real rational function $\tilde{g}(s) = p(s)/q(s)$ of degree n such that,

$$\tilde{g}^{(i)}(s_k) = g^{(i)}(s_k) \quad \text{for } 1 \le k \le N, \quad 0 \le i \le m_k$$

where $g^{(i)}$ denotes the i-th derivative of g. When g is a transfer function of a DPS, $\tilde{g}$, with $s_k = \pm\imath\omega_k$ and $\omega_k \in \mathbb{R}$, is frequently used as a reduced order model of g. The method is closely related to Padé approximation.

Padé approximation. The Padé approximant is an approximation of a complex- or real-valued function g by a rational function $\tilde{g}$ of given order $[N, m]$:

$$\tilde{g}(s) = \frac{b_0 + b_1 s + b_2 s + \cdots a_m s^m}{1 + a_1 s + a_2 s + \cdots a_N s^N},\ N, m \geq 0,\ \text{where}\ \begin{cases} g(0) & = \tilde{g}(0) \\ g'(0) & = \tilde{g}'(0) \\ & \vdots \\ g^{(N+m)}(0) & = \tilde{g}^{(N+m)}(0) \end{cases}$$

Control engineers have widely recognized this method as a powerful tool for the approximation of a transfer function in frequency domain, see for instance the textbook [94]. The Laplace transform of an infinite dimensional system yields a trancendental transfer function in the Laplace variable s, i.e. a non-algebraic function of s. Padé approximation is therefore well suited for 'late' reduction of transfer functions from linear infinite dimensional systems to rational transfer functions and subsequent inverse Laplace transformation.

Balanced truncation. Strictly speaking, balanced truncation is not a DPS reduction method. However, it can be used to efficiently reduce the model order of a (Fourier) series approximation of an infinite-dimensional system, see for example [41]. First, a realization is found such that the controllability and observability Gramians of the realization are equal and diagonal matrices, see [43] for an overview. The corresponding system is called a balanced realization. The Gramians contain ordered positive diagonal entries σ_i, called the Hankel singular numbers. Then, a reduced model is obtained by removing the states of a balanced realization which correspond to the least observable and controllable modes, i.e. $\sigma_i \leq \varepsilon$, with ε a certain treshold.

1.4 System identification

System identification comprises the problem of determining a system description (model), say Σ, from available input-output (i/o) data $\{u, y\}_\tau$, typically obtained from experiments with the system $\mathscr{S}$ within a finite time (or frequency) interval $\tau = [0, N]$. The associated identification procedure is characterized by a model characterization phase, a (parameter/state) estimation and/or prediction phase and a validation phase. Most often, (optimal) input and experimental design issues related to the identification experiments are also considered to take part in the identification procedure. The identification procedure of a parametric model $\Sigma(\theta)$, with θ the parameter vector, may be schematically represented by Figure 1.2.

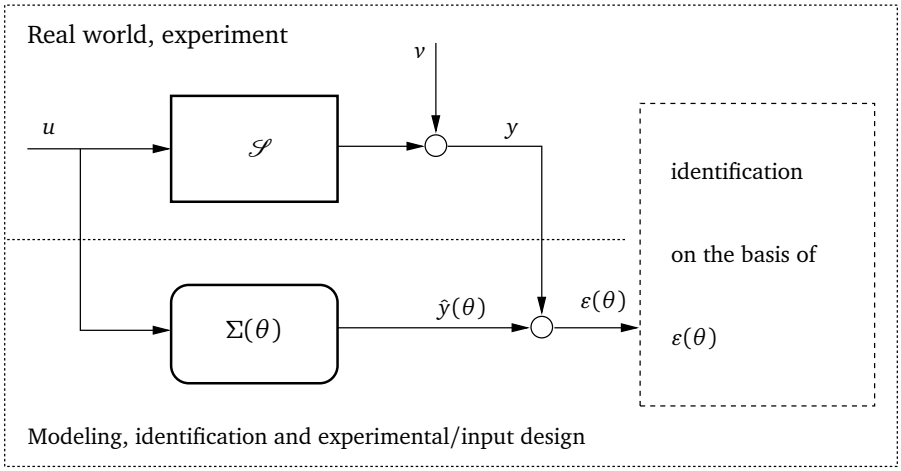


Figure 1.2: A paradigm for system identification. Adopted from Fig. 1.6 [46].

1.4.1 Model characterization

In general, it is not immediately clear how the structure of a given rational transfer function model from u to y can be linked to a physical model Σ, i.e. how to find a state space realization with physical interpretation ('bottom up' or synthesis approach). Alternatively, a 'top down' identification approach is defined as selecting a 'grey' or 'white' box model structure a priori on the basis of physical reasoning and subsequent identification.

1.4.2 Parameter estimation

Parameter estimation is considered as an important and indispensable aspect of system identification. Parameter estimation may be formulated within a prediction error framework as a minimization problem [59], i.e.:

$$\hat{\theta} = \min_{\theta \in \Theta} V_{N_k}(t; \theta, \varepsilon(\theta))$$

where $\theta \in \Theta$ are the unknown parameters in the parameter set Θ, V_{N_k} the objective function over N_k data points (e.g. a least squares norm over the prediction errors with respect to time) and ε the prediction errors $y - \hat{y}$, with $\hat{y}$ model predictions of the observations y.

As a consequence of physical modeling, the parameters in Σ often appear non-linearly in an input-output model structure. Consequently, in general, non-linear optimization solvers with (costly) iterative procedures are used, where it frequently occurs that the parameter search gets stuck in local minima, see e.g. [30]. The prob-

lem gets particularly more complex when dealing with infinite dimensional systems. In that case, it is a common approach to determine a minimal basis in order to solve the estimation problem in a finite-dimensional space. See e.g.: Galerkin approximation schemes [4, 6], rational approximations [82], collocation methods in diffusion systems [92], chemical reactors [57], minimal finite element approximations [17] and subsequent finite element reduction and subspace identification, see e.g. vibration structure examples in civil engineering [81]. Interestingly, subspace identification (see [69, 78, 99, 100, for identification routines of finite dimensional systems]) has the advantage over classical prediction error techniques of the absence of non-linear parameter optimizations.

1.4.3 Linear regressive prediction techniques

As mentioned previously, sooner or later there is a need for a model reduction step. In addition, there is the desire to avoid non-linear parametric optimizations. Linear regression[§] methods are an attractive alternative to subspace-based techniques if the underlying physically interpretable model structure is to be preserved. Inspired by reparametrizations of dynamic systems in the work of [60, 62], Doeswijk and Keesman exploited these reparametrizations to obtain a linear regressive equivalent form of the nominal model for the purpose of estimation and prediction. Implementation of this linear regressive method can be found for a non-linear storage model and for rational biokinetic functions like the Michaelis-Menten model in [30, 31].

1.4.4 State reconstruction and observer synthesis

Observation theory for finite dimensional systems is a more or less established research field. There are also many contributions in the area of linear infinite–dimensional system theory which are relevant to practical situations of state reconstruction or state observation, see e.g. [9, 20, 113]. For example, an interesting approach until now only applied to linear infinite–dimensional systems, is to use backstepping techniques to construct an observer [91].

In the last decade, asymptotically strong observers for bilinear infinite-dimensional systems have been reported, [see e.g. 10, 114]. These systems are of particular interest in (bio)chemical processes involving reactions. For example, in the CDR equations of the UV disinfection application, a bilinear term appears when the deactivation reaction term by the UV-lamp is considered for control.

In addition, as is also the case with the climate storage room and the UV disinfection process, measurement and control actions take place at pre-specified points. However, with point sensing and actuation, the corresponding operators map out of state space and introduce extra technicalities. For bilinear systems, [11] handled the ‘unboundedness’ of the observation and/or control operators by investigating

[§]Some background about linear regression and least squares is given in Appendix B.2.

admissibility and appropriate regularity assumptions. Such an investigation has also been done for linear regular systems, see e.g. [13, 111].

1.5 Research scope

Motivated by the above, the framework in which this research is performed will be further specified.

We concentrate on the *estimation and prediction of systems characterized by CDR models, on the basis of point measurement data while preserving physical knowledge to a large extend*. Hence a 'top-down' model selection approach is considered.

Furthermore, the following issues will be addressed:

1. *Given a late reduction technique, is it possible to preserve physical knowledge of the nominal CDR model in terms of the parameter estimate θ, when a discrete time LTI model is identified with output-error techniques?*

   The problem is how physical (possibly lumped) parameters of the approximate model can be recovered from the identified (output-error) model and discrete measurement data.

2. *Given an early reduction technique, is it possible to rewrite the estimation and prediction problem of a CDR system into a linear regression, and if so, how?*

   In [30, 31], a reparametrization technique is worked out for (non-)linear models which can be extended to the DPS case. Some difficulties to be addressed are: existence of a reparametrization and the formulation of a prediction method which is transparently related to the physical model.

3. *Without model order reduction, is it possible to obtain design conditions for a static observer of a CDR system, under the condition that it is (approximately) observable?*

   Systems described by point sensing or actuation typically have 'unbounded' control and/or observation operators, see for instance [p. 11, Chapter 1 of 21]. Is it possible to obtain (simple) concepts for observability and detectability for these particular CDR systems?

4. *Given a late reduction technique, is it possible to obtain a dynamic observer which is robust to disturbances at both the (boundary) input as well as the point observation, and if so, how?*

   Important questions are: well-posedness of the observer-feedback problem, influence of the model reduction and influence of typical CDR parameters on the observer synthesis.

1.6 Thesis outline

The thesis is outlined as follows:

Chapter 2. The climate storage in a closed room and the UV disinfection process are introduced as case studies.

Chapter 3. Parameter estimation of agricultural produce in a climate storage room is worked out. The aim is to preserve physical meaning captured by the parameters θ of a reduced order model. The climate controlled storage room is used to illustrate the method. *(C.f. research issue 1 in Section 1.5).*

Chapter 4. A method to obtain output predictions $\hat{y}_k$ of CDR systems on the basis of input-output data and linear regressive techniques is proposed. The technique is illustrated by a diffusion example with various boundary conditions. *(C.f. research issue 2 in Section 1.5).*

Chapter 5. This chapter studies the synthesis of an observer for the state variable z in CDR systems. It is divided in two parts. In the first part, infinite dimensional theoretical concepts like observability and detectability are formulated for boundary control systems. These concepts are further worked out for a static observer of the UV disinfection case study. The second part describes a robust observer synthesis approach for the same case study, but now under input and output disturbances. Results of both approaches are illustrated by numerical simulations of the UV disinfection case. *(C.f. research issues 3 and 4 in Section 1.5).*

Chapter 6. Some final remarks and conclusions are given.

Hence, the main contributions of this thesis are described in Chapters 3–5. The 'paths' which are used to obtain the desired (physical) parameter estimates θ, output predictions $\hat{y}_k$ and state observations z are schematically depicted in Figure 1.3.

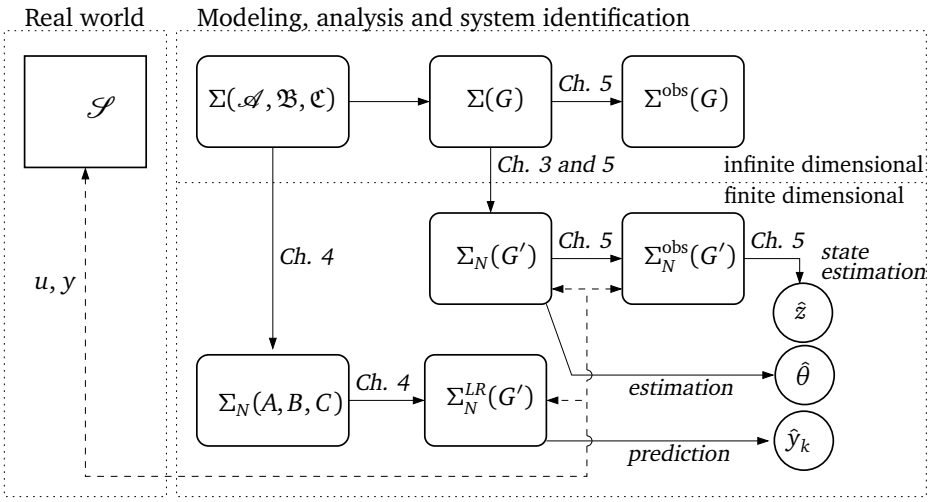


Figure 1.3: Schematic view of the trajectories from system $\mathscr{S}$ to prediction, parameter and state estimation, together with the thesis chapters. Used notation: Σ:=nominal DPS, Σ_N:=finite-dimensional system, obs:=observer, LR:=linear regression, Σ_d:=finite-dimensional discrete time system, G, G':=transfer function and its approximation, respectively, $\mathscr{A}$:=infinitesimal generator of a semigroup, $\mathfrak{B}$ and $\mathfrak{C}$ boundary and point observation operator respectively, A, B, C, system –, input – and observation matrix, respectively.

2

Case studies

Abstract

In this thesis, two examples of convection–diffusion–reaction (CDR) processes are studied to illustrate the parameter and state estimation methods. One example, referred to as Case A, deals with a food storage process in a store room for agricultural produce. The other example deals with a UV disinfection process and is referred to as Case B. In this chapter, models for Case A and B are proposed and outlined. The modeling aim is estimation and prediction of these CDR processes for use in control.

2.1 Conservation principles

Before elucidating on the two applications where flow phenomena occur, we briefly consider some general conservation principles from physics and deduce the convection-diffusion-reaction equation (without boundary conditions).

In many physical phenomena, the dependent variables of interest (say, $\psi \in \mathbb{R}^3$) obey a generalized conservation principle. This dependent variable may be involved in heat and mass transfer, turbulent flow and related phenomena. The general partial differential equation describing this conservation is given as [79]:

$$\frac{\partial}{\partial t}(\rho\psi) + \operatorname{div}(\rho\mathbf{v}\psi) = \operatorname{div}(\alpha \operatorname{grad}\psi) + S$$

or, written in Cartesian tensor form:

$$\frac{\partial}{\partial t}(\rho\psi) + \frac{\partial}{\partial \xi_j}(\rho v_j \psi) = \frac{\partial}{\partial \xi_j}(\alpha \frac{\partial \psi}{\partial \xi_j}) + S \tag{2.1}$$

with real-valued constants and variables Further, the flow field should satisfy an

ψ	mass fraction of chemical species, enthalpy or temperature, velocity component, turbulent kinetic energy, ...
ρ	density
$\mathbf{v} \in \mathbb{R}^3$	velocity vector
α	(turbulent/heat/mass) diffusion coefficient
S	source term (e.g. reactions, heat production)
ξ_j	j-th spatial coordinate ($j \in \{1,2,3\}$)
t	time.

additional constraint, namely, the mass-conservation or continuity equation (here in Cartesian tensor form):

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial \xi_j}(\rho v_j) = 0 \tag{2.2}$$

This mass conservation constraint reduces the differential equation in eqn. (2.1) to:

$$\rho \frac{\partial \psi}{\partial t} + \rho_j v_j \frac{\partial \psi}{\partial \xi_j} = \frac{\partial}{\partial \xi_j}(\alpha \frac{\partial \psi}{\partial \xi_j}) + S \tag{2.3}$$

under compatible initial and boundary conditions.

Notice that, without explicitly writing the initial and boundary conditions, eqn. (2.3) already describes CDR phenomena. For heat transfer, we denote $T := \psi$, whereas for mass transfer $c := \psi$. In this thesis, we will restrict ourselves to con-

sider only the simpler, one-dimensional form of the CDR process eqn. (2.3) (i.e., $\xi = \xi_j$, $j = 1$), by making several assumptions in the modeling process. A discretized version of eqn. (2.1) will be considered for linear regressive estimation and prediction techniques in Chapter 4.

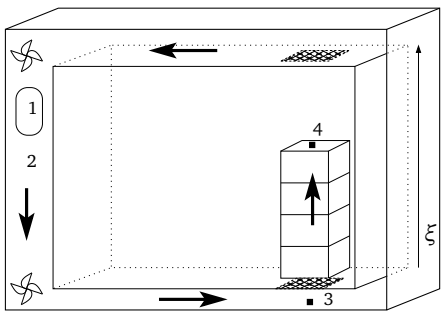
The modeling phase for the bulk storage application and the UV disinfection will be described case by case in the subsequent sections.

2.2 Case Study A: Storage room climate control

Modeling of physical processes in climate control applications of agricultural produce, amounts most times to the incorporation of heat transfer, moisture transport and fluid mechanics. As a result, models of food storage rooms describe time *and* spatially dependent heat and mass transport [40, 51, 63] and also detailed fluid mechanics in the resulting partial differential equations (see e.g. [28, 97, 98, 115]). Hence, such a modeling approach for storage rooms leads to high computational effort when solving these equations.

In [70, 71], a simplified physical model for a bulk storage room is proposed for designing control laws which are explicitly dependent on (lumped) physical parameters. In that work, the ultimate control objective is to achieve a spatial temperature distribution over the agricultural produce which is as uniform as possible, while keeping the (ventilation) energy costs as low as possible. In this thesis, it is tried to identify that part of the model which describes the spatial temperature distribution of the bulk product temperature.

As a reference, the complete model of the air temperature of the inner cell, of the shaft behind the cooler, and of the product temperature is presented in the next section and denoted as Σ^{room}. A schematic drawing of the complete climate room is shown in Figure 2.1. The submodel that describes the heat transfer over the bulk, referred to as Σ^{bulk}, is outlined in Section 2.2.4 and is used as the to-be-identified nominal model in Chapter 3.



1: cooling device
2: shaft behind the cooling device, T_s
3: air temperature sensor, $T_a(\xi = 0)$
4: air temperature sensor, $T_a(\xi = L)$

Figure 2.1: Schematic drawing of climate room.

In the following section, first the modeling of Σ^{room} is briefly outlined. A full account of the modeling phase of Σ^{room} is given in [72].

2.2.1 Modeling assumptions

The following assumptions are made with respect to the geometry and physical properties of the *room* and *products*:

A1. The walls are perfectly insulated.

A2. The products are spherical, having radius R.

Assumptions with respect to *heat* and *moisture* transport:

A3. The air- and product temperature are assumed uniform with respect to the width of the climate room, i.e., they only vary in the ξ-direction (see Figure 2.1).

A4. The temperature dynamics of the air between the top of the bulk and the fan are not incorporated.

A5. Moisture transport is not modeled. Nevertheless, the heat capacity of air is adjusted for a high humidity.

A6. Diffusion in the air is neglected. In addition, there is no heat exchange between the products (no 'bulk conduction').

A7. The product skin has the same heat conduction as the product interior.

A8. The whole product surface is exposed to air. Contact with other surface is accounted for by reducing the total product surfaces.

A9. Heat transport inside a product at height ξ is modeled by diffusion.

A10. Heat production of the product is assumed linear with its temperature. The heat production is considered as the source term S of the CDR process as in eqn. (2.3).

A11. The effectiveness of the cooler β_c (dimensionless), is assumed constant.

In [72, chapter 2], the modeling approach and the assumptions are discussed in detail. At first sight, it seems that the neglection of moisture transport and the absence of bulk conduction seem the most restrictive. The main motivation in [72] for these assumptions is to come up with a model that is simple enough to get explicit expressions for a controller.

2.2.2 Complete climate room model

The complete climate room model is given as follows:

$$\Sigma^{\text{room}} := \left\{ \begin{array}{rcl} M_0 \frac{\partial}{\partial t} T_s & = & v_s \left(\left(1-\beta_c\right) T_a(L,t) + \beta_c T_c(t) - T_s(t) \right) \\ \frac{\partial}{\partial t} T_a(\xi,t) & = & -v_p \frac{\partial}{\partial \xi} T_a(\xi,t) + M_4 \left(T_p(\xi,t) - T_a(\xi,t) \right) \\ \frac{\partial}{\partial t} T_p(r,\xi,t) & = & \frac{M_1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} T_p(r,\xi,t) \right) + M_2 T_p(r,\xi,t) \\ T_a(0,t) & = & T_s(t) \\ \frac{\partial}{\partial r} T_p(0,\xi,t) & = & 0 \\ \frac{\partial}{\partial r} T_p(R,\xi,t) & = & \frac{h(v_p)}{\lambda_p} \left(T_a(\xi,t) - T_p(R,\xi,t) \right) \end{array} \right. \tag{2.4}$$

with (detailed listings are given in Section C.3):

$v_s = F/V_a$	air velocity inside the shaft
$v_p = F/V_a$	superficial flow velocity in the bulk, volumetric flow rate devided by air volume
β_c	effectiveness constant of cooler
$M_0 = \rho_a c_a V_a$	lumped parameter
$M_1 = \lambda_p/(\rho_p c_p)$	diffusive heat transfer coefficient of product
$M_2 = a_p/c_p$	lumped parameter in heat production term
$M_4 = h(v_p)A_p/(\gamma \rho_a c_a)$	lumped parameter in heat exchange term
T_s, T_a, T_p	temperature in shaft behind the cooler, of air in the inner cell and of the agricultural produce respectively.

2.2.3 Approximation of bulk storage part

In [72], one of the goals is to obtain an explicit result for a switching control law. To this aim, two approximations are made:

1. Timescale separation of eqn. (2.4), by neglecting the fast dynamics within the product. As a consequence, the spatial distribution of the product temperature T_p is not taken into account;

2. Laplace transformation and subsequent Padé–[0,1] approximation of the heat transfer between the air and product surface.

These approximations lead to:

$$\widehat{T}_p(R,\xi,s) = \underbrace{\frac{B}{-A+s}}_{G_p(s)} \widehat{T}_a(\xi,s)$$

which after inverse Laplace transform reads

$$\frac{\partial}{\partial t} T_p(\xi, t) = AT_p(\xi, t) + BT_a(\xi, t) \tag{2.5}$$

The approximation for product temperature as in eqn. (2.5) leads to the following approximate storage room model $\overline{\Sigma}^{\text{room}}$:

$$\overline{\Sigma}^{\text{room}} = \begin{cases} \frac{M_0}{\nu_s} \frac{\partial}{\partial t} T_s = \left(1 - \beta_c\right) T_a(L, t) + \beta_c T_c(t) - T_s(t) \\ \frac{1}{\nu_p} \frac{\partial}{\partial t} T_a(\xi, t) = -\frac{\partial}{\partial \xi} T_a(\xi, t) + \frac{M_4}{\nu_p} \left(T_p(\xi, t) - T_a(\xi, t)\right) \\ T_a(0, t) = T_s(t) \\ \frac{\partial}{\partial t} T_p = AT_p(\xi, t) + BT_a(\xi, t). \end{cases} \tag{2.6}$$

On the basis of time-scale separation, another approximation is made by neglecting the fast dynamics of T_a and T_s. This results in the right-hand side of the first two equalities of $\overline{\Sigma}^{\text{room}}$ being equal to zero. Given the approximation of $\overline{\Sigma}^{\text{room}}$ after time-scale decomposition, the part that only describes the transfer over the bulk is denoted as Σ^{bulk} in the following subsection.

2.2.4 Nominal model

A critical part of the identification of $\overline{\Sigma}^{\text{room}}$, is the parameter estimation of the heat transfer over the bulk. To this aim, an isolated, small-scale climate room has been used for temperature measurements below (i.e. a system input $u := T_a(0)$) and on top (i.e. a system output $y := T_a(L)$) of a stack of crates filled with agricultural produce. Consequently, for identification of that part of the model, we are interested in estimation and validation of:

$$\Sigma^{\text{bulk}} := \begin{cases} \frac{\partial}{\partial \xi} T_a(\xi, t) = \frac{M_4}{\nu_p} \left(\left(T_p(\xi, t) - T_a(\xi, t)\right)\right), \quad T_a(\xi, 0) = T_a^0(\xi) \\ \frac{\partial}{\partial t} T_p = AT_p(\xi, t) + BT_a(\xi, t) \\ u(t) := T_a(0, t) = T_s(t) \\ y(t) := T_a(L, t). \end{cases} \tag{2.7}$$

The non–rational transfer function $G(s)$ from u to y obtained from Laplace transformation of Σ^{bulk}, reads

$$G(s) = e^{-M_5\left(-1 + \frac{B}{-A+s}\right)}, \tag{2.8}$$

where $M_5 := M_4 L/\nu_p$ is a dimensionless number denoting the ratio between chemical reaction rate and convection, i.e. a heat transfer variant to the Damköhler number [72].

The transfer function G as in Eqn. (2.8) is our starting point for validation of the heat transfer over the bulk of agricultural produce in Chapter 3.

2.3 Case Study B: UV disinfection

The UV disinfection case is a practical example of a CDR system where, typically, sensors and actuators are placed at prespecified points or at the boundary. UV light is, amongst others, applied in fluid (water/juice) treatment processes to deactivate (pathogenic) micro-organisms. In the food process industry, in (waste)water treatment and in greenhouse technology industries (see some examples in [32, 42, 56, 66]). UV disinfection techniques have gained more attention since they do not leave traces of chemical reagents, in constrast to e.g. water disinfection by chlorination.

Typically, the UV lamp intensity is *merely* controlled by the transmittance of the fluid to be treated, see e.g. chapter 14 in [64]. This is a rather conservative and indirect approach since the actual active pathogenic biomass may differ from the *a priori* assumed amount. In order to efficiently cut lamp energy costs, we would like to implement an observer (or 'software sensor') that uses one or more direct biomass measurements. If properly designed, such an observer allows us to monitor the (most resistant) pathogen concentration at any point in the reactor. In this thesis, the UV disinfection process in an annular reactor—generally used in greenhouse drain water infestation and in disinfection of fluid food products—is chosen an interesting case with respect to this aim.

2.3.1 Modeling assumptions

First, the inputs (control variables) and outputs (measurements) of our annular UV system model are specified. An important candidate for the UV disinfection process control variable is to manipulate the intensity of the UV lamp, referred to as $f_0(t)$. A second candidate is the inlet concentration $c_{\text{in}}(t)$. We furthermore assume that measurements of the active biomass concentration are available[†] at the inlet of the reactor, i.e. $c_{\text{in}}(t)$, and at some point along the main flow direction in the reactor $c(\xi^*, t)$, with $\xi^* \in [0, L]$ and L the length of the reactor. A schematic drawing of an annular UV disinfection reactor is shown in Figure 2.2.

In addition, the following assumptions are made with respect to *flow* and *dispersion*:

A1: In the axial direction, flow and dispersion is modeled by forced plug flow advection plus diffusion, where the diffusion coefficient α is constant. Ideal mixing is assumed in the radial direction $r \in [r_0, R]$ of the reactor.

[†]Such a 'smart' sensor for *on-line* measuring of the metabolic status of cells is still in development, see for instance [87].

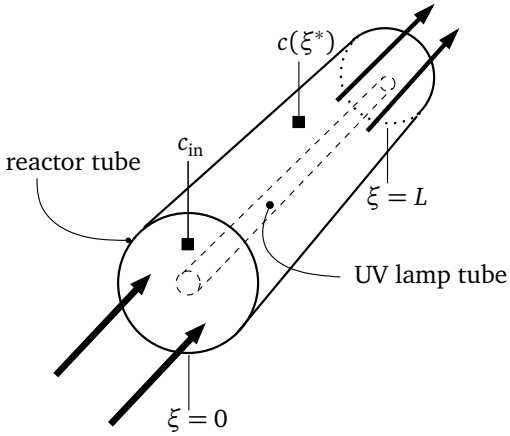


Figure 2.2: Schematic overview of an annular UV disinfection reactor

A2: Solids[‡] are homogenuosly suspended in the medium, thus the absorbance of the medium is constant with respect to the spatial coordinates.

A3: Only one species of micro-organisms is considered, thus $c := \psi$ is a scalar.

A4: The biomass concentration at the inlet point of the reactor obeys a Robin-type boundary condition. In other words, immediately at $\xi = 0$, dispersion of the inlet fluid with the fluid contained in the reactor takes place. At the reactor outlet, the concentration fulfills a Neumann-type condition. The combination of these boundary conditions is well-known under chemical engineers and is called Danckwerts-type, see the original work of [23] or [80, 90] for technical notes. The Danckwerts condition implies that no UV irradiation, nor dispersion takes place at the reactor outlet.

The *source* term S in eqn. (2.3) can be directly related to the inactivation of micro-organisms. The assumptions with respect to S are as follows:

A5: Deactivation of pathogenic organisms is assumed to obey first order reaction kinetics, with constant reaction constant κ per light intensity unit f. See [44, 88].

A6: UV light is emitted radially from the entire surface of the lamp.

A7: The effects of refraction and reflection in the suspended fluid is negligible.

A8: UV irradiation does not penetrate through the reactor walls and the reactor walls do not reflect.

[‡]That is, all the particles in the medium except for micro-organisms.

A9: The reaction term is of mild form, implying that a smooth solution of the model exists.

The assumptions A5–A9 on modeling the inactivation of micro-organisms by UV-light leads to the use of Lambert's law in the source term S of eqn. (2.3)[53], i.e.:

$$S = \kappa f c \qquad \text{subject to} \qquad \begin{cases} \frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r}(rf) = -Ef \\ f_0(r) := f(r_0) \end{cases} \tag{2.9}$$

where the arguments are dropped for notational convenience and where

f	UV irradiation
E	monochromatic absorbance of fluid
r	radius of reactor tube
r_0	outer radius of UV lamp tube.

Due to ideal mixing assumption in the radial direction (A1), the source term S as in eqn. (2.3) can be reduced to an average deactivation term f across a cross section of the tube:

$$S(\xi,t) = \kappa f(t) c(\xi,t) \quad \text{with} \quad f(t) = f_0(t)\kappa \frac{\int_{r_0}^{R} \frac{r_0}{r} \exp\left(-E(r-r_0)\right) \mathrm{d}r}{R - r_0} \tag{2.10}$$

where κ is the susceptibility constant of the micro-organism species to UV irradiance.

2.3.2 Nominal model

The general differential equation for CDR processes under the assumptions A1–A10, lead to the following model for UV disinfection:

$$\Sigma^{\mathrm{UV}} := \begin{cases} \frac{\partial}{\partial t} c(\xi,t) &=& \alpha \frac{\partial^2}{\partial \xi^2} c(\xi,t) - v_f \frac{\partial}{\partial \xi} c(\xi,t) - \kappa f(t) c(\xi,t) \\ c(\xi,0) &=:& c_0(\xi) \\ -\alpha \frac{\partial c}{\partial \xi}(0,t) + v c(0,t) &=:& v c_{in}(t) \\ \frac{\partial}{\partial \xi} c(L,t) &=& 0, \end{cases} \tag{2.11}$$

where $\xi \in [0, L]$ and the input and output observations are,

$$u(t) = c_{\mathrm{in}}(t), \qquad y(t) = c(\xi^*, t). \tag{2.12}$$

The non-dimensional form of model Σ^{UV} as outlined in Appendix D.1 will be used as an illustrative example of a CDR equation for the methods outlined in Chapter 5.

2.4 Summarizing remarks

Two engineering applications where flow, dispersion and reactions take place are modeled by CDR equations. The modeling phase for these applications is described case by case as summarized below.

Case A: storage room climate control. The control purpose of the storage room is to control the spatial distribution and level of the product temperature T_p by manipulating the air velocity in the shaft and temperature of the cooling device. Physical model based control depends thus heavily on the model describing the temperature distribution over the bulk of agricultural produce. The nominal model Σ^{bulk} as in eqn. (2.7), and its equivalent transfer function description G from $T_a(0)$ (below the bulk) to $T_a(L)$ (top of the bulk) in eqn. (2.8), describes this temperature distribution.

Case B: UV disinfection. The control purpose here is to exercise control over the outlet concentration $c(L)$ by manipulating the UV lamp irradiation f and/or the input concentration c_{in}. Convection and diffusion equations in the main flow direction are proposed to account for flow and dispersion effects. Furthermore, it is assumed that first order deactivation kinetics take the disinfection of biomass by UV light into account. The complete nominal model for UV disinfection control is given as Σ^{UV} in eqs. (2.11) and (2.12).

3

Parameter Estimation

This chapter is based on:

D. Vries et al. *Physical parameter estimation in a distributed parameter system: a food storage case*. Submitted to Journal of Process Control. 2007

Abstract

DISCRETE TIME MODELS of air temperature dynamics inside a bulk of agricultural produce for different flow conditions are identified, using an OE method. It is shown that the performance gain of higher order models compared to a first-order model is negligible. The parameters of this first order model are linked to the physical parameters in an analytically derived approximation of a diffusion-convection-reaction type of model. Local sensitivity analysis shows to what extend the physical parameters affect the empirical parameters in the identified model. Subsequently, the values of the most sensitive physical parameters in the approximate physical model are recovered from the identified model.

3.1 Introduction

In industry, there is a strong demand to incorporate physical knowledge into models for analysis, prediction and control. This preference for *physical* model-based prediction and control, while incorporating experimental data, needs an appropriate system identification approach [e.g. 59, 76].

In [110], it has been noted that there has been little attention to, and experience of, system identification for model-based control in food industry applications. In these specific applications often heat transfer plays a fundamental role. Starting from physical principles and after Laplace transformation, in the frequency domain usually *non-rational* transfer models result, which need special attention. However, some researchers, as e.g. in [28, 110], have chosen to neglect this phenomenon and simply restrict their identification to rational transfer functions in which the parameters are determined by the data. Hence, no physical information in terms of the original heat transfer processes is incorporated in their models. Consequently, it is hard to make predictions for systems other than the particular experimental setup. As an alternative approach to this direct rational transfer function modeling approach, in [58, 84] it has been suggested to retain the fractional powers of the Laplace variable in the transfer function descriptions. However, only modest progress in this approach has been made to preserve physical laws, see [83] for an application to diffusion systems. Other significant attempts in system identification to preserve prior knowledge have been made by using so-called grey-box or semi-physical modeling approaches, see e.g. [8, 89].

In the semi-physical modeling approaches to the food storage case, first principle knowledge that is available about heat transfer, moisture transport, fluid mechanics, etc., is incorporated in the modeling phase. As a result, models of food storage rooms describe time *and* spatially dependent heat and mass transport [40, 51, 63]. Moreover, these flow driven transport phenomena lead to nonlinearities in the resulting partial differential equations (PDE). Consequently, and even more so when including detailed fluid mechanics (see [15, 16, 28, 97, 98, 115]), the resulting simulation model requires a high computational effort, making parameter estimation of the full model almost inappropriate. Hence, approximate lower order models are needed. One way is to consider a (rough) spatial discretisation of the PDE(s) into compartments. As an example of this, for estimation and prediction in compartmental diffusive systems we refer to e.g. [7, 105]. However, in general, the information on the physical parameters of the original physical model gets lost after an approximation step. All these complications may contribute to the lack of penetration of system identification methods for control in the food industries.

In this chapter, an attempt is made to identify discrete-time transfer function models from noisy data while preserving physical insight of a specific distributed parameter system to a large extend. To circumvent the numerical obstacles mentioned before, we show how to recover physical parameters from low order transfer function models using Padé approximation, output-error (OE) identification and local

sensitivity analysis of the physical parameters with respect to the identified transfer function parameters. The goal of this chapter is to demonstrate this procedure to a specific food storage facility using experimental data.

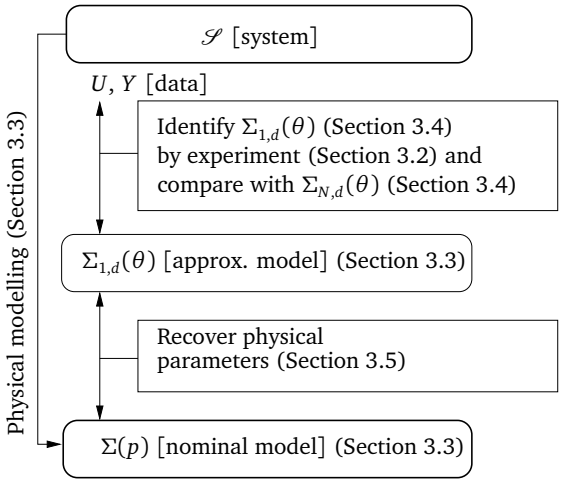
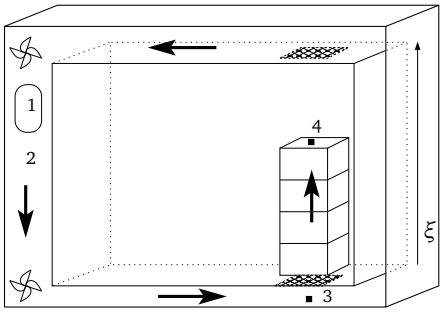


Figure 3.1: Chapter outline, where $\Sigma_{1,d}(\theta)$ is the (to be) identified first order model with parameters θ, $\Sigma_{N,d}(\hat{\theta})$ are identified OE-models with order N and $\Sigma(p)$ is the nominal, physical model with physical parameters p. Boxes which are thickly outlined indicate the starting point, boxes with rectangle corners: procedures/algorithms, rounded corners: models/system.

The chapter is organized as schematically depicted in Figure 3.1, where the numbers between parentheses indicate the section numbers, rectangle boxes denote procedures, rounded boxes mathematical entities and the arrows denote the interaction by procedure(s). Section 3.2 describes the experimental setup of the case study. In Section 3.3, modeling of the system $\mathscr{S}$ with $\Sigma(\theta)$ and its model order reduction to $\Sigma_{1,d}(\theta)$ is presented. Section 3.4 outlines the identification procedure. Section 3.5 illustrates how a subset of (uncertain) physical parameters can be recovered from the identified model. In the last section, concluding remarks are given.

3.2 Case study: experimental conditions

In this section, the experimental setup is outlined. Figure 3.2 shows a schematic side view of the experimental storage room (GTI Zephyr 39892). The dimensions of the room are $3.6 \times 2.9 \times 2.1\ m$. A stack of crates filled with potatoes is placed on the floor. The air temperatures are measured at the bottom and at the top of the stack. The air temperatures were measured at the top and at the bottom of the



1: cooling device
2: shaft behind the cooling device, T_s
3: air temperature sensor, $T_a(\xi = 0)$
4: air temperature sensor, $T_a(\xi = L)$

Figure 3.2: Schematic drawing of climate room.

stack, with two temperature Pt-100 4–wired sensors. A fan enforces air circulation through the stack, which is sealed with thick plastic foil and paper, for minimal air and temperature loss through the stack walls. For three settings of the fan the corresponding flux (m^3/h) through the bulk is measured by a calibrated flux meter above the stack of crates. This resulted in three average fluxes: $\Phi_1 = 211\ m^3/h$, $\Phi_2 = 710\ m^3/h$ and $\Phi_3 = 1200\ m^3/h$. Details about the dimensions, potato weight, etc., can be found in chapter 2 of [72].

In three experiments, step responses were measured. In Figure 3.3 the measured input signal and output responses are depicted. It can be seen that, in this experiment, dead time is indeed* small and therefore it is neglected.

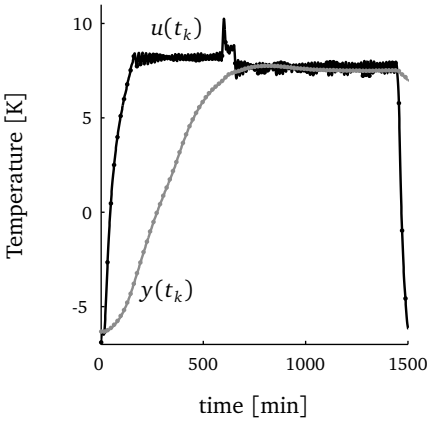
It should be noted that the second half (with respect to time) of the data set is chosen as calibration set, while the other half is used for validation.

The air temperature was set to 293 K, and was subsequently brought down to 278 K in three days. The relative air humidities were measured in the center of the bottom and top of the stack with a sensor of type Rotronic Hydroclip Campbell CR10 data. The system was monitored with a sample time of one minute. The relative humidities varied less than 10% in all the experiments. We assumed therefore, that the moisture transport between air and products is small and will thus have a small influence on the heat transport. This observation supports the use of the approximate model in the next section, that only describes temperature dynamics.

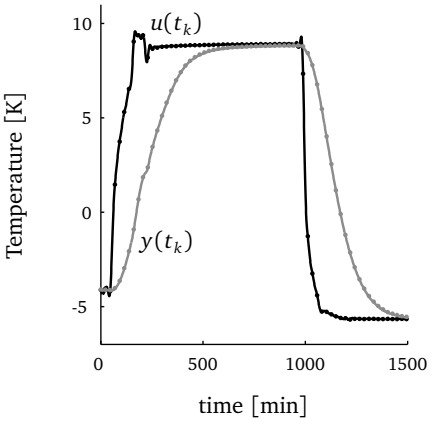
3.3 Physical model

In [70, 71], a physical model that describes the temperature dynamics and spatial distribution in a food storage room is presented (see also Chapter 2, Section 2.2 and Appendix C.3 for used notation). In that work, explicit expressions as a function of physical parameters have been derived. To ensure good controller performance, the identification of the unknown parameters is needed.

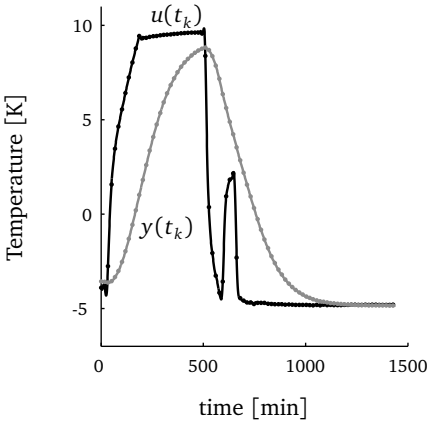
*from a priori physical knowledge, the ratio L/v in M_5 is small, see Section C.3 in the Appendices.



(a) Ventilator setting 59%



(b) Ventilator setting 85%



(c) Ventilator setting 100%

Figure 3.3: Measured air temperature below ($T_a(0, t_k)$, [**—**]) and above ($T_a(L, t_k)$, [—]) the stack of crates, sampled points denoted by [·]. $T_a(\xi, 0) \equiv 281$ [K].

To this aim, parameter estimation of an important part of that model is carried out, i.e. the submodel Σ^{bulk} which describes the spatial air temperature distribution from below $T_a(0)$ to the top $T_a(L)$ of the stack of crates, see eqn. (2.7). We assume that this model of the air temperature dynamics over a bulk of agricultural produce has the same transport dynamics as the stack of crates in our case study. Since the crates are fully filled, i.e., there are no large air spaces inside or between the crates, this assumption is defendable.

The model $\Sigma(p)^{\dagger}$ (see also Chapter 2, Section 2.2.4) with physical parameters p is given as:

$$\Sigma(p) := \begin{cases} \frac{\partial}{\partial \xi} T_a(\xi,t) = \frac{M_4(p)}{p_1}\Big((T_p(\xi,t) - T_a(\xi,t)\Big), \quad T_a(\xi,0) = T_a^0(\xi) \\ \frac{\partial}{\partial t} T_p = A(p)T_p(\xi,t) + B(p)T_a(\xi,t) \\ u(t) := T_a(0,t) \\ y(t) := T_a(L,t) \end{cases} \tag{3.1}$$

where v_p denotes the air velocity in the bulk and is related to the ventilator setting, see Appendix C, Section C.3 for details. Furthermore, the lumped physical parameters M_4, A and B are dependent on the physical parameters p and v_p, see also Appendix C.3. In the same work of Mourik, a first order approximate model that facilitates calibration and fast simulation is derived by Padé approximation of the Laplace transform of the transfer function of Σ (see eqn. (2.8)) to yield:

$$\widehat{y}(s) = \underbrace{\frac{\tilde{b}}{-\tilde{a}+s}}_{G^{[0,1]}(s)} \widehat{u}(s) \tag{3.2}$$

where, For notational convenience the argument p of $\tilde{a}$, $\tilde{b}$, A, B and M_5 have been

$\widehat{y}$	Laplace transform of y
$\widehat{u}$	Laplace transform of u
$\tilde{a} = -\frac{A^2}{M_5 B}$	inverse of the dominant time constant
$\tilde{b} = \frac{-A^2}{M_5 B} e^{M_5(-1-B/A)}$	the influence of the input
$M_4 := \frac{h_v A_p}{\gamma_p \rho_a c_a}$	lumped constant for heat transfer resistance between bulk and air (physical parameters are listed in Appendix C.3)
$M_5 := \frac{M_4 L}{v_p}$	Damköhler number, i.e. a dimensionless ratio between convection and reaction rate, see [72].

dropped. Notice, that in eqn. (3.2), it is assumed that (a) pure time delay due to transport phenomena and (b) the influences of heat exchange through the walls of the bulk and moisture transport are negligible.

Inverse Laplace transformation of eqn. (3.2) gives a first order approximate model (the model order is denoted in the subscript of Σ) in the continuous time

†Since in this chaper, only the model for the bulk storage case is considered, the superscript notation 'bulk' has been dropped for eligibility reasons, i.e. here $\Sigma := \Sigma^{\text{bulk}}$ with Σ^{bulk} as in eqn. (2.7).

domain:

$$\Sigma_1(p): \begin{cases} \frac{\mathrm{d}}{\mathrm{d}t} T_a(L,t) = \tilde{a}(p) T_a(L,t) + \tilde{b}(p) T_a(0,t), \quad T_a(\xi,0) = T_a^0(\xi) \\ u(t) := T_a(0,t) \\ y(t) := T_a(L,t) \end{cases} \tag{3.3}$$

A priori estimates for p have been found in [72, 109]. The parameters which were most difficult to estimate are: N_β (number of produce exposed to the air), γ_p (bulk porosity), A_p (surface per volume of product ratio) and β (efficiency of cooler). The efficiency of the cooler has been determined in a seperate experiment, see [72, chapter 3].

3.4 Identification procedure

Since we deal with sampled input and output data in the time domain, we choose to transform the continuous time model eqn. (3.3) to a discrete time one for further identification and validation instead of transforming the data to the continuous time domain. The transformation from continuous to discrete time, can be found in standard control textbooks, see *e.g.* [3].

The transformation of $\Sigma_1(p)$ with zero-order hold for u and y and equidistant time sampling $\Delta_t \equiv t_{k+1} - t_k$ leads to the discrete time form $\Sigma_{1,d}$:

$$\Sigma_{1,d}(\theta): y(t_k) = e^{\tilde{a}\Delta_t} y(t_{k-1}) + \frac{\tilde{b}}{\tilde{a}} \left(1 - e^{\tilde{a}\Delta_t}\right) u(t_k), \tag{3.4}$$

Closer inspection of $\tilde{a}$ and $\tilde{b}$ in eqn. (3.3) shows that for very small $M_3 = R\sqrt{a_p}/\sqrt{\lambda_p/\rho_p}$ (see also Appendix C.3), that is, for small products with very small product heat production and normal mass densities: $\tilde{b} \approx -\tilde{a}$. This implies that, for realistic a priori given physical parameter values, we get approximately unit gain: $-\tilde{b}/\tilde{a} \approx 1$. Consequently, $\tilde{b}$ and $\tilde{a}$ are not independently identifiable. We will come back to this result later.

As a first step and in line with the approximate model, we select a first order output-error (OE) model structure and evaluate the structure by OE-identification techniques with the given input/output data. The choice for an OE method will be further elucidated in the next section.

3.4.1 Output error modelling

It is common practice to minimize the output prediction errors $\varepsilon(t_k) = \hat{y}(t_k|\theta) - y(t_k)$ at time instants t_k under some norm. Here θ is the parameter vector, $y(t_k)$ contains the observations and $\hat{y}(t_k|\theta)$ is the model output prediction at time instant t_k.

The set of OE models is a suitable model class to characterize the specific physical discrete time model $\Sigma_{1,d}$. Discrete time OE models are characterized by the following definition:

DEFINITION 3.4.1. *An OE model set is determined by two polynomials:*

$$B(\theta,q) = b_0 + b_1 q^{-1} + b_2 q^{-2} + \ldots + b_{n_b} q^{-n_b}$$
$$F(\theta,q) = 1 + f_1 q^{-1} + f_2 q^{-2} + \ldots + f_N q^{-N}$$

with

$$\theta := \big(\underbrace{f_1 \ \cdots \ f_N}_{\theta^F} \ \underbrace{b_0 \ \cdots \ b_{n_b}}_{\theta^B}\big)$$

such that

$$\Sigma_{N,d} : \ y(t_k) = G_N(q,\theta)u(t_k) + e(t_k) \tag{3.5}$$

with $\Sigma_{N,d}$ belonging to the collection of predictor models $\mathscr{M} = \{ \ G_N(q,\theta) \mid G(q,\theta) = B(\theta,q)/F(\theta,q), \ \theta \in \mathbb{R}^{n_b+N} \ \}$. Furthermore, q is the so-called forward shift operator, i.e. $qw(t_k) = w(t_{k+1})$ and $e(t_k)$ represents a Gaussian white noise term.

Given $\Sigma_{1,d}$ as in eqn. (3.4), we assume that the following OE–model describes the real system behavior under a white noise perturbed output accurately enough:

$$\Sigma_{1,d} : \ y(t_k) = G_1(\theta,q)u(t_k) + e(t_k) \tag{3.6}$$

with

$$\theta = \big(\underbrace{-e^{\tilde{a}\Delta_t}}_{\theta^F} \ \underbrace{\frac{\tilde{b}}{\tilde{a}}\left(1 - e^{\tilde{a}\Delta_t}\right)}_{\theta^B}\big) \in \mathbb{R}^{n_b+N} = \mathbb{R}^2, \qquad u \in U = \mathbb{R} \quad \text{and} \quad y \in Y = \mathbb{R}. \tag{3.7}$$

Clearly, our model eqs. (3.6) and (3.7) is put in a natural way into an 'output error' form [59], since it has its origin in the physically interpretable state space representation eqn. (3.3). As a result, the parameter vector θ is linked to the physical parameters p via $\tilde{a}$ and $\tilde{b}$. The disadvantage of using an OE–model structure as opposed to e.g. equation error structures, is that the parameters appear non-linearly in the output. Fortunately, appropriate OE–methods have been developed to deal with this (see *e.g.* [59, 76]). Moreover, the disadvantage is outweighted by the advantage that, in general, OE models fit nicely in the low frequency region. A good low-frequency fit is in particular beneficial for simulation and long–term prediction.

3.4.2 Model structure selection

We will now define two conditions for which we call our model sufficiently accurate and consider the model structure as validated,

(i) the cross correlation of the residuals sequence, $\varepsilon(t_k) = \hat{y}(t_k|\hat{\theta}) - y(t_k)$ with the inputs $u(t_k)$, should be zero with confidence 99%;

(ii) the final prediction error (FPE) should be small compared to the measurement variance, i.e. $F_{p.e.} \leq 0.5$.

The final prediction error in the above is defined as Akaike's prediction error function. This function is related to the loss function V_{N_k} (see also Chapter 1, Section 1.4.2) and the number of parameters as follows [59]:

$$F_{\text{p.e.}}(\mathscr{M}) := V_n(\hat{\theta}, Z)\frac{1+(n_b+N)/N_k}{1-(n_b+N)/N_k} \quad \text{with} \quad V_n(\hat{\theta}, Z) = \frac{1}{N}\sum_{k=1}^{N_k}\tfrac{1}{2}\varepsilon(t_k, \hat{\theta}) \tag{3.8}$$

where N_k is the number of data points in the calibration or validation data set with the data set belonging to the real-valued space $Z = U \oplus Y$.

For an appropriate sampling rate f_s, with $f_s \in [0.1,\ 0.5]\tau_p$, we need a good estimation of the time constant $\tau_p = |1/\tilde{a}|$. But, in order to estimate the time constant a correct sampling rate is required. In the next section, we illustrate how we try to tackle this circular reasoning.

Furthermore, although θ^F and θ^B can be freely estimated, it is likely that $\tilde{b} \approx -\tilde{a}$ (see previous section) and thus $\theta^B \approx 1 + \theta^F$. Hence, we choose to only estimate one physical parameter from either θ^F or θ^B. Different parameters can be estimated (individually) when more than one experimental data set becomes available. This is done by using the estimate obtained in the previous set for the newly available set. Notice furthermore that for a single experiment, more physical parameters could be retrieved from θ by higher order (Padé) approximate models. The latter is under the assumption that enough information is contained in the data. Experimental data is presented in the next section.

3.5 Physical parameter estimation

As substantiated in Section 3.4.2, we first estimate the time constant for the bulk storage room $\tau_p := 1/\tilde{a}$ from $\hat{\theta}_F$ using OE–methods with different sample rates $f_s = \Delta_t^{-1}$. Then, we are able to calculate (unknown/uncertain) physical parameters like γ_p, N_β and others from τ_p.

3.5.1 Estimation time-constant

We reconstruct the dominant characteristic time $\tau_p = |1/\tilde{a}|$ via a data set processed with different (re)sample rates $f_s = 1/\Delta_t$. Basically, $\tilde{a}$ is obtained via calculation of the slope of $\ln|\hat{\theta}_F|$ versus Δ_t, see for details [109] and Appendix C. After calculation of the time constant, it is checked whether the model characterization conditions (i) and (ii) are fulfilled.

Table 3.1 presents these estimations ($\hat{\tau}_p$) for all ventilator settings, together with τ_p using physical parameter values (see Appendix C.3) from literature. These results will be further worked out in Section 3.5.4. It suffices to note that the results

Table 3.1: Estimated ($\hat{\tau}_p$) and time constant calculated from literature (τ_p) [min.] for different flow rates.

Ventilator rate	59%	85%	99%
$\hat{\tau}_p$ [min.]	178	133	117
τ_p [min.]	112	78	66

of $\hat{\tau}_p$ are roughly a factor 1.5 higher than the *a priori* calculated τ_p.

3.5.2 Model structure selection checks

In this section, we primarily inspect the final prediction errors ($F_{\text{p.e.}}$ for three different cases.

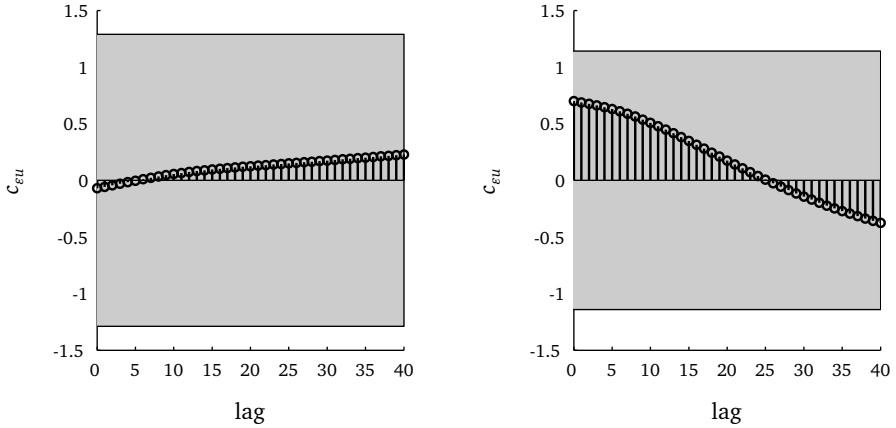
- $\Sigma_{1,d}(\hat{\theta})$: validated first order ($N = 1$) discrete-time OE-model as in eqn. (3.4) with calibrated parameters;
- $\Sigma^{\text{opt}}_{N,d}(\hat{\theta})$: validated discrete-time OE-model as in eqn. (3.5), with selected structure order N under Akaike's identification criterion as in eqn. (3.8);
- $\Sigma_{1,d}(\theta_{\text{lit}})$: first order model as in eqn. (3.6), with parameters $\theta_{\text{lit}}(p)$ calculated from literature values of p.

Final prediction errors $F_{\text{p.e.}}$ are presented in Table 3.2 for each case with the different ventilator rates.

Next, we check the cross correlation of the innovations $\varepsilon(t_k) = \hat{y}(t_k) - y(t_k)$ and inputs $u(t_k)$ at different time lags to find whether still some structural information is left in the data. Results for the cases 1 and 3 are presented in Figure 3.4, because case 2 showed similar results as case 1. For each case the sample time was chosen as $\Delta_t = (\hat{\tau}_p/8)$ min. to achieve an acceptable trade-off between capturing fast model dynamics and noise. Furthermore, we only present the case where the ventilator rate was set at 99%, $\Delta_t = 15$ min. For other ventilator rates, similar results have been obtained. The $F_{p.e.}$ associated with $\Sigma_{1,d}(\hat{\theta})$ is small and the

Table 3.2: $F_{\text{p.e.}}$ in $[K^2]$ for three model structure cases. Percentages between brackets denote the ventilator settings, numbers between square brackets the optimal model order N.

	$\Sigma_{1,d}(\hat{\theta})$	$\Sigma^{\text{opt}}_{N,d}(\hat{\theta})$	$\Sigma_{1,d}(\theta_{\text{lit}})$
$F_{\text{p.e.}}$ (59%)	0.25	0.00075 [6]	37.9
$F_{\text{p.e.}}$ (85%)	0.34	0.021 [2]	46.4
$F_{\text{p.e.}}$ (99%)	0.48	0.00040 [3]	25.2



(a) Correlation coefficients for $\Sigma_{1,d}(\hat{\theta})$

(b) Correlation coefficients for $\Sigma_{1,d}(\theta_{\text{lit}})$

Figure 3.4: Normalized cross correlation functions ($c_{\varepsilon u}$), where 99%–Confidence regions are shaded in gray.

correlation between the innovations does not improve much with increasing model order. Therefore, the first order identified model $\Sigma_{1,d}(\hat{\theta})$ is found suitably accurate for simulation or model predictive control purposes. We may call the first order OE–model a validated black box model if we stop at this point.

In the next section, we will analyze the 'strength' of the link between the OE-parameters θ and the physical parameters p using local parametric sensitivity analysis of the reduced-order model $\Sigma_{1,d}$.

3.5.3 Local sensitivity analysis

In order to check how the physical parameters, p_i, $i = \{1, 2, \ldots n_p\}$ in model eqn. (3.3) contribute to θ^F, local sensitivities of θ^F with respect to p are calculated. The sensitivities are normalized to compensate for their nominal parameter values, i.e. the initial guesses p_i° as found in literature.

The local, parametric sensitivities are defined as:

$$\theta^F_{p_i} := \left(\sum_{i=1}^{n_p} \left| p_i^\circ \cdot \frac{\partial |\theta^F|}{\partial p_i}(p_i^\circ) \right| \right)^{-1} p_i^\circ \frac{\partial \theta^F}{\partial p_i} \tag{3.9}$$

Note, that the sum of all the normalized local sensitivities $\theta^F_{p_i}$ as defined in eqn. (3.9) evaluated at $p = p^\circ$ equals 1. We can easily see by the absolute magnitude of $\theta^F_{p_i}$, whether a physical parameter p_i affects the parameter θ^F significantly ($0 \ll \theta^F_{p_i} < 1$) or not ($0 < \theta^F_{p_i} \ll 1$). All the significant normalized parametric sensitivities calculated at p° are shown in Table 3.3.

Table 3.3: Local sensitivities ($\times 10^2$) of θ^F w.r.t. some physical parameters p_i.

$\theta^F_{p_i}$	value
θ^F_R	23.3
θ^F_L, $\theta^F_{N_\beta}$, $\theta^F_{\rho_p}$, $\theta^F_{c_p}$	7.72
$\theta^F_{c_{\text{air}}}$, $\theta^F_{\rho_{\text{air}}}$, $\theta^F_{\gamma_p}$, $\theta^F_{v_p}$, $\theta^F_{V_{\text{cr}}}$	7.69

Most of the parameters p_i introduce comparable parametric sensitivities of θ^F. Other parameters have sensitivities in the order of 10^{-6} and are not shown in Table 3.3. The most important parameters ($\theta^F_{p_i} \gtrsim .05$) are: R (radius bulk produce), L (height of the bulk), N_β (number of products exposed to the air), ρ_p (density of product), c_{p} and c_{air} (heat capacities of product and air, respectively), ρ_{air} (density of air), γ_p (porosity bulk), v_p (flow velocity) and V_{crate} (volume of crate). Of course, the sampling interval Δ_t also contributes significantly to the value of θ^F, but we do not regard it as a physical parameter and exclude it from further analysis.

Furthermore, due to the equal magnitude of the normalized sensitivities with respect to a particular parameter, each contribution of a parameter could easily cause a different heat resistance and consequently a different estimated τ_p then what would be expected from literature values. This phenomenon will be shown in the next section.

3.5.4 Parameter estimation

Time constant

As mentioned before, given the first order model eqn. (3.3) with $\tilde{b} \approx -\tilde{a}$, only one physical parameter p_i can be recovered uniquely due to practical unidentifiability. Table 3.4 shows the estimated p_i from the estimated time-constant, $\hat{\tau}_p$ (shown in

Table 3.1), while leaving the other parameters p_j, $j \neq i$ unchanged at their nominal value: $p_j = p_j^\circ$. From the available data and the first order approximate model, it is possible to estimate three parameters subsequently. In the same table, an example of this estimation is shown by the updated values in bold face type: here, first R, then on the basis of R, N_β and finally γ_p is estimated.

Table 3.4: Physical parameters individually calculated from estimated time constant. Bold faced values are estimated subsequently. Nominal parameter values are shown between parentheses.

Ventilator rate [%]	$\hat{\tau}_p$ [min.]	R (0.29) $[m]$	L (1.47) $[m]$	N_β (234) $[-]$	γ_p (0.031) $[\frac{m^3}{m^3}]$	v (19.3) (27.6) (32.5) $[\frac{m}{min}]$
59	178	.033	2.28	**363**	**.44**	12.4
85	133	**.030**	1.71	**293**	**.59**	16.5
99	117	.029	1.50	**239**	**.68**	18.9

From the sensitivity analysis we already have that for the first order effects, the bulk properties (R, N_β, L, γ_p, ρ_p, c_p, V_{crate}), flow velocity (v) and air properties (ρ_{air}, c_{air}) contribute almost equally to the first order dynamics of the system. Since the densities ρ_p, ρ_{air} and heat coefficients c_p and c_{air} are widespread available in literature, we are mainly interested in the estimation of γ_p, N_β, R, L and v. The height of the bulk L is included to compare the measured height with the estimated one.

Table 3.4 shows that the air velocity inside the bulk, estimated from the time constant, has increased by a factor 1.33 which is almost in agreement with the measured ventilator setting increase of 1.44 from 59% to 85%. A similar agreement is found for lower air velocities and ventilator settings.

Gain

In Section 3.4.2, it was argued that it is practically unfeasible to estimate two physical parameters from θ, since $\tilde{b} \approx -\tilde{a}$. However, we may check if a unit gain is (approximately) achieved or not, by obtaining the slope from the line $\hat{\theta}^B$ versus $\hat{\theta}^F + 1$, see also eqn. (3.7). This gain, i.e. $K := -\tilde{b}/\tilde{a}$, is calculated for different ventilator rates and shown in Table 3.5. Table 3.5 shows that unit gain is approximately achieved, which indicate that there is no heat leakage through the foils.

Table 3.5: Comparison K calculated from $\hat{\theta}^F$ and $\hat{\theta}^B$ for different ventilator rates.

Gain	Ventilator rate		
	59 %	85 %	99 %
$K(\hat{\theta})$	0.98	1.02	1.02

3.5.5 Discussion

From the identification and physical parameter estimations, we summarize some points of discussion:

- generally, a higher order (OE) model is needed to achieve better $F_{\text{p.e.}}$ and to decrease the correlation between the innovations if inspecting the data at one particular ventilator rate. In order to establish a link between physical parameters and OE-model parameters in eqn. (3.3), a first order model $\Sigma_{1,d}$ is selected a priori. From Table 3.2 we conclude that a first order model provides small final prediction errors for different ventilator rates.

  For reference, a second order Padé approximate model is proposed in Appendix C, Section C.1. However, the mapping between the OE-model parameters θ and p becomes increasingly complex with higher model order N, and, moreover, this mapping also becomes non-unique.

- the time constant τ_p can be recovered quite accurately from a calibrated first order model.

- $F_{\text{p.e.}}$ for the first and optimal order model is considerably smaller than the $F_{\text{p.e.}}$ calculated with nominal parameter guesses. The differences between calibration and the a priori estimates can be attributed to:

  a. $\hat{\tau}_p > \tau_p$, can be caused by the higher heat resistance of the (agricultural) produce.

  b. $\hat{\tau}_p < \tau_p$, can be caused by the occurence of lower resistance, e.g. 're-circulating flows' that result in a smaller effective 'active mixing volume' due to turbulence.

  c. $\hat{K} < K$, can be caused by heat leakage to the environment,

  d. $\hat{K} > K$, can be caused by an unmodeled heat source or, more likely, underestimated heat production.

  Notice, from Table 3.1 that in our case study: $\hat{\tau}_p > \tau_p$ and $\hat{K} \approx K$.

- based on the estimated time constant, a sample time of approximately 10–20 minutes is sufficient to identify the model.

It should be noted that the model derived and identified in this chapter is based upon first experiments with agricultural produce within a storage room, while in general, (optimal) experiment design and identification also amounts to a (re)iterative process. For recent and more advanced experiment designs, the reader is referred to, e.g., [39, 95].

3.6 Conclusions

An identification approach with output–error modelling techniques showed that a heat transfer process in a typical food storage setup can be suitably modelled by first order models. Basically, the *rate of heat exchange* is an important property for climate cooling of agricultural produce. We capture this heat transfer rate by calibrating the dominant time constant τ_p in a first order OE-model under different ventilator settings.

In addition, local parametric sensitivity analysis was carried out to inspect how physical parameters affect this time constant. We conclude that important physical parameters can be individually recovered from the calibrated OE–model parameters. With the use of more experimental data sets under differing flow conditions (e.g. air velocities), it is shown that more than one physical parameter can be recovered.

4

Explicit linear regressive structures for prediction

This chapter is based on:

- D. Vries, K. J. Keesman, and H. Zwart. “Explicit linear regressive model structures for estimation, prediction and experimental design in compartmental diffusive systems”. In: *Proceedings of 14th IFAC Symposium on System Identification*. March 29–31, Newcastle, Australia 2006.
- D. Vries, K. J. Keesman, and H. Zwart. “Linear regressive model structures for estimation and prediction of compartmental diffusive systems”. In: *Proceedings of 5th International IMACS Symposium on Mathematical Modelling*. Vienna, Austria 2006.
- D. Vries, K. J. Keesman, and H. Zwart. *Estimation and prediction of convection–diffusion systems using linear regression*. In preparation.

Abstract

COMPARTMENTAL, OR DISTRIBUTED parameter systems which are discretized, belong generally to the class of linear structured systems. As a consequence, such a system can be reparametrized into linear regressive form which is suitable for parameter estimation and subsequent prediction. An approach for such a transformation is shown. Results are obtained for *a priori* determining parameter sensitivity and identifiability on the basis of the linear regressive form. Furthermore, explicit linear regressive structures are obtained for the case that the system matrix is characterized by a tridiagonal symmetric matrix, as is the case with discrete CDR systems.

4.1 Introduction

In Section 1.4, already some motivation is given for the use of linear regressive estimation and prediction. In short, linear regression techniques avoid non-linear parameter optimizations. This chapter addresses the parameter estimation and prediction problem of CDR-systems.

Typically, after discretization with central finite differences of an original linear partial differential equation (pde), one arrives at a state space system description $\Sigma_d(\mathbf{A}(\vartheta), \mathbf{B}(\vartheta), \mathbf{C}(\vartheta), \mathbf{D}(\vartheta))$ with mappings determined by physical parameters ϑ. An important issue is how to rephrase the estimation and prediction problem into linear regressive form. Or, in other words, how to obtain a linear regressive realization* Σ_d^{LR} from a state space system description Σ_d.

More specifically, we are seeking *explicit* results as a function of number of states n, actuator position and sensor position. Because, if this can be done without knowing the actual observations, we will be able to:

- *a priori* determine the effects of number of compartments, i.e. the influence of the number of states n in Σ_d if the original pde is defined with a scalar state variable;
- *a priori* decide on optimal sensor/actuator placement with respect to excitation of parametric regression output sensitivities;
- *a priori* determine the optimal number of states for prediction on the basis of the linear regressive structure, while keeping numerical conditioning difficulties to a minimum.

Given the structure of Σ_d after applying central difference techniques, it is not a coincidence that the inverse of a symmetric, tridiagonal matrix receives much attention in literature and explicit results exist, see for instance [47, 48, 67, 85] and references therein.

The idea in this chapter is to lump ϑ into θ in such a way that unique parameter estimates of θ from a physically interpretable single-input-single-output (SISO) structure Σ_d are within reach. It is then possible to use these parameter estimates for further (output) prediction with the aid of the linear regressive system realization Σ_d^{LR} while preserving knowledge of the original parameters ϑ to a large extend. As will be shown later, θ is defined as a vector of the polynomial variables ϑ. Note furthermore that, we do *not* imply to directly estimate ϑ.

Another problem is whether the parameters ϑ are structurally identifiable, i.e. whether the parameters can in principle be determined uniquely from the data [45, 60]. It turns out that the linear regressive form of Σ_d also allows a simple identifiability test.

*Notice, that we use the notion 'realization' here, since, analogously to a state space realization, the linear regressive system representation will generally not be unique.

Chapter outline

This chapter is organized as follows.

Section 4.2 Some definitions and notation are introduced for a linear regressive system in the case where $\mathbf{A}$ is affine in ϑ.

Section 4.3 In this section, it is explained how a particular linear structured system (which have often their origin in physical modeling) can be reparametrized into a form suitable for linear regressive estimation and prediction. Also some results on parametric sensitivity analysis and identifiability are obtained.

Section 4.4 Explicit expressions are given for two discrete-time diffusion systems. In these expressions, a distinction is made between physically interpretable parameters and 'parameters' related to the discretization scheme. It is also remarked how these expressions may be used for discretized CDR-systems.

Section 4.5 Concluding remarks are given.

4.2 Definitions and problem formulation

In this section, formal definitions are presented and the realization problem is posed.

4.2.1 Linear, structured systems in discrete time

Our starting point is a SISO physical system, it is natural to start from the following state space description of a linear, structured system (see e.g. [45]) in discrete time Σ_d (notice that the subscript of Σ is used to denote a discrete system):

$$\Sigma_d(\mathbf{A},\mathbf{B},\mathbf{C},\mathbf{D}) \qquad := \begin{cases} \mathbf{x}_{k+1} & = \mathbf{A}(\vartheta)\mathbf{x}_k + \mathbf{B}(\vartheta)\mathbf{u}_k, \quad \mathbf{x}(0) = \mathbf{x}_0(\vartheta) \\ \mathbf{y}_k & = \mathbf{C}(\vartheta)\mathbf{x}_k + \mathbf{D}(\vartheta)\mathbf{u}_k \end{cases} \tag{4.1}$$

and

$$\mathbf{A} : \Theta \mapsto \mathbb{R}^{n\times n}, \qquad \mathbf{B} : \Theta \mapsto \mathbb{R}^{n\times m,} \qquad \mathbf{C} : \Theta \mapsto \mathbb{R}^{1\times n}, \qquad \mathbf{D} : \Theta \mapsto \mathbb{R}.$$

$\gamma,\ \gamma'$	response functions
$\phi,\ \phi'$	regressors
$\mathbf{Z}_{k-n\cdots k} \subseteq Z \in \mathbb{R}^{2(n+1)}$	i/o (input-output) data in Z during time interval $[k-n,k]$
$Z = U_{\text{adm}} \oplus Y$	i/o space
$\mathbf{Z}_{k\cdots k+n} = \begin{pmatrix}\mathbf{U}_{k\cdots k+n}\\ \mathbf{Y}_{k\cdots k+n}\end{pmatrix} \in Z$	input-output data set
$\mathbf{U}_{k\cdots k+n} := \begin{pmatrix}u_k & u_{k+1} & \cdots & u_{k+n}\end{pmatrix}^\top$	input data set
$\mathbf{Y}_{k\cdots k+n} := \begin{pmatrix}y_k & y_{k+1} & \cdots & y_{k+n}\end{pmatrix}^\top$	output data set
$\Xi \subseteq \mathbb{R}^r$	space with known constants appearing in ϕ, ϕ', γ and γ'. □

The following notation is used:

$\mathbf{x} \in X = \mathbb{R}^n$	states in state space
$\mathbf{y} \in Y = \mathbb{R}^m$	output in observation space
$\mathbf{u} \in U_{\text{adm}} \subset \mathbb{R}$	input variable in the admissible input space
$\vartheta \in \Theta \subset \mathbb{R}^p,\ p \geq 1$	parameters in parameter space
$\mathbf{A}$	system matrix
$\mathbf{B}$	input matrix
$\mathbf{C}$	observation matrix
$\mathbf{D}$	feedthrough matrix

4.2.2 Linear regressive systems

The main motivation for finding a system in linear regressive form lies in the well-known optimal properties of linear regressions, i.e. *unique* solutions of parameter estimates. The following definition for a linear regressive prediction system will be used in the sequel.

DEFINITION 4.2.1. *The following system Σ_d^{LR} is called a SISO,* linear regressive *system in discrete time:*

$$\Sigma_d^{LR} := \begin{cases} \theta^\top \cdot \phi(\mathbf{Z}_{k-n\cdots k}) &= \gamma(\mathbf{Z}_{k-n\cdots k}) \\ y_k &= \theta^\top \cdot \phi'(\mathbf{Z}_{k-n\cdots k-1}) - \gamma'(\mathbf{Z}_{k-n\cdots k-1}) \end{cases} \tag{4.2}$$

and

$$\gamma, \gamma' : \Xi \times Z \mapsto \mathbb{R}, \qquad \phi, \phi' : \Xi \times Z \mapsto \mathbb{R}^r.$$

As in eqn. (4.1), y_k is again the observation of the external behaviour of Σ_d. In what follows, we need some additional conditions on Σ_d for the realization of Σ_d

as in eqn. (4.1) to Σ_d^{LR}:

A1: $\mathbf{x}_0 \equiv 0$;

A2: $\mathbf{D} = \mathbf{0}$, $\mathbf{B} \neq \mathbf{0}$ and $\mathbf{C} \neq \mathbf{0}$;

A3: The matrices are affine in ϑ, i.e., they can be permuted as follows,

$$\mathbf{A}(\vartheta) = \overline{\mathbf{A}} + \sum_{i=1}^{p} \widetilde{\mathbf{A}}_i \vartheta_i, \quad \mathbf{B}(\vartheta) = \overline{\mathbf{B}} + \sum_{i=1}^{p} \widetilde{\mathbf{B}}_i \vartheta \tag{4.3}$$

$$\mathbf{C}(\vartheta) = \overline{\mathbf{C}} + \sum_{i=1}^{p} \widetilde{\mathbf{C}} \vartheta_i. \tag{4.4}$$

with $\overline{\mathbf{A}}$, $\overline{\mathbf{B}}$ and $\overline{\mathbf{C}}$ assumed to be known.

And to obtain explicit results, we need:

A4. system matrix $\mathbf{A}$ being tridiagonal and symmetric as in the following two types:

Type I:

$$\mathbf{A} := \begin{pmatrix} \vartheta_2 & 1 & 0 & \cdots & 0 \\ 1 & \vartheta_1 & 1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & 1 & \vartheta_1 & 1 \\ 0 & \cdots & 0 & 1 & \vartheta_1 \end{pmatrix}_{n \times n}, \qquad \text{with} \quad n \in \mathbb{N}_+ \tag{4.5}$$

$$\vartheta_1 = \vartheta_2, \quad \vartheta_1, \vartheta_2 \in \mathbb{R}_+; \tag{4.6}$$

Type II: $\mathbf{A}$ as is defined in eqn. (4.5), but with:

$$\vartheta_1 \neq \vartheta_2, \quad \vartheta_1, \vartheta_2 \in \mathbb{R}_+. \tag{4.7}$$

□

Note that, it is straightforward to extend the results for the case $\mathbf{D} \neq 0$ if it is linear in ϑ. The realization problem of transforming Σ_d into Σ_d^{LR} is sketched in the following section.

4.2.3 Problem formulation

Given the preliminaries discussed so far, the realization problem is formulated as finding a method to obtain a linear regressive realization Σ_d^{LR}, given Σ_d in eqn. (4.1) under Assumptions A1–A3 and input-output data. Observe from Σ_d^{LR} in eqn. (4.2), that the model output $\hat{y}_k$ is based upon input-output data $\mathbf{Z}$, the estimates $\hat{\theta}$ and the regressor ϕ' and response function γ'. Furthermore, it is well-known how to obtain the input-output transfer function G from the state space model Σ_d. Hence, the realization problem can be phrased more specifically as:

How to obtain the relationship:

(i) between the input-output transfer function G of Σ_d as in eqn. (4.1) and the regressor ϕ, ϕ' and response functions γ, γ' of Σ_d^{LR} under Assumptions A1–A3;

(ii) between ϑ in Σ_d and θ in Σ_d^{LR}.

Furthermore, the explicit expressions found for tridiagonal matrices in e.g. [47, 48, 67, 85] suggests the realization problem under Assumptions A1–A4 can be tackled in some more detail.

Hence, the third issue of our realization problem becomes:

(iii) How to write the linear regressive realization Σ_d^{LR} with ϕ, ϕ', γ and γ' being explicit functions on the number of compartments, sensor position and actuation position, in the case that $\mathbf{A}$ can be written in the form of eqs. (4.5) *and* (4.6) *or eqs.* (4.5) *and* (4.7)*?*

In what follows, analysis of the proposed form of Σ_d^{LR} reveals an identifiability check and, in the case that the regression scheme is in explicit form, sensitivities of the response functions with respect to the parameters θ can be made explicit.

4.3 Realization to linear regressive form

In this section, a procedure is given for obtaining Σ_d^{LR} from Σ_d.

4.3.1 Realization method

Denote the transfer function G of Σ_d as $\mathbf{C}\mathbf{M}^{-1}\mathbf{B}$, with $\mathbf{M}(\vartheta, q) := q\mathbf{I} - \mathbf{A}$. Notice that $\mathbf{M}$ is a function of ϑ and the forward-shift operator q^1, $q^1 y_k = y_{k+1}$ (see e.g. [3] or a brief introduction in Appendix B, Section B.1.1). Hence, the key to the realization problem is to determine $\mathbf{M}^{-1}(\vartheta, q)$, such that a linear regressive set of equations in some parameter vector θ can be written. Or, in other words, obtain the linear regressive form of Σ_d by means of suitably defining the parameter (vector) function

$\varphi : \vartheta \mapsto \theta$, $\theta \in \mathbb{R}^r$. Relations between the $\mathbf{M}^{-1}$ of Σ_d and the regressor ϕ, ϕ' and response functions γ, γ' of Σ_d^{LR} will then follow.

We discern three steps for the realization Σ_d to $\Sigma^e xtLR_d$, namely:

Transfer function decomposition. The transfer function G of Σ_d is written into a parameter and time-shift operator part;

Transfer function reparametrization. The decomposition of G in the previous step is followed by a reparametrization into a linear regression form and;

Definition of output predictor. The linear regression is rewritten to linear regressive prediction by rearrangement of terms.

Transfer function decomposition

Recall that $\mathbf{M} = q\mathbf{I} - \mathbf{A}$. Denote the rational transfer function G of $\Sigma_d(\mathbf{A}, \mathbf{B}, \mathbf{C})$ as a fraction with numerator g_N and denominator g_R, i.e.,

$$y_k = \underbrace{\frac{g_N(\vartheta, q)}{g_R(\vartheta, q)}}_{G(\vartheta,q)} u_k \tag{4.8}$$

with g_N and g_R polynomials in ϑ and q.

For instance, Pintelon [82] and many others, split g_N and g_R in functions of the polynomial variable q (or the Laplace variable s) and the parameter vector ϑ, such that

$$\underbrace{\tilde{g}_N(\vartheta) n(q)}_{g_N} u_k = \underbrace{\tilde{g}_R(\vartheta) d(q)}_{g_R} y_k.$$

It is common to treat the entries of the vectors $\tilde{g}_N$ and $\tilde{g}_R$ as black-box parameters for further estimation and prediction. In this thesis, it is tried to prevent the loss of physically interpretable model structure by decomposition of the polynomials g_N and g_R not only in a shift operator dependent, but also in a (physical) parameter dependent part so to obtain polynomial coefficient matrices $\mathbf{N}$ and $\mathbf{R}$. This is done by defining the following vectors with polynomial variables partitioned in a p-dimensional space as follows:

$$\varphi(\vartheta_i) = \begin{pmatrix} \vartheta_i^r \\ \vartheta_i^{r-1} \\ \vdots \\ \vartheta_i \\ 1 \end{pmatrix} \begin{matrix} \Big\} \varphi_1(\vartheta_i) \\ \\ \\ \\ \} \varphi_2 \end{matrix}, \; n \leq r \leq n+1 \tag{4.9a}$$

with $i \in \{1, \ldots, p\}$, and

$$\psi(q) = \begin{pmatrix} 1 \\ q \\ \vdots \\ q^{n-1} \\ q^n \end{pmatrix} \begin{matrix} \Big\} \psi_1(q) \\ \\ \\ \\ \Big\} \psi_2(q) \end{matrix} \tag{4.9b}$$

Hence, the transfer function can then be written as a rational function of 'decomposed' polynomials g_N and g_R (the superscript $\top i$ denotes the ith direction of the transpose of a vector),

$$\begin{aligned} G(\vartheta, q) &= \mathbf{C}\mathbf{M}^{-1}\mathbf{B} \\ &= \frac{g_N(\vartheta, q)}{g_R(\vartheta, q)} \\ &= \frac{\varphi^\top(\vartheta_1) \cdots \left[\varphi^{\top i}(\vartheta_i) \cdots \left[\varphi^{\top p-1}(\vartheta_{p-1}) \left[\varphi^{\top p}(\vartheta_p) \cdot \widetilde{\mathbf{N}} \right] \right] \right] \cdot \psi(q)}{\varphi^\top(\vartheta_1) \cdots \left[\varphi^{\top i}(\vartheta_i) \cdots \left[\varphi^{\top p-1}(\vartheta_{p-1}) \left[\varphi^{\top p}(\vartheta_p) \cdot \widetilde{\mathbf{R}} \right] \right] \right] \cdot \psi(q)} \end{aligned} \tag{4.10}$$

where

$$\varphi \in \mathbb{R}^{(r+1)}, \psi \in \mathbb{R}^{(n+1)} \quad \text{and} \quad \widetilde{\mathbf{N}},\ \widetilde{\mathbf{R}} \in \mathbb{R}^{(r+1)^p \times (n+1)}. \tag{4.11}$$

Note from eqn. (4.10), that the notation $\varphi^\top$, $\varphi^{\top 2}$, $\varphi^{\top p}$ is used for a horizontal vector in 2D, 3D and p-dimensional space, respectively. Notice also that, $\widetilde{\mathbf{N}}$ and $\widetilde{\mathbf{R}}$ are coefficient matrices which map the polynomial variables q and ϑ to g_N and g_R, respectively.

Transfer function reparametrization

A reparametrization of the parameters in G is needed to obtain a linear regression. From eqn. (4.10), it can be observed that cross products of ϑ_i are formed, like $\vartheta_i^v \vartheta_j^w$, $i \neq j$ and with $v,\ w \leq n+1$. For n large, it is a laborious task to find the minimal number of lumped parameters, such that a set of linear regression equations is formed. As an alternative, it is proposed here to estimate parameters ϑ_i^r, $r \leq n+1$ iteratively; i.e. start with initial estimates of $\{\vartheta_2, \cdots, \vartheta_p\}$ in order to estimate ϑ_1.

Hence for the following, we define the coefficient matrices:

$$\mathbf{N} := \left[\varphi^{\top 2}(\vartheta_2) \cdots \left[\varphi^{\top p-1}(\vartheta_{p-1}) \left[\varphi^{\top p}(\vartheta_p) \cdot \widetilde{\mathbf{N}} \right] \right] \right] \tag{4.12}$$

$$\mathbf{R} := \left[\varphi^{\top 2}(\vartheta_2) \cdots \left[\varphi^{\top p-1}(\vartheta_{p-1}) \left[\varphi^{\top p}(\vartheta_p) \cdot \widetilde{\mathbf{R}} \right] \right] \right] \tag{4.13}$$

where

$$\mathbf{N},\ \mathbf{R} \in \mathbb{R}^{r \times n} \quad \text{and} \quad \{\vartheta_2, \ldots, \vartheta_p\} \text{ known.}$$

Before we continue, a simple example is shown to illustrate the above.

EXAMPLE 4.3.1. *Let a discrete-time, LTI state space system* $\Sigma_d(\mathbf{A}, \mathbf{B}, \mathbf{C})$ *be defined with:*

$$\mathbf{A} = \begin{pmatrix} \vartheta_2 & 1 \\ 1 & \vartheta_1 \end{pmatrix},\ \mathbf{B} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \textit{and} \quad \mathbf{C} = \begin{pmatrix} 1 & 0 \end{pmatrix}.$$

Then the transfer function of this system reads $G(\vartheta, q) = (q - \vartheta_1)/(q^2 - (\vartheta_1 + \vartheta_2)q + \vartheta_1\vartheta_2 - 1)$ *with the cross term* $\vartheta_1\vartheta_2$ *as indicated above. After fixing* ϑ_2*, we obtain the coefficient matrices* **N** *and* **R***:*

$$\mathbf{N} = \begin{pmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad \textit{and} \quad \mathbf{R} = \begin{pmatrix} 0 & 0 & 0 \\ \vartheta_2 & -1 & 0 \\ -1 & -\vartheta_2 & 1 \end{pmatrix}.$$

□

Eqn. (4.10) is now rewritten into the form $\theta^\top \phi(\cdot) = \gamma(\cdot)$ as in eqn. (4.2). **N** and **R** are partitioned as follows (the dimensions of the matrix entries are shown on top and on the right side),

$$\mathbf{N} = \begin{array}{c} \begin{matrix} n & 1 \end{matrix} \\ \left(\begin{array}{cc} \mathbf{N}_{11} & \mathbf{N}_{12} \\ \mathbf{N}_{21} & \mathbf{N}_{22} \end{array} \right) \begin{matrix} r \\ 1 \end{matrix} \end{array} = \begin{pmatrix} \mathbf{N}_1 \\ \mathbf{N}_2 \end{pmatrix}_{(r+1)\times(n+1)} \tag{4.14a}$$

and

$$\mathbf{R} = \begin{array}{c} \begin{matrix} n & 1 \end{matrix} \\ \left(\begin{array}{cc} \mathbf{R}_{11} & \mathbf{R}_{12} \\ \mathbf{R}_{21} & \mathbf{R}_{22} \end{array} \right) \begin{matrix} r \\ 1 \end{matrix} \end{array} = \begin{pmatrix} \mathbf{R}_1 \\ \mathbf{R}_2 \end{pmatrix}_{(r+1)\times(n+1)} \tag{4.14b}$$

Recognize that by Definition 4.2.1, we have

$$\mathbf{U}_{k\cdots k+n} = \psi(q)u_k \quad \text{and} \quad \mathbf{Y}_{k\cdots k+n} = \psi(q)y_k. \tag{4.15}$$

Define vector functions ϕ and γ,

$$\phi(\mathbf{Z}_{k\cdots k+n}) := \begin{pmatrix}\mathbf{N}_1 & -\mathbf{R}_1\end{pmatrix}\mathbf{Z}_{k\cdots k+n} \tag{4.16a}$$

$$\gamma(\mathbf{Z}_{k\cdots k+n}) = \begin{pmatrix}-\mathbf{N}_2 & \mathbf{R}_2\end{pmatrix}\mathbf{Z}_{k\cdots k+n} \tag{4.16b}$$

For the remaining step, let the parameter vector be defined by φ_1 in eqn. (4.9a):

$$\theta = \theta_1 := \begin{pmatrix}\vartheta_1^r \\ \vartheta_{r-1} \\ \vdots \\ \vartheta_1\end{pmatrix} = \varphi_1(\vartheta_1). \tag{4.17}$$

Then substitution of eqs. (4.16a), (4.16b) and (4.17) in G eqn. (4.10) with coefficient matrices as in eqs. (4.12) and (4.13) and $\varphi(\vartheta_2), \ldots, \varphi(\vartheta_p)$ known, leads to,

$$\begin{aligned}
0 &= \varphi^\top(\vartheta)\cdot\begin{pmatrix}\mathbf{N} & -\mathbf{R}\end{pmatrix}\cdot\mathbf{Z}_{k\cdots k+n} \\
&= \begin{pmatrix}\varphi_1(\vartheta_i)^\top & \varphi_2\end{pmatrix}\cdot\begin{pmatrix}\mathbf{N}_1 & -\mathbf{R}_1 \\ \mathbf{N}_2 & -\mathbf{R}_2\end{pmatrix}\cdot\mathbf{Z}_{k\cdots k+n} \\
&= \begin{pmatrix}\theta^\top & 1\end{pmatrix}\cdot\begin{pmatrix}\phi(\mathbf{Z}_{k\cdots k+n}) \\ \gamma(\mathbf{Z}_{k\cdots k+n})\end{pmatrix} \\
&= \theta^\top\phi(\mathbf{Z}_{k\cdots k+n}) - \gamma(\mathbf{Z}_{k\cdots k+n})
\end{aligned} \tag{4.18}$$

Hence, in eqn. (4.18) we have arrived at the linear regression equation $\theta^\top\phi = \gamma$ of Σ_d^{LR}.

Definition of output prediction

In this subsection, we study the deduction of a linear regressive output predictor $\hat{y}(k|\hat{\theta}; \mathbf{Z}_{k-n\cdots k-1})$.

Eqn. (4.18), obtained in step (2), is in fact the linear regressive equation suitable for parameter estimation in eqn. (4.2). By rearranging terms, this same equation can be used for prediction as well. Therefore, γ is first written out as follows:

$$\begin{aligned}
\gamma(\mathbf{Z}_{k\cdots k+n}) &= \begin{pmatrix}-\mathbf{N}_{21} & \mathbf{R}_{21}\end{pmatrix}\mathbf{Z}_{k\cdots k+n-1} + \begin{pmatrix}-\mathbf{N}_{22} & \mathbf{R}_{22}\end{pmatrix}\mathbf{Z}_{k+n} \\
&= \begin{pmatrix}-\mathbf{N}_{21} & \mathbf{R}_{21}\end{pmatrix}\mathbf{Z}_{k\cdots k+n-1} + \begin{pmatrix}-\mathbf{N}_{22} & \mathbf{R}_{22}\end{pmatrix}\begin{pmatrix}u_{k+n} \\ y_{k+n}\end{pmatrix} \\
&= \begin{pmatrix}-\mathbf{N}_{21} & \mathbf{R}_{21}\end{pmatrix}\mathbf{Z}_{k\cdots k+n-1} - \mathbf{N}_{22}u_{k+n} + \mathbf{R}_{22}y_{k+n}
\end{aligned} \tag{4.19}$$

Finally, assume $\mathbf{R}_{22}$ is invertible and define:

$$\phi' := \left(\mathbf{R}_{22}\right)^{-1} \phi \tag{4.20a}$$

$$\gamma' := \left(\mathbf{R}_{22}\right)^{-1} \left(\begin{pmatrix} -\mathbf{N}_{21} & \mathbf{R}_{21} \end{pmatrix} \mathbf{Z}_{k\cdots k+n-1} - \mathbf{N}_{22} u_{k+n} \right) \tag{4.20b}$$

In what follows, it will become clear that $\mathbf{N}_{22} = 0$. Consequently, by substituting eqs. (4.20a) and (4.20b) in eqn. (4.18), eqn. (4.18) reads:

$$0 = \theta^\top \phi'(\mathbf{Z}_{k\cdots k+n}) - \gamma'(\mathbf{Z}_{k\cdots k+n-1}) - y_{k+n-1} \tag{4.21}$$

After multiplication of eqn. (4.21) by q^{-n} (i.e. a backward time shift is applied) and rearrangement of terms, the predictor equation in eqn. (4.2) is obtained.

Properties of linear regressive realization

Until now, we have obtained the linear regression equation and the output predictor equation in Σ_d^{LR} from the state space description Σ_d. The existence and precise formulation of a realization Σ_d^{LR} of Σ_d is given in the following proposition.

PROPOSITION 4.3.1. *Given system Σ_d as in (4.1) and $\vartheta \in \mathbb{R}$. Then,*

(i) exact expressions of $g_N(\vartheta, q)$ and $g_R(\vartheta, q)$ as in eqn. (4.8) as a function of n exist, if the system matrix $\mathbf{A}$ can be written in the form of eqs. (4.5) *and* (4.6) *or eqs.* (4.5) *and* (4.7).

(ii) Σ_d eqn. (4.1) can be written in the form of Σ_d^{LR} defined in eqn. (4.2) with $\theta_i = \varphi(\vartheta_i)$, a vector consisting of polynomials and:

$$\phi'(\mathbf{Z}_{k\cdots k+n}) = \phi = \begin{pmatrix} \mathbf{N}_1 & -\mathbf{R}_1 \end{pmatrix} \mathbf{Z}_{k-n\cdots k} \tag{4.22a}$$

$$\gamma(\mathbf{Z}_{k\cdots k+n}) = \begin{pmatrix} -\mathbf{N}_{21} & \mathbf{R}_{21} \end{pmatrix} \mathbf{Z}_{k\cdots k+n-1} + y_k \tag{4.22b}$$

$$\gamma'(\mathbf{Z}_{k\cdots k+n}) = \begin{pmatrix} -\mathbf{N}_{21} & \mathbf{R}_{21} \end{pmatrix} \mathbf{Z}_{k-n-1\cdots k-1} \tag{4.22c}$$

Proof.

(i) Let $\mathbf{M} = q\mathbf{I} - \mathbf{A}$ and the determinant of $\mathbf{M}$ be denoted by $M = \det(\mathbf{M})$. Let $\mathbf{M}_{i'j'}$ and $\mathbf{A}_{i'j'}$ denote the submatrix of $\mathbf{M}$ and $\mathbf{A}$ respectively, both resulting from the deletion of row i and j. By Laplace expansion, the determinant of $\mathbf{M}$ is given by

$$\begin{aligned} M &= \sum_{j=1}^{n} (-1)^{i+j} m_{ij} M_{i'j'} \\ &= \sum_{i=1}^{n} (-1)^{i+j} m_{ij} M_{i'j'} \end{aligned}$$

$\forall i \leq n,\ j \leq n$ with $M_{i'j'} = \det(\mathbf{M}_{i'j'})$ assumed to be known. For any choice of row or column either expansion yields the determinant. The classical adjoint of $\mathbf{M}$ is defined by the transposed matrix of co-factors denoted by $\operatorname{adj}\mathbf{M}$, with entries $\operatorname{adj}\mathbf{M} = \sum_{i=1}^{n}(-1)^{i+j}\mathbf{M}_{ij'}$. The inverse of $\mathbf{M}$ is $\mathbf{M}^{-1} = \operatorname{adj}\mathbf{M}/M$, iff $\mathbf{M}$ non-singular. Because $\mathbf{A}$ is linear in ϑ (assumption A3, eqn. (4.3)), $\mathbf{M}_{i'j'} = q\mathbf{I}_{i'j'} - \overline{\mathbf{A}}_{i'j'} - \sum_{i=1}^{p}\widetilde{\mathbf{A}}_{i'j'}\vartheta_i$. From induction it follows that the determinant of $\mathbf{M}$ will be a polynomial in q and ϑ with maximal order n and $\operatorname{adj}\mathbf{M}$ a matrix filled with polynomials in q and ϑ with maximal order $n-1$. Since $G = \mathbf{C}(q\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} = \mathbf{C}\mathbf{M}^{-1}\mathbf{B}$ and both $\mathbf{B}$ and $\mathbf{C}$ linear in ϑ, G becomes a rational function of polynomials in q and ϑ. It follows that $g_N(\vartheta, q)$ and $g_D(\vartheta, q)$ can be decomposed as in eqn. (4.10).

(ii) Given $G(\vartheta, q) = g_N(\vartheta, q)/g_R(\vartheta, q)$, the transfer function of Σ_d. Therefore, $g_R(\vartheta, q)y_k = g_N(\vartheta, q)u_k$. Reparametrization of the numerator and denominator polynomials as in step (2) is possible under the condition that ϑ is a scalar. From part (i) it follows that the polynomial in the numerator of the transfer function G of Σ_d, i.e. $g_N(\vartheta, q)$, has maximal order $n-1$ in q. Hence, the column vector $\mathbf{N}_{\bullet 2} = \mathbf{0}^{(r\times 1)}$ in the coefficient matrix $\mathbf{N}$ is zero, i.e. $\mathbf{N}_{22} = 0$ also. With rearrangement of terms as outlined in step (2), one readily obtains $\theta^\top \phi_k = \gamma_k$ with $\theta_i = \varphi(\vartheta_i)$. Since $\mathbf{N}_{22} = 0$, γ and γ' in eqs. (4.16b) and (4.20b), respectively, change accordingly. Furthermore, it is always possible to make $g_R(\vartheta, q)$ monic, thus $\mathbf{R}_{22} \equiv 1$. Consequently, ϕ as in eqn. (4.16a) equals ϕ'. In step (3) of the reparametrization, $\mathbf{R}_{22}$ in eqs. (4.20a) and (4.20b) is always a non-zero constant, since $\mathbf{B} \neq \mathbf{0}$ and $\mathbf{C} \neq \mathbf{0}$. Notice that γ_k contains the output at the last time instant, y_{k+n}, and the equivalent form Σ_d^{LR} (4.2) of the state-space system Σ_d is obtained after multiplication with q^{-n}.

□

REMARK 1 [CONSEQUENCES OF PROPERTIES G OF Σ_d]. *From the proof of Proposition 4.3.1(i) it follows that the polynomial degree of $g_N(\vartheta, q)$ in q is determined by* $\mathbf{B}$ *or* $\mathbf{C}$ *and* $\operatorname{adj}\mathbf{M}$. *Let the nonzero entry of* $\mathbf{B}$ *be* b_i *and the nonzero entry of* $\mathbf{C}$ *be* c_j. *Define* $l = \min_{l\in\{1,\dots n\}}(i, j)$. *Then, from part (i) of the proof of Proposition 4.3.1, the maximum degree of polynomials in q is $n-l$, $1 \leq l \leq n$. Hence, the last l columns* $\mathbf{N}$ *are filled with zeros. The upper bound on l is not equal to n since G is always strictly proper by strict causality of Σ^d.*

Whenever the system matrix $\mathbf{A}$ is a tridiagonal matrix as in Assumption A4 (see eqns. (4.5)–(4.7)), explicit expressions of n of the matrix entries in $\mathbf{M}$ exist [47, 48, 67, 85]. Consequently, Proposition 4.3.1 suggests that the entries of the polynomial coefficient matrices $\mathbf{N}$ and $\mathbf{R}$ in eqn. (4.10) can be written by explicit relations on n. We come to the following remark to accomplish Σ_d^{LR} when starting from a discretized (CDR) state space description Σ_d with $\mathbf{A}$ being of Type I or Type II.

REMARK 2 [EXPLICIT FORM OF Σ_d^{LR}]. *For (discrete-time) CDR–systems for* $\mathbf{A}$ *being of*

Type I: *that is, where* $\vartheta_1 = \vartheta_2$ *as in eqn. (4.6) one may write:*

$$\mathbf{M} = q\mathbf{I} - \mathbf{A}(\vartheta_1) = \begin{pmatrix} \tilde{\vartheta} & 1 & 0 & \cdots & 0 \\ 1 & \tilde{\vartheta} & 1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & 1 & \tilde{\vartheta} & 1 \\ 0 & \cdots & 0 & 1 & \tilde{\vartheta} \end{pmatrix}, \text{ where } \tilde{\vartheta} := q - \vartheta_1 \qquad (4.23)$$

Explicit expressions for $\mathbf{M}^{-1}$*, with* $\mathbf{M}$ *as in eqn. (4.23) are reported in [47]. With mathematical software tools like MAPLE or Mathematica, these expressions can be easily written in the time-shift and parameter 'decomposed' form of the transfer function eqn. (4.10).*

Type II: *that is, where* $\vartheta_1 \neq \vartheta_2$ *in eqn. (4.5). By close examination of the determinant in Type II (*$\vartheta_1 \neq \vartheta_2$*), it appears that an iterative procedure —as suggested in step (2) of the general case— for parameter estimation is not needed. Denote* $\mathbf{M}_I^{n\times n}$ *for Type I and* $\mathbf{M}_{II}^{n\times n}$ *for Type II, with* $\mathbf{A} \in \mathbb{R}^{n\times n}$*. Furthermore, let the corresponding determinants be* $M_I^n := \det(\mathbf{M}_I^{n\times n})$ *and* $M_{II}^n := \det \mathbf{M}_I I^{n\times n}$*. The determinant* M_{II}^n *can be written in terms of the under-determinants* M_I^{n-1} *and* M_I^{n-2}*, namely,*

$$M_{II}^n = \vartheta_2 M_I^{n-1} - (1)^2 M_I^{n-2} = \vartheta_2 M_I^{n-1} - M_I^{n-2} \qquad (4.24)$$

Hence, the numerator and denominator of $G_{II} = \mathbf{C}\mathbf{M}_{II}^{-1}\mathbf{B}$ *can be written in determinants and adjoints of* $\mathbf{M}_I$ *and a corresponding* θ *is straightforwardly found.*

Remark 2 is further illustrated by two diffusion examples in Section 4.4. Usually, the sub diagonals of $\mathbf{A}$ as in eqn. (4.5) are not equal to one, i.e. $a_{i\pm1,i} \neq 1$. In that case, the matrix $\mathbf{A}$ can be suitably scaled by $a_{i\pm1,i}$. In cases where $a_{i\pm1,i}$ is an unknown (physical) parameter ϑ_i, one is restricted to an iterative parameter estimation procedure.

Before continuing, it is important to check whether the regression parameters θ are identifiable, i.e. whether θ can be determined *uniquely* from the data. This can be checked a priori by the concepts worked out in [45, 60]. Alternatively, the linear regressive form of Σ_d allows a simple check of identifiability if θ is estimated by least-squares techniques.

4.3.2 Parameter estimation by Least-Squares

It is straightforward to write the linear regression equations in Σ_d^{LR} as a convex parameter estimation problem, i.e. the well-known least squares formulation (see

also Appendix B.2). Define to this aim $\mathbf{Z}_1 := \mathbf{Z}_{1\cdots n}$, $\mathbf{Z}_2 = \mathbf{Z}_{2\cdots n+1}$, ..., and:

$$\mathbf{B} \quad \text{with} \quad b_{i*} \neq 0, \quad \mathbf{C} \quad \text{with} \quad c_{j*}, \quad i^*, j^* \in \{1, \ldots, n\} \tag{4.25}$$

Let furthermore,

$$\Phi = \begin{bmatrix} \phi(\mathbf{Z}_1) & \phi(\mathbf{Z}_2) & \cdots & \phi(\mathbf{Z}_N) \end{bmatrix} \tag{4.26}$$

$$\Gamma = \begin{bmatrix} \gamma(\mathbf{Z}_1) & \gamma(\mathbf{Z}_2) & \cdots & \gamma(\mathbf{Z}_N) \end{bmatrix} \tag{4.27}$$

with γ and ϕ as in eqn. (4.22). Then, the least-squares (LS) estimate $\hat{\theta}$ of the equality $\theta^\top \Phi = \Gamma$, with Φ as in eqn. (4.26) and Γ as in eqn. (4.27) is given by: $\hat{\theta} = (\Phi\Phi^\top)^{-1}\Phi\Gamma^\top$. For the existence of a LS estimate, it is necessary that $\ker(\Phi^\top) = 0$ and $\Gamma \neq \mathbf{0}^{1\times N}$.

REMARK 3 [IDENTIFIABILITY FOR $\vartheta \in \mathbb{R}$]. $\Sigma_d(\mathbf{A}, \mathbf{B}, \mathbf{C})$ *under assumptions A1–A4 and affine in a scalar* ϑ *and* $\mathbf{B}$, $\mathbf{C}$ *as in eqn. (4.25), is identifiable if* $\text{rank}(\mathbf{R}) = n$. *This follows from the condition* $\ker(\Phi^\top) = 0$.

4.3.3 Sensitivity Analysis

So far, we have derived (i) the model structure representation $\Sigma_d^{\text{LR}}(\theta)$, with θ depending on the physical ϑ, and, (ii) an identifiability test. Let us now, on the basis of this, investigate its sensitivities. The sensitivities of the response function γ with respect to the regression parameters θ and the inputs $\mathbf{U}$ allows a *practical* identifiability check. The result of this check could be, that for very small parameter sensitivities, one may reconsider the relumping of the parameters, sensor placement, number of compartments, or the input signal, since the LS estimation of θ will in that case suffer from bad convergence.

We start by realizing that the regression model of Σ_d exhibits an input-output mapping from $\mathbf{U}_k$ to the regression response γ_k. This becomes clear by substituting the convolution sum, i.e. in matrix form: $\mathbf{Y}_{k-n\cdots k} = \mathbf{H} * \mathbf{U}_{k-n\cdots k}$, in eqn. (4.16b) with $\mathbf{H} \in \mathbb{R}^{n_H \times n_H}$. The entries of the Hankel matrix $\mathbf{H}$ contain the *Markov* parameters and can be directly found by an impulse response as a function of the real parameters as θ_0. Consequently, for $n_H = n$ (notice that n_H should be large in order to satisfactorily approximate the infinite impulse response, so that n will then be limiting):

$$\phi \cdot \mathbf{Z}_{k-n\cdots k} = \begin{pmatrix} \mathbf{N}_1 & -\mathbf{R}_1 \end{pmatrix} \cdot \begin{pmatrix} \mathbf{U}_{k-n\cdots k} \\ \mathbf{H} * \mathbf{U}_{k-n\cdots k} \end{pmatrix} \tag{4.28}$$

Further, denote the real (unknown) parameters θ_0. We obtain the parametric sensitivities of the response γ by substitution of eqn. (4.28) in eqn. (4.2) and differenti-

ation, which gives

$$\gamma_\theta := \frac{\mathrm{d}\gamma}{\mathrm{d}\theta} = \phi = \begin{pmatrix}\mathbf{N}_1 & -\mathbf{R}_1\end{pmatrix}\begin{pmatrix}\mathbf{I}\\ \mathbf{H}\end{pmatrix} * \mathbf{U}_{k-n\cdots k} \tag{4.29}$$

$$\gamma_{\theta U} := \frac{\mathrm{d}\gamma_\theta}{\mathrm{d}\mathbf{U}} = \begin{pmatrix}\mathbf{N}_1 & -\mathbf{R}_1\end{pmatrix}\begin{pmatrix}\mathbf{I}\\ \mathbf{H}\end{pmatrix} = \mathbf{N}_1 - \mathbf{R}_1\mathbf{H}. \tag{4.30}$$

An explicit solution for γ_θ and $\gamma_{\theta U}$ is found for a compartmental diffusion system in the next section.

4.4 Application to discretized diffusion systems

A discrete-time compartmental system with unknown (lumped) physical parameters ϑ belongs to the model set Σ_d. First, we present explicit expressions for the coefficient matrices $\mathbf{N}$ and $\mathbf{R}$ involved in the transfer function G with $\mathbf{A}$ being of Type I or II (see also Assumption A4). These results will then be used in a subsequent section where a diffusion example for Type I is worked out in a parameter estimation problem.

4.4.1 Diffusion example I and II

We use the same naming conventions (Example I and II) for diffusion systems characterized by $\mathbf{A}$ of Type I and II, respectively. First, these examples are introduced.

Example I: boundary control system with Dirichlet conditions. Consider an infinite- dimensional system Σ_I^e of parabolic type on $[0,\infty)\times[0,\infty)$, see e.g. [21].

$$\Sigma_I^e : \begin{cases} \frac{\partial z}{\partial t}(\xi,t) & = \alpha^2 \frac{\partial^2 z}{\partial \xi^2}(\xi,t), \quad z(\xi,0) = z_0(\xi), \\ z(0,t) & = u(t) \\ y(t) & = z(\xi^*,t) \end{cases} \tag{4.31}$$

where $z_0(\xi) \in L_2(0,\infty)$, $\xi^* \in [0,\infty)$ and $U \in \mathbb{R}$.

The solution to this problem when applying a step input $u(t) = 1_{[0,\infty)}(t)$ is well-known and is given by the model output at $\xi = \xi^*$,

$$z(\xi^*,t) = \operatorname{erfc}\left(\frac{\xi^*}{2\alpha\sqrt{t}}\right) 1_{[0,\infty)}(t) \tag{4.32}$$

where α^2 can be interpreted as the diffusion constant. With this system some estimations of α are worked out in a subsequent section.

Example II: boundary control system with Neumann condition. This case is only to show how a Neumann condition affects the matrix $\mathbf{M}_{II}$.

We can approximate the following distributed parameter system,

$$\Sigma^e_{II}: \begin{cases} \frac{\partial}{\partial t}z(\xi,t) & = \alpha^2 \frac{\partial^2}{\partial \xi^2}z(\xi,t), \quad z(\xi,0)=z_0(\xi), \\ \frac{\partial}{\partial \xi}z(0,t) & = \frac{1}{\kappa}\left(z(0,t)-u(t)\right) \\ \frac{\partial}{\partial \xi}z(\infty,t) & = 0 \\ y(t) & = z(\xi^*,t) \end{cases} \tag{4.33}$$

where $z_0(\xi) \in L_2(0,\infty)$, $\xi^* \in [0,\infty)$ and $U=\mathbb{R}$, by a discrete-time compartmental system. This is shown in what follows.

In addition to Example I, we have a second parameter κ, which may correspond to a specific heat capacity divided by a thermal conductivity.

REMARK 4 [CDR SYSTEMS WITH POINT ACTUATION/OBSERVATION]. *Although this example is not further worked out, it is important to realize that CDR systems* where convection and reaction is included *and with one point actuation and measurement can be described by eqn. (4.1) with* $\mathbf{A}$ *as given in eqns. (4.5)–(4.7) via the use of centralized finite differences.*

4.4.2 Explicit structures

Here, we present explicit expressions for Example I and II.

Type I

Using centralized finite differences, we get for Example I (number of (physical) parameters $p=1$) the following system representation:

$$\Sigma^e_{d,I} := \begin{cases} \mathbf{x}_{k+1} & = \mathbf{A}(\vartheta)\mathbf{x}_k + \mathbf{B}(\vartheta)\mathbf{u}_k, \quad \mathbf{x}_0 = \mathbf{0} \\ y_k & = \mathbf{C}\mathbf{x}_k \end{cases} \tag{4.34}$$

where the discretized states are represented by $\mathbf{x}$ and the lumped parameter $\vartheta = \Delta\alpha^2$ consisting of the discretization parameter $\Delta := \Delta_t/\Delta_\xi^2$ and a diffusion coefficient α^2. The noise-free observation $y(t)=z(\xi^*,t)$ is approximated by $\mathbf{C}\mathbf{x}_k$, with $\mathbf{C}$

a matrix mapping a 'point' observation at the j^*-th compartment (i.e. $c_{j^*} = 1$).

$$\mathbf{A}(\vartheta) = \begin{pmatrix} 1-2\vartheta & \vartheta & 0 & \cdots & 0 \\ \vartheta & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \vartheta & 1-2\vartheta & \vartheta \\ 0 & \cdots & 0 & \vartheta & 1-2\vartheta \end{pmatrix} \tag{4.35a}$$

$$\mathbf{B}(\vartheta) = \begin{pmatrix} \vartheta & 0 & \cdots & 0 \end{pmatrix}^\top \quad \text{and} \quad \mathbf{C} = \begin{pmatrix} \mathbf{0}^{1\times j-1} & 1 & \mathbf{0}^{1\times n-j} \end{pmatrix} \tag{4.35b}$$

and with $j \in \{1, 2, \ldots, n\}$ and $\mathbf{0}^{1\times 0}$ an empty vector.

The resolvent of $\mathbf{A}$ is written as $q\mathbf{I} - \mathbf{A} = \mathbf{M}_I$. On the main diagonal of $\mathbf{M}_I$ we encounter $m_{ii} = \big((q-1)/\vartheta + 2\big)\,\vartheta$. Notice that a grid with n points directly leads to n states, because we have started with one state variable and one spatial direction in the PDE model Σ_I^e.

Hence, $\mathbf{M}_I = q\mathbf{I} - \mathbf{A}$ can be written in the form of eqn. (4.23) with $\tilde{\vartheta} = (q-1)/\vartheta + 2$. Notice furthermore that Σ_I^e satisfies Assumptions A1–A5. With respect to Assumption A3, it is briefly shown here that $\mathbf{A}$ and $\mathbf{B}$ are affine in ϑ:

$$\mathbf{A}(\vartheta) := \overline{\mathbf{A}} - \widetilde{\mathbf{A}}\vartheta, \quad \text{with} \quad \overline{\mathbf{A}} = \mathbf{I} \quad \text{and} \quad \widetilde{\mathbf{A}}_1 = -2\mathbf{I} + \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 1 & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & 1 & 0 & 1 \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix} \quad \text{and} \quad \mathbf{B} := \begin{pmatrix} 1 \\ \vdots \\ 0 \end{pmatrix} \vartheta. \tag{4.36}$$

The vectors φ_I, ψ_I and ϑ_I for Type I is given as follows:

$$\varphi_I := \begin{pmatrix} \theta_I \\ 1 \end{pmatrix}, \quad \text{with} \quad \theta_I := \begin{pmatrix} \vartheta^n \\ \vartheta^{n-1} \\ \vdots \\ \vartheta \end{pmatrix} \qquad \text{and} \qquad \psi_I := \begin{pmatrix} 1 \\ q \\ \vdots \\ q^n \end{pmatrix}. \tag{4.37}$$

Denote a shorthand notation to indicate the dimensions of square coefficient matrices $\mathbf{N}$ and $\mathbf{R}$ by using the superscript, i.e. $\dim \mathbf{R}^{(n+1)} = (n+1)\times(n+1)$. We can now write the coefficient matrices $\mathbf{N}$ and $\mathbf{R}$ in eqn. (4.10) for this diffusion example as follows.

PROPOSITION 4.4.1. *Given the symmetric 3-banded Toeplitz matrix $\mathbf{M}_I(\vartheta, q) = q\mathbf{I} - \mathbf{A}$ with $\mathbf{A}$ as in eqn. (4.35a), a newly defined parameter vector ϑ_I as in eqn. (4.37), then*

the entries for the coefficient matrices $\mathbf{N}_I$ *and* $\mathbf{R}_I$ *of* $\mathbf{M}_I$ *are given by,*

$$\mathbf{R}_I^{(n+1)}(i,j) = \begin{cases} (-1)^{i+j-2}\binom{n+i}{n-i+1}\binom{i-1}{j-1} & \text{if} \quad i \geq j \\ 0 & \text{elsewhere} \end{cases}$$

$$\mathbf{N}_I^{(n+1)}(i^*,j^*) = \sum_{s=0}^{\min(\imath,\jmath)-1} \mathbf{Q}_I^{[s,i^*,j^*]},$$

where,

$$\imath = \begin{cases} i^* & \text{if} \quad i^* \leq \frac{n}{2} \\ n-i^*+1 & \text{elsewhere} \end{cases}, \qquad \jmath = \begin{cases} j^* & \text{if} \quad j^* \leq \frac{n}{2} \\ n-j^*+1 & \text{elsewhere} \end{cases}$$

$$l_{[s,i^*,j^*]} = l^* + 2s + 1, \qquad \text{with} \qquad l^* = \begin{cases} |i^*-j^*| & \text{if} \quad i^* < \frac{n}{2} \\ |j^*-i^*| & \text{elsewhere} \end{cases}$$

$$\mathbf{Q}_I^{[s,i^*,j^*]} = \begin{pmatrix} \mathbf{0}^{(1\times k_2)} & \mathbf{0}^{(1\times(k_1-k_2))} \\ \mathbf{R}_I^{k_2} & \mathbf{0}^{(k_2\times(k_1-k_2))} \\ \mathbf{0}^{((k_1-k_2)\times k_2)} & \mathbf{0}^{((k_1-k_2)\times(k_1-k_2))} \end{pmatrix}$$

with s a counter index and

$$k_1 = (n - l_{[s,i^*,j^*]}) \in \{1,\cdots,n+1\}, \qquad k_2 = l_{[s,i^*,j^*]}.$$

Proof. The structure of **R** follows from the proof of [47]. The relations for **N** follow from the proof of Proposition 4.3.1 and straightforward algebraic calculations. □

Type II

In Type II (number of (physical) parameters $p = 2$), the discretized system becomes:

$$\Sigma_{d,II}^e := \begin{cases} \mathbf{x}_{k+1} & = \mathbf{A}(\vartheta)\mathbf{x}_k + \mathbf{B}(\vartheta)\mathbf{u}_k, \quad \mathbf{x}_0 = \mathbf{0} \\ y_k & = \mathbf{C}\mathbf{x}_k \end{cases} \tag{4.38}$$

where

$$\mathbf{A} = \begin{pmatrix} 1+\vartheta_2-2\vartheta_1 & \vartheta_1 & 0 & \cdots & 0 \\ \vartheta_1 & 1-2\vartheta_1 & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \vartheta_1 & 1-2\vartheta_1 & \vartheta_1 \\ 0 & \cdots & 0 & \vartheta_1 & 1-2\vartheta_1 \end{pmatrix} \tag{4.39a}$$

and

$$\mathbf{B} = \begin{pmatrix} \vartheta & 0 & \cdots & 0 \end{pmatrix}^\top \quad \text{and} \quad \mathbf{C} = \begin{pmatrix} \mathbf{0}^{1\times j-1} & 1 & \mathbf{0}^{1\times n-j} \end{pmatrix}. \tag{4.39b}$$

Again, $j \in \{1,\dots,n\}$ and $\mathbf{0}^{1\times 0}$ is defined as an empty vector, $\vartheta_1 = \Delta\alpha^2$ a lumped parameter with discretization parameter $\Delta = \Delta_t/\Delta_\xi^2$, diffusion coefficient α^2. $\vartheta_2 = (\Delta_\xi/\kappa - 1)\Delta$ is another lumped parameter with the physical parameter κ. The input mapping matrix $\mathbf{B}$ is similar as in [92]. Further, φ is compatibly defined, i.e. it is similar as in 4.41 and scaled with $1/\kappa$ due to the definition of $\mathbf{B}$.

On the main diagonal of $\mathbf{M}_{II}$, with $\mathbf{M}_{II} = qI - \mathbf{A}$, we now encounter: $m_{11} = \big((q-1+\vartheta_2)/\vartheta_1 + 2\big)\,\vartheta_1$ and $m_{ii} = \big((q-1)/\vartheta_1 + 2\big)\,\vartheta_1$, where $i > 1$. The subdiagonals of $\mathbf{M}_I$ read $m_{i\pm1} = -\vartheta_1$. Again, $\mathbf{M}_{II}$ can be written similar to the form of eqn. (4.23) to obtain $m_{i\pm1} = 1$. Notice that Σ_{II}^e satisfies Assumptions A1–A4 again (not shown).

In order to write the determinant of $\mathbf{M}_{II}$ as being a function of subdeterminants of $\mathbf{M}_I$ according to eqn. (4.24), we define a new lumped parameter:

$$\vartheta_2' = -\frac{1}{\vartheta_1} + \frac{\vartheta_2}{\vartheta_1^2} + 2 \tag{4.40}$$

The vectors φ_{II}, ψ_{II} and ϑ_{II} are then given as follows:

$$\varphi_{II} := \begin{pmatrix} \theta_{II} \\ 1 \end{pmatrix}, \qquad \theta_{II} := \begin{pmatrix} \frac{\vartheta_2'}{\vartheta_1}\varphi_I \\ \hline \theta_I \end{pmatrix} = \begin{pmatrix} \frac{\vartheta_2'}{\vartheta_1}\theta_I \\ \frac{\vartheta_2'}{\vartheta_1} \\ \hline \theta_I \end{pmatrix} \qquad \text{and} \qquad \psi_I := \begin{pmatrix} 1 \\ q \\ \vdots \\ q^n \end{pmatrix} \tag{4.41}$$

For the inverse of $\mathbf{M}_{II}$, we propose the following (the proof goes analogously to the proof of Proposition 4.4.1):

PROPOSITION 4.4.2. *Given the symmetric 3-banded Toeplitz matrix $M_{II}(\vartheta, q)$ with $\mathbf{A}$ as in eqn. (4.39a) and vectors φ_{II}, ψ_{II} and ϑ_{II} as in eqn. (4.41), then the entries for the coefficient matrices $\mathbf{R}_{II}$ and $\mathbf{N}_{II}$ are given by,*

$$\mathbf{R}_{II}^{n+1} = \left(\begin{array}{cc} 0 & \mathbf{0}^{1\times(n-1)} \\ -\mathbf{R}_I^n & \mathbf{0}^{n\times 1} \\ \hline \end{array} \right) \left(\begin{array}{cc} 0 & \mathbf{0}^{1\times n} \\ \mathbf{0}^{n\times 1} & \mathbf{R}_I^n \end{array} \right) - \left(\begin{array}{cc} \mathbf{R}_I^{n-1} & \mathbf{0}^{(n-1)\times 2} \\ \mathbf{0}^{2\times(n-1)} & \mathbf{0}^{2\times 2} \end{array} \right)$$

$$\mathbf{N}_{II}^{n+1}(i^*,j^*) = \begin{cases} \left(\begin{array}{c} 0 \\ \hline -\mathbf{N}_I^{n+1}(1,1) \end{array}\right) & \text{if } i^*=1 \vee j^*=1 \\ \left(\begin{array}{cc} 0 & \mathbf{0}^{1\times n} \\ \mathbf{N}_I^n(1,j^*-1) & \mathbf{0}^{n\times 1} \\ \hline 0 & \mathbf{0}^{1\times n} \\ \mathbf{0}^{n\times 1} & -\mathbf{N}_I^n(i^*-1,j^*-1) \end{array}\right) & \text{if } i^*=2 \wedge j^*\geq 2 \\ \left(\begin{array}{cc} 0 & \mathbf{0}^{1\times n} \\ \mathbf{N}_I^n(i^*-1,j^*-1) & \mathbf{0}^{n\times 1} \\ \hline \multicolumn{2}{c}{\mathbf{Q}_{II}^{n+1}(i^*,j^*)} \end{array}\right) & \text{if } i^*>2 \vee j^*\geq 2 \end{cases}$$

with

$$\mathbf{Q}_{II}^{n+1}(i^*,j^*) := \begin{pmatrix} 0 & \mathbf{0}^{1\times n} \\ \mathbf{0}^{n\times 1} & -\mathbf{N}_I^n(i^*-1,j^*-1) \end{pmatrix} + \begin{pmatrix} \mathbf{0}^{2\times(n-1)} & \mathbf{0}^{2\times 2} \\ \mathbf{N}_I^{n-1}(i^*-2,j^*-2) & \mathbf{0}^{2\times(n-1)} \end{pmatrix}$$

where the coefficient matrices $\mathbf{N}_{II}$ *and* $\mathbf{R}_{II}$ *are divided in blocks which correspond to the blocks in* φ_{II}*, see eqn. (4.41).*

4.4.3 Estimation

We assume the value of ϑ_1 in Example I to be unknown. In this study, we fix $i^* = j^* = 1$ for **B** and **C** vectors, so that for $n = r = 4$, the coefficient matrices read:

$$\mathbf{N} = \begin{pmatrix} 4 & 0 & 0 & 0 & 0 \\ -10 & 10 & 0 & 0 & 0 \\ 6 & -12 & 6 & 0 & 0 \\ -1 & 3 & -3 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \mathbf{R} = \begin{pmatrix} 6 & 0 & 0 & 0 & 0 \\ -35 & 35 & 0 & 0 & 0 \\ 56 & -112 & 56 & 0 & 0 \\ -36 & 108 & -108 & 36 & 0 \\ 10 & -40 & 60 & -40 & 10 \end{pmatrix}$$

We estimate our original parameter α by simulating the system with $\alpha = 0.2$.

(i) non-linear least squares estimation from the disturbed classical solution of Σ_I^e after applying a step input (see the noise-free solution in eqn. (4.32)) and,

(ii) ordinary (linear) least squares estimation given the linear regression form of $\Sigma_{d,I}^e$ according to eqn. (4.2), with additive disturbances.

In the latter case, we approximate α as a weighted average $\hat{\alpha}_{LR} \approx |\frac{1}{n}\sum_{i=1}^{n} \hat{\vartheta}_i^{\frac{1}{2(n-i)}}|$. The system $\Sigma_{d,I}^e$ as in eqn. (4.31) is in case (ii) approximated with 4 compartments ($n = r = 4$). In case (i), the output data $y(t)$ is generated by the solution of Σ_I^e under excitation of a step input $u(t) = 1(t)$, $t > 0$, and sampling the output, i.e. $y(t_k) = z(\xi^*, t_k) + d + e(t_k)$. Here, $z(\xi^*, t_k)$ is calculated from eqn. (4.32), the bias has a range $d \in [-0.2, 0.2]$ and e is a simulated Gaussian noise sequence. In both

cases, $\alpha = 0.2$. The results of the parameter estimations (i) and (ii) are shown in tables 4.1 and 4.2, respectively. It is shown in tables 4.1 and 4.2, that the estimates

Table 4.1: Non-linear LS estimates $\hat{\alpha}_{\text{NLS}}$ of the erfc-model eqn. (4.32) under different bias d and noise variance σ_d after applying a step input.

	$\hat{\alpha}_{\text{NLS}}$				
	$d=-0.2$	$d=-0.1$	$d=0$	$d=0.1$	$d=0.2$
$\sigma_d = 0$:	0.13	0.16	0.20	0.26	0.34
$\sigma_d = 0.01$:	0.13	0.16	0.20	0.26	0.34

Table 4.2: Linear regressive LS estimates $\hat{\alpha}_{\text{LR}}$ of Σ_I^{LR} under different bias d and noise variance σ_d after applying a step input.

	$\hat{\alpha}_{\text{LR}}$				
	$d=-0.2$	$d=-0.1$	$d=0$	$d=0.1$	$d=0.2$
$\sigma_d = 0$:	0.15	0.14	0.20	0.14	0.14
$\sigma_d = 0.01$:	0.62	0.63	0.64	0.68	0.64

of α in the linear regressive compartmental model are rather insensitive to bias, but very sensitive to added Gaussian white noise, when the system is driven by a step input. Interestingly, Table 4.3 illustrates that a pseudo random binary (PRBS) input sequence improves our linear regression results for the noiseless case.

Table 4.3: Linear regressive LS estimates $\hat{\alpha}_{\text{LR}}$ of Σ_I^{LR} under different bias d and noise variance σ_d after applying a PRBS signal e, with switching probability $P = 0.5$.

	$\hat{\alpha}_{\text{LR}}$				
	$d=-0.2$	$d=-0.1$	$d=0$	$d=0.1$	$d=0.2$
$\sigma_e = 0$:	0.20	0.20	0.20	0.20	0.20
$\sigma_e = 0.01$:	0.27	0.24	0.27	0.27	0.27

We observe the opposite when non-linear least squares procedure is used for estimating α given eqn. (4.32), that is, it is sensitive to bias and weakly sensitive to additive measurement errors. Iterative pre-filtering (see e.g. [93]) or more advanced least squares techniques may compensate the colored disturbances in the linear regressive model structure.

4.4.4 Sensitivity results

Evaluation of the sensitivity matrix $\gamma_{\theta U}$ gives us, for a given impulse response as a function of the nominal parameter vector θ_0, valuable information about sensitivi-

ties, independent of θ or $\mathbf{U}$. These sensitivities can be used for numerical identifiablity tests (γ_θ) or input design ($\gamma_{\theta U}$).

For the diffusion system Σ_I^e with $\alpha = 0.2$, $n = n_H = 8$, $\gamma_{\theta U}$ has been calculated at different sensor positions c_{j^*}, $j^* \in \{1, \ldots, n\}$ and the matrix values of $\gamma_{\theta U}$ are depicted in Figure 4.1. Strong peaks indicate high sensitivities. The column matrix entries at the j-axis correspond to time shifting whereas the i-axis correspond to polynomial variables to the power i, i.e. ϑ^{n+1-i}.

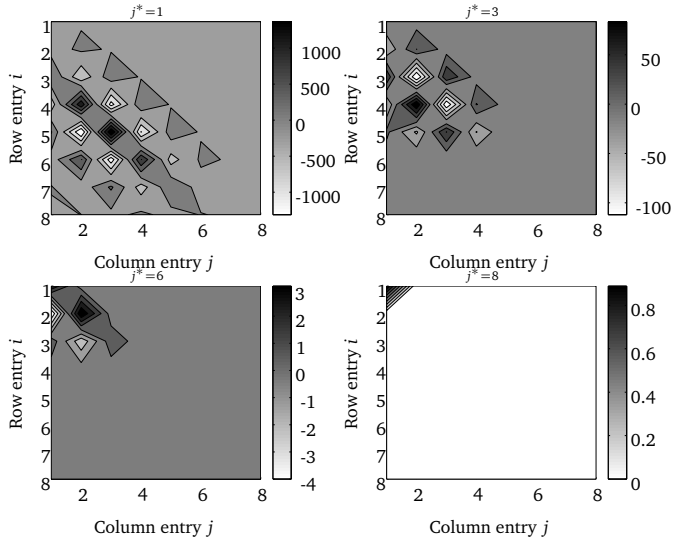


Figure 4.1: Contour plots of sensitivity matrix $\gamma_{\theta U}$ at different sensor positions j^*

Although seeming counter-intuitively, Figure 4.1 shows that it is recommended to put the sensor at the first compartment which is closest to the boundary input, see also [52]. In this compartment, *all* sensitivities are excited, whereas the magnitude rapidly diminishes if the sensor is placed further away.

The eigenvalues in the right graph of Figure 4.2 confirm this finding. The left graph of Figure 4.2 shows the parametric sensitivity γ_θ when applying a step input, which is obtained by summation of the columns of $\gamma_{\theta U}$. Again, it seems favorable to place a sensor in the first compartment.

4.5 Concluding remarks

Summarizing, the sketched methodology is attractive for estimation and prediction as the parameters in the transformed regression space can be determined uniquely. Furthermore, since we keep track of the physical parameters, physical knowledge in the model structure is preserved throughout analysis and estimation.

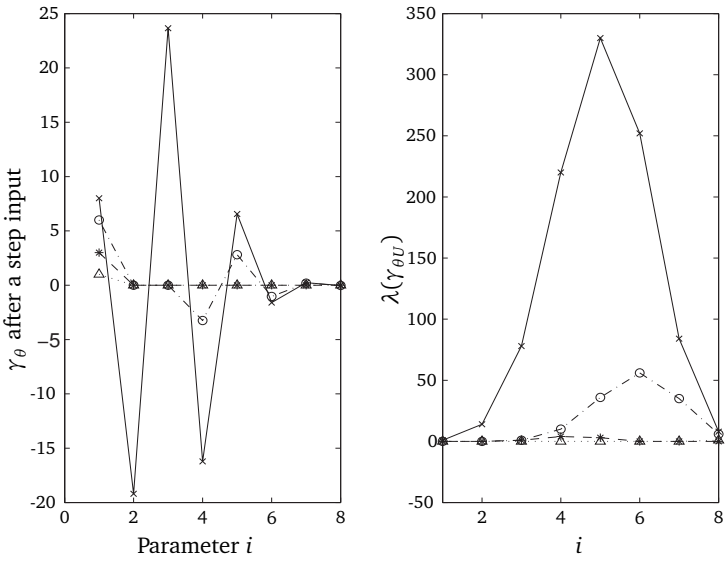


Figure 4.2: Parametric sensitivities γ_θ for a step input and eigenvalues of the sensitivity matrix $\gamma_{\theta U}$, the $(\times)$, $(\circ)$, $(*)$ and $(\triangle)$-marked lines correspond to a sensor at $j^* = 1, 3, 6$ and 8 respectively.

Linear regressive realizations Σ^{LR} have been derived for two discrete time diffusion models with different boundary conditions. The linear regressive structures **N** and **R** have been explicitly expressed in terms of number of compartments n, input and sensor location. These structures are also applicable to (non-dimensionalized) discretized CDR-systems.

A priori parametric sensitivity analysis can be made on the regression vector ϕ which allows for numerical identifiability tests and input design. Also, a simple identifiability test on the basis of a rank test of **R** is obtained.

In practice, however, there will be measurement noise deteriorating both the inputs and outputs of the linear regressive system Σ^{LR}. It is well known that application of ordinary least squares to these types of errors-in-variables problems will lead to bias. This unwanted effect is subject of further study, but can, in principle, be tackled by existing techniques such as Total Least Squares.

5

Boundary Observer Synthesis

This chapter is based on:

- D. Vries, K. J. Keesman, and H. Zwart. “A Luenberger observer for an infinite dimensional bilinear system: a UV disinfection example”. In: *Proceedings of IFAC Symposium on Systems, Structure and Control*. Foz de Iguassu, Brazil 2007
- D. Vries, K. J. Keesman, and H. Zwart. “An H_∞–observer at the boundary of an infinite dimensional system”. In: *Book of Abstracts of 5th IFAC Workshop on Distributed Parameter Systems*. Namur, Belgium 2007
- D. Vries, K. J. Keesman, and H. Zwart. *Luenberger Boundary Observer Synthesis for Sturm-Liouville Systems*. In preparation
- D. Vries, K. J. Keesman, and H. Zwart. *Robust dynamic filter for a Convection-Diffusion-Reaction Process using Point Measurements*. In preparation

Abstract

THE OBSERVER SYNTHESIS OF a typical bilinear system with point measurements and *boundary control* actions is studied. Approximate observability and the existence of a mild solution is proved for this system Σ in boundary control form. Furthermore, two observer design approaches are worked out in the case that Σ contains a Sturm-Liouville operator A. First, detectability and design results for a static boundary observer is presented. Second, a numerical synthesis procedure for a dynamic observer is proposed which ensures robust performance under input and output disturbances. Both approaches are compared and evaluated using the UV disinfection process as a case study.

5.1 Introduction

In this chapter, observer synthesis of a typical bilinear system with *point* measurements and *boundary control* actions is studied. The literature on observers with point measurement has already been discussed in Chapter 1, Section 1.4.4.

The motivation here is, to treat the observer design of a CDR–system with infinite dimensional system theory concepts. Since the observer correction will be formulated at the boundary, the theoretical framework developed for boundary control [33, 36] suffices and provides an elegant and mathematically simple approach for observer design as well. As far as we know, there has been little attention to apply this theory to convection–diffusion–reaction (CDR) type of problems where boundary or point measurements are used for observations. Instead, for analysis of CDR systems, one usually considers a control or observation on a small interval $[0, w]$; see e.g. [25, 112, 114].

With point measurements and boundary control in mind, we make the following choice regarding the observer design: *the output estimation error* $\big(y(t) - \hat{y}(t)\big)$ *is manipulated by the observer at the boundary of the error system.* We will call such an observer a *boundary* observer.

We focus on the boundary observer design problem with $u(t)$ given. To clarify the idea, we present the bilinear system and its observer in abstract boundary control form as follows:

$$\Sigma := \left\{ \begin{array}{rcll} \dot{z}(t) & = & \mathfrak{A}z(t) - b_1 u_1(t) z(t); & z(0) = z_0 \\ \mathfrak{B}z(t) & = & u_2(t) + v_1(t) & \\ \mathfrak{C}z(t) & = & y(t) + v_2(t). & \end{array} \right. \tag{5.1}$$

And, similar to eqn. (5.1), we define the observer system as:

$$\Sigma^{\text{obs}} := \left\{ \begin{array}{rcll} \dot{\hat{z}}(t) & = & \mathfrak{A}\hat{z}(t) - b_1 u_1(t) \hat{z}(t); & \hat{z}(\eta, 0) = \hat{z}_0 \\ \mathfrak{B}\hat{z}(t) & = & u_2(t) + L(t) * \mathfrak{C}(z(t) - \hat{z}(t)) & \\ \mathfrak{C}\hat{z}(t) & = & \hat{y}(t) & \end{array} \right. \tag{5.2}$$

with $z, \hat{z}$ in the Hilbert space Z, differential operator $\mathfrak{A}$ with $\mathfrak{A} : D(\mathfrak{A}) \subset Z \mapsto Z$ and $D(\cdot)$ denoting the domain of an operator. The $*$ denotes the convolution product.

Furthermore, we have the scalar control $u_1 \in \mathbb{R}$, and also the vector $u_2 \in U$, with $U = \mathbb{R}^m$, $m \geq 1$ since we have two boundaries. The vector Robin-type boundary control operator $\mathfrak{B}$ and observation operator $\mathfrak{C}$ should be interpreted in the sense of definition 3.3.2 in [21], where $\mathfrak{B} : D(\mathfrak{B}) \subset Z \mapsto U$ satisfies $D(\mathfrak{A}) \subset D(\mathfrak{B})$ and $\mathfrak{C} : D(\mathfrak{A}) \subset D(\mathfrak{C}) \subset Z \mapsto \mathbb{R}^q$, $q \geq 1$. To comply with the condition in their definition, we further assume that A is given by:

$$Az = \mathfrak{A}z \quad \text{for} \quad z \in D(A) = D(\mathfrak{A}) \cap \ker(\mathfrak{B}) \tag{5.3}$$

generates a C_0-semigroup[†]. We specify $L(t)$ in what follows.

This chapter is basically divided into two parts. In the first part, necessary concepts for boundary control systems are introduced which will be used for the dynamical analysis of CDR-systems and for a static, boundary observer design. This static boundary observer design will be referred to as *Case I*. In the second part a numerical procedure is outlined for the synthesis of a dynamic boundary observer which is robust to disturbances on input and output signals. This approach will be referred to as *Case II*. A schematic overview for both model–observer design cases is presented in Figure 5.1. The observer operator $L \in \mathscr{L}(Y,U)$, where $\mathscr{L}(Y,U)$

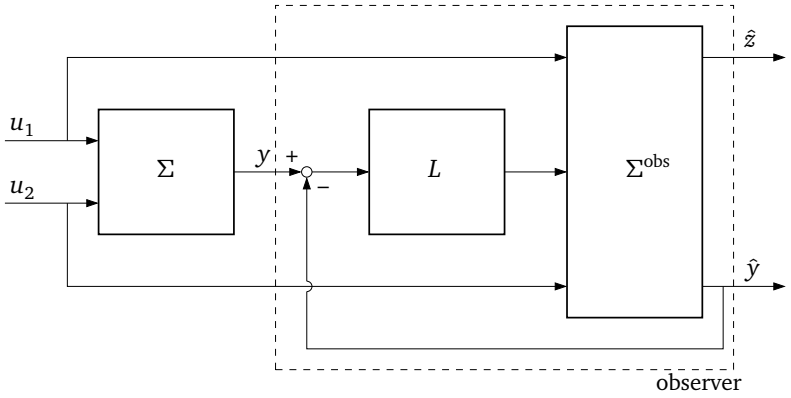


Figure 5.1: Schematic overview of the process model Σ and its observer

denotes the space of bounded linear operators from Y to U, comprises of $(m \times q)$ observer gains. Some differences between Case I and II should be noted:

Case I: $L(t) \equiv \mathbf{L}\delta(t)$ and there are no disturbances, i.e. $v_1 = 0 = v_2$. So the second equation in eqn. (5.2) becomes $\mathfrak{B}\dot{\hat{z}}(t) = u_2(t) + \mathbf{L}\mathfrak{C}(z(t) - \hat{z}(t))$. Furthermore, $\mathfrak{C}$ will be specified as a *boundary operator*. The reason for the latter will be clarified in what follows in Section 5.3.

***Case II*:** $L(t)$ will be found after solving a disturbance rejection problem by formulating it as an H_∞–problem. In addition, $\mathfrak{C}$ will be specified as an operator which specifies point observations.

We will further elucidate on eqn. (5.2) and design conditions on L in what follows. Subsequently, an approach for synthesizing a dynamic observer is proposed and numerical results are shown.

[†]See Appendix B, Definition B.1.4 for a definition

Chapter outline

The complete outline of this chapter is as follows:

Section 5.2 Some preliminary characteristics with respect to a system Σ as defined in eqn. (5.1) are specified and the concept of observability for a boundary control system is introduced. Also, the system operator A is further specified as a Sturm-Liouville (S-L) type operator (typical for convection–diffusion–reaction processes) and some characteristics are given for this class of systems.

Section 5.3 Detectability and design rules with respect to a boundary static observer as in eqn. (5.2), with A being a S-L operator and $\mathbf{L} \in \mathbb{R}^{m\times q}$, are evaluated. The results are illustrated by an observer design for a UV disinfection process with boundary measurements and a boundary control action.

Section 5.4 A dynamic boundary observer with disturbances on the measurements and on a boundary condition is synthesized. The same UV process example as in Section 5.3 is worked out.

Sections 5.5 and 5.6 Observer performance results of both observer configurations are shown for the UV-disinfection process. Some discussion about the merits and pitfalls of formulating a boundary observer are discussed.

Section 5.7 Some final remarks and conclusions are given.

5.2 Preliminaries

In the following, we show that a bilinear system as in eqn. (5.1) with A given in eqn. (5.3), permits a mild solution and we define approximate observability. Recall that, $\mathfrak{B} : D(\mathfrak{B}) \subset Z \mapsto U$ satisfies $D(\mathfrak{A}) \subset D(\mathfrak{B})$ and $\mathfrak{C} : D(\mathfrak{A}) \subset D(\mathfrak{C}) \subset Z \mapsto \mathbb{R}^q$, $q \geq 1$. Furthermore, by assumption A is a densely defined linear differential operator on the Hilbert space Z, which generates the C_0-semigroup $T(t)$. We now characterize the solution of eqn. (5.1).

5.2.1 Mild solution

THEOREM 5.2.1. *For eqn. (5.1), where A as in eqn. (5.3), there exists a mild solution, with mild solution operator* $U(t,s)z_0 = T(t-s)e^{\int_s^t b_1 u_1(\tau)d\tau} z_0$, *where* T *is the* C_0*-semigroup generated by* $(A,\ D(A))$.

Proof. The proof originates from the work of Jean Bernoulli on ordinary differential equations for the scalar case. First, let $\mathscr{L}(Z)$ be a short-hand notation for a bounded linear operator from Z onto Z. Furthermore, w is a scalar valued function. Write $z = vw$, then with $v = T(t-s)v_0$ and with z subject to $\dot{z} - Az = -b_1 u_1 z$, we get $v\dot{w} = -b_1 u_1 vw$. It follows that $w = w_0 \exp(-\int_s^t b_1 u_1(\tau)d\tau)$. Substituting $z_0 = v_0 w_0$

gives the result. Further, according to Definition 3.2.4 [21], $U(t,s) : \Lambda(\tau) \to \mathscr{L}(Z)$ is a mild solution operator with $\Lambda(\tau) = \{(t,s);\, 0 \le s \le t \le \tau\}$, since,

a. $U(s,s) = I$, $s \in [0,\tau]$ holds,

b. A is an infinitesimal generator of a C_0-semigroup, hence:

$$\begin{aligned} U(t,r)U(r,s)z_0 =& T(t-r)e^{\int_r^t b_1 u_1(\tau)d\tau} T(r-s)e^{\int_s^r b_1 u_1(\tau)d\tau} z_0 \\ =& T(t,s)e^{\int_r^t b_1 u_1(\tau)d\tau + \int_s^t b_1 u_1(\tau)d\tau} z_0 \end{aligned}$$

which equals $U(t,s)$, $0 \le s \le t \le \tau$,

c. It is standard to show that $U(\cdot,s)$ is strongly continuous on $[s,\tau]$ and that $U(t,\cdot)$ is strongly continuous on $[0,t]$.

□

5.2.2 Approximate observability

In [21], controllability and observability results are only derived for *bounded* B and C operators. In earlier work, stabilizability and detectability results are obtained for parabolic distributed systems in the case of unbounded B and C operators using a modal approach [19]. Stability and observability results, again in the case where B and C are bounded operators, are deduced for the Sturm-Liouville class of systems [25, 112].

In this section, a generalization with respect to the (approximate) observability of Σ with A as in eqn. (5.3) and $\mathfrak{B}$ and $\mathfrak{C}$ as in eqn. (5.1) is presented. Instead of depending heavily on (very) technical notions of admissibility and regularity [10, 11, 111], we present an approximate observability result which closely resembles the results in [21]. In a subsequent section on the observer design, we deal with the detectability of S-L systems, typically encountered in CDR processes. For the observability result, we need the following concepts.

DEFINITION 5.2.1 [SEMIGROUP INVARIANCE]. *Let V be a subspace of the Hilbert space Z and let $T(t)$ be a C_0-semigroup on Z. We say that V is $T(t)$–invariant if for all $t \ge 0$: $T(t)V \subset V$.*

DEFINITION 5.2.2 [ADMISSIBILITY]. *Let $\mathfrak{C} : D(A) \mapsto Y$. Then $\mathfrak{C}$ is* admissible *if $\forall z \in D(A)$,*

$$\int_0^{t_1} \|\mathfrak{C}T(t)z\|^2 \mathrm{d}t \le m(t_1)\|z\|^2.$$

for some arbitrary constant $m(t_1)$.

DEFINITION 5.2.3 [APPROXIMATE OBSERVABILITY]. *Let a system $\Sigma(A,-,\mathfrak{C})$ as in eqn. (5.1) be defined with A as in eqn. (5.3) an infinitesimal generator of a C_0-semigroup $T(t)$, $b_1 = 0$ and $\mathfrak{C}$ admissible. The* observability map *of $\Sigma(A,-,\mathfrak{C})$ on $[0,\tau]$, $\tau < \infty$, is the bounded linear map $\mathscr{C}^\tau : Z \to L_2([0,\tau],Y)$ defined by:*

$$\mathscr{C}^\tau z := \mathfrak{C}T(\cdot)z.$$

The non-observable *subspace of $\Sigma(A,-,\mathfrak{C})$ is the subspace of all initial states producing a zero output for almost all $t \geq 0$:*

$$\mathscr{N} := \{z \in Z \mid \mathfrak{C}T(t)z = 0 \text{ for almost all } t \geq 0\} = \bigcap_{\tau>0} \ker \mathscr{C}^\tau.$$

$\Sigma(A,-,\mathfrak{C})$ *is* approximately observable *if the only initial state producing the output zero on $[0,\infty)$ is the zero state, i.e., if $\mathscr{N} = \{0\}$.*

We now characterize $\mathscr{N}$ with respect to our system Σ.

LEMMA 5.2.1 [PROPERTIES NON-OBSERVABLE SUBSPACE]. *The non-observable subspace $\mathscr{N}$ has the following characterization with respect to $\Sigma(A,-,\mathfrak{C})$ as in eqn. (5.1):*

$\mathscr{N}$ of $\Sigma(A,-,\mathfrak{C})$ is the largest closed $T(t)$–invariant subspace contained in $\ker \mathfrak{C}(rI-A)^{-1}$*, with $r > \omega_0$ and ω_0 the* growth bound *on $T(t)$, i.e. $\forall \omega > \omega_0$, $\exists M$ such that $\forall t \geq 0$, $\|T(t)\| \leq Me^{\omega t}$.*

Proof. First, let $\omega_0 < 0$ so that the operator A is invertible. For $z_1 := A^{-1}z$, we have $\mathfrak{C}T(t)A^{-1}z = \mathfrak{C}A^{-1}T(t)z$.

Let,

$$\mathscr{N}_0 = \{z \mid \mathfrak{C}T(t)z = 0, \text{ almost everywhere}\}$$
$$\mathscr{N}_1 = \{z \mid \mathfrak{C}A^{-1}T(t)z = 0, \ \forall t\}.$$

We now prove that $\mathscr{N}_0 = \mathscr{N}_1$ and we begin by showing that $\mathscr{N}_1 \subset \mathscr{N}_0$. Suppose $z_1 \in \mathscr{N}_1$, then $A^{-1}z_1 \in \mathscr{N}_0$, since $z_1 \in Z$ and $A^{-1}z_1 \in D(A)$. Consequently, $A^{-1}z_1 \in \mathscr{N}_0 \cap D(A)$. Since $\mathscr{N}_0$ is $T(t)$-invariant, $A\left(\mathscr{N}_0 \cap D(A)\right) \subset \mathscr{N}_0$; see Exercise 2.31 in [21]. Hence, $A(A^{-1})z_1 \in \mathscr{N}_0$, i.e. $\mathscr{N}_1 \subset \mathscr{N}_0$.

Now take $z \in \mathscr{N}_0$ and consider $\mathfrak{C}A^{-1}T(t)z = \mathfrak{C}T(t)A^{-1}z = \mathfrak{C}T(t)z_1$. Since $z \in \mathscr{N}_0$, $A^{-1}z \in \mathscr{N}_0$ by Lemma 2.5.6 in [21]. In other words, $\mathscr{N}_0$ is closed and A^{-1}–invariant. We also have $\mathfrak{C}T(t)A^{-1}z = 0$ and $z \in \mathscr{N}_1$. Consequently, $\mathscr{N}_0 \subset \mathscr{N}_1$. By Definition 5.2.3, we also have $\mathscr{N} \subseteq \mathscr{N}_0$. The result follows from Lemma 4.1.18 in [21].

If the operator A is not invertible, then replace in the above A with $A - rI$ where $r > \omega_0$, with ω_0 the growth bound on $T(t)$. □

5.2.3 S-L system characteristics

Previously, we have shown that Σ as in eqn. (5.1) permits a mild solution and we characterized the non-observable subspace with respect to this system. Here, we specify A in Σ further as an S-L (Sturm-Liouville) operator and summarize some properties of A. As a consequence, we will denote $\Sigma(A, \mathfrak{B}, \mathfrak{C})$ with A a S-L operator as $\Sigma_{S.L.}(A, \mathfrak{B}, \mathfrak{C})$.

The operator A

As also pointed out in [26], in many physical systems (e.g. vibration/diffusion problems or convection–dispersion in chemical reactor models) A or $-A$ is a S-L operator. Hence, the UV disinfection process along with our modeling assumptions and written as a CDR model in Chapter 2, is also a S-L–type system. As such, we are motivated to inspect the properties of A being of Sturm-Liouville type.

Let us first define the differential operator in eqn. (5.1) as,

$$\mathfrak{A}z = \frac{1}{w}\left(\frac{\mathrm{d}}{\mathrm{d}\eta}\left(p\frac{\mathrm{d}z}{\mathrm{d}\eta}\right) - qz\right) \tag{5.4a}$$

with

$$p(\eta), w(\eta) \in \mathbb{R}_+, \text{ both } C^1\text{-continuous functions} \tag{5.4b}$$

and

$$q(\eta) \in \mathbb{R} \text{ on } \left[\eta_1, \eta_2\right]. \tag{5.4c}$$

Furthermore, the domain $D(\mathfrak{A})$ is given by,

$$D(\mathfrak{A}) = \left\{ z \in Z \mid z, \frac{\mathrm{d}z}{\mathrm{d}\eta} \text{ absolutely continuous}, \frac{\mathrm{d}^2}{\mathrm{d}\eta^2} z \in Z \right\} \tag{5.4d}$$

and boundary (control) operator

$$\mathfrak{B}z := \begin{pmatrix} \beta_1 z(\eta_1) + \gamma_1 \frac{\mathrm{d}z}{\mathrm{d}\eta}(\eta_1) \\ \beta_2 z(\eta_2) + \gamma_2 \frac{\mathrm{d}z}{\mathrm{d}\eta}(\eta_2) \end{pmatrix} = u_2 \tag{5.4e}$$

with β_i, γ_j real constants satisfying $|\beta_1| + |\gamma_1| > 0$ and $|\beta_2| + |\gamma_2| > 0$. In Section 5.2.4, β_i and γ_i are further specified for the UV disinfection process.

We define observations at points η_1^* and η_2^* as

$$\mathfrak{C}z := \begin{pmatrix} z(\eta_1^*) \\ \frac{dz}{d\eta}(\eta_1^*) \\ z(\eta_2^*) \\ \frac{dz}{d\eta}(\eta_2^*) \end{pmatrix} = y \tag{5.5}$$

and the weighted inner product with $w(\eta)$ as given in eqn. (5.4) as:

$$\langle z_1, z_2 \rangle_w = \int_{\eta_1}^{\eta_2} z_1(\eta)\overline{z_2(\eta)}w(\eta)d\eta. \tag{5.6}$$

Now, we turn to some characteristics of the S-L system. Recognize from eqn. (5.4) that, $-(A, D(A))$ with $Az = \mathfrak{A}z$ for $z \in D(\mathfrak{A}) \cap \ker(\mathfrak{B})$ and $\mathfrak{A}$, $\mathfrak{B}$ as in eqn. (5.4), is a Sturm-Liouville operator, self-adjoint in a weighted inner product $\langle \cdot, \cdot \rangle_w$ and closed on Z, [see also 21, exercise 2.10].

We mention the following result from [26].

LEMMA 5.2.2 [DELATTRE, DOCHAIN, AND WINKIN]. *Let A be the negative of a Sturm-Liouville operator defined on its domain $D(A)$ given by eqn. (5.3). Then,*

(i) A is a Riesz spectral operator,

(ii) A is the infinitesimal generator of a C_0-semigroup of bounded linear operators on $L_2(\eta_1, \eta_2)$,

(iii) A has compact resolvent.

As a consequence of Lemma 5.2.2, $\Sigma_{S.L.}(A, \mathfrak{B}, \cdot)$, with A and $\mathfrak{B}$ further specified as in eqs. (5.3) and (5.4), $\Sigma_{S.L}$ has a mild solution. See Theorem 5.2.1 for details of this solution.

In the observer design section, it will become clear that it is convenient to check whether A in $\Sigma_{S.L.}$ is negative.

LEMMA 5.2.3 [POSITIVITY OF OPERATOR $-A$]. *The Sturm-Liuoville operator $-A$, with A as in eqn. (5.4) and with positive real-valued continuous functions $p(\eta)$, $w(\eta)$ and*

$q(\eta)$, is positive as defined in Appendix A for $z \neq 0$:

$$
\begin{aligned}
&\textit{if} \quad \frac{\beta_2}{\gamma_2} \geq 0, \quad \frac{\beta_1}{\gamma_1} \leq 0 \quad \textit{and} \quad |\beta_1| + |\beta_2| > 0 \quad \textit{for} \quad \gamma_1, \gamma_2 \neq 0 \\
&\textit{if} \quad \gamma_1 = 0: \quad \frac{\beta_2}{\gamma_2} \geq 0 \quad \textit{and} \quad \gamma_2 \neq 0 \\
&\textit{if} \quad \gamma_2 = 0: \quad \frac{\beta_1}{\gamma_1} \leq 0 \quad \textit{and} \quad \gamma_1 \neq 0 \\
&\textit{if} \quad \gamma_1 = 0 = \gamma_2.
\end{aligned}
$$

Proof. It is sufficient to check the time derivative of the weighted norm of z, $\frac{\mathrm{d}}{\mathrm{d}t}\|z\|_w^2 = \langle z, -Az\rangle_w \geq 0$, using the inner product eqn. (5.6). It follows that,

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\|z(\eta,t)\|_w^2 &= \int_{\eta_1}^{\eta_2} -\left(\frac{\mathrm{d}}{\mathrm{d}\eta}\left(p\frac{\mathrm{d}z}{\mathrm{d}\eta}\right) + qz\right)\cdot z \mathrm{d}\eta \\
&= -p(\eta)\frac{\mathrm{d}z(\eta,\cdot)}{\mathrm{d}\eta}z(\eta,\cdot)\Big|_{\eta_1}^{\eta_2} + \int_{\eta_1}^{\eta_2} p(\eta)\left(\frac{\mathrm{d}z(\eta,\cdot)}{\mathrm{d}\eta}\right)^2 + q(\eta)z(\eta,\cdot)^2\mathrm{d}\eta \\
&= -\left[p(\eta_2)\frac{\mathrm{d}z}{\mathrm{d}\eta}(\eta_2,\cdot)z(\eta_2,\cdot) - p(\eta_1)\frac{\mathrm{d}z}{\mathrm{d}\eta}(\eta_1,\cdot)z(\eta_1,\cdot)\right] + \dots \\
&\quad \int_{\eta_1}^{\eta_2} p(\eta)\left|\frac{\mathrm{d}z(\eta,\cdot)}{\mathrm{d}\eta}\right|^2 + q(\eta)z(\eta,\cdot)^2\mathrm{d}\eta \\
&= -\left[p(\eta_2)\left(-\frac{\beta_2}{\gamma_2}z(\eta_2,\cdot)^2\right) - p(\eta_1)\left(-\frac{\beta_1}{\gamma_1}z(\eta_1,\cdot)^2\right)\right] + \dots \\
&\quad \int_{\eta_1}^{\eta_2} p(\eta)\left|\frac{\mathrm{d}z(\eta,\cdot)}{\mathrm{d}\eta}\right|^2 + q(\eta)z(\eta,\cdot)^2\mathrm{d}\eta
\end{aligned}
$$

Hence, given $p, w > 0$ and $q \geq 0$, the (sufficient) conditions directly follow. □

Approximate observability for $\Sigma_{S.L.}$

From Definition 5.2.3 and Lemma 5.2.1 it is straightforward to check whether Σ of eqn. (5.1) is *approximately observable* when A is a Riesz-spectral operator. This is shown in the following corollary.

COROLLARY 5.2.1. *Let $\mathcal{N}$ be the unobservable subspace. Then, $\Sigma(A, -, \mathfrak{C})$ as in eqn. (5.1) with A a Riesz-spectral operator on the Hilbert space Z and with the orthonormal representation $Az = \sum_{n=1}^{\infty}\lambda_n\langle z, \phi_n\rangle_w\phi_n$ is approximately observable if and only if*

$\mathscr{N} = \{0\}$, *i.e. iff* $\mathfrak{C}\phi_n \neq 0$ *for all* n.

Proof. So far, it has been proved in Lemma 5.2.1 that $\mathscr{N}$ is a closed subspace and $T(t)$–invariant. In [26], it is proven that A is a Riesz–spectral operator with the Riesz basis of orthonormal eigenvectors $\{\phi_n,\ n \geq 1\}$ and associated semigroup $T(t)$. Consider,

$$\begin{aligned}\mathbb{J} &= \{n \in \mathbb{N} \mid \mathfrak{C}A^{-1}\phi_n = 0\} \\ &= \{n \in \mathbb{N} \mid \mathfrak{C}\frac{\phi_n}{\lambda_n} = 0\} \\ &= \{n \in \mathbb{N} \mid \mathfrak{C}\phi_n = 0\}.\end{aligned}$$

Hence, by the above and Lemma 2.5.8. in [21], we have that $\mathscr{N} = \overline{\operatorname{span}}_{n\in\mathbb{J}}\{\phi_n\}$. □

Hence, it follows from Lemma 5.2.2 and Lemma 5.2.3 that Corollary 5.2.1 is applicable to $\Sigma_{S.L.} := \Sigma(A, -, \mathfrak{C})$, with A being an S-L-operator. We now turn to our example.

5.2.4 UV disinfection process in boundary control form

In Section 2.2, it is explained under which assumptions a UV disinfection process may be described as in eqn. (2.11). Here, these equations are written in a dimensionless, boundary control form.

First, the UV disinfection model is transformed into a dimensionless form by Buckinghams theorem [12], see for details Section D.1. The resulting dimensionless variables r_i together with their interpratation are shown in Table 5.1. With these dimensionless variables and parameters, we can next derive the full equations of the UV disinfection process in boundary control form. We have

$$\mathfrak{A}z = \frac{1}{w}\left(\frac{\mathrm{d}}{\mathrm{d}\eta}\left(p\frac{\mathrm{d}z}{\mathrm{d}\eta}\right) - qz\right) \tag{5.7}$$

where for the UV process, p, w and q are specified as,

$$p(\eta) = e^{-\mathrm{p_e}\eta}, \qquad w(\eta) = \mathrm{p_e}e^{-\mathrm{p_e}\eta} \quad \text{and} \quad q = 0. \tag{5.8}$$

With p, w, q as in eqn. (5.8) and boundary conditions specified with

$$\beta_1 = 1, \qquad \gamma_1 = -1/\mathrm{p_e}, \qquad \beta_2 = 0 \quad \text{and} \quad \gamma_2 = 1 \tag{5.9}$$

Table 5.1: Dimensionless variables and parameters with their interpreatation

r_i	—	Interpretation
$z(\cdot,t)$	$c(\cdot,t)/\bar{c}$	state variable for the concentration of living micro-organisms scaled with the mean concentration at the inlet
t	$\tau v/L$	time, scaled with the time it takes for the fluid to travel from the inlet to the outlet of the reactor
η	ξ/L	spatial coordinate scaled with axial length of reactor
$\mathrm{p_e}$	$v_f \cdot L/\alpha$	ratio of convective transfer and diffusion
b_1	$\kappa f_{\max} L/v$	scaled susceptibility of micro-organisms to UV radiation
$u_1(t)$	$f(t)/f_{\max}$	UV lamp intensity scaled with the maximal intensity of the lamp
$\tilde{u}_2(t)$	$c_{\mathrm{in}}(t)/\bar{c}$	inlet concentration scaled with mean pathogenic micro-organisms concentration

see also Appendix D.1, eqn. (D.1). This leads to the UV system description[‡]:

$$\Sigma_I : \begin{cases} \dot{z} &= \mathfrak{A}z - b_1 u_1 z \\ &= \frac{1}{\mathrm{p_e}} \frac{\mathrm{d}^2}{\mathrm{d}\eta^2} z - \frac{\mathrm{d}z}{\mathrm{d}\eta} - b_1 u_1 z, \qquad z(0) = z_0 \\ \mathfrak{B}z &= \begin{pmatrix} z(0) - \frac{1}{\mathrm{p_e}} \frac{\mathrm{d}z}{\mathrm{d}\eta}(0) \\ \frac{\mathrm{d}z}{\mathrm{d}\eta}(1) \end{pmatrix} \end{cases} \tag{5.10}$$

and observation $y = \mathfrak{C}z$ specified as in eqn. (5.5).

To distinguish this system from the dynamic observer design Case II, the subscript I in Σ is used. Notice also that Σ_I inhibits one Robin and one Neumann boundary and is therefore known to have Danckwerts boundary conditions [23]. Danckwerts conditions appear naturally in transport phenomena in reactors, such as we have seen in the UV disinfection example.

The aim is to design an observer for the disinfection process eqn. (5.10) as schematically depicted in Figure 5.1. We first derive some static gain design results for the general CDR_1 class of systems and then apply these results to the UV disinfection case.

[‡]Note that the superscript *UV* of Σ as in Chapter 2 has been dropped, since in this chapter we only refer to the UV disinfection case.

5.3 Case I: static, boundary observer

Now that the system characteristics of S-L boundary control systems and the UV disinfection process in particular have been laid out, we can investigate the total bilinear model–boundary observer setup as in eqs. (5.1) and (5.2). More specifically, we focus on the decay of the estimation error $\varepsilon(\eta, t)$ for a Sturm-Liouville type of estimation problem and give conditions under which the design of $\mathbf{L}$ leads to a stable error system. Therefore, the full error system is written out.

5.3.1 Error system

Let us first inspect the estimation error system $\Sigma^{\varepsilon}(A, -, \mathfrak{C})$, with A as in eqn. (5.4) and $\mathfrak{C}$ as in eqn. (5.5) with $\nu_1 = 0 = \nu_2$. To comply with the Sturm-Liouville framework, we describe the dynamics of the estimation error $\varepsilon := z - \hat{z}$ by,

$$\Sigma^{\varepsilon} := \left\{ \begin{array}{rcl} \dot{\varepsilon} & = & \mathfrak{A}\varepsilon - b_1 u_1 \varepsilon, \quad \varepsilon(0) = \varepsilon_0 \\ \mathfrak{B}\varepsilon & = & \begin{pmatrix} \beta_1 \varepsilon(\eta_1) + \gamma_1 \frac{\mathrm{d}\varepsilon}{\mathrm{d}\eta}(\eta_1) \\ \beta_2 \varepsilon(\eta_2) + \gamma_2 \frac{\mathrm{d}\varepsilon}{\mathrm{d}\eta}(\eta_2) \end{pmatrix} \quad = \quad \mathbf{L}\mathfrak{C}\varepsilon(\eta) \end{array} \right. \tag{5.11a}$$

with $\mathfrak{C}$ as in eqn. (5.5) mapping the states to point observations at η_1^* and η_2^* and,

$$\mathbf{L}\mathfrak{C}\varepsilon = \begin{pmatrix} L_{11}\varepsilon(\eta_1^*) + L_{12}\frac{\mathrm{d}\varepsilon}{\mathrm{d}\eta}(\eta_1^*) + L_{13}\varepsilon(\eta_2^*) + L_{14}\frac{\mathrm{d}\varepsilon}{\mathrm{d}\eta}(\eta_2^*) \\ L_{21}\varepsilon(\eta_2^*) + L_{22}\frac{\mathrm{d}\varepsilon}{\mathrm{d}\eta}(\eta_2^*) + L_{23}\varepsilon(\eta_1^*) + L_{24}\frac{\mathrm{d}\varepsilon}{\mathrm{d}\eta}(\eta_1^*) \end{pmatrix} \tag{5.11b}$$

The above shows that it is not straightforward to arrive at design conditions for $\mathbf{L}$ such that the closed loop configuration remains stable whenever the measurement is *not* taken at the boundary, i.e., whenever $z(\eta_i^*, t) \neq z(\eta_i)$, $i, j \in \{1, 2\}$. For this reason, we only consider boundary measurements for the remainder of this section.

Hence, we write the following operator,

$$\mathfrak{C}^b z = \mathfrak{C} z \quad \text{for} \quad \eta_i^* = \eta_i,\ i \in \{1, 2\}. \tag{5.12}$$

For the UV process, we assume to only have the availability of boundary measurements and furthermore we let $L_{i3} = 0 = L_{i4}$, $i \in \{1, 2\}$. With misuse of notation, we write

$$\mathbf{L} := \begin{pmatrix} L_{11} & L_{12} \\ L_{21} & L_{22} \end{pmatrix}$$

To check detectability of Σ^{ε} in the next subsection, it is convenient to introduce

the operator $\mathfrak{B}^L$:

$$\mathfrak{B}^L\varepsilon := \left(\mathfrak{B} - \mathbf{L}\mathfrak{C}^b\right)\varepsilon = \begin{pmatrix} \beta_1^L\varepsilon(\eta_1) + \gamma_1^L\frac{\mathrm{d}\varepsilon}{\mathrm{d}\eta}(\eta_1) \\ \beta_2^L\varepsilon(\eta_2) + \gamma_2^L\frac{\mathrm{d}\varepsilon}{\mathrm{d}\eta}(\eta_2) \end{pmatrix} \tag{5.13}$$

where,

$$\beta^L := \begin{pmatrix} \beta_1 - L_{11} \\ \beta_2 - L_{21} \end{pmatrix}, \quad \text{and} \quad \gamma^L := \begin{pmatrix} \gamma_1 - L_{12} \\ \gamma_2 - L_{22} \end{pmatrix}. \tag{5.14}$$

5.3.2 Observer design conditions

We now study under which conditions the design of a *static, boundary observer* would lead to a detectable error system $\Sigma^\varepsilon(A, \mathfrak{B}, \mathfrak{C}^b)$. Detectability will be defined as follows.

DEFINITION 5.3.1 [DETECTABILITY]. *Whenever there exists an $\mathbf{L} \in \mathscr{L}(\mathbb{R}^q, \mathbb{R}^m)$, such that A^L with $A^L\varepsilon = \mathfrak{A}\varepsilon$ for $\varepsilon \in D(\mathfrak{A}) \cap \ker(\mathfrak{B}^L)$ generates an exponentially stable C_0-semigroup; then we say that $\Sigma^\varepsilon(A^L)$ is detectable.*

COROLLARY 5.3.1 [EXISTENCE BOUNDARY STATIC OBSERVER]. *Suppose that A^L with $\mathfrak{A}$ as in eqn. (5.3) and $\mathfrak{B}^L$ as in eqn. (5.13), generates an exponentially stable C_0-semigroup. Then for some gain matrix $\mathbf{L} \in \mathscr{L}(\mathbb{R}^q, \mathbb{R}^m)$, the system $\Sigma(A^L)$ is detectable.*

Proof. Consider the estimation error $\varepsilon = z - \hat{z}$. From the proof of Lemma 5.2.3, it follows that it is sufficient to check whether $-A^L$ is a positive Sturm-Liouville operator, i.e. $\langle \varepsilon, -A^L\varepsilon\rangle > 0$. Since $\beta^L\varepsilon + \gamma^L\frac{\mathrm{d}\varepsilon}{\mathrm{d}\eta} = 0$, we can always choose a suitable L such that $|\beta^L| + |\gamma^L| > 0$ and the boundary conditions in $D(A^L)$ are such that $-A^L$ is positive. Since $-A^L$ is a Sturm-Liouville operator, it has compact resolvent and therefore generates an exponentially stable semigroup [25]. □

The following corollary immediately follows.

COROLLARY 5.3.2 [STATIC OBSERVER DESIGN]. *The system $\Sigma^\varepsilon(A^L, \mathfrak{B}^L)$ is detectable:*

$$\begin{aligned}
&\textit{if} \quad \frac{\beta_2^L}{\gamma_2^L} \geq 0, \quad \frac{\beta_1^L}{\gamma_1^L} \leq 0 \quad \textit{and} \quad |\beta_1^L| + |\beta_2^L| > 0 \quad \textit{for} \quad \gamma_1^L, \gamma_2^L \neq 0 \\
&\textit{if} \quad \gamma_1^L = 0: \quad \frac{\beta_2^L}{\gamma_2^L} \geq 0 \quad \textit{and} \quad \gamma_2^L \neq 0 \\
&\textit{if} \quad \gamma_2^L = 0: \quad \frac{\beta_1^L}{\gamma_1^L} \leq 0 \quad \textit{and} \quad \gamma_1^L \neq 0 \\
&\textit{if} \quad \gamma_1^L = 0 = \gamma_2^L.
\end{aligned}$$

For many processes, it is not practical to implement an observer where $\gamma^L \neq 0 \neq \gamma$, since a calculation or measurement of the spatial derivative of $z(\eta_i)$ is needed. Hence, if possible, a boundary observer matrix $\mathbf{L}$ should be chosen so that derivative terms of y in the error system are cancelled.

5.3.3 Application of a static, boundary observer

We illustrate the design of a boundary observer for the UV disinfection process and also show its influence on the error dynamics. In the UV disinfection process, it is desired to have a good estimate of the concentration at the outlet of the reactor. Furthermore, in the previous section we have obtained conditions for a static observer when we have measurements at the boundary, i.e. $\mathfrak{C}^b = \mathfrak{C}$ for $\eta^* = \eta_i$, $i \in \{1, 2\}$. We now continue to investigate how to choose $\mathbf{L}$.

Design of L for the UV process

In what follows in the eigenvalue analysis of the system Σ_I, it is shown that the calculation of eigenvalues of the system dynamics is less straightforward (and is numerically more involved) whenever there are Danckwerts boundary conditions. Hence, we choose to have a Dirichlet-type condition at η_1 to simplify further analysis and ease the eigenvalue calculations of the error system. To this aim, the first row of $\mathbf{L}$ is set to $\mathbf{L}_{1\bullet} = (0 \;\; \gamma_1)$ such that the derivative term in the boundary of the error system cancels. Furthermore, from an engineering point of view we prefer to tune our observer with just one scalar gain, so we let $L_{11} = 0 = L_{22}$.

These prerequisites lead to:

$$\mathbf{L} = \begin{pmatrix} 0 & \gamma_1 \\ L_{21} & 0 \end{pmatrix} = \begin{pmatrix} 0 & -\frac{1}{\mathrm{p_e}} \\ L_{21} & 0 \end{pmatrix} \qquad \text{with} \qquad L_{21} \neq \beta_2 \qquad \text{to be chosen.} \tag{5.15}$$

With γ and β as in eqn. (5.10). Consequently,

$$\beta^L = \begin{pmatrix} 1 \\ -L_{21} \end{pmatrix} \quad \text{and} \quad \gamma^L = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

The error system Σ_I^ε for this example with L as in eqn. (5.15) then reads:

$$\Sigma_I^\varepsilon := \left\{ \begin{array}{rcl} \dot{\varepsilon} & = & \mathfrak{A}\varepsilon - b_1 u_1 \varepsilon \\ & = & \frac{1}{\mathrm{p_e}} \frac{\mathrm{d}^2}{\mathrm{d}\eta^2}\varepsilon - \frac{\mathrm{d}\varepsilon}{\mathrm{d}\eta}, \quad \varepsilon(0) = \varepsilon_0 \\ \mathfrak{B}^L\varepsilon & = & \begin{pmatrix} \varepsilon(0) \\ -L_{21}\varepsilon(1) + \frac{\mathrm{d}\varepsilon}{\mathrm{d}\eta}(1) \end{pmatrix} \end{array} \right. \tag{5.16}$$

For observer design, L_{21} can be tuned by pole placement of the error system. By Corollary 5.3.2 we also know that,

REMARK 5 [CONDITION FOR L_{21} IN THE UV DISINFECTION PROCESS]. *Σ_I^ε as in eqn. (5.16) is detectable, whenever the sufficient condition $\frac{\beta_2^L}{\gamma_2^L} \geq 0$ holds, i.e., whenever $L_{21} \leq 0$.*

Eigenvalue analysis of the UV process error dynamics

We would like to know how L_{21} influences the error dynamics. Therefore, we calculate the eigenvalues λ of A^L, where $A^L\varepsilon = \mathfrak{A}\varepsilon$ for $\varepsilon \in D(\mathfrak{A}) \cap \ker(\mathfrak{B}^L)$.

For $\Sigma_I^\varepsilon(A^L)$ we obtain the following Lemma.

LEMMA 5.3.1. *Let $\Sigma_I^\varepsilon(A^L)$ be detectable with $L_{21} \leq 0$. Then the spectrum of the operator A^L consists of isolated eigenvalues with finite multiplicities given by,*

$$\sigma(A^L) = \sigma_p(A^L) = \{\lambda_k^L : k \geq 0\} \subset (-\infty, 0)$$

where $\sigma_p(A^L)$ denotes the point spectrum of A^L. The eigenvalues λ_k^L, $k \geq 0$ are simple, real and given by,

$$\lambda_k^L = -\tfrac{1}{\mathrm{p_e}} \left(\varsigma_k^L\right)^2 - \tfrac{1}{4}\mathrm{p_e} \tag{5.17}$$

where ς_k^L, $k \geq 0$ is the set of all the solutions to the resolvent equation

$$\tan(\varsigma_k^L) = -\frac{\varsigma_k^L}{\frac{1}{2}\mathrm{p_e} - L_{21}} \quad \textit{and} \quad \varsigma_k^L > 0 \tag{5.18}$$

such that

$$0 < \varsigma_k^L < \varsigma_{k+1}^L, \quad \forall k > 0. \tag{5.19}$$

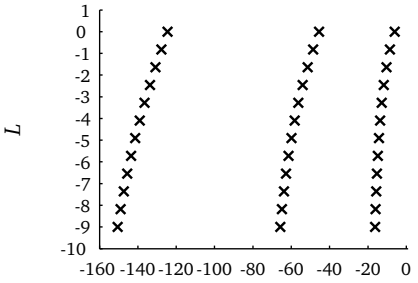
Hence, $\lambda_k^L < \frac{1}{4}\mathrm{p_e} < 0$, $\lambda_k^L \to -\infty$ as $k \to \infty$ and $|\lambda_{k+1}^L - \lambda_k^L| \to \infty$ as $k \to \infty$. The associated eigenvectors $\phi_k^L \in D(A^L)$, $k \geq 0$ are given for all $\eta \in [0, 1]$ and for all $k \geq 1$ by

$$\phi_k(\eta) = \exp\left(\tfrac{1}{2}\mathrm{p_e}\eta\right) \sin(\varsigma_k^L \eta). \tag{5.20}$$

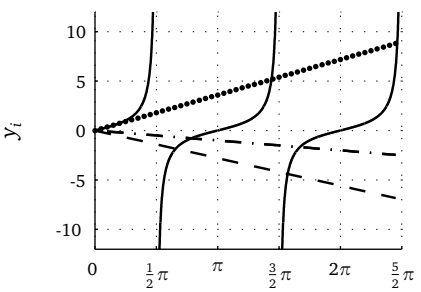
Proof. The detectability condition follows from Corollary 5.3.2 and Remark 5. The derivation of the spectral properties of Σ_I^ε are shown in Appendix D.2. □

As a consequence of Lemma 5.3.1, ς^L behaves like $\pm\pi k$ for $k \to \infty$, and for k finite, we have to obtain ς_k^L numerically. To illustrate, Figure 5.2 shows some intersection points ς_k^L, $k = \{1, 2, 3\}$ and their corresponding eigenvalues λ_k^L have been calculated for $\mathrm{p_e} = 0.5$ and different values of L_{21}.

REMARK 6 [CASE WHERE $q := b_1 u_1$]. *In the case that the control u_1 is given and constant, the operator A^L changes. In that case, A^L is specified with $q := b_1 u_1$, $q \geq 0$,*



(a) From left to right: λ_2^L, λ_1^L, λ_0^L for varying L_{21}



(b) $\tan(\varsigma^L)$ [–], $f(\varsigma^L, 1)$ [•], $f(\varsigma^L, -1)$ [––], $f(\varsigma^L, -3)$, [–•] with $f(\varsigma^L, L_{21}) := \varsigma^L / (L_{21} - 0.5)$

Figure 5.2: (a) λ_k^L and (b) the intersection between f and the tangent function in eqn. (5.18) for different values of L_{21}.

and the eigenvalues change accordingly. Here, we only give the result, since the calculation goes similar as in Lemma 5.3.1. We impose detectability of Σ_I^ε with $L_{21} \leq 0$. Then, the spectrum of A^L with $q := b_1 u_1$ consists of isolated eigenvalues λ_k^q with finite multiplicities given by,

$$\sigma(A^L) = \sigma_{p,q}(A^L) = \{\lambda_k^q : k \geq 0\} \subset (-\infty, 0)$$

where $\sigma_{p,q}(A^L)$ denotes the point spectrum of A^L with $q := b_1 u_1$. The eigenvalues λ_k^q, $k \geq 0$ are simple, real and given by,

$$\lambda_k^q = -\tfrac{1}{p_e}(\varsigma_k^L)^2 - \tfrac{1}{4}p_e - q < \tfrac{1}{4}p_e - q < 0 \tag{5.21}$$

where ς_k^L, $k \geq 0$ are all the solutions to the resolvent equation

$$\tan \varsigma_k^L = -\frac{\varsigma_k^L}{\frac{1}{2}p_e - L_{21}}, \quad \varsigma_k^q > 0 \tag{5.22}$$

such that

$$0 < \varsigma_k^L < \varsigma_{k+1}^L, \quad \forall k > 0. \tag{5.23}$$

In the time-varying case we have that the evolution operator $U(t,s)$ as in Theorem 5.2.1

is bounded from above by:

$$\|U(t,s)\| \leq \exp\left((\omega_0 - q_{\min})(t-s)\right)$$

with ω_0 *the growth bound*[§] *on* $T(t)$ *and* $q \geq q_{\min} \geq 0$ *the lower bound on the lamp strength.*

Until now, we have obtained results for the eigenvalues λ of the error system Σ_I^ε for a particular $\mathbf{L}$. Since the goal is to design a boundary observer with detectable error dynamics, the calculation of λ of Σ_I^ε suffices and calculation of λ of the UV disinfection system Σ_I is not needed. For reference, the eigenvalues λ and associated eigenvectors ϕ_k of Σ_I as calculated in Appendix D.3 are given here as well:

$$\lambda_k = -\tfrac{1}{4}\mathrm{p_e} - \frac{1}{\mathrm{p_e}}\varsigma_k^2 \tag{5.24}$$

with ς_k, $k \geq 0$ the set of all solutions to the resolvent equation:

$$\tan(\varsigma_k) = \frac{2\mathrm{p_e}\varsigma_k}{\varsigma_k^2 - \left(\frac{1}{2}\mathrm{p_e}\right)^2} \tag{5.25}$$

and orthonormal associated eigenvectors:

$$\phi_k = C_0 \exp\left(\tfrac{1}{2}\mathrm{p_e}\eta\right)\left[\frac{\mathrm{p_e}}{\varsigma_k}\sin(\varsigma_k\eta) + 2\cos(\varsigma_k\eta)\right], \qquad C_0 > 0 \tag{5.26}$$

Note that the solution of this eigenvalue problem is numerically more involved due to the presence of the Danckwerts conditions.

We will discuss how to design L_{21} with some margin on the growth bound of $T(t)$, i.e. $\lambda_0 := \sup_{k\in\mathbb{N}} \lambda_k$, in Section 5.5 (the results section). Some remarks on the upper bounds of performance increase are given in the next section.

5.3.4 Remarks

We complete this section by some remarks on the eigenvalue analysis of Σ_I^ε, observability and the solution of the whole system.

Eigenvalue limits and performance

REMARK 7. *From eqn. (5.18), it follows that*

- *for fixed* k, $\lim_{L\to-\infty} \varsigma_k^L = \pm k\pi$, *and therefore* $\lim_{L\to\infty} \lambda_k^L = -\frac{1}{4}k^2\pi^2/\mathrm{p_e}$;

[§]The definition of a growth bound ω_0 of a Riesz spectral operator A is given in Appendix A, Theorem A.3.1.

- *for fixed* $L \in [-\infty, 0]$*, we get* $\varsigma_k \in \left[(k-\frac{1}{2})\pi, k\pi\right]$*, and therefore the eigenvalues* $\lambda_k^L \in -\frac{1}{\mathrm{p_e}}\left[(k-\frac{1}{2})^2\pi^2, k^2\pi^2\right]$.

Indeed, Remark 7 reveals what would be suspected from Figure 5.2, i.e. the magnitude of the distance $|\omega_k - \omega_{k-1}|$ if $L_{21} \to -\infty$ or if $k \to \infty$.

Furthermore, it may be interesting to deduce the difference between the growth bounds of $T^\varepsilon(t)$ with different L_{21}, say,

$$\Delta_\lambda := \lambda_0 - \lambda_0^L, \qquad \text{with growth bound} \qquad \lambda_0 := \sup_{k\in\mathbb{N}} \lambda_k. \tag{5.27}$$

We will call Δ_λ the *performance increase* of the error dynamics of the observer. For the UV disinfection case, Remark 7 and the eigenvalues λ^L as in eqn. (5.17) tells us that $-\frac{1}{4}\mathrm{p_e}^2 - \pi^2 < \mathrm{p_e}\lambda_0^L < -\frac{1}{4}\mathrm{p_e}^2$ for all L_{21}, hence there is a maximal performance increase $\Delta_\lambda^{\max} = \pi^2/\mathrm{p_e}$. Similarly, if we only allow $L_{21} \le 0$, then $-\frac{1}{4}\mathrm{p_e}^2 - \pi^2 < \mathrm{p_e}\lambda_0 < -\frac{1}{4}\mathrm{p_e}^2 - \frac{1}{4}\pi^2$ and the performance increase can maximally be $\frac{3}{4}\pi^2/\mathrm{p_e}$.

Observability

Given system $\Sigma_I(A, -, \mathfrak{C})$ as in eqs. (5.10) and (5.5), we can calculate when the UV system is approximately observable. We propose the following.

PROPOSITION 5.3.1. *Given system* $\Sigma_I(A, -, \mathfrak{C})$ *as in eqs.* (5.10) *and* (5.5)*, i.e. with observations on the interval* $\eta^* \in [0, 1]$*, then*

(i) by considering $y = z(\eta^*)$ *as the only observation,* Σ_I *is* approximately observable *if,*

$$\varsigma_k \neq -\tfrac{1}{2}\mathrm{p_e}\eta^* \tan(\varsigma_k\eta^*), \qquad k \ge 0 \tag{5.28}$$

(ii) by considering $y = \frac{\mathrm{d}z}{\mathrm{d}\eta}(\eta^*)$ *as the only observation,* Σ_I *is* approximately observable *if*

$$\tan(\varsigma_k\eta^*) \neq -\frac{\varsigma}{\mathrm{p_e}}, \qquad k \ge 0 \quad \textit{and} \quad \eta^* \neq 1. \tag{5.29}$$

Proof. From Lemma 5.2.1, we should have that $\mathfrak{C}\phi_k(\eta^*) \neq 0$, $k \ge 0$ and $\mathfrak{C}$ admissible to obtain approximate observability. First admissibility is checked. Introduce the

short-hand notation. Consider,

$$y(t) = \mathfrak{C}T(t)z_0 = \mathfrak{C}e^{At}\sum_k z_k\phi_k$$
$$= \mathfrak{C}\sum_k^\infty z_k e^{\lambda_k t}\phi_k = \sum_k z_k e^{\lambda_k t}\underbrace{\mathfrak{C}\phi_k}_{c_k}.$$

By Cauchy-Schwarz and orthonormality of the eigenvectors of A (see Appendix D.3):

$$\|y(t)\|^2 \leq \sum_{k=1}\left|c_k e^{\lambda_k t}\right|^2 \sum_k \left|z_k\right|^2 = \sum_k e^{2\lambda_k t}|c_k|^2\|z_0\|^2.$$

By Definition 5.2.2, $\int_0^\infty \|y\|^2 \mathrm{d}t \leq m\|z_0\|^2$ should hold. The above leads to

$$\int_0^\infty \|y(t)\|^2 = \int_0^\infty \left|\sum_k c_k e^{\lambda_k t}\right|^2 \mathrm{d}t \leq \int_0^\infty \sum_k \left|c_k e^{\lambda_k t}\right|^2 \sum_k \left|z_k\right|^2 \mathrm{d}t$$
$$\leq \left(\int_0^\infty \sum_k e^{2\lambda_k t}|c_k|^2 \mathrm{d}t\right)\|z_0\|^2 \Leftrightarrow$$
$$\sum_k \int_0^\infty \frac{c_k^2}{-2\lambda_k}\mathrm{d}t\|z_0\|^2 \leq m\|z_0\|^2$$

with m an arbitrary positive constant. Now, consider

(i) $y = z(\eta^*)$, thus $\mathfrak{C}\phi = \phi(\eta^*)$, $\eta^* \in [0,1]$ and the eigenvalues of A, i.e. λ_k in eqn. (5.24) with their associated eigenvectors ϕ_k in eqn. (5.26). By the above $\mathfrak{C}$ is admissible, since for $k \to \infty$, λ_k behaves like $k^2\pi^2/(4\mathrm{p_e})$ so that there exists indeed a value of m which makes the inequality true.

(ii) $y = \dot{z}(\eta^*)$, thus $\mathfrak{C}\phi = \dot{\phi}(\eta^*) \neq 0$, $k \geq 0$. Hence, if $\eta^* = 1$, $\dot{\phi}_k(1) = 0$ so the system is not (approximately) observable. Again, the eigenvectors ϕ_k read as eqn. (5.26), thus we obtain the condition that $\dot{\phi}_k(\eta) = \mathrm{p_e}\sin(\varsigma_k\eta) + \varsigma\cos(\varsigma_k\eta) \neq 0$ for $\eta^* < 1$. The admissibility check goes analogously to the proof of (i).

□

As a consequence of Proposition 5.3.1, Σ_I is always approximately observable for $\mathfrak{C}z = \mathfrak{C}^b z := z(1)$, since we get the condition that $\varsigma \neq \pm\imath\frac{\sqrt{3}}{2}\mathrm{p_e}$ which is always true since $\varsigma \in \mathbb{R}_+$. It is easy to see that for observations $y = z(0)$, the system is not observable. For a point observation in the interval $(0,1)$ the approximate observability has to be checked by eqn. (5.28).

It has already been mentioned that only the estimate $\hat{y} = \hat{z}(1)$ is desired. The non-observability for $\Sigma_I(\cdot,\cdot,\mathfrak{C}^b)$ when $\frac{\mathrm{d}}{\mathrm{d}\eta}z(1)$ or $z(0)$ is involved, is not a problem if only estimates of $z(1)$ are needed due to the degrees of freedom in the choice of $\mathbf{L}$.

Mild solution in Riesz bases

The mild solution of the system Σ_I with $-A$ an S-L operator and the error system Σ^ε as in eqn. (5.16), can be directly written in orthogonal Riesz bases. By Theorem 5.2.1:

$$z(\cdot,t) = U(t,0)z_0(\cdot,t) = \sum_{k=1}^{\infty} e^{\lambda_k t}\phi_k\langle z_0,\phi_k\rangle e^{-\int_0^t b_1 u_1(t)d\tau}$$

with λ_k given in eqn. (5.26) satisfying eqn. (5.25) and with associated eigenvectors ϕ_k as in eqn. (5.26) for the UV disinfection model Σ_I. Similarly, for the error system Σ^ε, λ_k^L is given in eqn. (5.17) satisfying the equation eqn. (5.18), for all $L_{21} \leq 0$, and has associated eigenvectors ϕ_k in eqn. (5.20) for the error dynamics system Σ_I^ε.

5.4 Case II: boundary robust dynamic observer

In the previous section, design conditions for a static gain boundary observer are derived for a CDR process described by $\Sigma_I(A,\mathfrak{B},\mathfrak{C}^b)$, as in eqs. (5.1), (5.4) and (5.12). These design results are illustrated with the observer design for a UV disinfection process. Although this static gain observer has the attractive property of keeping insight in the error dynamics of the closed loop observer system, it requires engineering skills to optimally choose a static gain when there are e.g. disturbances on the input and output signals. As an alternative, we present a boundary observer which is *robust* to these signal perturbations. As a side effect, the design of such a dynamic observer relies heavily on numerical procedures and consequently, insight in dynamics is less transparent (explicit) then with the design of a static gain boundary observer.

In Section 5.3, it is argued that for many processes, it is not practical to implement a boundary observer which is dependent on an observation $\frac{\mathrm{d}z}{\mathrm{d}\eta}(\eta_i)$ for some η_i, $i = \{1,2\}$. For the application of the procedure here, we take the same UV disinfection process as a model case and take the same observer gain matrix $\mathbf{L}$ as in eqn. (5.15), but now with dynamic L_{21}

Furthermore, the assumption that we have only boundary measurements available is weakened, i.e. now $z(\eta^*)$ is the undisturbed observation, where $\eta_1 \leq \eta^* \leq \eta_2$. Hence,

5.4.1 Problem formulation

The aim is to design a dynamic boundary observer which is robust to disturbances on two signals entering the UV disinfection model–observer system, i.e. u_2 (inlet biomass fraction) and y:

$$\mathfrak{B}z := \begin{pmatrix} z(0) \\ \frac{\mathrm{d}}{\mathrm{d}\eta} z(1) \end{pmatrix} = \begin{pmatrix} u_2 + v_1 \\ 0 \end{pmatrix} \quad \text{and} \quad \mathfrak{C}z := z(\eta^*) = y + v_2. \tag{5.30}$$

To be more specific, focus goes to an unbiased estimate of $z(\eta^*)$ with a *causal* filter, such that the disturbances v have a minimal effect on the estimation error at the outlet of the reactor $\varepsilon(1,t)$. Hence, we have to find $L_{21}(t)$ from

$$\inf_{L_{21}\ \text{causal}} \sup_{v} \frac{\|\varepsilon(1)\|_2}{\|v\|_2} < \gamma_H, \quad \varepsilon(\eta, 0) = 0 \tag{5.31}$$

with $\varepsilon = z - \hat{z}$ driven by the error system Σ^ε as in eqn. (5.11), the structure of $\mathbf{L}$ as in eqn. (5.15) and with $\hat{y} = \hat{z}(\eta^*)$. Given disturbances on $z(0)$ and $z(\eta^*)$, the error system becomes

$$\Sigma^\varepsilon_{II} := \left\{ \begin{array}{rcl} \dot{\varepsilon} & = & \frac{1}{\mathrm{p_e}} \frac{\mathrm{d}^2}{\mathrm{d}\eta^2} \varepsilon - b_1 u_1 \varepsilon, \qquad \varepsilon(\cdot, 0) = \varepsilon_0 \\ \begin{pmatrix} \varepsilon(0) \\ \frac{\mathrm{d}\varepsilon}{\mathrm{d}\eta}(1) \end{pmatrix} & = & \begin{pmatrix} v_1 \\ L_{21} * \varepsilon(\eta^*) + L_{21} * v_2 \end{pmatrix} \end{array} \right. \tag{5.32}$$

To solve problem eqn. (5.31), the following sequence of steps is proposed:

- Calculate the transfer function $G_{v\varepsilon}$ from v to $\varepsilon(1)$ and obtain the poles of $G_{v\varepsilon}(s)$.
- Put the problem in a *Linear Fractional Transformation* (LFT) framework,
- Apply approximation techniques to obtain a rational, finite dimensional system and for implementation, reduce the modal approximated model by balanced truncation,
- Synthesize the observer with robust control tools.

Each step in the design procedure will be explained shortly and illuminated with the UV disinfection case.

5.4.2 Transfer function of error system

In this case, the error system for the UV process is defined with a boundary dynamic observer. In other words, Σ^ε_{II} is similar to Σ^ε_I eqn. (5.16), but now with $\gamma_1 = 0$, L time-varying and $(z(0), z(\eta^*))$ disturbed by (v_1, v_2) as in eqn. (5.30). In what

follows, the subscript of L_{21} is dropped and no new notation is introduced for the Laplace transforms of signals, since the meaning will be clear from the context.

By Laplace transformation of the equations and signals in Σ_{II}, the transmission function $G_{v\varepsilon}(s)$ from v to $\varepsilon(1)$ reads

$$G_{v\varepsilon}(s) := \begin{pmatrix} G_1(s) & G_2(s) \end{pmatrix} \tag{5.33a}$$

with

$$\begin{aligned} G_1(s) &= \frac{\varsigma(s)\exp(\frac{1}{2}\mathrm{p_e}) + \exp(\frac{1}{2}\mathrm{p_e}\eta^*)\sinh\left(\varsigma(s)(1-\eta^*)\right)L(s)}{\varsigma(s)\cosh(\varsigma(s)) + \frac{1}{2}\mathrm{p_e}\sinh(\varsigma(s)) - \exp\left(\frac{1}{2}\mathrm{p_e}(-1+\eta^*)\right)\sinh(\eta^*\varsigma(s))L(s)} \\ &=: \frac{R_0(s)}{P(s)-Q(s)L(s)} + R_1(s)\frac{L(s)}{P(s)-Q(s)L(s)} \end{aligned} \tag{5.33b}$$

$$\begin{aligned} G_2(s) &= \frac{\sinh(\varsigma(s))L(s)}{\varsigma(s)\cosh(\varsigma(s)) + \frac{1}{2}\mathrm{p_e}\sinh(\varsigma(s)) - \exp\left(\frac{1}{2}\mathrm{p_e}(-1+\eta^*)\right)\sinh\left(\eta^*\varsigma(s)\right)L(s)} \\ &=: \frac{R_2(s)}{\varsigma(s)}\frac{L(s)}{P(s)-Q(s)L(s)} \end{aligned} \tag{5.33c}$$

with,

$$\varsigma(s) := \sqrt{\left(\tfrac{1}{2}\mathrm{p_e}\right)^2 + \mathrm{p_e}s}. \tag{5.33d}$$

Note that we have used the following shorthand notation with respect to eqs. (5.33b) and (5.33c)

$$R_0(s) := \varsigma(s)\exp\left(\tfrac{1}{2}\mathrm{p_e}\right) \tag{5.34a}$$

$$R_1(s) := \exp\left(\tfrac{1}{2}\eta^*\mathrm{p_e}\right)\sinh\left((1-\eta^*)\varsigma(s)\right) \tag{5.34b}$$

$$R_2(s) := \varsigma(s)\sinh(\varsigma(s)) \tag{5.34c}$$

$$P(s) := \varsigma(s)\cosh(\varsigma(s)) + \tfrac{1}{2}\mathrm{p_e}\sinh(\varsigma(s)) \tag{5.34d}$$

$$Q(s) := \exp\left(\tfrac{1}{2}\mathrm{p_e}(-1+\eta^*)\right)\sinh\left(\eta^*\varsigma(s)\right). \tag{5.34e}$$

$G_{v\varepsilon}$ is also called a *transfer function* on the right half plane $\mathbb{C}_+^\omega := \{s \in \mathbb{C} \mid \mathrm{Re}(s) > \omega\}$, where ω is the growth bound of the semi-group of A^L, [see 118]. The gain $L(s)$ has to be found such that it fulfills the equivalent form of eqn. (5.31) in Laplace domain.

Notice, that for $\eta^* = 1$, eqns. (5.34a)–(5.34e) simplify greatly.

5.4.3 LFT of filtering problem

The estimation problem can be schematically represented as a filtering problem. In this problem, an open-loop error transfer function M and a filter transfer function F is formulated. More specifically, M is defined as mapping from the disturbances v to the states z and an output y_M, i.e. $M : v \mapsto (z, y_M)$. F is formulated as a filter of y_M to an estimate of the state $\hat{z}$, i.e. $M : y_M \mapsto \hat{z}$. A diagram of this filtering problem is shown in Figure 5.3.

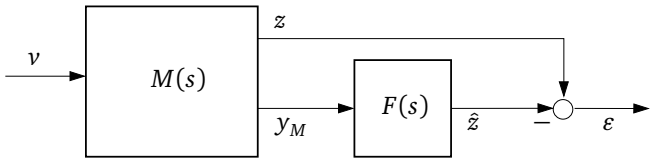


Figure 5.3: Estimation problem presented as a filtering problem of v to ε. M is the open-loop error system, F the to be synthesized filter, and z, $\hat{z}$ and y_M the state, to be estimated state and open-loop system output signal, respectively.

In the closed–loop configuration from v to ε, $G_{v\varepsilon}$ can be related to M and F as in Figure 5.3. This is also known as a *L*inear *F*ractional *T*ransformation. In the sequel we make use of Definition 8.3.11 and Lemma 8.3.12 in [21] for a lower LFT, where $\varepsilon(1)$ will be evaluated in the H_∞–problem of the case study. A well-defined LFT is defined as follows.

DEFINITION 5.4.1. *A lower LFT between two complex-valued matrix functions F and $M = \left(\begin{smallmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{smallmatrix}\right)$, is defined by:*

$$\mathscr{F}_l(M,F) := M_{11} + M_{12}F(I - M_{22}F)^{-1}M_{21} \tag{5.35}$$

Furthermore, $\mathscr{F}_l(M,F)$ is said to be well-defined *(or well-posed) if $\big(I - M_{22}(s)F(s)\big)$ is invertible.*

The configuration of the lower LFT for the estimation problem as shown in Figure 5.3 is represented in Figure 5.4 and is frequently used in H_∞–filtering problems, see *e.g.* section 17.5, p. 462 of [117] or the original paper of Nagpal and Khargonekar [73].

In H_∞–filtering, M is partitioned as follows,

$$\begin{pmatrix} \varepsilon \\ y_M \end{pmatrix} = \underbrace{\begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}}_{M} \begin{pmatrix} v \\ \hat{z} \end{pmatrix}$$

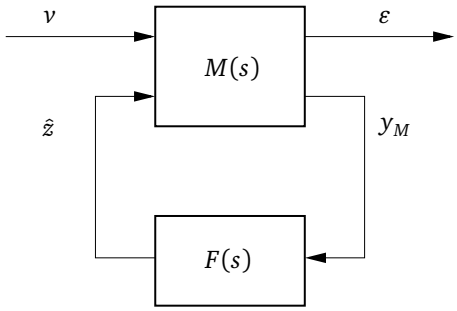


Figure 5.4: Lower fractional transformation for a model–robust-filter configuration as in Figure 5.3, with v, ε, y_F the disturbances, estimation error, input signal and output signal, respectively

and in our case study, with two disturbances:

$$\begin{pmatrix} \varepsilon(1) \\ y_M \end{pmatrix} = \underbrace{\begin{pmatrix} M_{111} & M_{112} & M_{12} \\ M_{211} & M_{212} & M_{22} \end{pmatrix}}_{M} \begin{pmatrix} v_1 \\ v_2 \\ \hat{z} \end{pmatrix} \qquad \text{and with } y_M \text{ to be specified.} \quad (5.36)$$

It is always possible to write a configuration as in Figure 5.3 in an LFT framework. Therefore, the loop from $v \mapsto \varepsilon$ is closed by $G_{v\varepsilon}(M,F) = M_{11} + M_{12}F(I - M_{22}F)^{-1}M_{21}$, whenever M and F are of compatible dimensions and $\det(I - M_{22}F) \neq 0$, see also Definition 5.4.1.

The H_∞-filtering problem as formulated in [117], implies that $M_{12} = -I$ and $M_{22} = 0$. In that case,

$$G_{v\varepsilon}(M,F) = M_{11} - FM_{21} \quad (5.37)$$

We will now seek for an explicit expression of M and F. We could start by looking for a state state representation of M. However, the problem of this route is that the resulting filter F is not guaranteed to be proper. Another way is to directly use the closed loop transfer function $G_{v\varepsilon}$ and try if to split it up in fractions such that the fraction containing L, is proper. This latter route will be explored in the following subsection.

In the sequel, the following generalization of the notions 'proper' and 'strictly proper' are used.

DEFINITION 5.4.2. *The transfer matrix H is said to be* proper *on $\overline{\mathbb{C}_+}$ if for sufficiently large ρ, $\beta > 0$, and* $\mathrm{Re}(s) > \beta$,

$$\sup_{\{s \in \overline{\mathbb{C}_+^\beta} \,||\, |s| \geq \rho\}} |G(s)| < \infty.$$

The transfer matrix H is said to be strictly proper *on $\overline{\mathbb{C}_+}$ if for sufficiently large ρ,*

$$\lim_{\rho\to\infty}\left[\sup_{\{s\in\overline{\mathbb{C}_+^\beta}\,||s|\geq\rho\}}|G(s)|\right]=0.$$

Fraction splitting

The closed-loop transfer function of our example system, eqs. (5.33) and (5.34a)–(5.34e). G_1 can be split into fractions, i.e.

$$G_1(s)=\frac{R_0(s)+R_1(s)L(s)}{P(s)-Q(s)L(s)}=\frac{R_0(s)}{P(s)}+\left(R_1(s)+\frac{Q(s)R_0(s)}{P(s)}\right)\frac{L(s)}{P(s)-Q(s)L(s)} \tag{5.38}$$

$$G_2(s)=\frac{1}{\varsigma(s)}\frac{R_2(s)L(s)}{P(s)-Q(s)L(s)} \tag{5.39}$$

First, we need the following lemma to show that our example system $\Sigma_{II}^{\varepsilon}$ can be written as a well-defined H_∞–filtering problem.

LEMMA 5.4.1. *Let $G:=\mathscr{F}_l(M,F)$ be a given LFT with $M=\left(\begin{smallmatrix}M_{11}&M_{12}\\M_{21}&M_{22}\end{smallmatrix}\right)$ and F transfer functions. Then G is proper if M and F are proper with* $\det(I-M_{22}F)(s)\neq 0$.

Proof. Follows immediately from Definition 5.4.1. □

We propose to define F and M as follows.

PROPOSITION 5.4.1. *$\Sigma_{II}^{\varepsilon}$ with closed loop transfer function $G_{v\varepsilon}$ as in eqn. (5.33) can be written in a well-defined lower LFT form, with*

$$\underbrace{\begin{pmatrix}M_{111}&M_{112}&M_{12}\\M_{211}&M_{212}&M_{22}\end{pmatrix}}_{M}=\begin{pmatrix}R_0P^{-1}&0&-I\\QR_0(PR_2)^{-1}+R_1R_2^{-1}&\varsigma^{-1}&0\end{pmatrix} \tag{5.40}$$

and,

$$F=-R_2\frac{L}{P-QL}=-G_2\varsigma \tag{5.41}$$

if and only if L is proper.

Proof. Clearly, $M_{11}=R_0/P(s)$ acts as a feedthrough of v_1 to $\varepsilon(1)$, see eqs. (5.38) and (5.39). Hence, M_{11} reads:

$$M_{11}:=\begin{pmatrix}M_{111}&M_{112}\end{pmatrix}=\begin{pmatrix}\frac{R_0}{P(s)}&0\end{pmatrix} \tag{5.42}$$

In the case of an H_∞–filtering problem, $I - M_{22}F := I$ and is obviously invertible. Furthermore, following the line of [117], we set $\varepsilon(1) = M_{11}\nu - \hat{z}$ which leads to $M_{12} = -I$. According to eqn. (5.37), $G_{\nu\varepsilon} = M_{11} - F\begin{pmatrix} M_{211} & M_{212}\end{pmatrix}$. As a consequence of the above,

$$M = \begin{pmatrix} R_0P^{-1} & 0 & -I \\ M_{211} & M_{212} & 0 \end{pmatrix}$$

and by setting

$$F := -R_2L/(P - QL),$$

M_{212} reduces to ς^{-1} and M_{211} reduces to $QR_0(PR_2)^{-1} + R_1R_2^{-1}$. Hence, M equals eqn. (5.40).

We now check if M and F are proper. Clearly $M_{212} = (\frac{1}{4}\mathrm{p_e}^2 + \mathrm{p_e}s)^{-1/2}$ is strictly proper. $M_{111} = R_0/P$ can be written as $a[\cosh(\varsigma) + \mathrm{p_e}\sinh(\varsigma)/(2\varsigma)]^{-1}$, with constant $a = \exp(\mathrm{p_e}/2)$. M_{111} converges to 0, in the least favorable scenario that $s = \imath\omega$ and $\omega \to \infty$. Thus $M_{111}(s)$ is strictly proper. In M_{211}, the term $R_1R_2^{-1}$ is clearly strictly proper. The first term $QR_0(PR_2)^{-1}$ is also strictly proper since $\sinh(\eta^*\varsigma)/(\varsigma\sinh(\varsigma)$ is strictly proper and is multiplied by the strictly proper function $\exp(-\frac{1}{2}(1-\eta^*))/P(s)$. Hence, we conclude that M is proper.

Finally, we check whether F is proper. Assume that L is proper. Now, divide the numerator and denominator of F by R_2. Consequently,

$$F = -\frac{L}{\coth(\varsigma) + \frac{\mathrm{p_e}}{2\varsigma} - \left(\exp\left(-\frac{1}{2}\mathrm{p_e}(1-\eta^*)\right)\frac{\sinh(\eta^*\varsigma)}{\varsigma\sinh(\varsigma)}\right)L}$$

Recognize that, in the denominator of F

$$\underbrace{\frac{\coth(\varsigma)}{I}}_{\text{proper}} + \underbrace{\frac{\mathrm{p_e}}{2\varsigma}}_{\text{strictly proper}} + \underbrace{\exp\left(-\frac{1}{2}\mathrm{p_e}(1-\eta^*)\right)\frac{\sinh(\eta^*\varsigma)}{\varsigma\sinh(\varsigma)}}_{\text{strictly proper}}\underbrace{\frac{L}{I}}_{\text{proper}}.$$

Furthermore its limit equals 1 and so F is proper if and only if L is proper. □

From eqn. (5.41), we find $L = PF/\left(-R_2 + QF\right)$.

With the above results, the H_∞–problem formulation as given in eqn. (5.31) can be straightforwardly translated into an equivalent formulation in frequency domain.

PROPOSITION 5.4.2 [H_∞–FILTERING PROBLEM IN FREQUENCY DOMAIN]. *Given $\Sigma_{II}^{\varepsilon}$ as in eqn. (5.32), then the H_∞–filtering problem as in eqn. (5.43) is equivalently described by the H_∞ optimization:*

$$\inf_{F\ proper} \|\mathscr{F}(M,F)\|_\infty < \gamma_H \tag{5.43}$$

with $\mathscr{F}_l(M,F)$ an LFT between $M(s)$ and $F(s)$, with $M(s)$ given in eqn. (5.40) and $F(s)$ related to $L(s)$ as in eqn. (5.41).

The above used H_∞-norm is given in the Appendices, Lemma A.4.1.

So, all that remains is to calculate $F^{\text{opt}}(s)$ in the H_∞–problem (see Proposition 5.4.2) to obtain $L^{\text{opt}}(s)$. This calculation is only possible by solving the H_∞-problem by a numerical routine, in which $M(s)$ is first approximated to a rational transfer function. The approximation of M and analysis of the residual error convergence is described in Appendix D.4. It follows that, in most cases, a modal truncation up to $N = 25$, Padé-$[1,2]$ approximation of M_{212} and subsequent balanced truncation to around $m = 2$ Hankel singular values (except for small Péclet numbers, i.e. $\mathrm{p_e} \ll 1$ where we choose $m = 8$), results in suitable observer filters F and L of low order. These results are presented in Section 5.6. But first, we turn to the results of Case I, i.e. the static observer design.

5.5 Numerical results of Case I

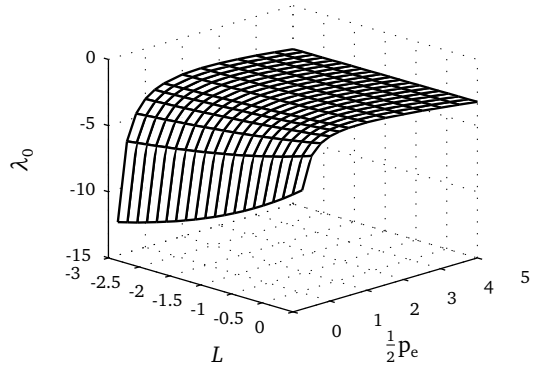
Now that we have formulated the system and its observer, it is possible to study the influence of the observer gain L_{21} on the eigenvalues of the error system. With slight abuse of notation, we omit the sub- and superscript of L in the figures.

5.5.1 Parametric influence on growth bound

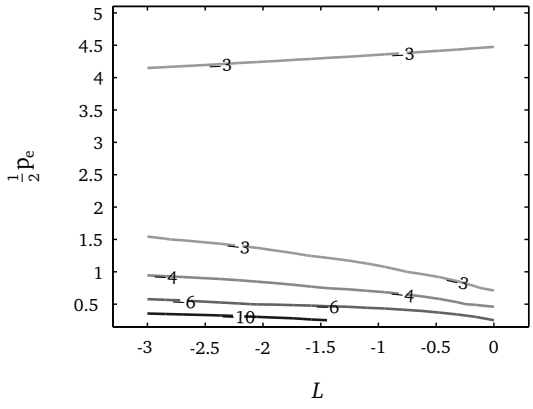
Figure 5.5(a) and its equivalent contour plot Figure 5.5(b), show the behavior of the growth bound λ_0 for $\mathrm{p_e} \in (0,10]$ and $L \in [-3,0]$. Indeed, the larger the $\mathrm{p_e}$-numbers, the larger the growth bound and the lesser the effect of the observer gain. Notice also that for smaller L, the growth bound tends to zero and the stability margin becomes smaller.

5.5.2 Parametric influence on growth bound improvement

For (control) engineering applications, it may be more interesting to find out the magnitude of L_{21} for a given performance *increase* $\Delta_\lambda^{\max} := \lambda_0 - \lambda_0^L$. In eqn. (5.27), bounds on Δ_λ are given with the aid of Remark 7. In addition, we now calculate $L^\star := L_{21}$ for several *Péclet*-numbers at which a given Δ_λ is obtained. The results are depicted in Figure 5.6. We see that, for increasing $\mathrm{p_e}$, $L^\star$ increases rapidly for some Δ_λ. For reference, the line $L = 0$ is also shown in Figure 5.6(a). Note that the performance gain Δ_λ can only be achieved for a certain range of Péclet-numbers, see Section 5.3.4 and Figure 5.2(b).

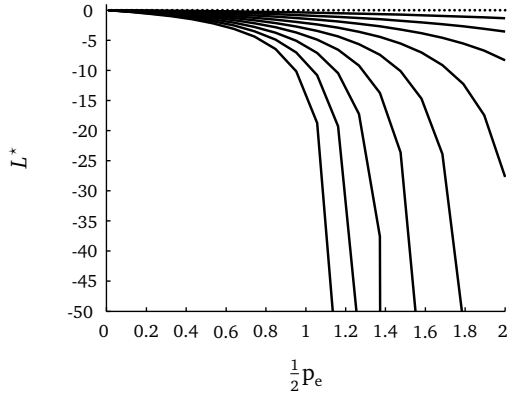


(a) 2D plot of λ_0

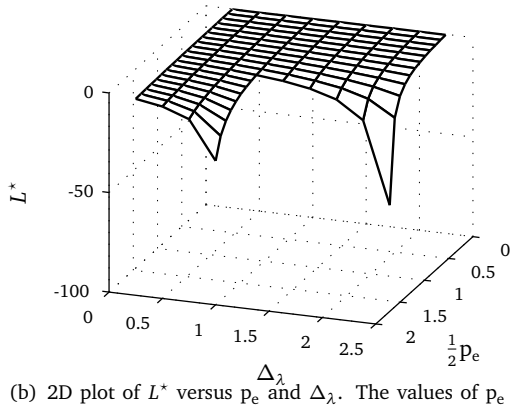


(b) Contour lines of λ_0

Figure 5.5: λ_0 versus $-3 < L < 0$ and $0 < \mathrm{p_e} \leq 10$



(a) $L = 0$ (dashed) and $L^\star$ for $0 < p_e \leq 4$ and varying Δ_λ, *i.e.* from upper to lowest line $\Delta_\lambda = 0.25$ to 2.25 (solid) with steps of 0.25.



(b) 2D plot of $L^\star$ versus p_e and Δ_λ. The values of p_e for which the performance gain Δ_λ, is not feasible (i.e. no solutions exist), is depicted by empty space.

Figure 5.6: $L^\star$ versus Δ_λ and p_e

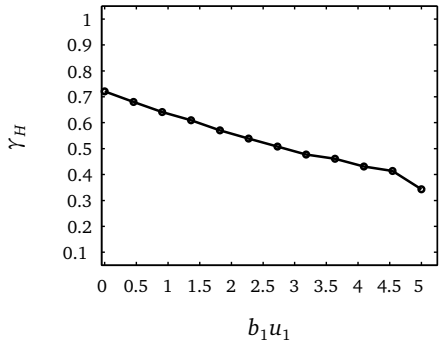
5.6 Numerical results of Case II

In this section, we investigate the influence of different parameter/variable values on the γ_H–iteration value, F^{opt} and $G_{\nu\varepsilon}$. We start with γ_H–sensitivities.

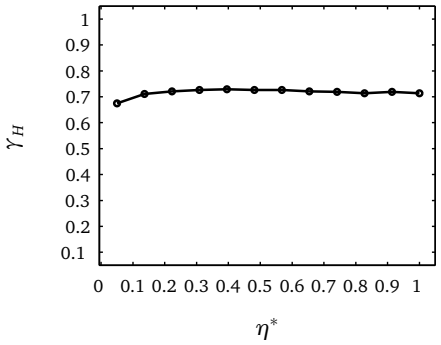
5.6.1 Parametric influence on observer optimality

γ_H–Sensitivies to different parameters

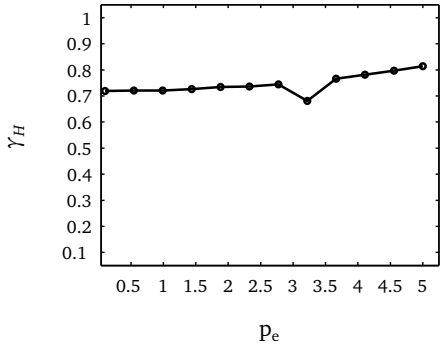
Notice from eqn. (5.43) that a constant γ_H for different parameter values means, that the same disturbance rejection is achieved and, consequently, the optimization of F is relatively insensitive to this parameter. A varying γ_H with different parameter values, means it is sensitive to this parameter. Figure 5.7 depicts the γ_H–values under different parameter cases. Simulations run with $N = 25$ and $m = 17$.



(a) γ_H under varying b_1u_1. (p_e=0.5 and η^*=0.5).



(b) γ_H under varying η^*. (p_e=0.5 and b_1u_1=0).



(c) γ_H under varying p_e. (b_1u_1=0 and η^*=0.5).

Figure 5.7: γ_H–Iteration values under different b_1u_1, η^* and p_e values.

Bode diagrams

Here, we show the influence of different values of $b_1 u_1$, $\mathrm{p_e}$, and η^* on the synthesis of L^{opt} and F^{opt} in Figure 5.8. The bode amplitude graphs illustrate the low frequency sensitivity of the filters F and L for the lamp strength and $\mathrm{p_e}$–number, which is exactly what we would expect from the eigenvalue analysis in Section 5.3 (see also Figure 5.6 for the influence of $\mathrm{p_e}$). Notice also that the amplitude of L remains more or less at the same level across the frequency domain, irrespectable of the varied parameter. It should be noted that for small $\mathrm{p_e}$-values, the truncation number m had to be increased to $m = 8$ to have acceptable approximation errors. For the other parameters, a truncation to $m = 2$ yielded an acceptable error trade-off.

By inspecting Figure 5.8(c)–(d) and (e), but also Figure 5.7(b), it can be seen that the synthesized filter is hardly sensitive to a varying sensor position. The filter synthesis seems to be more responsive to changes in the Péclet number, especially in the lower frequency regions, see Figure 5.8.

This is also the case for the UV lamp strength. In the latter case, the eigenvalues change linearly with $b_1 u_1$, affecting the infemum γ_H of the optimization problem, see also Figure 5.7(a).

5.6.2 Parametric influence on open-loop performance

In Figure 5.9, we only show the influence of the Péclet-number and the sensor position on the open-loop performance of the plant M, since the influence of the lamp strength is obvious.

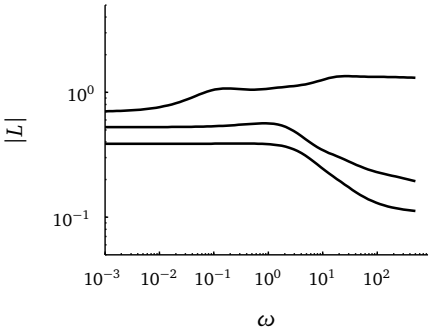
With increasing Péclet number, the amplitude decreases for M_{111}, M_{211} and M_{212}. The absolute magnitude of M_{111} is only affected in the higher frequency regions, whereas the other open loop magnitude plots show amplitude changes over the whole frequency range.

Whenever the sensor is getting closer to the outlet of the reactor, the magnitude plot of M_{211}, which is indeed a function of η^*, shows a decreasing amplitude in the high frequency region. This can be contributed to the eigenvector characteristics which in M_{211} are dependent on η^*.

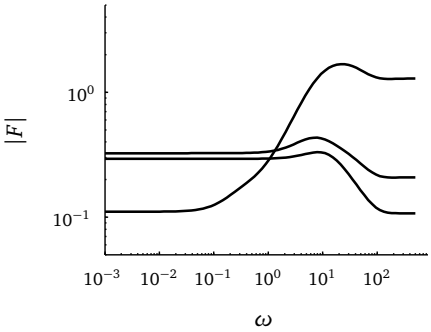
5.6.3 Parametric influence on closed-loop performance

The closed loop response from $v \mapsto \varepsilon(1)$ is interesting, since it reflects the error suppression originating from an inexact boundary estimate of the incoming concentration of micro-organisms and the initial boundary effect (M_{111}).

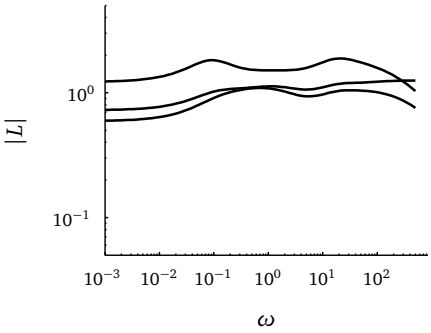
In Figure 5.10 both the closed loop responses from v_1 and v_2 to $\varepsilon(1)$ have been depicted. The closed loop responses under optimal filtering (i.e. with F^{opt}) and under a naive filter $F = I$ are compared. Notice that for almost every parameter variation, the robust filter performs better, especially in the lower frequencies.



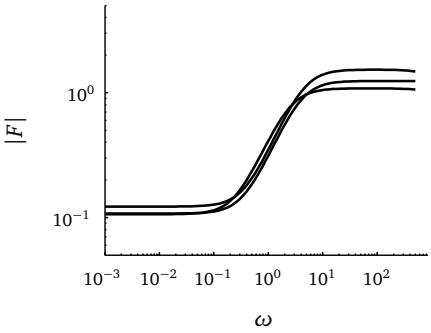
(a) L^{opt} for $b_1u_1 \in \{0, 2.5, 5\}$. Amplitude decreases with increasing b_1u_1.



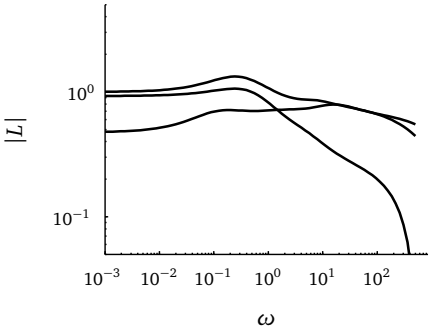
(b) F^{opt} for $b_1u_1 \in \{0, 2.5, 5\}$. Amplitude decreases with increasing b_1u_1 in high frequency region.



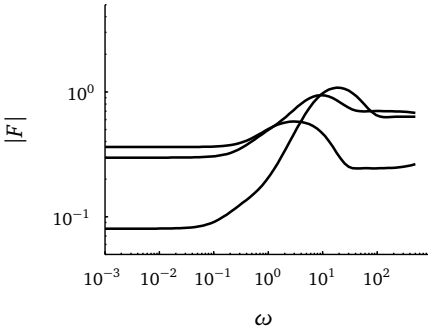
(c) L^{opt} for $\eta^* \in \{0.25, 0.60, 1\}$. Amplitude decreases with increasing η^*.



(d) F^{opt} for $\eta^* \in \{0.25, 0.60, 1\}$. Amplitude increases slightly with increasing η^*.

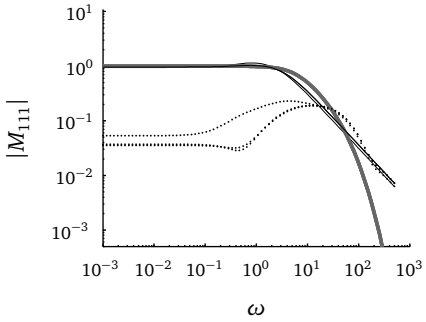


(e) L^{opt} for $\mathrm{p_e} \in \{0.5, 1.25, 2\}$. Amplitude increases and rotates slightly clockwise with increasing $\mathrm{p_e}$.

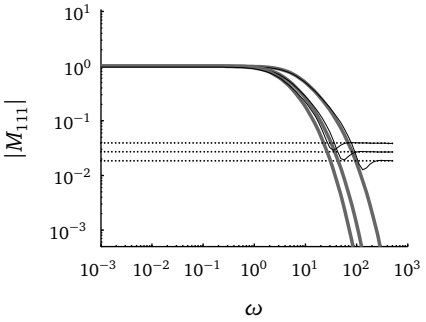


(f) F^{opt} for $\mathrm{p_e} \in \{0.50, 1.25, 2\}$. Amplitude increases and rotates slightly counterclockwise with increasing $\mathrm{p_e}$.

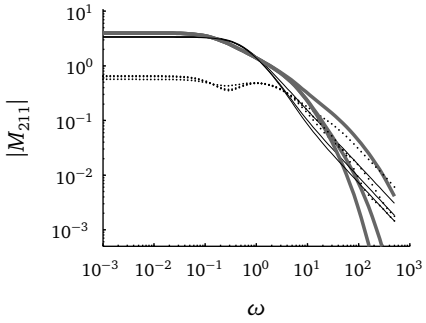
Figure 5.8: Optimal filter $F(s)$ and L^{opt} for different b_1u_1, η^* and $\mathrm{p_e}$ values.



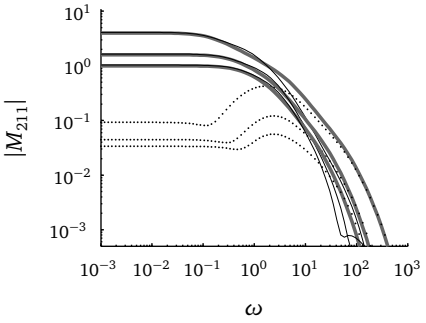
(a) M_{111} with $\eta^* \in \{.25, .60, 1\}$. No difference in changing η^*.



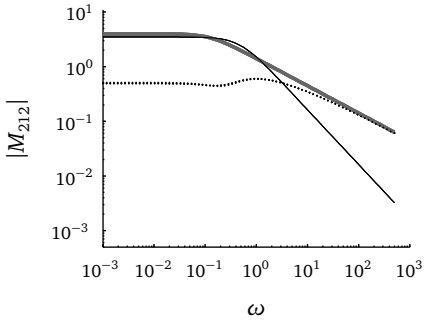
(b) M_{111} with $p_e \in \{0.50, 1.25, 2\}$. Lower line, largest p_e



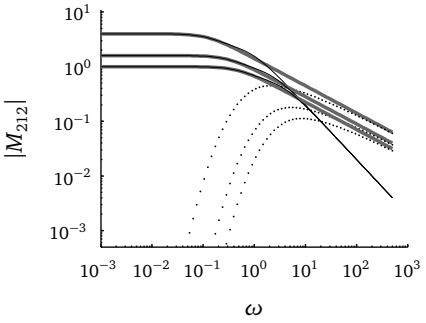
(c) M_{211} with $\eta^* \in \{.25, .60, 1\}$. With same linestyle, upper line: lowest η^*, lower line: largest η^*.



(d) M_{211} with $p_e \in \{0.50, 1.25, 2\}$. Lower line, largest p_e.

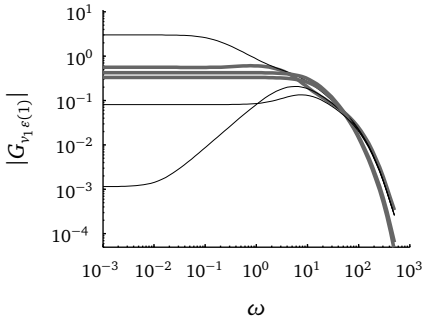


(e) M_{212} with $\eta^* \in \{.25, 0.60, 1\}$. No difference in changing η^*.

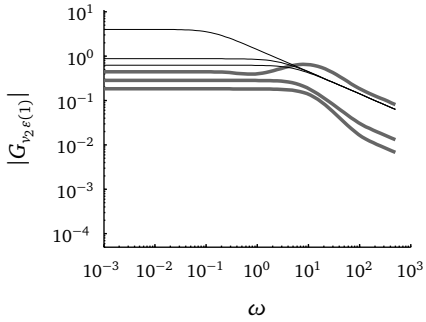


(f) M_{212} with $p_e \in \{0.50, 1.25, 2\}$. Lower line, largest p_e.

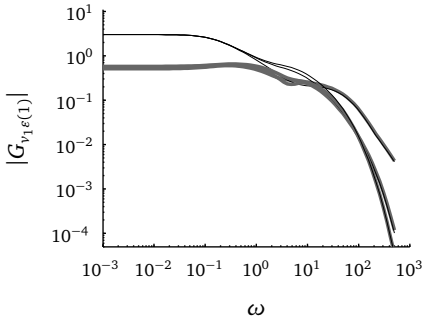
Figure 5.9: M [—] with truncated approximation M^m [–] and approximation error [$\cdot\cdot$] for different values of η^* (left; $m = 2$) and p_e (right; $m = 17$).



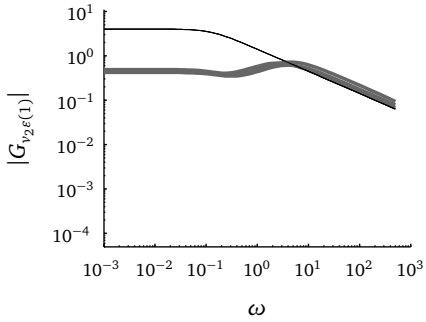
(a) $G_{v_1\varepsilon}$ for $b_1u_1 \in \{0, 2.5, 5\}$. Same linestyle, from upper to lowest line: increasing b_1u_1.



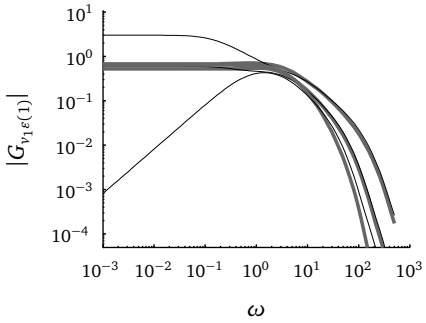
(b) $G_{v_2\varepsilon}$, $b_1u_1 \in \{0, 2.5, 5\}$. Same linestyle, from upper to lowest line: increasing b_1u_1.



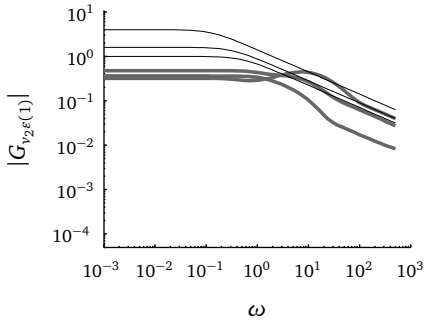
(c) $G_{v_1\varepsilon}$, $\eta^* \in \{.25, .60, 1\}$. Same linestyle, from upper to lowest line: increasing η^*.



(d) $G_{v_2\varepsilon}$, $\eta^* \in \{.25, .60, 1\}$. Same linestyle, from upper to lowest line: increasing η^*.



(e) $G_{v_1\varepsilon}$, $\mathrm{p_e} \in \{.5,\ 1.25, 2\}$. Same linestyle, from upper to lowest line: increasing $\mathrm{p_e}$.

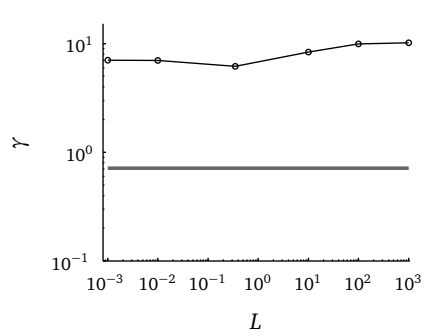


(f) $G_{v_1\varepsilon}$, $\mathrm{p_e} \in \{.5, 1.25, 2\}$. Same linestyle, from upper to lowest line: increasing $\mathrm{p_e}$.

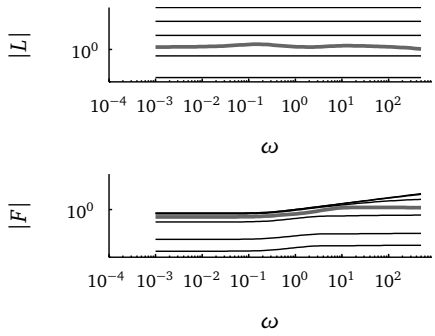
Figure 5.10: Amplitudes of $G_{v_1\varepsilon}$ (left) and $G_{v2\varepsilon}$ (right) for $F^{\mathrm{opt}}(s)$ [—] and $F = I$ [-] with different b_1u_1 (a) and (b) η^* (c) and (d) and $\mathrm{p_e}$ values (e) and (f).

5.6.4 Performance comparison Case I and II

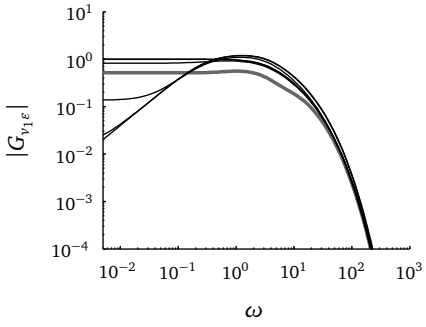
What might perhaps be more interesting, is how values of a constant gain matrix L affects the norm on $G_{v\varepsilon}$ and how this compares to an calculated F^{opt}. Results of these calculations are shown in Figure 5.11, including the γ_H–iteration value of the H_∞–problem. The following parameter values were used for the simulation: $\text{p}_\text{e} = 0.50$, number of modes $N = 25$ truncated to $m = 8$ Hankel modes, $\eta^* = 1$ and $b_1 u_1 = 0$.



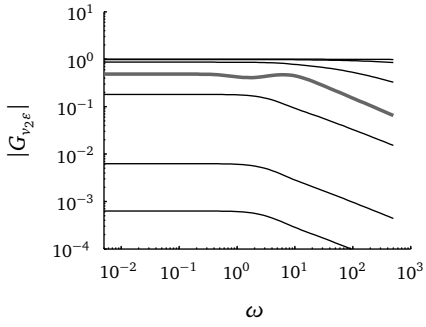
(a) γ_H, given F^{opt} and varying L.



(b) F^{opt} and L^{opt} and F based upon varying L.



(c) $G_{v_1\varepsilon}$ based upon F^{opt} and varying L.



(d) $G_{v_2\varepsilon}$ based upon F^{opt} and varying L.

Figure 5.11: Influence of optimal, dynamic F^{opt} [—] and non-optimal static gain filter F [–] based upon $L \in \{-0.01, -0.35, -10, -100, -1000\}$. From upper ($L = -1000$) to lowest line ($L = -0.01$), L decreases.

5.7 Concluding remarks

We organize some concluding remarks per case and the comparison of the cases.

Case I (static): Inspired by CDR processes in food and water treatment industry, we analyzed approximate observability and detectability in the case where only boundary measurements are available. With the analysis, rules have been derived for the design of a static, Luenberger boundary observer with guaranteed detectability of the resulting estimation error system. These design rules have been implemented and tested in a numerical study of a UV disinfection process case. In this example case, we come to the following conclusions:

- From the eigenvalue analysis and numerical calculations, it follows that for mild Péclet-numbers, ($\mathrm{p_e} \ll 1$, hence a low convection–diffusion ratio), there is more room to obtain a performance gain with a suitable observer gain L_{21}. For large Péclet numbers, fast process dynamics already push the estimation error to zero. In this case one may decide to choose a small positive $L_{21} > \frac{1}{2}\mathrm{p_e}$ as a smoothing filter.
- The growth bound of the error system is pushed to higher absolute magnitudes whenever the lamp strength is stronger, i.e. for increasing $b_1 u_1$.

Furthermore, analyzing the system in the infinite dimensional setting gives a good impression how the error system will behave, *independent* of some choice of discretization or approximation method.

Case II (dynamic): It has been shown, that it is possible to design an H_∞–observer on the basis of a estimation correction at the boundary of the estimation error system. A sequence of steps is performed for the observer synthesis: Laplace transformation of $\Sigma(A^L, \mathfrak{B}, \mathfrak{C})$ to obtain the closed-loop transfer function G, formulate the H_∞–problem in an LFT framework and calculate the optimal H_∞–filter after approximation of the open-loop model M and its balanced truncation. From the closed-loop solution G and the appropriate choice of M, the relationship of F with L can be deduced. We show some results on the robust boundary observer synthesis for the UV disinfection process under a disturbance rejection H_∞–criterion.

Comparison static L with dynamic L^{opt}: As expected, a dynamic robust observer synthesis approach is needed if disturbance rejection is desired to achieve the robust criterion. However, a static gain observer has the clear advantage that the observer gain preserves system insight, while a dynamic observer synthesis is based upon numerical routines. For the UV disinfection example, we see that a static $L_{21} \to 0$ pushes the maximal magnitude of the transfer function $G_{v2\varepsilon}$ to zero, but it does not minimize the maximal magnitude of $G_{v1\varepsilon}$, optimally.

6

Conclusions and Remarks

In this work, different procedures with respect to estimation and prediction of systems characterized by convection, diffusion and reactions on the basis of point measurement data, have been studied. In Chapter 1, the following research issues with respect to these procedures have been postulated, namely:

1. *Given a late reduction technique, is it possible to preserve physical knowledge of the nominal CDR model in terms of the parameter estimate θ, when using a typical identification output-error technique?*

2. *Given an early reduction technique, is it possible to rewrite the estimation and prediction problem of a CDR system into a linear regression, and if so, how?*

3. *Without model order reduction, is it possible to obtain design conditions for a static observer of a CDR system, under the condition that it is (approximately) observable?*

4. *Given a late reduction technique, is it possible to obtain a dynamic observer which is robust to disturbances at both the (boundary) input as well as the point observation?*

The above questions have been related to two applications which are introduced in Chapter 2. The questions are answered case by case in the following section.

6.1 Concluding remarks

1. In Chapter 3, an identification approach based on linear time-invariant discrete time OE models, showed that a heat transfer process in a typical food storage setup can be suitably modelled by first order, approximate model. The identification amounts to the calibration of the dominant time constant and gain under different ventilator settings.

   In addition, local parametric sensitivity analysis was carried out to inspect how physical parameters affect this time constant. Furthermore, for all ventilator settings the gain of the bulk storage process was close to one, indicating that there was no unexpected heat loss or gain. The conclusion is that important physical parameters can be individually recovered from the calibrated OE model parameters.

2. In general, CDR models are written in (Hilbert) state-space form. Via early model reduction by finite differences techniques a CDR model can be discretized to a discrete-time, linear state space model belonging to the class of discrete-time, linear structured systems Σ_d. This system Σ_d is the starting point in Chapter 4 for obtaining a linear regressive realization Σ_d^{LR} which, when it exists, largely facilitates parameter estimation and prediction.

It is shown that under parameter-affine conditions, Σ_d^{LR} can be found via reparametrizations of the transfer function of Σ_d. In this realization, a linear regressive mapping from newly defined parameters to input-output data is formulated and characterized by parameter independent coefficient matrices **N** and **R**. These coefficient matrices have been explicitly expressed in terms of number of compartments n, actuator and sensor location for two discretized diffusion examples. Under specific boundary conditions, these explicit expressions are also applicable to dimensionless representations of discretized CDR-systems.

Besides convexity of the resulting parameter estimation problem, another advantage is the ability to analyze *a priori* the parametric sensitivities on the regression vector ϕ based on the explicit results. Furthermore, a simple identifiability test on the basis of a rank test of **R** has been obtained. When ordinary least squares is used to solve the linear regression rpoblem, a disadvantage turns up. As the regressors contain input and output data, they are corrupted by measurements noise, thus deteriorating the estimation result. Hence, more advanced least-squares techniques are needed for the estimation of the parameters.

3. The UV disinfection process is used as an illustrating example for the design of a static, Luenberger-type observer for a CDR process in Section 5.3 of Chapter 5. Such a process is typically described by a Sturm-Liouville operator. Consequently, detectability and approximate observability results are obtained for systems characterized by a Sturm-Liouville operator and boundary actuation and/or sensing. The results are used in deducing pole placement design rules for an observer of a UV disinfection model using boundary data. By numerical simulation, the observer characteristics are further illustrated (Section 5.5).

   This method has the advantage of obtaining results about observer (or state estimator) properties as a function of the underlying dynamic system properties, independent of some approximation or model reduction technique. However, in practice, a finite dimensional form of such an observer has to be implemented and caution must be taken to generalize infinite-dimensional observer results to the finite-dimensional case. Furthermore, it requires engineering skills to choose an optimal observer gain under given (practical) circumstances. Last, the observer performance in the CDR application is highly dependent on the CDR system dynamics.

4. In Section 5.4 of Chapter 5, it is shown that it is possible to design an H_∞–observer aiming at disturbance rejection on the basis of an estimation correction at the boundary of the error system. Hereto, a sequence of steps is performed for the observer synthesis: (i) Laplace transformation of the estimation error system to obtain the closed-loop transfer function G, (ii) formulation of the H_∞–problem in a well-posed linear fractional transformation

framework, (iii) approximation and subsequent balanced truncation of the open-loop model M and (iv) the calculation of the optimal H_∞–filter with the aid of the approximation of M.

Numerical results on the robust filter, the H_∞–synthesis performance and closed loop transfer function of the estimation error in a UV disinfection model under disturbance rejection are shown in Section 5.6. Also, a numerical comparison with the static observer obtained in the first part of the Chapter gives a strong indication of the superiority of the H_∞-observer with respect to disturbance rejection.

Another advantage of the numerical procedure to obtain the dynamic observer, is that it is generally applicable to linear, distributed parameter systems where only point measurements and actuation is available. A disadvantage is that a finite-dimensional approximation of the open-loop transfer function of the estimation error system has to be made. In general, an approximation leads to a loss in physical insight.

6.2 General comments and future work

While executing this research, several issues were encountered that may be equally important and therefore deserve further study:

- The fundamental model characterization problem. In other words, whether CDR equations satisfactorily describe the processes (like e.g. flow patterns) encountered in the case studies. There are numerous model candidates available, varying from rational (finite-dimensional) input-output systems, rational input-output systems with fractional representation or more complex infinite dimensional systems which have to be identified with experimental data. In this thesis, CDR models in the most simple form are used and several procedures have been discussed. As an alternative, it would be interesting to start with grey-box identification where one of the goals is to come up with a model selection criterion where the (physical) model reduction error is weighted against prediction errors and number of (lumped physical) parameters in some norm.

- The model reduction problem. That is, whether a model reduction step should be in an early stage or in a later stage. In most cases, a finite-dimensional implementation of a state or parameter estimator has to be used in practice. Optimal controller or observer order reduction projections have been reported for the finite dimensional case and are also developed for the infinite dimensional case. Adaptive (and possibly recursive) reduction methods depending on (some statistical) measure of information content in input-output data may be an attractive alternative.

- The coupled model reduction–estimation problem applied to CDR processes. To our knowledge, no coherent inventory has been made of estimation methods that are applicable to (model reduced) CDR systems under some specific identification criterion. To this aim, several combinations of model order reduction methods and system identification techniques have to be investigated, like in the work of Söderström and Bhikkaji for parameter estimation of diffusion systems [92, 93].

- The application of recursive estimation techniques. The study of using the estimation methods outlined in this thesis in a recursive fashion, is still open. For example, it is straightforward to implement the linear regressive technique in Chapter 4 recursively. Given a recursive implementation, it may be worthwile to implement the proposed methods, possibly in combination with an optimal input calculation routine in tandem as in [95].

6.3 Epilogue

The less technical, but interested reader may still wonder what the added value is when using one of the proposed estimation and/or prediction techniques in practice. We try to briefly summarize in what we think is valuable.

First of all, the aim has been to work out techniques for dynamic models that are at the basis of on-line process control. Hence, simple models are proposed or taken from literature to allow fast simulation and avoid the numerical burden of simulation of (complicated) distributed parameter models.

Secondly, mathematical concepts like *identifiability (i.e., is it possible to recover parameter values from data, given the used estimation method and model?)*, *detectability (does the evolution of the estimation error reside within bounds, or roughly said: will it not 'explode'?)* and *observability (when some measurements are available, is the answer about the underlying state like temperature or concentration, the only one?)* are important checks for engineers when using an estimation or prediction method.

Thirdly, it seems irrational to ignore the widely available physical knowledge reported in literature. Therefore the desire is to use (approximate) physical model-based control and identify these models. All the proposed estimation and prediction techniques in this thesis, preserve physical insight in some way, namely:

- via recovery of physical parameters using powerful black-box identification techniques;

- by use of linear regression for estimation and prediction. This way, the physically interpretable structures appear also in the estimation problem and can even be written as algebraic relations for CDR systems, so that e.g. suitable input signals can be calculated beforehand;

- via the design rules of the static boundary observer, the link to the observer performance with respect to physical parameters is transparent. Physical knowledge is preserved in a lesser extend in the case of the dynamic observer, although the dynamic observer possesses the same characteristics as the physical model.

Appendices

A

Basic functional analysis concepts

The results in this chapter are well known, but repeated here for the comfort of the reader. Sources are the Appendices in [21, 26].

A.1 Normed linear spaces

DEFINITION A.1.1 [VECTOR SPACE]. *Let X be a nonempty linear vector space and let $\mathbb{K}$ be a set of scalars (i.e. $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$). Furthermore, let $(x, y) \mapsto x + y \in X$ (addition) and $(\alpha, x) \mapsto \alpha x \in X$ (scalar multiplication) for all $x, y \in X$ and $\alpha \in \mathbb{K}$. Then, $\forall x, y, z \in X$, the vector space X over $\mathbb{K}$ has the following properties:*

a. $x + y = y + x$ *(commutative)*

b. $(x + y) + z = y + (x + z)$ *(associative)*

c. $\exists 0_X$, $0_X + y = y$ *(existence zero)*

d. $1x = x$, $1 \in \mathbb{K}$ *(one)*

e. $\exists x \in X$, $x + (-x) = 0$ *(existence inverse)*

f. $\alpha, \beta \in \mathbb{K}$, $\alpha(\beta u) = (\alpha\beta)u$

g. $\alpha, \beta \in \mathbb{K}$, $\alpha(x + y) = \alpha x + \alpha y$ *and* $(\alpha + \beta)x = \alpha x + \beta x$ *(distributitive).*

DEFINITION A.1.2 [NORMED SPACE]. *Let X be a linear vector space. Then a norm is a nonnegative set function on X, $\|\cdot\| : X \mapsto \mathbb{R}_+ = [0, \infty)$, such that*

a. $\|x\| = 0$ *if and only if* $x = 0$,

b. $\|\alpha x\| = |\alpha|\|x\|$, $\forall x \in X$ *and* $\forall \alpha \in \mathbb{K}$.

c. $\|x + y\| \leq \|x\| + \|y\|$

The linear vector space X induced under the norm $\|\cdot\|$ is called a normed vector space.

DEFINITION A.1.3 [DENSE SPACE]. *A subset V of a normed linear space X is* dense *in X if its closure is equal to X.*

This property means that we may approximate every $x \in X$ as closely as we like by some $v \in V$, i.e., for any $x \in X$ and $\varepsilon > 0$, $\exists v \in V$ such that $\|v - x\| < \varepsilon$.

DEFINITION A.1.4 [SEPERABLE SPACE]. *A normed linear space $(X, \|\cdot\|)$ is* separable *if it contains a dense subset that is countable.*

DEFINITION A.1.5 [CAUCHY SEQUENCE]. *Let $(X, \|\cdot\|_X)$ be a normed linear space. A sequence $x_n \subseteq X$ is a Cauchy sequence if*

$$\lim_{m,n\to\infty} \|x_m - x_n\| = 0$$

Every Cauchy sequence in $\mathbb{R}$ is convergent. This does not hold for general normed linear spaces. Spaces in which Cauchy sequences always have a limit are called *complete* normed spaces.

DEFINITION A.1.6 [BANACH SPACE]. *A normed space is complete, whenever every Cauchy sequence has a limit in X. A Banach space is a complete, normed linear space.*

For the definition of a Hilbert space, we need the following notion.

DEFINITION A.1.7 [INNER PRODUCT]. *An* inner product *on a linear vector space X defined over the scalar field $\mathbb{K}$ is a map $\langle\cdot,\cdot\rangle : X \times X \mapsto \mathbb{K}$ such that $\forall x,\ y \in X$ and $\forall \alpha,\ \beta \in \mathbb{K}$ it holds that*

a. $\langle \alpha x + \beta y, z\rangle = \alpha\langle x, z\rangle + \beta\langle y, z\rangle$;

b. $\overline{\langle x, y\rangle} = \langle y, x\rangle$;

c. $\langle x, x\rangle \geq 0$ *and* $\langle x, x\rangle = 0$ *if and only if* $x = 0$.

DEFINITION A.1.8 [HILBERT SPACE]. *A Hilbert space is an inner product space that is complete as a normed linear space under the induced norm,*

$$\|x\| = \sqrt{\langle x, x\rangle}$$

DEFINITION A.1.9 [ORTHONORMAL BASIS OF HILBERT SPACES]. *Let H be a separable Hilbert space under the inner product $\langle\cdot,\cdot\rangle$ and let ϕ_n be an orthonormal basis. Then an orthogonal subset $\{\phi_n\}_{n\geq 1}$ of H is called an* orthonormal basis *if the following conditions hold:*

a. $\{\phi_n\}_{n\geq 1}$ *is maximal, i.e.,* $\overline{\operatorname{span}_{n\geq 1}\phi_n} = H$.

b. for $\{\phi_n\}_{n\geq 1} \notin \emptyset$, *the inner product* $\langle \phi_n, \phi_m\rangle = \delta_{n,m} := \begin{cases} 1 & \text{if } n = m \\ 0 & \text{if } n \neq m \end{cases}$

In other words, $(\phi_n)_{n\geq 1}$ are mutually orthogonal unit vectors. Note that in $\mathbb{R}^n$, any element can be expressed as a linear combination of any set of n mutually orthogonal elements (i.e. a basis) as in condition a. of Definition A.1.9.

DEFINITION A.1.10 [FOURIER EXPANSION]. *Let H be a separable Hilbert space under the inner product $\langle\cdot,\cdot\rangle$. Then for any $x \in H$, we have the Fourier expansion*

$$\forall x \in H, \quad x = \sum_{n=1}^{\infty} \langle x, \phi_n\rangle \phi_n$$

The terms $\langle x, \phi_n\rangle$ are called the Fourier coefficients of x with respect to ϕ_n. Furthermore, we have the important Parseval equality. *Any two vectors $x, y \in H$ satisfy:*

$$\langle x, y\rangle = \sum_{n=1}^{\infty} \langle x, \phi_n\rangle \overline{\langle y, \phi_n\rangle}.$$

In particular, for $x = y$ we have $\|x\|^2 = \sum_{n=1}^{\infty} |\langle x, \phi_n \rangle|^2$.

DEFINITION A.1.11 [L_p SPACES]. *For any fixed real number $p \in [1, \infty)$, the Banach space $L_p[a, b]$ is the completion of the normed space which consists of all continuous real-valued functions on $[a, b]$, $-\infty \leq a < b \leq \infty$, and the norm*

$$\|x\| = \left(\int_a^b |x(t)|^p \mathrm{d}t \right)^{\frac{1}{p}}$$

with $\int_a^b |x(t)|^p \mathrm{d}t$ finite.

A.2 Operators

In this section, transformations Q from one normed linear space X to another Y are treated. Usually, X and Y will be either Banach or Hilbert spaces and Q will be linear.

In the following theorem, the mapping Q does not have to be linear. Notice that the operator U in Chapter 5 is a contraction.

THEOREM A.2.1 [CONTRACTION MAPPING THEOREM]. *Let X be a Banach space, Q a mapping from X to X, $m \in \mathbb{N}$ and $\alpha < 1$. Suppose that Q satisfies $\|Q^m(x_1) - Q^m(x_2)\| \leq \alpha \|x_1 - x_2\|$ for all $x_1, x_2 \in X$ Then, there exists a unique $x^* \in X$ such that $Q(x^*) = x^*$; x^* is the fixed point of Q. Furthermore, for any $x_0 \in X$, the sequence $\{x_n, n \geq 1\}$ defined by $x_n := Q^n(x_0)$ converges to x^* as $n \to \infty$.*

Proof. See [55, 152, theorem 5.4-3]. □

DEFINITION A.2.1 [BOUNDED LINEAR OPERATOR]. *Let X and Y be normed linear spaces and Q a linear operator from $D(Q) \subset X \mapsto Y$. Q is a* bounded linear operator *or Q is* bounded, *if $\exists \alpha > 0$ such that for all $x \in X$,*

$$\|Qx\|_Y \leq \alpha \|x\|_X.$$

Continuity and boundedness are equivalent concepts for linear operators.

THEOREM A.2.2. *Let X and Y be normed linear spaces. If $Q : D(Q) \subset X \mapsto Y$ is a linear operator, then:*

a. Q is continuous if and only if Q is bounded;

b. if Q is continuous at a single point, it is continuous on $D(Q)$.

Proof. See [55, theorem 2.7-9]. □

DEFINITION A.2.2 [GRAPH]. *Let X and Y be normed linear spaces and $Q : D(Q) \subset X \to Y$ a linear operator. The* graph $\mathcal{G}$ *is the set*

$$\mathcal{G}(Q) = \{(x, Qx) \mid x \in D(Q)\}$$

in the product space $X \times Y$.

DEFINITION A.2.3 [CLOSED OPERATOR]. *An operator Q is said to be* closed *if its graph $\mathcal{G}(Q)$ is a closed subspace of $X \times Y$. Alternatively, Q is closed whenever*

$$x_n \in D(Q), n \in \mathbb{N} \quad \text{and} \quad \lim_{n\to\infty} x_n = x, \lim_{n\to\infty} Qx_n = y,$$

it follows that $x \in D(Q)$ and $Qx = y$.

DEFINITION A.2.4 [COMPACT OPERATOR]. *Let X and Y be normed linear spaces. An operator $Q \in \mathcal{L}(X, Y)$ is said to be a* compact operator *if Q maps bounded sets of X onto relatively compact sets of Y. An equivalent definition is that Q is linear and for any bounded sequence $\{x_n\}$ in X, $\{Qx_n\}$ has a convergent subsequence in Y.*

DEFINITION A.2.5. *A self-adjoint operator A on the Hilbert space Z is* nonnegative *if*

$$\langle Az, z\rangle \geq 0 \text{ for all } z \in D(A);$$

A is positive *if,*

$$\langle Az, z\rangle > 0 \text{ for all nonzero } z \in D(A);$$

and A is coercive *if $\exists \varepsilon > 0$ such that*

$$\langle Az, z\rangle \geq \varepsilon \|z\|^2$$

A.3 Riesz spectral operators

Here, a convenient representation for large classes of linear partial differential systems of both parabolic and hyperbolic types is introduced. Riesz spectral operators allow for non-self-adjoint operators whose eigenvectors may not be orthogonal, but do form a Riesz basis.

DEFINITION A.3.1 [RIESZ BASIS]. *A sequence of vectors $\{\phi_n\}_{n\geq 1}$ in a Hilbert space Z forms a* Riesz basis *for Z if the following conditions hold:*

a. $\overline{\text{span}\{\phi_n\}}_{n\geq 1} = Z$;

b. $\exists m > 0,\ M > 0$ *such that for arbitrary $N \in \mathbb{N}$ and arbitrary scalars α_n, $n =$*

$1, \ldots, N$, *such that,*

$$m \sum_{n=1}^{N} |\alpha_n|^2 \leq \| \sum_{n=1}^{N} \alpha_n \phi_n \|^2 \leq M \sum_{n=1}^{N} |\alpha_n|^2. \tag{A.1}$$

DEFINITION A.3.2. *Suppose that A is a linear, closed operator on a Hilbert space Z with* simple *eigenvalues $\{\lambda_n\}_{n\geq 1}$ and suppose that the corresponding eigenvectors $\{\phi_n\}_{n\geq 1}$ form a Riesz basis in Z. If the closure of $\{\lambda_n\}_{n\geq 1}$ is totally disconnected, then we call A a* Riesz-spectral operator.

By totally disconnected, it is meant that no two points $\lambda, \mu \in \operatorname{span}\{\lambda_n\}_{n\geq 1}$ can be joined by a segment lying entirely in $\{\lambda_n\}_{n\geq 1}\}$. In other words, Definition A.3.2 covers the case where A has finitely many accumulation points.

The following characterization of a Riesz spectral operator is made in Theorem 2.3.5 in [21]:

THEOREM A.3.1 [RIESZ SPECTRAL OPERATOR]. *Suppose that A is a Riesz-spectral operator with simple eigenvalues $\{\lambda_n,\ n \geq 1\}$ and corresponding eigenvectors $\{\phi_n, n \geq 1\}$. Let $\{\psi_n,\ n \geq 1\}$ be the eigenvectors of A^* such that $\langle \phi_n, \psi_m \rangle = \delta_{nm}$. Then A satisfies:*

a. $\rho(A) = \{\lambda \in \mathbb{C} \mid \inf_{n\geq 1} |\lambda - \lambda_n| > 0\}$, $\sigma(A) = \overline{\{\lambda_n,\ n \geq 1\}}$, *and for* $\lambda \in \rho(A)$ $(\lambda I - A)^{-1}$ *is given by*

$$(\lambda I - A)^{-1} = \sum_{n=1}^{\infty} \frac{1}{\lambda - \lambda_n} \langle \cdot, \psi_n \rangle \phi_n;$$

b. A has the representation

$$Az = \sum_{n=1}^{\infty} \lambda_n \langle z, \psi_n \rangle \phi_n$$

for $z \in D(A)$, and $D(A) = \{z \in Z \mid \sum_{n=1}^{\infty} |\lambda_n|^2 |\langle z, \psi_n \rangle|^2 < \infty\}$;

c. A is the infinitesimal generator of a C_0-semigroup iff $\sup_{n\geq 1} \operatorname{Re}(\lambda_n) < \infty$ *and $T(t)$ is given by*

$$T(t) = \sum_{n=1}^{\infty} e^{\lambda_n t} \langle \cdot, \psi_n \rangle \phi_n;$$

d. The growth bound of the semigroup is given by

$$T(t) = \inf_{t>0} \left(\frac{1}{t} \log \|T\| \right) = \sup_{n\geq 1} \operatorname{Re}(\lambda_n). \tag{A.2}$$

A.4 Frequency domain spaces

For the following, we need the following concepts of complex function theory.

DEFINITION A.4.1. *Let Υ be a domain in $\mathbb{C}$ and let f be a function defined on Υ with values in $\mathbb{C}$. The function is holomorphic on Υ if $\frac{\mathrm{d}}{\mathrm{d}s}f(s_0)$ exists for every s_0 in Υ.*

Furthermore, the function is said to be entire *if it is holomorphic on $\mathbb{C}$. The function g is meromorphic on Υ if g can be expressed as $g = f_1/f_2$, where f_1 and f_2 are holomorphic on Υ.*

Note that some texts use the term *analytic*, instead of holomorphic. Examples of holomorphic functions are all polynomials and exponential powers. Rational functions are meromorphic on $\mathbb{C}$ and holomorphic on every domain not containing the zeros of the denominator.

DEFINITION A.4.2 [HARDY SPACES]. *For a Banach space X and a separable Hilbert space H, the following* Hardy spaces *are defined:*

$$H_\infty(X) := \left\{ G : \mathbb{C}_+^0 \mapsto X \mid G \text{ is holomorphic and } \sup_{\mathrm{Re}(s)>0} \|G(s)\| < \infty \right\};$$

$$H_2(H) := \left\{ G : \mathbb{C}_+^0 \mapsto H \mid G \text{ is holomorphic and } \right.$$
$$\left. \|f\|_2^2 = \sup_{\zeta>0} \left(\frac{1}{2\pi} \int_\infty^\infty \|f(\zeta + \imath\omega)\|^2 \mathrm{d}\omega \right) < \infty \right\}.$$

LEMMA A.4.1 [H_∞-NORM]. *If X is a Banach space, then $H_\infty(X)$ from Definition A.4.2 is a Banach space under the H_∞–norm:*

$$\|G\|_\infty := \sup_{\mathrm{Re}(s)>0} \|G(s)\|_X$$

Proof. See the proof of Lemma in A.6.16 [21]. □

B

Dynamic systems and linear regression

This chapter gives a short introduction to some dynamic system descriptions, some system theoretical concepts and also a short introduction to linear regression.

B.1 System theoretical concepts

In this section, it is tried to give a short background on some basic system theoretical concepts used throughout this thesis. The concepts introduced here are inspired by the textbooks [21, 117].

Complementary to this section, some background in functional analysis is given in Appendix A.

B.1.1 Finite dimensional linear systems

This section describes some basic system theoretical concepts regarding finite dimensional linear systems. In Section 4 a CDR process will be modeled by a state space system in finite dimensional state space form after discretization of the original partial differential evaluation. In chapters 3 and 5, finite dimensional continuous linear systems are obtained after model reduction of the nominal, CDR model formulated as an infinite dimensional system.

Continuous time systems

A *l*inear *t*ime-*i*nvariant (LTI), finite-dimensional system $\Sigma(A,B,C,D)$ on the state space X can be described by the following linear differential equations:

$$\Sigma(\mathbf{A},\mathbf{B},\mathbf{C},\mathbf{D}) := \begin{cases} \dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t)+\mathbf{B}\mathbf{u}(t),\ t\geq 0 \quad \mathbf{x}(t_0)=\mathbf{x}_0 \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t)+\mathbf{D}\mathbf{u}(t) \end{cases} \tag{B.1}$$

$$\mathbf{A}:\mathbb{R}^n \mapsto \mathbb{R}^n, \quad \mathbf{B}:\mathbb{R}^m \mapsto \mathbb{R}^n, \quad \mathbf{C}:\mathbb{R}^n \mapsto \mathbb{R}^l, \quad \mathbf{D}:\mathbb{R}^m \mapsto \mathbb{R}^l. \tag{B.2}$$

where

$\mathbf{x}(t)\in X=\mathbb{R}^n$, $\mathbf{x}_0$	system states and initial condition on the state space X, resp.
$\mathbf{u}(t)\in U=\mathbb{R}^m$	system inputs
$\mathbf{y}(t)\in Y=\mathbb{R}^l$	are system outputs
$\mathbf{A}$	linear system matrix (bounded)
$\mathbf{B}$	linear input matrix (bounded)
$\mathbf{C}$	observation matrix (bounded)
$\mathbf{D}$	feedthrough matrix (bounded)

Hence, we have that X, Y and U are finite-dimensional vector spaces and $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{D}$ are bounded linear maps.

A dynamical system with *S*ingle *I*nput ($m = 1$) and *S*ingle *O*utput ($m = 1$) is called a SISO system, otherwise it is called a *M*ultiple *I*nput and *M*ultiple *O*utput (MIMO) system.

Now, given an arbitrary initial condition $\mathbf{x}_0$ and the input $\mathbf{u} \in L_2([0,\tau];U)$, then the dynamical system response $\mathbf{x} \in \mathscr{C}([0,\tau];X)$ ($\mathscr{C}$ denotes the class of continuous functions for all $t \in [0,\tau]$ whose derivatives is again continuous on $[0,\tau]$) and $\mathbf{y} \in L_2([0,\tau];Y)$ are given by:

$$\begin{cases} \mathbf{x}(t) &= e^{\mathbf{A}(t-t_0)}\mathbf{x}(t_0) + \int_{t_0}^{t} e^{\mathbf{A}(t-\tau)}\mathbf{B}\mathbf{u}(\tau)d\tau \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) \end{cases}$$

If we let $\mathbf{x}_0 = 0$ in eqn. (B.1), then by taking Laplace transforms we obtain the corresponding frequency domain description $\tilde{y}$ as a function of $\tilde{u}$:

$$\tilde{y}(s) = \underbrace{\mathbf{C}(sI - \mathbf{A})^{-1}\mathbf{B}}_{G(s)}\tilde{u}(s) \tag{B.3}$$

In time domain, the input-output relationship with zero initial state (i.e. $\mathbf{x}_0 = 0$) can be described by the convolution equation:

$$\begin{aligned} \mathbf{y}(t) &= (g * \mathbf{u})(t) \\ &:= \int_{-\infty}^{\infty} g(t-\tau)\mathbf{u}(\tau)d\tau = \int_{-\infty}^{t} g(t-\tau)\mathbf{u}(\tau)d\tau \end{aligned}$$

with $g(t) = \mathbf{C}e^{\mathbf{A}t}\mathbf{B}$.

Observability, controllability, stability and detectability

DEFINITION B.1.1. $\Sigma(A,B,-)$ *is* controllable *if for some* $\tau > 0$ *the* controllability map $\mathscr{B}^\tau : L_2([0,\tau];U) \mapsto Z$ *has* Z *in its range, where,*

$$\mathscr{B}^\tau u := \int_0^\tau e^{A(\tau-s)}Bu(s)\mathrm{d}s.$$

It follows that $\Sigma(A,B,-)$ is controllable if and only if rank $\left[B : AB : \ldots : A^{n-1}B\right] = \dim(Z) = n$.

Similarly, $\Sigma(A,B,-)$ is observable if and only if rank $\left[C^* : A^*C^* : \ldots : (A^{n-1})^*C^*\right] = \dim(Z) = n$.

The following definition is given for stability.

DEFINITION B.1.2. *A is exponentially stable if there exists positive constant* M *and* α

such that

$$\|e^{at}z_0\|_Z \leq Me^{-\alpha t}\|z_0\|_z, \quad \forall t \geq 0.$$

A is called *antistable* if $-A$ is exponentially stable.

The following definition for stabilizability and detectability is adopted from [21]. Notice that the detectability definition introduced in Chapter 5 for boundary control systems more or less resembles the detectability definition for finite dimensional systems given in the following definition.

DEFINITION B.1.3. *$\Sigma(A, B-)$ is exponentially stabilizable if there exists an $F \in \mathscr{L}(Z, U)$ such that $A+BF$ is exponentially stable.* $\Sigma(A, -C)$ *is* exponentially detectable *if there exists a $L \in \mathscr{L}(Y, Z)$ such that $A+LC$ is detectable.*

Shift operator calculus in discrete time systems

In Chapter 3 and 4, discrete time system representations are used. A definition for a discrete time (structured) linear system is given in the latter chapter, so we only give a brief sketch of shift operator calculus.

Differential operator calculus is a a convenient tool for manipulating linear differential equations with constant coefficients. The class of signals where the shift operators act on is considered as doubly infinite sequences $\{f(k) : k = \ldots -1, 0, 1 \ldots\}$ [3]. Assume the sampling period is chosen as the time unit. Then the forward shift operator is denoted by q and has the property:

$$qf(k) = f(k+1)$$

For a norm of a signal $\|f\| = \sup_k |f(k)|$, or, $\|f\|^2 = \sum_{k=-\infty}^{\infty} f^2(k)$, it follows that the shift operator has unit norm. Hence, the calculus is then greatly simplified and, since it is an doubly infinite sequence, we may take the inverse and define the *backward-shift* operator or *delay operator* denoted by q^{-1}, i.e.:

$$q^{-1}f(k) = f(k-1).$$

The shift operator is used to simplify manipulation of higher order difference equations, as is done in Chapter 4.

B.1.2 Infinite–dimensional linear systems

In this section, a short overview of infinite-dimensional system descriptions is given. For a general introduction of linear systems theory concepts in infinite dimensions, the reader is referred to [21].

Let $\mathscr{L}(X)$ be a shorthand notation for the bounded linear map from X to X and $\mathscr{L}(X, Y)$ a bounded linear map from X to Y. The following class of infinite-

dimensional systems with input u and output y is considered:

$$\Sigma(A,B,C,D) := \left\{ \begin{array}{rcl} \dot{z}(t) & = & Az(t)+Bu(t) \\ y(t) & = & Cx(t)+Du(t) \end{array} \right. , \tag{B.4}$$

where $A \in \mathscr{L}(Z)$ generates a C_0-semigroup $T(t)$, $B \in \mathscr{L}(\mathbb{C},Z)$, $C \in \mathscr{L}(\mathbb{C},Y)$, $D \in \mathscr{L}(\mathbb{C},Y)$ and furthermore: The notion $A \in \mathscr{L}(Z)$ generates a strongly continuous

$z(t) \in Z$, z_0	system states and initial condition on Hilbert state space Z
$u(t) \in U = \mathbb{C}^m$	system inputs
$y(t) \in Y = \mathbb{C}^l$	are system outputs.

semigroup $T(t)$ will be introduced shortly.

Furthermore, it may seem artificial to choose $U = Y = \mathbb{C}$ in a physical example, but we choose to complexify it here for mathematical reasons. Notice also that operators B and C can be approximately modeled as 'point' actuators and 'point' sensors with shaping functions around the control and sensing point, e.g.:

$$b(\eta) = \frac{1}{2\varepsilon} 1_{[\eta_1-\varepsilon,\eta_1+\varepsilon]}(\eta), \quad c(\eta) = \frac{1}{2\nu} 1_{[\eta_2-\nu,\eta_2+\nu]}(\eta)$$

$$\text{where } 1_{[a,b]}(x) = \begin{cases} 1 & \text{for } \alpha \leq \eta \leq \beta \\ 0 & \text{elsewhere.} \end{cases}$$

The main problem with shaping functions $b(\eta)$ and $c(\eta)$ is that, the lacking of *a priori* information of the size and shape of the sensor and actuator used in practice. If, for example, the spatial dimension of the application under study is given by L and $2\varepsilon \ll L$ or $2\nu \ll L$, it would be more natural to consider point mappings B and C. A formulation which makes this possible is given in Section B.1.3.

First, the definition of a semi-group is introduced.

Semi-groups

The semigroup $T(t)$ can be seen as the infinite-dimensional extension of a set of (finite-dimensional) state transition matrices $e^{At}_{t\leq 0}$ under the condition that the state x satisfies the Hadamard well-posedness conditions. More precisely:

DEFINITION B.1.4 [DEFINITION 2.1.2 IN [21]]. *Let Z be a separable complex Hilbert space under the induced norm $\|\cdot\|$. A* strongly continuous semigroup *is an operator-valued function $T(t)$ from $\mathbb{R}_+$ to $\mathscr{L}(Z)$ that satisfies the following properties:*

(a) $T(t+s) = T(t)T(s)$ *for* $t,s \geq 0$*;*

(b) $T(0) = I$*;*

(c) $\|T(t)z - z\| \to 0$ *as* $t \to 0^+$ *for all* $z \in Z$*.*

$T(t)z$ is related to the solution of an abstract differential equation as in eqn. (B.4), in the following definition.

DEFINITION B.1.5 [DEFINITION 2.1.8 IN [21]]. *The infinite-dimensional generator of a C_0-semigroup on a Hilbert space Z is defined by*

$$Az = \lim_{t \to 0^+} \frac{1}{t}(T(t) - I)z$$

on its domain

$$D(A) = \{z \in Z \mid \lim_{t \to 0^+} \frac{T(t)z - z}{t} \text{ exists}\}$$

In the next theorem, some usefull properties of a C_0-semigroup generated by an infinitesimal generator A are introduced. The proof can be found in [21].

THEOREM B.1.1 [ADOPTED FROM THEOREM 2.1.10 IN [21]]. *$T(t)_{t \geq 0}$ is a C_0-semigroup on Z with infinitesimal generator A if it has the following properties:*

(a) For $z \in D(A)$, $\forall t \geq 0$, $T(t)z \in D(A)$ and $\frac{dT(t)z}{dt} = AT(t)z = T(t)Az$.

(b) For $z \in D(A)$, $\forall t \geq 0$, $T(t)z - z = \int_0^t T(s)Az\,ds$.

(c) A is a closed operator and furthermore, $D(A)$ is dense in Z.

B.1.3 Abstract Cauchy problem

In theorem 2.1.10 of [21], it is shown that if A is an infinitesimal generator of the C_0-semigroup $T(t)$, the solution of the abstract homogeneous Cauchy initial value problem

$$\dot{z}(t) = Az(t) + f(t), \quad t \geq 0, \quad z(0) = z_0 \in D(A)$$

is given by

$$z(t) = T(t)z_0 + \int_0^t T(t-s)f(s)\mathrm{d}s \tag{B.5}$$

assumed that $f \in C([0,\tau];Z)$. Eqn. (B.5) is also called an abstract evolution equation or abstract differential equation. We define the classical solution, with $C^1([0,\tau];Z)$ denoting the class of continuous functions on$[0,\tau]$ whose derivative is again continuous on $[0,\tau]$.

DEFINITION B.1.6 [CLASSICAL SOLUTION ([21])]. *Consider eqn. (B.5) on the Hilbert space Z. We call $z(t)$ a classical solution of eqn. (B.5) on $[0,\tau]$ if $z(t) \in C^1([0,\tau];Z)$, $z(t) \in D(A)$ $\forall t \in [0,\tau]$ and $z(t)$ satisfies eqn. (B.5) $\forall t \in [0,\tau]$.*

The function is a classical solution on $[0,\infty)$ if $z(t)$ is a classical solution on $[0,\tau]$ for every $\tau \geq 0$.

The conditions for a classical solution are in general too strong for control problems, due to the C^1-continuity assumption of $f(t)$. In the application of Chapter 5, a weaker concept of a solution of eqn. (B.5) is considered.

First consider the following abstract differential equation with perturbed system operator $A + D(\cdot)$:

$$\dot{z}(t) = (Az(t) + D(t))z(t) + f(t), \quad z(0) = z_0. \tag{B.6}$$

The solution for this equation is defined as follows.

DEFINITION B.1.7 [SOLUTIONS FOR EQN. (B.6)]. *Consider eqn. (B.6) with A the infinitesimal generator of a C_0-semigroup on the Hilbert space Z, $z_0 \in Z$, $D \in P_\infty([0,\tau];\mathscr{L}(Z)$ and $f \in L_p([0,\tau];Z)$, $p \geq 1$. The function is a classical solution of eqn. (B.6) on $[0,\tau$ if $z(\cdot) \in C^1([0,\tau];Z)$, $z(t) \in D(A)$ $\forall t \in [0,\tau]$ and $z(t)$ satisfies eqn. (B.5) $\forall t \in [0,\tau]$.*

If $f \in L_p([0,\tau];Z)$, $p \geq 1$ and $z_0 \in Z$, then the mild solution of eqn. (B.6) is defined as

$$z(t) = U(t,0)z_0 + \int_0^t U(t,s)f(s)\mathrm{d}s. \tag{B.7}$$

with mild evolution operator $U(t,s)$ generated by $A + D(\cdot)$.

Boundary control formulation

In this subsection, we consider the following class of *abstract boundary control problems*:

$$\dot{z}(t) = \mathfrak{A}z(t), \quad z(0) = z_0 \tag{B.8}$$

$$\mathfrak{B}z(t) = u(t) \tag{B.9}$$

where $D(\mathfrak{A}) \subset Z \mapsto Z$, $u(t) \in U$, a separable Hilbert space and the *boundary control operator* $\mathfrak{B} : D(\mathfrak{B}) \subset Z \mapsto U$ satisfies $D(\mathfrak{A}) \subset D(\mathfrak{B})$.

It is possible to reformulate eqn. (B.8) into an abstract form eqn. (B.5) with the following conditions.

DEFINITION B.1.8 [BOUNDARY CONTROL SYSTEM]. *The control system eqn. (B.8) is a* boundary control system *if the following hold:*

a. *The operator* $A : D(A) \mapsto Z$ *with* $D(A) = D(\mathfrak{A}) \cap \ker(\mathfrak{B})$ *and*

$$Az = \mathfrak{A}z \quad \textit{for} \quad z \in D(A)$$

is the infinitesimal generator of a C_0*-semigroup on* Z*;*

b. *There* $\exists\, B \in \mathscr{L}(U, Z)$ *such that* $\forall u \in U$, $Bu \in D(\mathfrak{A})$ *the operator* $\mathfrak{A}B \in \mathscr{L}(U, Z)$ *and*

$$\mathfrak{B}B = u, \quad u \in U. \tag{B.10}$$

Assuming that eqn. (B.8) is a boundary control problem for $u \in C^2([0, \tau]; U)$, the following abstract differential equation on Z is well-posed:

$$\dot{v}(t) = Av(t) - B\dot{u}(t) + \mathfrak{A}Bu(t), \quad v(0) = v_0 \tag{B.11}$$

The following theorem relates the classical solutions of eqn. (B.8) and the abstract Cauchy equation eqn. (B.6) (Theorem 3.3.3 in [21]).

THEOREM B.1.2. *Given the boundary control system eqn. (B.8) and the abstract Cauchy eqn. (B.11) and let* $u \in C^2([0, \tau]; U)\ \forall \tau > 0$. *Then, if* $v_0 = z_0 - Bu(0) \in D(A)$, *the classical solution of eqn. (B.8) and eqn. (B.11) are related by*

$$v(t) = z(t) - Bu(t).$$

Furthermore, the classical solution of eqn. (B.8) is unique.

The abstract equation eqn. (B.11) also have well defined mild solutions for $\dot{u} \in L_p([0, \tau]; U)$ for some $p \geq 1$, $v_0 \in Z$ and $z_0^e \in Z^e$, with $Z^e := U \oplus Z$ an extended space (not shown here); respectively. Under these weaker assumptions, $z(t)$

$$\begin{aligned} z(t) &= \begin{pmatrix} B & I \end{pmatrix} z^e(t) \\ &= Bu(t) - T(t)Bu(0) + T(t)z_0 - \int_0^t T(t-s)B\dot{u}(s)\mathrm{d}s \\ &\quad + \int_0^t T\,t-s)\mathfrak{A}Bu(s)\mathrm{d}s \end{aligned}$$

with

$$z^e(t) = \begin{pmatrix} (z^e(t))_1 \\ (z^e(t))_2 \end{pmatrix} = \begin{pmatrix} u(t) \\ v(t) \end{pmatrix}, \qquad (z_0^e)_1 = u(0) \quad \text{and} \quad (z_0^e)_2 = v(0)$$

is called a mild solution of the boundary control equation (see chapter 3 in [21] for details. eqn. (B.8).

Observability and detectability

In this thesis, only infinite dimensional systems in boundary form are discussed. Furthermore, the starting point is the use of point observations and boundary inputs. Hence, the concept for observability is tailored to these boundary form systems in Section 5.2. In the same section, a definition of detectability for CDR systems is given.

B.2 Linear regression

Suppose, the matrix equation $\mathbf{Ax} = \mathbf{b}$, $\mathbf{A} \in \mathbb{R}^{n\times m}$ is to be solved. We would like to know if there's a solution, and if so, how many solutions are possible. A well-known theorem is the *Fredholm alternative theorem*, found in many textbooks on applied mathematics, e.g. [50]:

THEOREM B.2.1 [FREDHOLM ALTERNATIVE THEOREM]. *The equation* $\mathbf{Ax} = \mathbf{b}$ *has a solution iff* $\langle \mathbf{b}, \mathbf{v} \rangle = 0$ *for every vector* v *satisfying* $\mathbf{A}^* v = 0$. *Furthermore, a solution of* $\mathbf{Ax} = \mathbf{b}$*, if it exists, is unique iff* $\mathbf{x} = \mathbf{0}$ *is the only solution of* $\mathbf{Ax} = \mathbf{0}$.

Typical engineering examples are overdetermined systems from curve fitting of collected data. These examples generally contain measurement data corrupted with noise. Consequently, the Fredholm alternative shows that these examples do not have an exact solution of $\mathbf{Ax} = \mathbf{b}$. Hence, the goal is to find an $\mathbf{x}$ that minimizes $\|\mathbf{Ax} - \mathbf{b}\|$ under some norm $\|\cdot\|$. A widely used norm is the Euclidean norm $\|\cdot\|_2$ which leads us to the least squares problem:

DEFINITION B.2.1 [LEAST SQUARES PROBLEM]. *For a matrix* $\mathbf{A} \in \mathbf{R}^{m\times n}$ *and a vector* $\mathbf{b}$*, the* least squares problem *is:*

$$\min_{\mathbf{x}\in\mathbb{C}^n} \|\mathbf{Ax} - \mathbf{b}\|_2$$

where $\|\cdot\|_2$ *denotes the Euclidean norm.*

C

Bulk storage model

C.1 Padé approximants of bulk storage model

Second order approximation

Similarly, if $G(s)$ as in eqn. (2.8) is approximated by a Padé–[1,2] approximation $\widetilde{G}^{[1,2]}(s)$, the following relationship is obtained:

$$\widetilde{G}^{[1,2]}(s) := \frac{b_0 + b_1 s}{a_2 s^2 + a_1 s + 1} \tag{C.1}$$

from which we can recover (lumped) physical parameters, i.e. a_1, a_2, b_0 and b_1.

We now give the results to calculate A and BM_5 from the estimates. The estimates of a_1, a_2, b_0 and b_1 can be obtained by the use of continuous time identifaction methods, see e.g. [86]:

$$\hat{A} = \frac{2\hat{b}_0\left(\hat{b}_1 - \frac{\hat{a}_1}{\hat{a}_0}\right)}{2(-\hat{a}_1^2\hat{b}_0 + \hat{a}_2\hat{b}_0 + \hat{a}_1\hat{b}_1) + \hat{b}_0\left(\hat{b}_1 - \frac{\hat{a}_1}{\hat{a}_0}\right)^2}$$

$$\widehat{BM_5} = \hat{A}^2\frac{\hat{b}_1 - \hat{a}_1\hat{b}_0}{\hat{b}_0}$$

C.2 Estimation of time constant

In figure C.1, linear regressions of $\ln\theta^F$ versus Δ_t are depicted for a ventilator setting of 85% of its capacity. Similar results were obtained with other ventilator settings. From the slope of these regressions, τ_p is reconstructed.

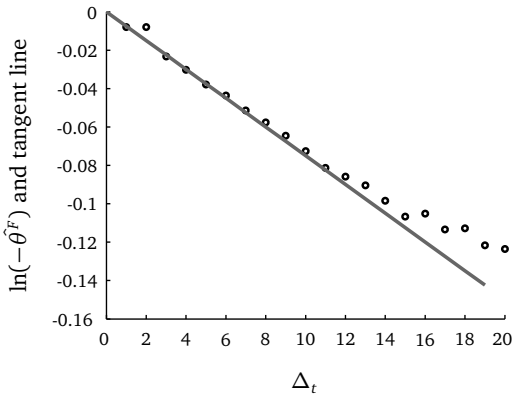


Figure C.1: Estimated θ^F with different sample times [∘] and tangent line [—] at $\Delta_t \to 0^+$.

C.3 Physical bulk storage model parameters

Table C.1: Variables and their units

T_a	K	air temperature in bulk
T_p	K	product temperature
Φ	$\frac{m^3}{s}$	air flow through shaft
v_p	$\frac{m}{s}$	$A_f \gamma_p \Phi$, air velocity inside bulk

Table C.2: Physical parameters p and their units

α_{th}	$\frac{m^2}{s}$	thermal diffusivity of air ($1.87\ 10^{-5}$)
γ_p	$\frac{m^3}{m^3}$	porosity air over bulk
λ_a	$\frac{W}{m\,K}$	conduction of air ($2.43 \cdot 10^{-2}$)
λ_p	$\frac{W}{m\,K}$	conduction of product
ν	$\frac{m^2}{s}$	kinematic viscosity of air ($1.35 \cdot 10^{-5}$
ρ_a	$\frac{kg}{m^3}$	air density (1.27)
ρ_p	$\frac{kg}{m^3}$	produce density
A_f	m^2	floor area of the bulk
N_β	–	βN_p amount of spherical products exposed to air per crate
V_{cr}	m^3	crate volume
V_p	m^3	product volume
R	m	product radius
L	m	bulk height
a_p	$\frac{W}{kg\,K}$	product heat production
c_a	$\frac{J}{kg\,K}$	heat capacity of air
c_p	$\frac{J}{kg\,K}$	heat capacity of produce
h_v	$\frac{W}{m^2\,K}$	heat transfer coefficient

Table C.3: Lumped physical parameters

A_p	$\frac{m^2}{m^3}$	$N_\beta V_p / V_{\text{cr}}$ produce surface per bulk volume
$\mathrm{b_i}$	—	Biot number $\frac{2h_v R}{\lambda_a}$
L_2	m	$R\gamma_p(1-\gamma_p)$, characteristic length
M_0	$\frac{J}{K}$	$\rho_a c_a V_a$, lumped parameter
M_1	$\frac{m^2}{s}$	$\frac{\lambda_p}{\rho_p c_p}$
M_2	$\frac{1}{s}$	$\frac{a}{c_p}$
M_3	—	$\sqrt{M_2/M_1}R$
M_4	$\frac{1}{s}$	$\frac{h_v A_p}{\gamma_p \rho_a c_a}$
M_5	—	$\frac{M_4 L}{\nu}$
$\mathrm{n_u}$	—	Nusselt number $\frac{2h_v R}{\lambda_a}$
$\mathrm{p_r}$	—	Prandtl number $\frac{\nu}{\alpha_{th}}$
$\mathrm{r_e}$	—	Reynolds number $\frac{v_p L_2}{\nu}$
A	$\frac{1}{s}$	inverse of time-constant $-\frac{2M_3\cot(M_3)-2+\mathrm{b_i}}{\frac{R^2}{M_1}\cot^2(M_3)+\frac{R^2}{M_1}-\frac{M_3}{M_2}\cot(M_3)}$
B	$\frac{1}{s}$	$\frac{\mathrm{b_i}}{\frac{R^2}{M_1}\cot^2(M_3)+\frac{R^2}{M_1}-\frac{M_3}{M_2}\cot(M_3)}$

D

UV disinfection model: analysis and model approximation

D.1 Dimensional analysis

The following is a central theorem in dimensional analysis and adopted from [68]. In this theorem only scalar variables and parameters are treated.

THEOREM D.1.1 [BUCKINGHAM]. *Consider a system with (scalar) variables $x_1, \ldots, x_k$ and scalar parameters $p_1, \ldots, p_l$. The associated dimensions are denoted by $d_1, \ldots, d_m$. Each relation*

$$f(x_1, \ldots, x_k, \ldots, p_1, \ldots, p_l) = 0$$

can be rewritten in the equivalent, dimensionless form,

$$f(r_1, \ldots, r_n) = 0$$

with $r_1, \ldots, r_n$ dimensionless products of (powers of) the x's and p's. The number is given by

$$n = k + l - m$$

Proof. See [12]. □

The homogeneous part of the UV disinfection equations eqn. (2.11) reveals that there are,

- 3 parameters, i.e.: ν, α and $\bar{c}$.
- 3 variables, i.e. where τ and ξ are independent, c is a dependent variable.
- 2 dimensions, i.e. the dimension of time t is denoted by $t \equiv [T]$, $\xi \equiv [L]$ and $c \equiv [\text{mol}/L]$.

Now Buckingham's theorem tells us that there are $3 + 3 - 2 = 4$ dimensionless variables. It is natural to scale the concentration c with the mean value of the inlet concentration $\bar{c} := \int_0^{\tau_f} c_{\text{in}}(\tau)\mathrm{d}\tau/\tau_f$, with τ_f large, i.e. this turns to $z := c/\bar{c}$. Another obvious choice is to scale the spatial coordinate with the length of the reactor, i.e. $\eta := \xi/L$. Hence 4 variables and parameters are left over. To find the other two dimensionless variables, we form the products

$$r = \tau^{s_1} \alpha^{s_2} \nu^{s_3}$$

The 'dimension' condition $[r] \equiv 0$ leads to:

$$\begin{aligned} s_1 - s_2 - s_3 &= 0 \\ 2s_2 + s_3 &= 0 \end{aligned}$$

We may choose one r freely. A natural choice is $(s_1, s_2, s_3) = (-1, 1, -2)$ since then $r = r_1 r_2 r_3 r_4 \equiv \alpha/(\nu^2 \cdot t)$. Notice that in this case, $r_3 \equiv \nu \cdot \tau/L$ is a scaled time variable and $r_4 \equiv \alpha/(\nu \cdot L)$ is the inverse of the familiar $\mathrm{p_e}$-number.

For the inhomogenous term we let $u_1 := f/f_{\max}$ and $b_1 := \kappa f_{\max} L/v$. Notice that b_1 corresponds to the Damköhler number [22], that is, a dimensionless number indicating the magnitude of the chemical reaction relative to the convective mass transfer. Furthermore, with proper scaling of the boundary conditions and $\tilde{u}_2 := c_{\text{in}}/\bar{c}$; the dimensionless form of eqn. (2.11) reads:

$$\begin{array}{rcl} \frac{\partial z}{\partial t}(\eta,t) & = & \frac{1}{\text{p}_\text{e}}\frac{\partial^2}{\partial \eta^2}z(\eta,t) - \frac{\partial z}{\partial \eta}(\eta,t) - b_1 u_1(t) z(\eta,t) \\ z(\eta,0) & =: & z_0(\eta) \\ z(0,t) - \frac{1}{\text{p}_\text{e}}\frac{\partial z}{\partial \eta}(0,t) & =: & \tilde{u}_2(t) \\ \left.\frac{\partial z}{\partial \eta}\right|_{(1,t)} & = & 0. \end{array} \tag{D.1}$$

with input and output observations,

$$u_2 = \begin{pmatrix} \tilde{u}_2 \\ 0 \end{pmatrix}, \quad \text{and,} \quad y = z(\eta^*, \cdot) \tag{D.2}$$

As also mentioned by [68], in most systems, the choice of r is not unique. For the UV disinfection model, $(s_1, s_2, s_3) = (1, -1, 2)$ is also possible, leading to another dimensionless form. Furthermore, sometimes more insight is gained when r contains more than one variable. In that case, the corresponding transformation is then sometimes called a *similarity* transformation [5].

D.2 Spectral analysis of error system Σ_I^ε

Eigenvalues

Proof ofLemma 5.3.1

$A_{(I)}^L$ is an S-L operator and therefore it has a spectrum with isolated eigenvalues with finite multiplicities, see Lemma 5.2.2. Furthermore it is self-adjoint in the inner product $\langle \cdot, \cdot \rangle_w$—see Section 5.2.3— and negative for $L_{21} \leq 0$ —see Remark 5. Consequently, $\lambda^L < 0$.

For this Sturm-Liouville type problem, we write

$$\phi^L = C_1^L \phi_1^L + C_2^L \phi_2^L \tag{D.3}$$

Furthermore, ϕ^L should satisfy the boundary conditions

$$\mathfrak{B}_1^L \phi^L := \phi^L(0) = 0 \tag{D.4}$$

$$\mathfrak{B}_2^L \phi^L := \dot{\phi}^L(1) - L_{21}\phi^L(1) = 0 \tag{D.5}$$

The eigenvalues λ^L can be found from $A_{(I)}^L \phi^L = \lambda^L \phi$. To this aim, consider the

following three cases where in

Case 1a: Let $\lambda^L > \frac{1}{4}\mathrm{p_e}$, i.e. $\phi_1^L = e^{\mu_1(\lambda^L)\eta}$ and $\phi_2^L = e^{\mu_1(\lambda^L)\eta}$ with:

$$\mu_1(\lambda^L) = \frac{\mathrm{p_e}}{2} - \varsigma_a^L, \quad \mu_2(\lambda^L) = \frac{\mathrm{p_e}}{2} + \varsigma_a^L \quad \text{and} \quad \varsigma_a^L = \sqrt{\left(\frac{\mathrm{p_e}}{2}\right)^2 + \mathrm{p_e}\lambda^L} \tag{D.6}$$

Case 1b: Let $\lambda < \frac{1}{4}\mathrm{p_e}$, i.e. $\phi_1^L = e^{\mu_1^{\lambda L}\eta}$ and $\phi_2^L = e^{\mu_1(\lambda^L)\eta}$ with:

$$\mu_1(\lambda^L) = \frac{\mathrm{p_e}}{2} - \varsigma_b^L, \quad \mu_2(\lambda^L) = \frac{\mathrm{p_e}}{2} + \varsigma_b^L \quad \text{and} \quad \varsigma_b^L = \imath\varsigma_L = \sqrt{\left(\frac{\mathrm{p_e}}{2}\right)^2 + \mathrm{p_e}\lambda^L} \tag{D.7}$$

Case 2: $\lambda^L = (\frac{1}{2}\mathrm{p_e})^2$, i.e. let $\phi_1^L = e^{\frac{1}{2}\mathrm{p_e}\eta}$ and $\phi_2^L = \eta e^{\frac{1}{2}\mathrm{p_e}\eta}$.

The eigenvalues λ^L of A^L exist if and only if the determinant $\Delta(\lambda^L)$ of the system of boundary equations eqn. (D.5) is zero, see Exercise 2.10b in [21], i.e.,

$$\Delta(\lambda^L) := \det\begin{pmatrix} \mathfrak{B}_1^L\phi_1^L & \mathfrak{B}_1^L\phi_2^L \\ \mathfrak{B}_2^L\phi_1^L & \mathfrak{B}_2^L\phi_2^L \end{pmatrix} = 0$$

Hence, case by case we get:

Case 1a:

$$\delta(\lambda^L) = \det\begin{pmatrix} 1 & 1 \\ \left(\mu_1(\lambda^L) - L_{21}\right)e^{\mu_1(\lambda^L)} & \left(\mu_2(\lambda^L) - L_{21}\right)e^{\mu_2(\lambda^L)} \end{pmatrix}$$

Hence, $\Delta(\lambda^L) = 1\cdot\left(\mu_2(\lambda^L) - L_{21}\right)e^{\mu_2(\lambda^L)} - 1\cdot\left(\mu_1(\lambda^L) - L_{21}\right)e^{\mu_1(\lambda^L)}$. Since for all L_{21} and for $\mathrm{p_e} > 0$, $\varsigma_L^a > 0$, this leads to:

$$e^{\varsigma_L^a} > e^{-\varsigma_a^L} > 0 \quad \text{and} \quad \left(\tfrac{1}{2}\mathrm{p_e} + \varsigma_a^L - L_{21}\right) > \left(\tfrac{1}{2}\mathrm{p_e} - \varsigma_a^L - L_{21}\right) > 0.$$

Consequently, $\Delta(\lambda) > 0$ and no solution for the eigenvalues λ^L can be found.

Case 1b: Analogous to Case 1a, we get:

$$\Delta(\lambda^L) = 1\cdot\left(\mu_2(\lambda^L) - L_{21}\right)e^{\mu_2(\lambda^L)} - 1\cdot\left(\mu_1(\lambda^L) - L_{21}\right)e^{\mu_1(\lambda^L)}.$$

However, for $\lambda^L < 0$ this reduces to:

$$\begin{aligned}\Delta(\lambda) &= e^{\frac{\mathrm{p_e}}{2}}\left[\left(\tfrac{1}{2}\mathrm{p_e} + \imath\varsigma^L - L_{21}\right)e^{\imath\varsigma^L} - \left(\tfrac{1}{2}\mathrm{p_e} - \imath\varsigma^L - L_{21}\right)e^{-\imath\varsigma_L}\right] \\ &= e^{\frac{\mathrm{p_e}}{2}}\left[2\imath\varsigma^L\cos(\varsigma^L) + 2\imath\left(\tfrac{1}{2}\mathrm{p_e} - L_{21}\right)\sin(\varsigma^L)\right]\end{aligned}$$

Hence for $\Delta(\lambda^L) = 0$, ς_k^L, $k \geq 0$ is the set of all solutions to the resolvent equation:

$$\tan(\varsigma_k^L) = -\frac{\varsigma_k^L}{\frac{1}{2}\mathrm{p_e} - L_{21}}. \tag{D.8}$$

Case 2:

$$\Delta(\lambda^L) = \det\begin{pmatrix} 1 & 0 \\ \left(\frac{1}{2}\mathrm{p_e} - L_{21}\right)e^{\frac{1}{2}\mathrm{p_e}} & \left(\frac{1}{2}\mathrm{p_e} + 1 - L_{21}\right)e^{\frac{1}{2}\mathrm{p_e}}\end{pmatrix}$$

Hence, $\Delta(\lambda^L) = 0$ if $L_{21} = \frac{1}{2}\mathrm{p_e} + 1$, since $\mathrm{p_e} > 0$.

The eigenvalues λ^L follow from eqn. (D.7), Case 1b or the rather exceptional Case 2. For a detectable system Σ_I^e with $L_{21} \leq 0$, Case 2 does not occur. Furthermore, recognize that $\lambda_k^L \to -\infty$ as $k \to \infty$ and $|\lambda_{k+1}^L - \lambda_k^L| \to \infty$ as $k \to \infty$.

Eigenvectors

From $\mathfrak{B}_1\phi$, i.e. the Dirichlet condition at $\eta_1 = 0$ and λ^L as in eqn. (D.7), we obtain for the eigenvectors ϕ as in eqn. (D.3):

$$\mu_1(\lambda^L)C_1^L + \mu_2(\lambda^L)C_2^L = 0, \quad C_1^L \neq 0 \neq C_1^L$$

Hence, the associated eigenvectors of A^L, i.e. $\phi_k^L \in D(A^L)$, $k \geq 0$ are given for all $\eta \in [0,1]$ and for all $k \geq 1$ by

$$\begin{aligned}&\phi_k^L = C_0\left[\left(\tfrac{1}{2}\mathrm{p_e} - \imath\varsigma_k^L\right)\exp\left(\tfrac{1}{2}\mathrm{p_e} - \imath\varsigma_k^L\right) - \left(\tfrac{1}{2}\mathrm{p_e} - \imath\varsigma^L\right)\exp\left(\tfrac{1}{2}\mathrm{p_e} + \imath\varsigma_k^L\right)\right] \Leftrightarrow \\ &\phi_k^L(\cdot) = C_0\exp\left(\tfrac{1}{2}\mathrm{p_e}\eta\right)\sin(\varsigma_k^L\eta) \quad \text{with} \quad C_0^L := C_1^L = -C_2^L\end{aligned}$$

D.3 Spectral analysis of system Σ_I

Eigenvalues

The eigenvalues of the system can be calculated similar to the proof of Lemma 5.3.1. In the proof of Proposition 2.4 of [24, pp. 58–61], the eigenvalues and associated

eigenvectors of exactly the same system are calculated. We include these results here (with slightly different notation).

Let A be given as in eqs. (5.7) and (5.8), i.e.

$$Az := \frac{1}{\mathrm{p_e}} \frac{\mathrm{d}^2}{\mathrm{d}\eta^2} z - \frac{\mathrm{d}z}{\mathrm{d}\eta}$$

We have that A is an S-L operator and therefore it has a spectrum with isolated eigenvalues with finite multiplicities, see Lemma 5.2.2.

More specifically, the spectrum of A is given by

$$\sigma(A) = \sigma_p(A) = \{\lambda_k : k \geq 0\} \subset (-\infty, 0).$$

where $\sigma_p(A)$ denotes the point spectrum of A. The eigenvalues λ_k, $k \geq 0$ are calculated as shown in what follows.

$-A$ is an S-L-operator, hence it is a Riesz operator [26], and we can write,

$$A\phi - \lambda\phi = \frac{1}{\mathrm{p_e}} \frac{\mathrm{d}^2}{\mathrm{d}\eta^2} \phi - \frac{\mathrm{d}\phi}{\mathrm{d}\eta} - \lambda\phi.$$

Furthermore, ϕ should satisfy the boundary conditions

$$\mathfrak{B}_1\phi := \frac{1}{\mathrm{p_e}} \dot{\phi}(0) - \phi(0) = 0 \tag{D.9}$$

$$\mathfrak{B}_2\phi := \dot{\phi}(1) = 0 \tag{D.10}$$

The eigenvalues λ can be found from $A_I\phi = \lambda\phi$. To this aim, consider the following three cases where in

Case 1a: Let $\lambda > \frac{1}{4}\mathrm{p_e}$, i.e. $\phi_1 = e^{\mu_1(\lambda)\eta}$ and $\phi_2 = e^{\mu_1(\lambda)\eta}$ with:

$$\mu_1(\lambda) = \frac{\mathrm{p_e}}{2} - \varsigma_a, \quad \mu_2(\lambda) = \frac{\mathrm{p_e}}{2} + \varsigma_a \quad \text{and} \quad \varsigma_a = \sqrt{\left(\frac{\mathrm{p_e}}{2}\right)^2 + \mathrm{p_e}\lambda} \tag{D.11}$$

Case 1b: Let $\lambda < \frac{1}{4}\mathrm{p_e}$, i.e. $\phi_1 = e^{\mu_1(\lambda)\eta}$ and $\phi_2 = e^{\mu_1(\lambda)\eta}$ with:

$$\mu_1(\lambda) = \frac{\mathrm{p_e}}{2} - \varsigma_b, \quad \mu_2(\lambda) = \frac{\mathrm{p_e}}{2} + \varsigma_b \quad \text{and} \quad \varsigma_b = \imath\varsigma = \sqrt{\left(\frac{\mathrm{p_e}}{2}\right)^2 + \mathrm{p_e}\lambda} \tag{D.12}$$

Case 2: $\lambda = (\frac{1}{2}\mathrm{p_e})^2$, i.e. let $\phi_1 = e^{\frac{1}{2}\mathrm{p_e}\eta}$ and $\phi_2 = \eta e^{\frac{1}{2}\mathrm{p_e}\eta}$.

The eigenvalues λ of A exist if and only if the determinant $\Delta(\lambda)$ of the system of

boundary equations eqn. (D.5) is zero, see Exercise 2.10b in [21], i.e.,

$$\Delta(\lambda) := \det \begin{pmatrix} \mathfrak{B}_1\phi_1 & \mathfrak{B}_1\phi_2 \\ \mathfrak{B}_2\phi_1 & \mathfrak{B}_2\phi_2 \end{pmatrix} = 0$$

Case by case we get:

Case 1a:

$$\Delta(\lambda) = \det \begin{pmatrix} \frac{1}{\mathrm{p_e}}\mu_1(\lambda) - 1 & \frac{1}{\mathrm{p_e}}\mu_2(\lambda) - 1 \\ \mu_1(\lambda)e^{\mu_1(\lambda)} & \mu_2(\lambda)e^{\mu_2(\lambda)} \end{pmatrix}$$

Hence,

$$\begin{aligned} \Delta(\lambda) &= \left(\frac{1}{\mathrm{p_e}}\mu_1(\lambda) - 1\right)\mu_2(\lambda)e^{\mu_2(\lambda)} - \left(\frac{1}{\mathrm{p_e}}\mu_2(\lambda) - 1\right)\mu_2(\lambda)e^{\mu_1(\lambda)} \\ &= e^{\frac{\mathrm{p_e}}{2}}\left[\frac{1}{\mathrm{p_e}}\left(\tfrac{1}{2}\mathrm{p_e} - \varsigma_a - \mathrm{p_e}\right)\cdot\left(\tfrac{1}{2}\mathrm{p_e} + \varsigma_a\right)e^{\varsigma_a} - \frac{1}{\mathrm{p_e}}\left(\tfrac{1}{2}\mathrm{p_e} + \varsigma_a - \mathrm{p_e}\right)\right. \\ &\qquad \left.\cdot\left(\tfrac{1}{2}\mathrm{p_e} - \varsigma_a\right)e^{-\varsigma_a}\right] \\ &= \frac{e^{\frac{\mathrm{p_e}}{2}}}{\mathrm{p_e}}\left[-\left(\tfrac{1}{2}\mathrm{p_e} + \varsigma_a\right)^2 e^{\varsigma_a} + \left(\tfrac{1}{2}\mathrm{p_e} - \varsigma_a\right)^2 e^{-\varsigma_a}\right] \end{aligned}$$

Take $\varsigma_a > 0$, then $\exp\varsigma_a > \exp -\varsigma_a > 0$ and furthermore, $\left(\frac{1}{2}\mathrm{p_e} + \varsigma_a\right)^2 > \left(\frac{1}{2}\mathrm{p_e} - \varsigma_a'\right)^2 > 0$ since $\mathrm{p_e} > 0$. Consequently, $\Delta(\lambda) < 0$ and no solution λ can be found.

Case 1b: Analogous to Case 1a, we get:

$$\begin{aligned} \Delta(\lambda) = e^{\frac{\mathrm{p_e}}{2}}\left[\frac{1}{\mathrm{p_e}}\left(\tfrac{1}{2}\mathrm{p_e} - \varsigma_b - \mathrm{p_e}\right)\cdot\left(\tfrac{1}{2}\mathrm{p_e} + \varsigma_b\right)e^{\varsigma_b}\ldots\right. \\ \left. - \frac{1}{\mathrm{p_e}}\left(\tfrac{1}{2}\mathrm{p_e} + \varsigma_b - \mathrm{p_e}\right)\cdot\left(\tfrac{1}{2}\mathrm{p_e} - \varsigma_b\right)e^{-\varsigma_b}\right]. \end{aligned}$$

However, for $\lambda^L < 0$ this reduces to:

$$\begin{aligned}
\Delta(\lambda) &= e^{\frac{\mathrm{p_e}}{2}} \cdot \ldots \\
&\quad \left[\frac{1}{\mathrm{p_e}} \left(\tfrac{1}{2}\mathrm{p_e} - \imath\varsigma - \mathrm{p_e} \right) \cdot \left(\tfrac{1}{2}\mathrm{p_e} + \imath\varsigma \right) e^{\imath\varsigma} \ldots \right. \\
&\quad \left. - \frac{1}{\mathrm{p_e}} \left(\tfrac{1}{2}\mathrm{p_e} + \imath\varsigma - \mathrm{p_e} \right) \cdot \left(\tfrac{1}{2}\mathrm{p_e} - \imath\varsigma \right) e^{-\imath\varsigma} \right] \\
&= \frac{e^{\frac{\mathrm{p_e}}{2}}}{\mathrm{p_e}} \left[- \left(\tfrac{1}{2}\mathrm{p_e} + \imath\varsigma \right)^2 e^{\imath\varsigma} + \left(\tfrac{1}{2}\mathrm{p_e} - \imath\varsigma \right)^2 e^{-\imath\varsigma} \right] \\
&= \frac{e^{\frac{\mathrm{p_e}}{2}}}{\mathrm{p_e}} \left[\left(- \left(\tfrac{1}{2}\mathrm{p_e} \right)^2 - \imath\mathrm{p_e}\varsigma + \varsigma^2 \right) e^{\imath\varsigma} + \left(\left(\tfrac{1}{2}\mathrm{p_e} \right)^2 - \imath\mathrm{p_e}\varsigma - \varsigma^2 \right) e^{-\imath\varsigma_\lambda} \right] \\
&= \frac{2\imath e^{\frac{\mathrm{p_e}}{2}}}{\mathrm{p_e}} \left[\left(\varsigma^2 - \left(\tfrac{1}{2}\mathrm{p_e} \right)^2 \right) \sin(\varsigma) - \mathrm{p_e}\varsigma\cos(\varsigma) \right]
\end{aligned}$$

Hence, for $\Delta(\lambda) = 0$, ς_k, $k \geq 0$ is the set of all solutions to the resolvent equation:

$$\tan(\varsigma) = \frac{2\mathrm{p_e}\varsigma}{\varsigma^2 - \left(\tfrac{1}{2}\mathrm{p_e} \right)^2} \tag{D.13}$$

Case 2:

$$\Delta(\lambda) = \det \begin{pmatrix} \frac{1}{2} e^{\frac{1}{2}\mathrm{p_e}} - 1 & \frac{1}{\mathrm{p_e}} \\ \frac{1}{2}\mathrm{p_e} e^{\frac{1}{2}\mathrm{p_e}} & \left(\frac{1}{2}\mathrm{p_e} + 1 \right) e^{\frac{1}{2}\mathrm{p_e}} \end{pmatrix}$$

Consequently, $\Delta(\lambda) = 0$ only if $\frac{3}{4}e^{\mathrm{p_e}} - \frac{1}{2}(\mathrm{p_e}+1)^2 e^{\frac{1}{2}\mathrm{p_e}}/\mathrm{p_e} = 0$ which does not hold for all $\mathrm{p_e} > 0$.

The eigenvalues λ follow from eqn. (D.12), Case 1b:

$$\lambda_k = -\tfrac{1}{4}\mathrm{p_e} - \frac{1}{\mathrm{p_e}}\varsigma_k^2 \quad \text{with} \quad \varsigma_k \text{ the solutions of } eqn.\ (\mathrm{D.13}) \tag{D.14}$$

Furthermore, recognize that $\lambda_k \to -\infty$ as $k \to \infty$ and $|\lambda_{k+1} - \lambda_k| \to \infty$ as $k \to \infty$.

Eigenvectors

The associated eigenvectors $\phi_k \in D(A)$, $k \geq 0$ are calculated for all $\eta \in [0,1]$ and for all $k \geq 1$ as follows. From $\mathfrak{B}_1\phi$, we obtain:

$$\begin{aligned}\left(\frac{1}{\mathrm{p_e}}\mu_1(\lambda)-1\right)C_1+\left(\frac{1}{\mathrm{p_e}}\mu_2(\lambda)-1\right)C_2=0 \Leftrightarrow\\ \left(\mu_1(\lambda)-\mathrm{p_e}\right)C_1+\left(\mu_2(\lambda)-\mathrm{p_e}\right)C_2=0 \Leftrightarrow\\ \left(-\tfrac{1}{2}\mathrm{p_e}-\varsigma_b\right)C_1+\left(-\tfrac{1}{2}\mathrm{p_e}+\varsigma_b\right)C_2=0.\end{aligned}$$

Now choose

$$C_1=-\frac{C_0}{\varsigma_b}\left(-\tfrac{1}{2}\mathrm{p_e}+\varsigma_b\right) \quad \text{and} \quad C_2=-\frac{C_0}{\varsigma_b}\left(\tfrac{1}{2}\mathrm{p_e}+\varsigma_b\right)$$

then

$$\begin{aligned}\phi_k(\cdot)&=-\frac{C_0}{\varsigma_b}\left[\left(-\tfrac{1}{2}\mathrm{p_e}+\varsigma_b\right)e^{(\frac{1}{2}\mathrm{p_e}-\varsigma_b)\eta}+\left(\tfrac{1}{2}\mathrm{p_e}+\varsigma_b\right)e^{(\frac{1}{2}\mathrm{p_e}+\varsigma_b)\eta}\right]\\ &=-\frac{C_0}{\imath\varsigma}e^{\frac{1}{2}\mathrm{p_e}}\left[\left(-\tfrac{1}{2}\mathrm{p_e}+\imath\varsigma\right)e^{-\imath\varsigma\eta}+\left(\tfrac{1}{2}\mathrm{p_e}+\imath\varsigma\right)e^{\imath\varsigma\eta}\right]\\ &=C_0\exp\left(\tfrac{1}{2}\mathrm{p_e}\eta\right)\left[\frac{\mathrm{p_e}}{\varsigma}\sin(\varsigma\eta)+2\cos(\varsigma\eta)\right]\end{aligned} \tag{D.15}$$

D.4 Approximation analysis of M

Here, whenever possible, we choose a modal approximation of M_{111} and M_{211}. In what follows, it is clarified that a modal approximation of M_{212} is not suitable. For this particular transfer function we use a Padé–$[1,2]$ approximation.

The entries of the transfer function matrix M which are described by trigoniometric functions, are approximated by using the residue theorem. Denote the rational transfer function $g(s)=a(s)/p(s)$ with $p(s)$ having roots in s_n. Then,

$$g_N=\sum_{n=1}^{N}\frac{a(s_n)}{p'(s_n)\left(s-s_n\right)} \quad \text{with} \quad p'(s_n)=\frac{\mathrm{d}p}{\mathrm{d}s}(s_n), \tag{D.16}$$

is an approximation of $g(s)$.

Note that, for $N \geq 1$, the approximation to M_{111}, M_{211} (and M_{212}) is bounded with L_∞–error (see also Exercise 8.21 from [21]),

$$\|M-M^N\| \leq \sum_{n=N+1}^{\infty}\frac{|a_n|}{|\mathrm{Re}(s_n)|}=\sum_{n=N+1}^{\infty}\frac{|a_n|}{|s_n|}$$

with a_n the numerator of M evaluated at s_n. Therefore, a modal approximation of $M_{212} = 1/R_2 = 1/\varsigma$ would lead to:

$$M_{212}^N = \sum_{n=1}^{N} \frac{1}{s - s_n} \frac{2\varsigma(s_n)}{\mathrm{p_e}} \tag{D.17}$$

Now, recall that $s_n = -\frac{1}{4}\mathrm{p_e} + \varsigma^2(s_n)/\mathrm{p_e}$ for $n \geq 0$, see eqn. (5.33d). Assume we have real-valued poles of M: $s_n < -\frac{1}{4}\mathrm{p_e}$. Consequently, $\varsigma(s_n) \in \imath\mathbb{R}$. In that case we get for large enough n that $\varsigma \to \pm\imath(n - \frac{1}{2})\pi$. Even in the case that the start assumption is $s_n \in \mathbb{C}$, there is only one other pole solution, namely, $s_n = -\frac{1}{4}\mathrm{p_e}$. We now check the upper bounds of the approximation error in the case that $s_n < -\frac{1}{4}\mathrm{p_e}$, leading to $\varsigma \to \pm\imath(n - \frac{1}{2})\pi$.

With the aid of eqn. (D.17), the L_∞–error analysis of M_{212} shows

$$\|M_{212} - M_{212}^N\| \leq \frac{2}{\mathrm{p_e}} \sum_{n=N+1}^{\infty} \frac{|\imath\left(n - \frac{1}{2}\right)\pi|}{\left|(n^2 - \frac{1}{4})\pi^2 - \left(\frac{1}{2}\mathrm{p_e}\right)^2\right|}.$$

Hence, from the above summation it appears that M_{212}^N does not converge to zero with increasing N. As an alternative, we estimate the function $R_2 = 1/\varsigma$ by Padé approximation.

The calculations of the derivatives of $M_{212} = 1/R_2$ with respect to the Laplace variable s and the unknowns of the Padé function lead to:

$$M_{212}^{[1,2]}(s) = \frac{\sqrt{1 + 4q/\mathrm{p_e}}s + 2q\sqrt{1 + 4q/\mathrm{p_e}} + \sqrt{1 + 4q/\mathrm{p_e}}}{\frac{1}{2}s^2 + \left(4q + \mathrm{p_e}\right)s + 4q^2 + 2\mathrm{p_e}q + \frac{1}{4}\mathrm{p_e}^2} \tag{D.18}$$

with $q := b_1 u_1$ in the case that u_1 is measured and constant. Recall that M_{211} also contains $R_2 = \varsigma$ in its denominator (see Chapter 5, eqn. (5.40). Hence, we choose to combine the Padé approximation as in eqn. (D.18) with the modal truncation of the trigoniometric part of M_{212}, i.e. QR_0/P, for the whole approximation of M_{211}.

The modal approximation of $M_{211} = R_0/P$ (see eqs. (5.34a) and (5.34d)) and of the trigoniometric part of $M_{212} = QR_0/P$ (see eqs. (5.34e), (5.34a) and (5.34d))

combined with the Padé approximation lead to:

$$M_{111}^N = \sum_{n=1}^{N} \frac{1}{s-s_n} \frac{2\varsigma^2(s_n)\exp\left(\frac{1}{2}\mathrm{p_e}\right)}{\left(1+\frac{1}{2}\mathrm{p_e}\right)\cosh\left(\varsigma(s_n)\right)+\varsigma(s_n)\sinh\left(\varsigma(s_n)\right)} \tag{D.19a}$$

$$M_{211}^N = \left(\sum_{n=1}^{N} \frac{1}{s-s_n} \frac{e^{\frac{1}{2}\eta^*\mathrm{p_e}}\left[\varsigma(s_n)\cosh\left((\eta^*-1)\varsigma(s_n)\right)-\frac{1}{2}\mathrm{p_e}\sinh\left((\eta^*-1)\varsigma(s_n)\right)\right]}{\left(1+\frac{1}{2}\mathrm{p_e}\right)\cosh\left(\varsigma(s_n)\right)+\varsigma(s_n)\sinh\left(\varsigma(s_n)\right)} \right) \cdot \frac{\sqrt{1+4q/\mathrm{p_e}}s+2q\sqrt{1+4q/\mathrm{p_e}}+\sqrt{1+4q/\mathrm{p_e}}}{\frac{1}{2}s^2+\left(4q+\mathrm{p_e}\right)s+4q^2+2\mathrm{p_e}q+\frac{1}{4}{\mathrm{p_e}}^2} \tag{D.19b}$$

With short-hand notation $\tilde{\eta} := 1-\eta^*$, and $\tilde{n} = n-\frac{1}{2}$ and assuming the Padé approximation $M_{212}^{[1,2]}$ is bounded with a positive error $m < \infty$, we continue the modal truncation error analysis with $N \gg 1$:

$$\|M_{111}-M_{111}^N\| \le \mathrm{p_e}\exp\left(\tfrac{1}{2}\mathrm{p_e}\right) \sum_{\tilde{n}=N+1/2}^{\infty} \frac{\left|(-1)^{\tilde{n}+1/2}\tilde{n}\pi\right|}{\left|\tilde{n}^2\pi^2-\frac{1}{2}\mathrm{p_e}\right|}$$

$$\|M_{211}-M_{211}^N\| \le 2m\exp\left(\tfrac{1}{2}\mathrm{p_e}(\tilde{\eta}+1)\right)\cdot \sum_{\tilde{n}=N+1/2}^{\infty} \left| \frac{\imath(-1)^{\tilde{n}+1/2}\tilde{n}\pi\left[\tilde{n}\pi\cos\tilde{n}\pi\tilde{\eta}\right)-\mathrm{p_e}\sin\tilde{n}\pi\tilde{\eta}\right)\Big]}{\tilde{n}^2\pi^2-\frac{1}{2}\mathrm{p_e}} \right| \left| \frac{1}{\tilde{n}\pi^2-\left(\frac{1}{2}\mathrm{p_e}\right)^2} \right|$$

so that the total approximation error in M_{211}^N behaves like

$$\|M_{211}-M_{211}^N\| \le m \qquad \text{for } \eta^* = 0$$

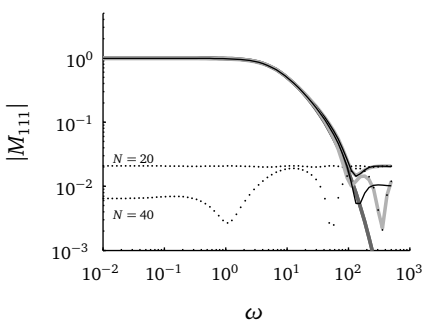
and

$$\|M_{211}-M_{211}^N\| \le me^{\frac{1}{2}\mathrm{p_e}} \sum_{n=N+1}^{\infty} \left| \frac{\imath(-1)^{\tilde{n}+1/2}\tilde{n}\pi}{\mathrm{p_e}} \right| \left| \frac{1}{\tilde{n}^2\pi^2-\frac{1}{2}\mathrm{p_e}} \right| \qquad \text{for } \eta^* = 1.$$

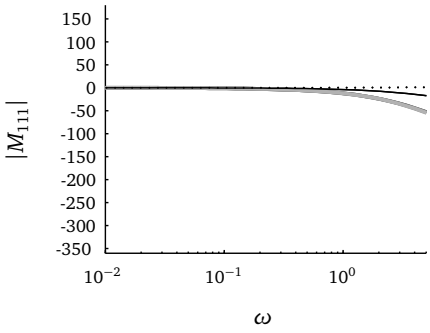
Thanks to the alternating sign, the error in M_{211}^N converges (slowly) to zero, depending on the sensor position. The same is true for the truncation error of M_{111}^N.

Graphs of approximation error

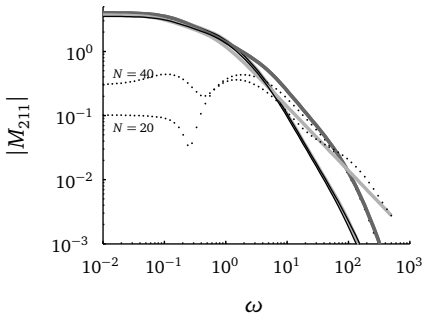
Here, it is shown what the Bode graphs of the approximation error $M(\jmath\omega)-M^N(\jmath\omega)$ is. Parameter values are: $\mathrm{p_e} = 1$, $m = 8$, $\eta^* = 0.5$ and $b_1u_1 = 0$.



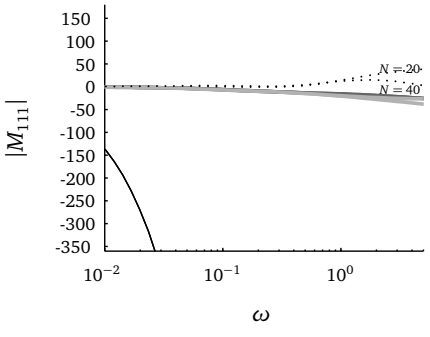
(a) $|M^N_{111}|$, with $N \in \{20, 40\}$



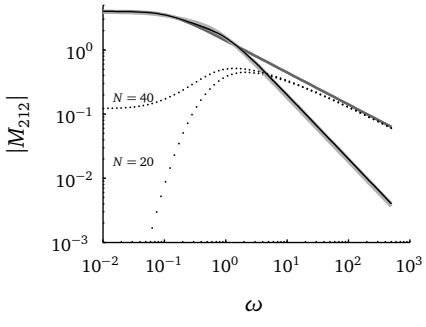
(b) $\arg M^N_{111}$ with $N \in \{20, 40\}$



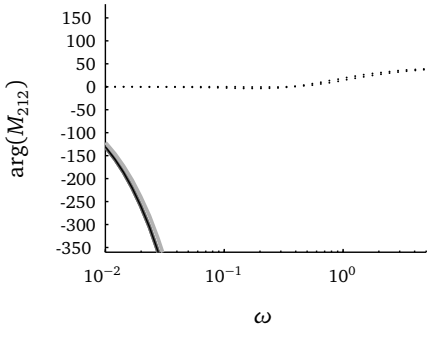
(c) $|M^N_{211}|$, with $N \in \{20, 40\}$



(d) $\arg M^N_{211}$, with $N \in \{20, 40\}$



(e) $|M^N_{212}|$, with $N \in \{20, 40\}$



(f) $\arg M^N_{212}$, with $N \in \{20, 40\}$

Figure D.1: Bode diagrams of M [—], M^N with varying N [-], balanced truncation M^8 [—] and balanced truncation error [··]. Left: amplitude diagrams, right: phase diagrams.

Bibliography

[1] Y. Adam. “Highly accurate compact implicit methods and boundary conditions”. In: *Journal of Computational Physics* 24(1) (1977). Pp. 10–22.

[2] D.C. Antoulas A.C. Sorensen and S. Gugercin. “A survey of model reduction methods for large-scale systems”. In: *Structured Matrices in Operator Theory, Numerical Analysis, Control, Signal and Image Processing, Contemporary Mathematics, AMS publications* 280 (2001). Pp. 193–219.

[3] K.J. Åström and B. Wittenmark. *Computer–Controlled Systems, Theory and Design*. 3rd ed. Prentice Hall, New Jersey, 1996.

[4] H.T. Banks and K. Kunisch. *Estimation Techniques for Distributed Parameter Systems*. Birkhäuser, Boston, 1989.

[5] G.I. Barenblatt. *Similarity, self-similarity, and intermediate asymptotics. Transl. from Russian by Norman Stein, ed. by Milton Van Dyke.* New York, London: Consultants Bureau, 1979.

[6] J. Baumeister et al. “On line parameter estimation for infinite–dimensional dynamical systems”. In: *SIAM Journal of Control and Optimization* 35(2) (1997). Pp. 678–713.

[7] B. Bhikkaji. “Model Reduction and Parameter Estimation for Diffusion Systems”. PhD thesis. Uppsala Universitet, 2004.

[8] T. Bohlin and S.F. Graebe. “Issues in nonlinear stochastic grey box identification”. In: *International Journal of Adaptive Control and Signal Processing* 9(6) (1995). Pp. 465–490.

[9] J. Bontsema and R.F. Curtain. “Perturbation Properties of a Class of Infinite-Dimensional Systems with Unbounded Control and Observation”. In: 5(4) (1988). Pp. 333–352.

[10] H. Bounit and H. Hammouri. “Observers for Infinite Dimensional Bilinear Systems”. In: *Eur. J. Control* 3.1 (1997). Pp. 325–339.

[11] H. Bounit and A. Idrissi. “Regular Bilinear Systems”. In: *IMA J. Math. Control Inf.* 22.1 (2005). Pp. 26–57.

[12] E. Buckingham. “On Physically Similar Systems; Illustrations of the Use of Dimensional Equations”. In: *Phys. Rev.* 4.4 (1914). Pp. 345–376.

[13] C.I. Byrnes et al. “Regular linear systems governed by a boundary controlled heat equation.”. In: *J. Dyn. Control Syst.* 8.3 (2002). Pp. 341–370.

[14] X.-C. Cai. “Additive Schwarz algorithms for parabolic convection-diffusion equations”. In: *Numerische Mathematik* 60(1) (1991). Pp. 41–61.

[15] M.K. Chourasia and T.K. Goswami. “Simulation of transport phenomena during natural convection cooling of bagged potatoes in cold storage, part I: fluid flow and heat transfer”. In: *Biosystems engineering* 94(1) (2006). Pp. 33–45.

[16] M.K. Chourasia and T.K. Goswami. “Simulation of transport phenomena during natural convection cooling of bagged potatoes in cold storage, part II: mass transfer”. In: *Biosystems engineering* 94(2) (2006). Pp. 207–219.

[17] D. Coca and S.A. Billings. “Identification of finite dimensional models of infinite dimensional dynamical systems”. In: *Automatica* 38(11) (2002). Pp. 1851–1865.

[18] J. Crank. *The Mathematics of Diffusion*. Oxford, Clarendon Press, 1957.

[19] R. Curtain. “Finite-dimensional compensator design for parabolic distributed systems with point sensors and boundary input”. In: *IEEE Transactions on Automatic Control* 27(1) (1982). Pp. 98–104.

[20] R.F. Curtain, M. A. Demetriou, and K. Ito. “Adaptive observers for structurally perturbed infinite dimensional systems”. In: *International Journal of Applied Mathematics and Computer Science* 13(4) (2003). Pp. 441–452.

[21] R.F. Curtain and H. Zwart. *An Introduction to Infinite–Dimensional Linear Systems Theory*. Springer-Verlag, New York, 1995.

[22] G. Damköhler. “Einflüsse der Strömung, Diffusion und des Wärmeübergangs auf die Leistung von Reaktionsöfen”. In: *Z. Elektrochemie* 42 (1936). Pp. 846–862.

[23] P.V. Danckwerts. “Theory of mixtures and mixing”. In: *Res. (London) 6* (1953). Pp. 355–361.

[24] C. Delattre. “Etude de l’observabilité de systèmes de Sturm-Liouville: application aux réacteurs biochimiques à paramètres répartis”. PhD thesis. Université Catholique de Louvain, 2003.

[25] C. Delattre, D. Dochain, and J. Winkin. “Observability analysis of nonlinear tubular (bio)reactor models: a case study”. In: *Journal of Process Control* 14(6) (2004). Pp. 661–669.

[26] C. Delattre, D. Dochain, and J. Winkin. “Sturm-Liouville systems are Riesz-spectral systems.”. In: *Int. J. Appl. Math. Comput. Sci.* 13.4 (2003). Pp. 481–484.

[27] T. Z. Desta et al. “CFD for model-based controller development”. In: *Building and Environment* 39.6 (2004). Pp. 621–633.

[28] T. Z. Desta et al. “Combining CFD and data-based mechanistic (DBM) modelling approaches”. In: *Energy and Buildings* 36.6 (2004). Pp. 535–542.

[29] D. Dochain. *Contribution to the analysis and control of distributed parameter systems with application to (bio)chemical processes and robotics*. Thèse d'Agrégation de l'Enseignement Supérieur. Université Catholique de Louvain, Louvain-la-Neuve, Belgium. July 1994.

[30] T.G. Doeswijk and K. J. Keesman. "Parameter Estimation and Prediction of a Nonlinear Storage Model: an algebraic approach". In: *2005 International Conference on Control and Automation (ICCA2005)*. IEEE. Budapest, Hungary 2005. Pp. 95–100.

[31] T.G. Doeswijk and K. J. Keesman. "Parameter estimation and prediction of nonlinear biological systems: some examples". In: *14th IFAC Symposium on System Identification (SYSID06)*. Newcastle, Australia 2006. Pp. 1156–1161.

[32] A. G. Duse, M. P. da Silva, and I. Zietsman. "Coping with hygiene in South Africa, a water scarce country". In: *Int. J. Environ. Heal. R.* 13 (2003). S95–S105.

[33] Z. Emirsajlow and S. Townley. "From PDEs with Boundary Control to the Abstract State Equation with an Unbounded Input Operator". In: *European Journal of Control* 6(1) (2000).

[34] R. Everson and L. Sirovich. "Karhunen-Loève procedure for gappy data". In: *Journal of Optical Society of America A* 12 (1995). Pp. 1657–1664.

[35] R.E. Ewing, R.D. Lazarov, and A.T. Vassilev. "Finite Difference Scheme for Parabolic Problems on Composite Grids with Refinement in Time and Space". In: *SIAM Journal on Numerical Analysis* 31(6) (1994). Pp. 1605–1622.

[36] H.O. Fattorini. "Boundary Control Systems". In: *SIAM Journal of Control* 6 (1968). Pp. 349–388.

[37] B.A. Finlayson. *Method of Weighted Residuals and Variational Principles*. Academic Press, London, 1972.

[38] G.F. Froment and K.B. Bischoff. *Chemical reactor analysis and design*. Wiley, New York, 1979.

[39] M. Gevers and X. Bombois. "Input design : from open-loop to control-oriented design". In: *IFAC Symp. on System Identification 2006 (SYSID 2006), Newcastle, Australia*. April 2006. Pp. 1329–1334.

[40] K. Gottschalk and W. Schwarz. "Klimaautomatisierung für Kartoffellager". In: *Landtechnik* 3 (1997). Pp. 132–133.

[41] G. Gu, P.P. Khargonekar, and E.B. Lee. "Approximation of infinite-dimensional systems". In: *IEEE Transactions on Automatic Control* 34(6) (1989). Pp. 610–618.

[42] J. A. Guerrero-Beltran and G. V. Barbosa-Canovas. "Review: Advantages and limitations on processing foods by UV light". In: *Food Sci. Technol. Int.* 10.3 (2004). Pp. 137–147.

[43] S. Gugercin and A.C. Antoulas. “A Survey of Model Reduction by Balanced Truncation and Some New Results”. In: *International Journal of Control* 77(8) (2004). Pp. 748–766.

[44] W. A. M. Hijnen, E. F. Beerendonk, and G. J. Medema. “Inactivation credit of UV radiation for viruses, bacteria and protozoan (oo)cysts in water: A review”. In: *Water Research* 40(1) (2006). Pp. 3–22.

[45] J.M. van den Hof. “Structural Identifiability of Linear Compartmental Systems”. In: *IEEE Transactions on Automatic Control* 43(6) (1998). Pp. 800–818.

[46] P. M. J. van den Hof and X. Bombois. *System Identification for Control*. DISC Lecture Notes. Mar. 2004.

[47] G.Y. Hu and R.F. O’Connell. “Analytical inversion of symmetric tridiagonal matrices”. In: *J. Phys. A: Math. Gen.* 29 (1996). Pp. 1511–1513.

[48] Y. Huang and W.F. McColl. “Analytical inversion of general tridiagonal matrices”. In: *J. Phys. A: Math. Gen.* 30 (1997). Pp. 7919–7933.

[49] F. Jauberteau, C. Rosier, and R. Temam. “The nonlinear Galerkin method in computational fluid dynamics”. In: *Applied Numerical Mathematics* 6 (1990). Pp. 361–370.

[50] J. P. Keener. *Principles of Applied Mathematics: Transformation andApproximation*. Perseus Books.

[51] K. J. Keesman, D. Peters, and L. J. S. Lukasse. “Optimal climate control of a storage facility using local weather forecasts”. In: *Control Engineering Practice* 11(5) (2003). Pp. 505–516.

[52] K. J. Keesman and J. D. Stigter. “On Compartmental Modelling Of Mixing Phenomena”. In: *15th IFAC World Congress on Automatic Control*. (CD-Rom). Barcelona, Spain 2002.

[53] K. J. Keesman et al. “Modeling and control of water disinfection process in annular photoreactors”. In: *Proceedings of European Control Conference*. Kos, Greece 2007.

[54] P. V. Kokotovic, H. K. Khalil, and J. O’Reilly. *Singular perturbation methods in control: analysis and design*. Academic Press, London, 1986.

[55] E. Kreyszig. *Introductory Functional Analysis with Applications*. John Wiley & Sons, 1978.

[56] V. Lazarova et al. “Advanced wastewater disinfection technologies: State of the art and perspectives”. In: *Water Sci. Technol.* 40.4-5 (1999). Pp. 203–213.

[57] L. Lefevre et al. “Optimal selection of orthogonal polynomials applied to the integration of chemical reactor equations by collocation methods”. In: *Computers & Chemical Engineering* 24(12) (2000). Pp. 2571–2588.

[58] J. Lin et al. “Parameter estimation of fractional systems : application to the modeling of a lead-acid battery”. In: *(CD-ROM) Proceedings IFAC SysId 2000, Santa Barbara, 21 - 23 June*. 2000.

[59] L. Ljung. *System Identification: Theory for the User*. Prentice Hall, Englewood Cliffs, New Jersey, 1987.

[60] L. Ljung and T. Glad. “On global identifiability for arbitrary model parametrizations”. In: *Automatica* 30(2) (1994). Pp. 265–276.

[61] F. J. Loge et al. “Ultraviolet disinfection of secondary wastewater effluents: prediction of performance and design”. In: *Water Environment Research* 68 (1996). Pp. 900–916.

[62] L. J. S. Lukasse, K. J. Keesman, and G. Van Straten. “Grey-box identification of dissolved oxygen dynamics in activated sludge processes.”. In: *Proceedings of 13th IFAC World Congress*. Vol. N. San Francisco, USA 1996. Pp. 485–490.

[63] L. J. S. Lukasse, J. E. de Kramer-Cuppen, and A. J. van der Voort. “A physical model to predict climate dynamics in ventilated bulk–storage of agricultural produce”. In: *International Journal of Refrigeration* 30(1) (2007). Pp. 195–204.

[64] J.P. Malley. *Control of Microorganisms in Drinking Water*. Ed. by S. Lingireddy. American Society of Civil Engineers, 2002.

[65] S. Matthews, E. O’Riordan, and G.I. Shishkin. “A numerical method for a system of singularly perturbed reaction-diffusion equations”. In: *Journal of Computational and Applied Mathematics* 145(1) (2002). Pp. 151–166.

[66] V. Mavrov, A. Fahnrich, and H. Chmiel. “Treatment of low–contaminated waste water from the food industry to produce water of drinking quality for reuse”. In: *Desalination* 113.2-3 (1997). Pp. 197–203.

[67] G. Meurant. “A review on the inverse of symmetric tridiagonal and block tridiagonal matrices”. In: *SIAM Journal on Matrix Analysis and Applications* 13.3 (1992). Pp. 707–728.

[68] J. Molenaar. “Mathematical modelling and dimensional analysis”. In: *Topics in engineering mathematics, modeling and methods*. Ed. by A. v.d. Burgh and J. Simonis. Kluwer Academic Publ. Dordrecht, 1992. Pp. 93–120.

[69] M. Moonen et al. “On- and off-line identification of linear state-space models”. In: *International Journal of Control* 49 (1989). Pp. 219–232.

[70] S. van Mourik, H. Zwart, and K. J. Keesman. *Analytical Control law for a food storage room*. Memorandum 1819. University of Twente, 2007.

[71] S. van Mourik, H. Zwart, and K. J. Keesman. *Switching control for a class of nonlinear SISO systems with an application to post-harvest food storage*. Memorandum 1853. University of Twente, 2007.

[72] S. van Mourik. “Modelling and control of systems with flow”. PhD thesis. University of Twente, 2008.

[73] K.M. Nagpal and P. Khargonekar. “Filtering and Smoothing in an H^∞ Setting”. In: *IEEE Transactions on Automatic Control* 36(2) (1991). Pp. 152–166.

[74] A.J. Newman. *Model reduction via the Karhunen-Loeve expansion Part I: An exposition*. Tech. rep. T.R. 1996-32. Inst. Systems Research, 1996.

[75] A.J. Newman. *Model Reduction via the Karhunen-Loeve Expansion Part II: Some Elementary Examples*. Tech. rep. T.R. 1996-33. Inst. Systems Research, 1996.

[76] J.P. Norton. *An Introduction to Identification*. Academic Press, London and New York, 1986.

[77] E. O’Riordan, M.L. Pickett, and G.I. Shishkin. “Singularly perturbed problems modeling Reaction-Convection-Diffusion Processes”. In: *Computational Methods in Applied Mathematics* 3 (2003). Pp. 424–442.

[78] P. van Overschee. “Choice of State-Space Basis in Combined Deterministic-Stochastic Subspace Identification”. In: *Automatica* 31(12) (1995). Pp. 1877–1883.

[79] S.V. Patankar. *Numerical Heat Transfer and Fluid Flow*. Washington, D.C: Hemisphere, 1980.

[80] J.R.A. Pearson. “A note on the “Danckwerts” boundary conditions for continuous flow reactors”. In: *Chemical Engineering Science* 10(4) (1959). Pp. 281–284.

[81] B. Peeters and G. de Roeck. “Reference-based stochastic subspace identification for output-only modal analysis”. In: *Mechanical Systems and Signal Processing* 6 (1999). Pp. 855–878.

[82] R. Pintelon et al. “Analysis, development and applications of TLS algorithms in frequency domain system identification”. In: *SIAM J. Matrix Anal. Appl.* 19(4) (1998). Pp. 983–1004.

[83] R. Pintelon et al. “Diffusion systems: Stability, modeling, and identification”. In: *IEEE Transactions on Instrumentation and Measurement* 54(5) (2005). Pp. 2061–2067.

[84] T. Poinot and J.C. Trigeassou. “Identification of fractional systems using an output-error technique”. In: *Nonlinear dynamics* 38(1–4) (2004). Pp. 133–154.

[85] P. Schlegel. “The explicit inverse of a tridiagonal matrix”. In: *Mathematics of Computation* 24(111) (1970). P. 665.

[86] J. Schoukens, Pintelon R., and J. Renneboog. “A Maximum Likelihood Estimator for Linear and Nonlinear Systems—a Practical Application of Estimation Techniques in Measurement Problems”. In: *IEEE Transactions on Instrumentation and Measurement* 37(1) (1988). Pp. 10–17.

[87] K.C. Schuster. “Monitoring the Physiological Status in Bioprocesses on the Cellular Level”. Ed. by B. Sonnleiter. In: *Adv. Biochem. Eng. Biot.* 66 (2000). Pp. 185–208.

[88] B.F. Severin et al. “Inactivation kinetics in a flow-through UV reactor”. In: *Joumal of the Water Pollution Control Federation* 56(2) (1984). Pp. 164–169.

[89] J. Sjöberg et al. “Nonlinear Black-box Modeling in System Identification: a Unified Overview”. In: *Automatica* 31(12) (1995). Pp. 1691–1724.

[90] R. Smith. “Entry and Exit Conditions for Flow Reactors”. In: *IMA Journal of Applied Mathematics* (1988).

[91] A. Smyshlyaev and M. Krstic. “Backstepping observers for a class of parabolic PDEs.”. In: *Syst. Control Lett.* 54.7 (2005). Pp. 613–625.

[92] T. Söderström and B. Bhikkaji. “Reduced Order Models for Diffusion Systems via Collacation Methods”. In: *IFAC 12th Symposium on System Identification*. 6. (CD-Rom). Santa Barbara, California 2000.

[93] T. Söderström and S. Remle. “Parameter Estimation and model approximation for diffusion models”. In: *IFAC 12th Symposium on System Identification*. (CD-Rom). Santa Barbara, California 2000.

[94] G. Stephanopoulos. *Chemical Process Control: An introduction to theory and practice*. Ed. by N.R. Amundson. Prentice Hall, Englewood Cliffs, New Jersey, 1984.

[95] J.D. Stigter, D Vries, and K.J. Keesman. “On adaptive optimal input design: A bioreactor case study”. In: *American Institute of Chemical Engineers Journal* 52(9) (2006). Pp. 3290–3296.

[96] J.W. Thomas. *Numerical Partial Differential Equations: Finite Difference Methods*. Springer, New York, 1995.

[97] P. Verboven et al. “Modelling transport phenomena in refrigerated food bulks, packages and stacks: basics and advances”. In: *International Journal of Refrigeration* 29 (2006). Pp. 985–997.

[98] G.J.C. Verdijck and G. van Straten. “A modelling and control structure for product quality control in climate–controlled processing of agro–material”. In: *Control Engineering Practice* 10(5) (2002). Pp. 533–548.

[99] M. Verhaegen and P. Dewilde. “Subspace model identification Part 1: The output-error state-space model identification class of algorithms”. In: *International Journal of Control* 56(5) (1992). Pp. 1187–1210.

[100] M. Verhaegen and P. Dewilde. “Subspace model identification Part 2. Analysis of the elementary output-error state-space model identification algorithm”. In: *International Journal of Control* 56(5) (1992). Pp. 1211–1241.

[101] D. Vries, K. J. Keesman, and H. Zwart. “A Luenberger observer for an infinite dimensional bilinear system: a UV disinfection example”. In: *Proceedings of IFAC Symposium on Systems, Structure and Control*. Foz de Iguassu, Brazil 2007.

[110] L. Wang et al. “Indirect approach to continuous time system identification of food extruder”. In: *Journal of Process Control* 14(6) (2004). Pp. 603–615.

[111] G. Weiss and R.F. Curtain. “Dynamic stabilization of regular linear systems.”. In: *IEEE Trans. Autom. Control* 42.1 (1997). Pp. 4–21.

[112] Joseph J. Winkin, Denis Dochain, and Philippe Ligarius. “Dynamical analysis of distributed parameter tubular reactors.”. In: *Automatica* 36(3).3 (2000). Pp. 349–361.

[113] C.-Z. Xu, J. Deguenon, and G. Sallet. “Infinite Dimensional Observers for Vibrating Systems”. In: *45th IEEE Conference on Decision and Control*. 2006. Pp. 3979–3983.

[114] C. Z. Xu, P. Ligarius, and J. P. Gauthier. “An Observer for Infinite–Dimensional Dissipative Bilinear Systems”. In: *Computers & Mathematics with Applications* 29(7).7 (1995). Pp. 13–21.

[115] Y. Xu and D. Burfoot. “Predicting condensation in bulks of foodstuffs”. In: *Journal of food engineering* 40 (1999). Pp. 121–127.

[116] J. Zhang. “An explicit fourth-order compact finite difference scheme for three-dimensional convection-diffusion equation”. In: *Communications in Numerical Methods in Engineering* 14(3) (1998). Pp. 209–218.

[117] K. Zhou. *Robust and Optimal Control*. Ed. by J.C. Doyle and K. Glover. Prentice-Hall, New Jersey, 1996.

[118] H. Zwart. “Transfer functions for infinite-dimensional systems”. In: *Systems and Control Letters* 52 (2004). Pp. 247–255.

Notation

α	Diffusion constant, p. 15
$*$	$h * g$, Convolution product of h and g, p. 55
β	Multiplicative constant on $z(a)$, with a a boundary of the spatial domain (only this chapter), p. 76
β	Weight of weighted residuals method (only this chapter), p. 5
β_c	Efficiency constant of cooling device, p. 18
$\delta(t)$	Delta operator, p. 68
Δ_t	Time discretization constant, p. 32
Δ_ξ	Spatial discretization parameter, p. 58
$\dot{z}$	Derivative of z w.r.t. time or space (depending on the context), p. 5
η	Spatial coordinate, p. 5
Γ	Vector of response function values, p. 55
γ	Multiplicative constant on $z(a)$, with a a boundary of the spatial domain (only this chapter), p. 76
γ	Response function containing input and output data (only this chapter), p. 45
γ'	Response function in predictor of linear regression system, p. 45
γ_H	Infemum of H_∞–optimization, p. 86
$\hat{x}$	Estimate of x, p. 7
$\imath$	Imaginary number, p. 84
$\imath$	Index limit (only this chapter), p. 59
$\langle \cdot, \cdot \rangle$	$\langle u, v \rangle$, Inner product of u and v., p. 73
$\jmath$	Index limit (only this chapter), p. 59
κ	Susceptibility constant of micro-organisms w.r.t. UV irrandiance, p. 22

$\ker T$ Kernel of T, p. 55

λ Eigenvalue, p. 75

λ_p Conduction of product, p. 18

$\mathbb{C}_+^\omega$ All complex numbers with real part larger than ω, p. 87

$\mathbb{J}$ Index set, p. 75

$\mathbb{N}_+$ Set of positive integers, p. 46

$\mathbb{R}$ Set of real numbers, p. 15

$\mathbb{R}_+$ Set of positive (i.e. larger than zero) real numbers, p. 46

$\mathbf{A}$ System matrix, p. 46

$\mathbf{B}$ Input matrix, p. 46

$\mathbf{C}$ Observation matrix, p. 46

$\mathbf{D}$ Feedthrough matrix, p. 46

$\mathbf{H}$ Hankel matrix, p. 55

$\mathbf{L}$ Observer gain matrix, p. 68

$\mathbf{M}$ Inverse of the resolvent of system matrix $\mathbf{A}$, p. 53

$\mathbf{N}$ Coefficient matrix in numerator of transfer function, p. 49

$\mathbf{Q}$ (Auxiliary) coefficient matrix, p. 59

$\mathbf{R}$ Coefficient matrix in denominator of transfer function, p. 49

$\mathbf{U}_{k-n\cdots k}$ Input data set in a time interval $[k-n,k]$, p. 45

$\mathbf{v}$ Velocity vector, p. 15

$\mathbf{x}$ State vector, p. 44

$\mathbf{y}$ Observation/system output vector, p. 44

$\mathbf{Y}_{k-n\cdots k}$ Output data set in a time interval $[k-n,k]$, p. 45

$\mathbf{Z}_{k-n\cdots k}$ Input-output data set in a time interval $[k-n,k]$, p. 45

$\mathscr{C}^\tau$ Observability map on $[0,\tau]$, p. 71

$\mathscr{F}_L(P,Q)$ Lower linear fractional transformation, p. 92

$\mathscr{L}(X,Y)$ Bounded linear operator(s) from X to Y, p. 68

$\mathscr{M}$ Collection of prediction models, p. 33

$\mathscr{N}$ Non-observable subspace, p. 71

$\mathscr{S}$ Notion of real world system, p. 7

$\mathfrak{A}$ Differential operator, p. 67

$\mathfrak{B}$ Boundary (control) operator, p. 67

$\mathfrak{C}$ Observation operator, p. 67

$\mathfrak{C}^b$ boundary observation operator, p. 77

$\|\cdot\|$ $\|z\|$, Norm of z, p. 70

ω_0 Growth bound of semigroup, p. 71

∂ Partial derivative operator, p. 37

Φ Matrix of regressors, p. 55

ϕ Orthonormal eigenvector, p. 75

ϕ Regressor vector containing input and output data, p. 45

ϕ' Regressor in predictor of linear regression system, p. 45

ψ Stacked shift operator vector, p. 49

ψ non-dimensionless state variable, p. 15

ρ Density, p. 15

Σ System representation, p. 3

$\sigma(A)$ spectrum of A, p. 80

$\Sigma(A,B,C,D)$ State (bi)linear (infinite-dimensional) system, p. 12

$\Sigma(G')$ Input-output linear system with approximated transfer function, p. 12

$\Sigma(G)$ Input-output (infinite-dimensional) linear system, p. 12

$\Sigma^{\text{obs}}(\cdot)$ Observer system, p. 12

Σ^{room} Climate room model, p. 18

$\Sigma^e_{\cdot}$ Example model, p. 56

Σ_N^{opt} Identified model with respect to a particular model selection criterion., p. 35

Σ_d Discrete (time) model, p. 28

σ_d Noise variance, p. 62

$\Sigma_N(\cdot)$ Finite-dimensional system, p. 12

$\Sigma_{N,d}$ N-th order discrete time model, p. 28

τ_p Dominant time constant of process, p. 34

θ (To be estimated) parameter, p. 7

θ_B Parameter vector associated to B-polynomial of output-error model, p. 33

θ_F Parameter vector associated to F-polynomial of output-error model, p. 33

θ_p Local sensitivity of θ w.r.t. the physical parameter p, p. 37

$\tilde{a}$, $\tilde{b}$ Parameters related to first order model, p. 31

$\top$ Transpose of vector or matrix, p. 45

ε Residual or error, p. 5

φ Basis function, p. 5

φ Stacked parameter vector, p. 49

ς Constant closely related to eigenvalue, p. 80

ϑ Physical parameter, p. 44

$\widehat{\ }$ Laplace transform of signal, p. 19

$\widetilde{\mathbf{N}}$ Coefficient matrix in numerator of G, with three or more dimensions, p. 49

$\widetilde{\mathbf{R}}$ Coefficient matrix in denominator polynomial of G with three or more dimensions, p. 49

ξ Non-dimensionless spatial coordinate, p. 15

$^{\text{LR}}$ Linear regressive, p. 12

A Differential operator, p. 68

$B(\theta, q)$ Polynomial in output-error model, related to the system inputs, p. 33

c Concentration of living/pathogenic micro-organisms, p. 20

c_p heat capacity of produce, p. 18

c_{in} Concentration of (pathogenic) micro-organisms at inlet of reactor, p. 20

$c_{\varepsilon u}$ Cross correlation coefficient between residuals and inputs, p. 35

$D(\cdot)$ Domain of operator, p. 67

E Monochromatic absorbance of medium, p. 22

f Irradiation function, p. 22

$F(\theta, q)$ Polynomial in output-error model, related to the system outputs, p. 33

$F_{p.e.}$ Final prediction error functional, p. 34

G Transfer function, p. 12

g_N Numerator of transfer function, p. 48

g_R Denominator of transfer function, p. 48

h Heat transfer coefficient, p. 18

H_∞ Hardy space of bounded holomorphic functions on $\mathbb{C}_+^0$ with values in $\mathbb{C}$, p. 68

I Identity matrix, p. 53

K Gain constant of output error model, p. 39

L Length of object, p. 16

$L_p(a, b)$ Class of Lebesque measurable functions with $\int_a^b |f(t)|^p \mathrm{d}t < \infty$, p. 56

l_s Index limit of counters, p. 59

$M_.$ Lumped parameter in climate room model, with index from 1 to 5, p. 18

N Model order or degree of denominator polynomial of transfer function, p. 7

N_k Number of data points, p. 8

p Physical parameter, p. 31

q^{-1} One time step backward shift operator, p. 33

R Radius of UV disinfection reactor, p. 22

r Radius of agricultural produce, p. 18

r_0 Outer radius of UV lamp tube, p. 22

S Source term, p. 15

$T(t,s)$ C_0-semigroup (strongly continuous semi-group for $t,s \geq 0$), p. 70

T_a Air temperature, p. 16

T_c Temperature of cooling device, p. 16

t_k k-th time instant, p. 32

T_p Product temperature, p. 18

T_s Temperature of the shaft in the climate storage room, p. 16

U Input or (admissible) control space, p. 28

u System input, p. 7

V Identification cost function or loss function, p. 8

v Disturbance, p. 7

V_a Air volume, p. 18

v_f Velocity of fluid, p. 22

v_p Velocity of air in bulk, p. 18

v_s Velocity of air in shaft, p. 18

Y Observation space, p. 28

y System output, p. 7

z State variable, p. 5

g,$\tilde{g}$ Transfer function and its approximation, p. 6

Summary

In this thesis, various procedures with respect to estimation and prediction of systems characterized by convection, diffusion and reactions on the basis of point measurement data have been worked out. These systems are called infinite-dimensional, since the state variables (like temperature or concentration) are described by partial differential equations or delay difference equations in an infinite-dimensional manifold (or state space).

Two applications of these convection-diffusion-reaction (CDR) systems have been used as a show case of the proposed estimation and prediction methods. One is a climate room for bulk storage of agricultural produce[a] and the other is a UV disinfection process used in water treatment, food industry and greenhouse technologies[b].

Besides the occurence of flow, dispersion and source terms (reactions/heat production), other similarities in these applications are as follows:

(a) processes take place in an enclosure with fixed boundaries;

(b) only boundary and/or point measurements are available; and

(c) the modeling aim in both applications is process control.

In Chapter 2, the applications under study are introduced. In Case A, the dynamic, climate room storage model proposed in [72] is used as a starting point for the modeling of heat transfer over agricultural produce. In Case B, the UV disinfection process is modeled as a convection–diffusion process with first-order biomass deactivation kinetics in the main flow direction. The models obtained are called the nominal models.

The body of this thesis is basically divided into four parts, in which a parameter estimation, prediction and two state estimation methods for CDR systems are worked out. An essential step in the implementation of estimation and prediction is model reduction. Model reduction in the context of this thesis means the approximation of infinite-dimensional to finite dimensional systems. A distinction is made between early and late model reduction. In early model reduction, the approximation is performed *before* estimation and prediction is worked out. Late model reduction encompasses approximation of an infinite-dimensional system *after* an estimator (observer) or predictor is synthesized.

The estimation and prediction methods are described here consecutively.

Parameter estimation. Chapter 3 deals with the following research question:

[a] (Case A)
[b] (Case B)

> *'Given an early reduction technique, is it possible to preserve physical knowledge of the nominal CDR model in terms of the parameter estimate $\hat{\theta}$, when a typical identification (output-error) technique is used?'*

A procedure is proposed to recover physical (possibly lumped) parameters of the approximate physical model from the identified model and discrete measurement data. The nominal, physical model of heat transfer in a storage room is approximated via Padé approximation and calibrated via output-error identification. One of the identified parameters, in this case the dominant time constant, is then used to calculate the values of the original physical parameters in the approximate model after selecting a parameter subset through local parametric sensitivity analysis.

It is found that for all ventilator settings the gain of the bulk storage process was close to one, indicating that there was no unexpected heat loss or gain. In addition to this, important physical parameters (like porosity of the bulk, surface of the bulk exposed to air, height of the bulk, ...) can be individually recovered from the calibrated output-error model parameters.

Parameter estimation and output prediction. In Chapter 4, the estimation and prediction problem is approached from another perspective:

> *'Given an early reduction technique, is it possible to rewrite the estimation and prediction problem of a CDR system into a linear regression, and if so, how?'*

Inspired by the work of Doeswijk and Keesman in [30] for reparametrizations of non-linear systems, a reparametrized linear regressive system formulation is worked out for a discretized distributed parameter system in state-space form. The discretization is achieved via early model reduction by finite central differences techniques of the CDR model. The discretized form is referred to as Σ_d, and belongs to the class of discrete-time, linear structured systems.

It is shown that under parameter-affine conditions, a linear regressive formulation Σ_d^{LR} of the state space model Σ_d can be found via reparametrizations of the transfer function of Σ_d. In this realization, a linear regressive mapping from newly defined parameters to input-output data is formulated and characterized by parameter independent coefficient matrices. These coefficient matrices have been explicitly expressed in terms of number of compartments n, actuator and sensor location for two discretized diffusion examples under specific boundary conditions. The explicit results can also be applied to dimensionless discretized representations of CDR systems.

The formulation of a linear regressive estimation and prediction method allows (i) a transparent link to the original physical model, (ii) unique parameter estimates by least-squares techniques and (iii) a simple identifiability rank

test on the basis of a coefficient matrix. Furthermore, we show explicit results of the parametric sensitivities on the response vector for diffusion cases and specific CDR cases.

A disadvantage shows up in the practical implementation of linear regressive parameter estimation with ordinary least squares techniques. Since the regressors contain input and output data, they are corrupted by measurements noise, thus deteriorating the estimation result.

Static observer design for state estimation. In the first part of Chapter 5, the following question is addressed:

> *'Without model order reduction, is it possible to obtain design conditions for a static, boundary-type observer of a CDR system, under the condition that it is (approximately) observable?'*

Infinite-dimensional systems like the CDR system with point sensing and/or point actuation, typically have 'unbounded' control and/or observation operators [see for instance 21, Chapter 1] and introduce mathematical technicalities. It is hypothesized that (simple) concepts for observability and detectability for these particular CDR systems can be obtained by using abstract boundary system formulations and that these concepts can be worked out for static observer design.

To this aim, the UV disinfection process is used as an illustrating example for the design of a static, Luenberger-type observer. Since a CDR process is typically described by a Sturm-Liouville operator, detectability and approximate observability results were looked for and indeed obtained for boundary system formulations characterized by a Sturm-Liouville operator.

The results are used in deducing pole placement design rules for an observer of a UV disinfection model using only boundary data. By numerical simulation, the observer characteristics are further illustrated at the end of the chapter.

Although engineering skills are needed for choosing the observer gain, it is concluded that the proposed observer design procedure preserves physical insight and is independent of some approximation of the original infinite-dimensional system.

Dynamic observer design for state estimation. In the second part of Chapter 5, the following question is addressed:

> *'Given a late reduction technique, is it possible to obtain a dynamic observer which is robust to disturbances at both the (boundary) input as well as the point observation, and if so, how?'*

To this aim, a robust observer synthesis approach has been worked out on the basis of a reformulation of the estimation error system. Hereto, the following sequence of steps for observer synthesis is proposed:

(a) Laplace transformation of the estimation error system to obtain the closed-loop transfer function G;

(b) formulation of the H_∞–problem in a well-posed linear fractional transformation framework;

(c) modal and Padé approximation and subsequent balanced truncation of the open-loop model M, and

(d) calculation of the optimal H_∞–filter with the aid of the approximation of M.

The proposed procedure is applied to the UV disinfection process. Numerical results are illustrated by Bode diagrams of the transfer functions of the robust filter, the closed loop estimation errors and graphs of the H_∞–synthesis performance. A numerical comparison with the static observer obtained in the first part of Chapter 5 shows the superiority of the H_∞–observer with respect to disturbance rejection.

Although the necessary model approximation leads to a loss of physical insight, the proposed dynamic observer synthesis procedure is generally applicable to linear, infinite-dimensional systems with point actuation and/or point sensing.

Finally, conclusions and future perspectives are drawn in the last chapter.

Samenvatting

In dit proefschrift zijn diverse procedures uitgewerkt die betrekking hebben op schattingen en voorspellingen van systemen die door convectie, dispersie en reacties worden gekenmerkt. De procedures werken op basis van modellen en puntmetingen. Deze systemen worden oneindig-dimensionaal genoemd, omdat de toestandsvariabelen (bijvoorbeeld temperatuur of concentratie) worden beschreven door partiële differentiaalvergelijkingen of dode-tijd differentievergelijkingen in een oneindig-dimensionale 'manifold' (of ook wel toestandsruimte).

Twee toepassingen van deze convectie-diffusie-reactie (CDR) systemen zijn gebruikt als testvoorbeelden van de voorgestelde schatting- en voorspellingsmethodieken. Een daarvan is een klimaatruimte voor bulkopslag van landbouwproducten. De andere toepassing is een UV-desinfectieproces voor utilisatie in waterbehandeling, de levensmiddelenindustrie of in kasteelt.

Naast de aanwezigheid van stromingen, diffusie en brontermen als reacties of warmteproductie, zijn andere gelijkenissen tussen deze toepassingen:

(a) processen binnen afgesloten ruimtes met vastliggende geometrie;

(b) alleen rand- en/of puntmetingen zijn beschikbaar;

(c) de modelleerdoelstelling is procesbesturing.

De toepassingen zijn geïntroduceerd in hoofdstuk 2. Voor toepassing A wordt het klimaatruimtemodel zoals is voorgesteld in [72], gebruikt als startpunt voor het modelleren van warmtetransport over agrarische producten. Voor toepassing B wordt het UV-desinfectie proces gemodelleerd als een convectie-diffusie proces met eerste-orde biomassadeactivatie kinetiek in de richting van de hoofdstroming. De verkregen modellen worden aangeduid als nominale modellen.

De kern van dit proefschrift is onderverdeeld in vier delen. Parameterschatting, predictie en twee toestandschattingsmethoden voor CDR-systemen worden uitgewerkt. Bij de implementatie van schatting en predictie*, is modelreductie een essentiële stap. In dit proefschrift betekent modelreductie de benadering van een oneindig-dimensionaal systeem naar een eindig-dimensionaal systeem. Hierbij wordt onderscheid gemaakt tussen vroege en late modelreductie. Bij vroege modelreductie wordt de benadering gemaakt alvorens schattings- en predictiemethoden uit te werken. Late modelreductie behelst de benadering van een oneindig-dimensionaal systeem nadat een schatter (waarnemer) of voorspeller is gesynthetiseerd.

De schattings- en predictiemethoden worden hier opeenvolgend beschreven.

*Onder predictie wordt de wiskundige methodiek bedoeld, waarmee voorspellingen worden gegenereerd.

Parameterschatting. Hoofdstuk 3 behandelt de volgende onderzoeksvraag:

> *'Is het, gegeven een vroege reductietechniek, mogelijk om fysische kennis van het nominale CDR-model te behouden in de parameterschatting $\hat{\theta}$, indien een 'output-error'-identificatietechniek wordt gebruikt?'*

Een procedure is voorgesteld om fysische (mogelijk samengestelde) parameters van een benaderd fysisch model te achterhalen uit een geïdentificeerd model en discrete meetdata. Het nominale, fysische warmtetransportmodel voor een bewaarplaats is benaderd met behulp van Padé-approximatie en gecalibreerd via output-error–identificatie. Verder is uit de verzameling van fysische parameters een deelverzameling geselecteerd door analyse van lokale gevoeligheden van de tijdconstante van het benaderde model met betrekking tot de fysische parameters. Eén van de geïdentificeerde parameters, in dit geval de tijdconstante, is daarna gebruikt voor het berekenen van de originele fysische parameterwaarden uit de geselecteerde deelverzameling.

Voor elke ventilatorstand is een versterkingsfactor van ongeveer één gevonden, wat aangeeft dat er nauwelijks onverwachte warmteverliezen of warmteproductie optreedt. Verder is aangetoond dat belangrijke fysische parameters (porositeit van de bulk, bulkproductoppervlak dat blootstaat aan de lucht, bulkhoogte, ...) individueel kunnen worden achterhaald uit de gecalibreerde output-error–modellen.

Parameter schatting en predictie van de systeemuitgang. In hoofdstuk 4 is het schattings- en predictieprobleem benaderd vanuit een ander perspectief:

> *'Is het, gegeven een vroege reductietechniek, mogelijk om het schattings- en predictieprobleem van een CDR-systeem te herschrijven naar een lineaire regressie, en zo ja, hoe?'*

Geïnspireerd door het werk van Doeswijk en Keesman in [30] over niet-lineaire systeemreparametrisaties, is een gereparametriseerd, lineair regressie systeembeschrijving uitgewerkt voor gediscretiseerde, oneindig-dimensionale systemen in toestandsvorm. De discretisatie is tot stand gekomen via vroege modelreductie door eindige (centrale) differentietechnieken uitgeoefend op het CDR-model. De discrete vorm wordt aangeduid als Σ_d en behoort tot de klasse van discrete-tijd lineair gestructureerde systemen.

Er wordt aangetoond dat, onder parameter-affiene condities, een lineair regressieve formulering Σ_d^{LR} van het toestandsmodel Σ_d kan worden gevonden via reparametrisaties van de overdrachtsfunctie van Σ_d. In deze realisatie is een lineair regressieve afbeelding van nieuw gedefinieerde parameters naar systeemingangs-uitgangsdata beschreven en gekarakteriseerd door parameteronafhankelijke coëfficiëntmatrices. Deze coëfficiëntmatrices zijn expliciet

uitgedrukt als functie van het aantal compartimenten n, actuator- en sensorpositie voor twee gediscretiseerde diffusie voorbeelden met specifieke randvoorwaarden. De expliciete resultaten kunnen ook worden toegepast op dimensieloze representaties van CDR-systemen.

De formulering van een lineair regressieve schattings- en predictiemethode maakt het volgende mogelijk: (i) een transparante link tussen het originele fysische model, (ii) unieke parameterschattingen door kleinste-kwadratentechnieken en (iii) een eenvoudige identificeerbaarheidstest op basis van een coëfficiëntmatrix. Verder laten we expliciete resultaten zien met betrekking tot parametergevoeligheden op de responsievector voor diffusie- en specifieke CDR-gevallen.

Een nadeel wordt duidelijk bij de praktische implementatie van lineair regressieve parameterschatting met gewone kleinste-kwadratentechnieken. Omdat de regressoren ingangs- en uitgangsdata bevatten, zijn deze beïnvloed door meetruis. Dit heeft op zijn beurt als gevolg dat de schattingsresultaten verslechteren.

Ontwerp van een statische waarnemer voor toestandschatting. In het eerste deel van hoofdstuk 5, is de volgende vraag gesteld:

> *'Is het, zonder modelreductie toe te passen, mogelijk om ontwerpcondities voor een statische randwaarnemer van een CDR-systeem te ontwerpen, onder de voorwaarde dat het systeem (bij benadering) observeerbaar is?'*

Oneindig-dimensionale systemen–als het CDR-systeem met puntsensoriek en (of) puntactuatie–worden getypeerd door 'onbegrensde' waarnemingsoperatoren en (of) besturingsoperatoren, zie bijvoorbeeld [21, hoofdstuk 1], en veroorzaken wiskundige moeilijkheden. Er wordt verondersteld dat eenvoudige concepten voor observeerbaarheid en detecteerbaarheid van deze specifieke CDR systemen afgeleid kunnen worden door abstracte randsysteembeschrijvingen[†] te gebruiken. Deze abstracte concepten kunnen dan worden uitgewerkt tot een ontwerp van een statische waarnemer.

Hiertoe wordt een UV-desinfectieproces gebruikt als een illustratief voorbeeld voor het ontwerp van een statische, Luenberger-type waarnemer. Doorgaans worden CDR-processen getypeerd door een Sturm-Liouville systeemoperator. Vanwege deze reden worden resultaten verkregen betreffende detecteerbaarheid en observeerbaarheid (bij benadering) voor randsystemen die gekarakteriseerd worden door een Sturm-Liouville operator.

De verkregen resultaten zijn gebruikt in het afleiden van waarnemerontwerpregels voor poolplaatsing bij een UV-desinfectiemodel waarbij alleen data van

[†]Met randsysteem wordt een systeem bedoeld waar de sensoriek en sturing/verstoring op de rand van het systeem is beschreven.

de systeemranden beschikbaar is. Door numerieke simulatie worden de waarnemerkarakteristieken verder geïllustreerd aan het eind van het hoofdstuk.

Ook al zijn ingenieursvaardigheden benodigd om een geschikte versterkingsfactor van de waarnemer te kiezen, de voorgestelde waarnemerontwerpregels behouden het fysisch inzicht en zijn onafhankelijk van een willekeurige benadering van het originele oneindig-dimensionale systeem.

Ontwerp van een dynamische waarnemer voor toestandschatting. In het tweede deel van hoofdstuk 5, wordt de volgende vraag behandeld:

> *'Is het, gegeven een late reductietechniek, mogelijk om een dynamische waarnemer te verkrijgen die robuust is ten opzichte van verstoringen op zowel de randconditie als ook de puntwaarneming, en zo ja, hoe?'*

Met bovenstaande als doel, wordt een methodiek voor robuuste waarnemersynthese uitgewerkt op basis van een herformulering van het (schattings)foutensysteem. Hiervoor worden de volgende stappen in het waarnemerontwerpproces voorgesteld:

(a) Laplace transformatie van het foutensysteem om de gesloten-lus overdrachtsfunctie G te verkrijgen;

(b) formulering van het H_∞–probleem in een goedgesteld, lineaire fractionele transformatie raamwerk;

(c) modale en Padé-benadering en opeenvolgend een gebalanceerde truncatie op het open-lus model M, en

(d) berekening van het optimale H_∞–filter door middel van de benadering van M.

De voorgestelde procedure is toegepast op het UV-desinfectieproces. Numerieke resultaten worden geïllustreerd door Bode-diagrammen van de overdrachtsfuncties van het robuuste filter, de gesloten-lus schattingsfouten en grafieken van de H_∞–synthese prestaties. Een numerieke vergelijking met de statische waarnemer verkregen in het eerste deel van hoofdstuk 5 laat de superioriteit door de H_∞–waarnemer zien met betrekking tot verstoringsonderdrukking.

Alhoewel de benodigde modelbenaderingen leiden tot een verlies van fysisch inzicht, is de voorgestelde procedure voor ontwerp van een dynamische randwaarnemer algemeen toepasbaar op lineaire, oneindig-dimensionale systemen met puntsturing/verstoring en/of puntsensoriek.

In het laatste hoofdstuk worden conclusies getrokken en onderzoeksmogelijkheden voor de toekomst aangedragen.

Curriculum Vitae

D. (Dirk) Vries was born on the 8th of March 1979 in Paramaribo, Surinam. He completed his secondary school in 1997 at Develstein College in Zwijndrecht, The Netherlands. From 1997 he studied Bioprocess Technology at Wageningen University, The Netherlands. His first master's thesis was entitled *Observer-based Growth Rate Control of Bordetella Pertussis* and dealt with observer-based control of the growth rate of whooping cough vaccin cells. His second master's thesis, *On Adaptive Optimal Dynamic Input Design for Parameter Estimation*, describes a recursive algorithm for optimal input design in tandem with parameter estimation. The latter work led to a nomination for the 'Regeltechniekprijs 2003', a Dutch Control Engineering award for M.Sc. students. In 2004, he joined the Systems and Control Group of Wageningen University as a Ph.D. student on the project 'Modeling and Control of Flows'. Results of this work are presented in this thesis.

His research interests lie in system identification, distributed parameter systems and robust filtering and control; with focus on biological systems and applications.

Graduate School Certificate

This dissertation has been completed in fulfillment of the requirements of the *d*utch *i*nstitute of *s*ystems and *c*ontrol (disc) for graduate study.

The following courses have been completed:

- System Identification for Control*
- Control of distributed parameters systems*
- Design Methods for Control Systems*
- System and Control Theory of Nonlinear Systems*
- Summer School on Modeling, Estimation and Optimization of Complex Physical Systems* (2005)
- CFD II†
- Turbulence†

* disc
† J.M. Burgerscentrum

This research was financially supported by the Dutch Technology Foundation STW WWI.6345.

Printed by Universal Press, Veenendaal.

Typesetting by the author, using LaTeX 2ε (LPP license) and Vim (Vim charityware license). Schematic drawings are made with XFig (XFig license). Graphs are produced by using Matlab®, ©1994–2008 The Mathworks Inc.

Front artwork and invitiation by the author, using Inkscape (GNU GPL). Artwork made press-ready with help of Tim Jacobs using Adobe InDesign®.

Photography by the author. Photo visualizes a delta distribution response of syrup entering a flask filled with water.