

Set-membership Estimation, Analysis, and Experimental Design for Biological Systems

Dissertation

zur Erlangung des akademischen Grades
Doktoringenieur (Dr.-Ing.)

von

Steffen Borchers

geboren am 12. Dezember 1980 in Chemnitz

genehmigt durch die Fakultät für Elektrotechnik und Informationstechnik der
Otto-von-Guericke-Universität Magdeburg

Gutachter:

Prof. Dr. Rolf Findeisen

Prof. Dr. Vassily Hatzimanikatis

Prof. Dr. Fabian Theis

eingereicht am 10. Oktober 2012

Promotionskolloquium am 21. Juni 2013

**Contributions in Systems Theory and Automatic Control
Otto-von-Guericke-Universität Magdeburg**

Band 4

Steffen Borchers

**Set-membership Estimation, Analysis, and
Experimental Design for Biological Systems**

Shaker Verlag
Aachen 2013

Bibliographic information published by the Deutsche Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at <http://dnb.d-nb.de>.

Zugl.: Magdeburg, Univ., Diss., 2013

Copyright Shaker Verlag 2013

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the publishers.

Printed in Germany.

ISBN 978-3-8440-2143-1

ISSN 2192-2799

Shaker Verlag GmbH • P.O. BOX 101818 • D-52018 Aachen

Phone: 0049/2407/9596-0 • Telefax: 0049/2407/9596-9

Internet: www.shaker.de • e-mail: info@shaker.de

Acknowledgments

This thesis is based on my work as a research assistant at the Institute for Automation and Control of the Otto von Guericke University Magdeburg from 2008 to 2013. During that time, I was also a PhD student of the International Max Planck Research School for Analysis and Optimization of Chemical and Biochemical Processes Magdeburg. It would not have been possible to write this doctoral thesis without the help and support of the kind people around me, to only some of whom it is possible to give particular mention here.

Above all, I would like to thank Mari-Liis for her personal support and great patience at all times. My parents and brothers have given me their unequivocal support throughout, as always, for which for which I am very grateful.

I want to thank my supervisor Prof. Rolf Findeisen for motivating me to work on very interesting systems biology projects, for supervising my doctoral studies and the support at all levels. I am also very grateful to Prof. Dietrich Flockerzi, Dr. Friedrich von Haeseler, and PhD Sasa V. Rakovic for mentoring me during that time, particularly for all the insightful and encouraging personal and scientific discussions we had. I would like to express my gratitude to Prof. Eric Bullinger, who always inspired me with his enthusiasm and immense knowledge about systems biology.

I am most grateful to all my colleagues at the institute, particularly Philipp Rumschinski for the intense and fruitful cooperation in the first years, Benjamin Kern for all the support with the hardware and for being a friend, and Juan Pablo Zometa for all the nice time we spend together. I am also very grateful to my colleagues from other institutions, particularly Dr. Sandro Bosio, Susann Freund, Dr. Lars Küpfer, Monica Schliemann, and Dr. Stefan Streif. I would also like to express my thanks to Ulrike Thürmer for her kindness, patience and support, and to Dr. Barbara Witter for all the help.

Magdeburg 2013

Steffen Borchers

Contents

Abstract	V
Deutsche Kurzfassung	VII
1. Introduction	1
1.1. Research topics	3
1.2. Challenges	10
1.3. Contribution	10
1.4. Outline	11
2. Modeling and Data of Biosystems	13
2.1. Modeling biological processes	13
2.2. Data and uncertainty description	17
2.2.1. Measurement data	18
2.2.2. A priori data	22
2.2.3. Structural constraints	23
2.3. Summary	23
3. Set-Membership Framework for Falsification and Estimation	25
3.1. The feasibility, estimation, and optimization problems	26
3.2. Reformulation and relaxation	29
3.2.1. Reformulation	30
3.2.2. Semidefinite programming relaxation	31
3.2.3. Dual certificate	33
3.3. Estimation	34
3.3.1. Interval bounding	34
3.3.2. Homothetic bounding	35
3.3.3. Partitioning	37
3.3.4. Branch-and-bound optimization	38
3.4. Computational notes	39
3.4.1. Relaxation hierarchy	39
3.4.2. Decomposition & integration of multiple experiments	40
3.5. Summary	41

4. Model Invalidation and Parameter Estimation	43
4.1. Model invalidation	43
4.2. Parameter estimation	46
4.2.1. Parameter sensitivity	48
4.2.2. Parameter correlation analysis	48
4.2.3. Optimal parameter estimation	49
4.3. Summary and discussion	56
5. State Estimation, Reachability, and Outlier Analysis	59
5.1. State estimation	59
5.2. Reachability analysis	61
5.2.1. Open loop control	63
5.3. Outlier analysis	64
5.3.1. Combinatorial outlier detection approach	67
5.3.2. Reachability-based outlier detection	71
5.4. Summary	71
6. Design of Experiments	73
6.1. Review	73
6.2. Setup	75
6.3. Parameter identifiability	77
6.4. Limits of design	79
6.5. Optimal experimental design	82
6.6. Summary and conclusions	88
7. Application Example: Cell Growth in Batch	91
7.1. Introduction	91
7.2. Data and process description	92
7.3. Identification of growth phases	94
7.4. Phase of exponential cell growth	95
7.5. Phase of decreasing cell growth	98
7.6. Design of experiments	102
7.7. Conclusions	104
8. Conclusions	107
8.1. Discussion and conclusions	107
8.2. Outlook	109
A. Discretization	111
A.1. Forward Euler	111
A.2. Backward Euler	112
A.3. Sampling	113

B. Data Description	114
B.1. A priori data	114
B.2. Measurement data	115
B.3. Structural data	116
B.4. Summary	116
C. Relaxation	117
C.1. Monomial bounding sets	117
C.2. Feasible solution set	117
Bibliography	119

Abstract

Studying biological processes by using dynamical models is a key approach in systems biology. To obtain a reliable model or to improve an existing one, system theoretical methods are required which are suited, first and foremost, for nonlinear dynamical systems and uncertain data.

In this thesis, we develop new methods to address several theoretical and practical issues encountered when modeling biological processes. The proposed methods are based on describing uncertainty of the available data and disturbances by bounded sets, semidefinite programming relaxations, and set-membership estimation techniques to obtain set-valued outer-estimates of the parameters, states, or inputs. Overall, the proposed methods are applicable to polynomial dynamical systems of moderate size, they yield a robust perspective because uncertainties are taken explicitly into account, and they provide guaranteed and conclusive results.

Particularly, we derive a invalidity criterion for purpose of hypothesis falsification, and address the estimation of parameter confidence intervals, sets, and optimal parameter values. Based on this, we study parameter sensitivities, i.e. how variations of the parameters affect consistency of the model. A set-valued observer is developed to reconstruct missing state values.

In addition, we study how uncertainty in initial conditions and parameters propagates to the model outputs using reachability analysis. Besides, reachable sets are considered for identifying outliers in measurement data within a model-generic setting.

Focusing on polynomial systems which are linear in the parameters, we design robust optimal experiments and study the limits of the possible designs so as to obtain as good as possible estimates of the unknown parameters.

The methods are finally put into use for a genuine case, a cell growth process of a human cell line. We demonstrate applicability and utility of the methods for real world applications, as the findings provide new insights into the underlying mechanisms of cell growth and metabolism.

Deutsche Kurzfassung

Die Untersuchung biologischer Prozesse mittels dynamischer Modelle ist kennzeichnend für die Systembiologie. Modellierung und Modellanalyse ermöglichen einen integrativen und quantitativen Zugang zu den oft komplizierten Vorgängen, welche den Prozessen zugrunde liegen, und ergänzen dabei die klassische experimentelle Forschung. Die Findung aussagekräftiger Modelle und deren Analyse bedarf jedoch systemtheoretischer Methoden, welche auf die speziellen Herausforderungen biologischer Prozesse ausgerichtet sind, insbesondere durch Berücksichtigung von Messungenauigkeiten und die Nichtlinearität der betrachteten Systeme. Die Entwicklung entsprechender Methoden ist das Thema dieser Arbeit.

Eine Herausforderung in der Modellbildung von Stoffwechselprozessen, Signaltransduktionsvorgängen oder Genregulationmechanismen, ist nichtlineares Verhalten. Nur durch nichtlineare Modelle lassen sich viele typische Eigenschaften abbilden, so Multistabilität und -stationarität, allerdings wird dadurch unter anderem die Identifikation der Modellparameter und die Modellanalyse erheblich erschwert. Eine weitere grundsätzliche Schwierigkeit ist der Mangel an und die Ungenauigkeit von biologischen Messdaten für einen betrachteten Prozess. Nur selten sind alle Zustände eines Prozesses direkt messbar, fehlende Zustände müssen rekonstruiert werden. Messungenauigkeiten experimenteller Daten sind hier darüberhinaus im Allgemeinen nicht homogen, können sehr gross und möglicherweise fehlerhaft sein, das heisst Ausreisser aufweisen. Im Gegensatz zu rein technischen Prozessen resultiert Messungenauigkeit hier oftmals nicht nur aus der angewandten Messmethode, sondern auch aus inherenter Heterogenität. Aus diesem Grund kann die Ungenauigkeit der Daten nicht einfach außer Acht gelassen werden, da ansonsten wertvolle Information, beispielsweise über die Variabilität der betrachteten Systeme, verloren gehen kann. Im Umkehrschluss ist deswegen auch eine Systemanalyse für biologische Prozesse oftmals ungenügend, die sich allein auf ein nominelles anstatt robustes Systemverhalten stützt. Der Mangel an Prozesswissen führt darüberhinaus oft zu konkurrierenden Modellhypthesen, welche auf der Grundlage vorhandener, ungenauer Messdaten validiert bzw. falsifiziert werden müssen.

Unter Berücksichtigung dieser Schwierigkeiten entwickeln wir in dieser Arbeit neue Methoden für die Modellfalsifizierung, Identifikation der Parameter und Zustände, der Analyse, und der Experimentenplanung. Die Methoden sind speziell entwickelt für polynomielle dynamische Systeme moderater Größe, wobei Ungenauigkeiten der Messungen (Falsifikation, Identifikation, Experimentenplanung) beziehungsweise Un-

sicherheiten der Parameter (Analyse) unmittelbar berücksichtigt werden. Dadurch liefern die Methoden nicht nur nominell gültige Resultate, sondern auch eine robuste Perspektive auf die vielfältigen hier untersuchten Fragestellungen. Die Anwendbarkeit und Nutzen der hier entwickelten Methoden veranschaulichen wir an einem realen Beispiel aus der Bioprozesstechnik.

Die vorgeschlagenen Methoden sind mengenbasiert, das heisst Ungenauigkeiten der verfügbaren Daten und Störungen werden generell mittels kompakter Mengen beschrieben. Neben experimentellen Daten können auch a priori Wissen sowie qualitative Information berücksichtigt werden. In Kombination mit einer semidefiniten Relaxation und effizienten Eingrenzungsalgorithmen ermöglicht es dieser Ansatz, die Existenz gültiger Lösungen zu überprüfen, Optimallösungen zu bestimmen, und die Menge aller gültigen Parameter, Zustände, oder Eingänge je nach Fragestellung abzuschätzen. Bei der Ausgestaltung des hier verfolgten Ansatzes achten wir auf dessen praktische Anwendbarkeit, welche vor allem durch den erforderlichen Rechenaufwand begrenzt ist. Dazu schlagen wir mehrere Strategien vor, um eine dem Problem angemessene Balance zwischen der Präzision der Resultate einerseits und dem numerischen Aufwand andererseits zu gewährleisten, zum Beispiel durch die Wahl der Relaxation, des Eingrenzungsverfahrens, oder durch Anwendung eines Sukzessionsalgorithmus.

Nach der Beschreibung der betrachteten Modelle und der Daten (Kapitel 2) sowie der Ausarbeitung des mengenbasierten Ansatzes (Kapitel 3) betrachten wir die folgenden konkreten Fragestellungen.

In Kapitel 4 betrachten wir die Modellfalsifizierung und das Parameterschätzproblem. Dazu leiten wir ein hinreichendes Invalierungskriteriums her, und zeigen die Schätzung der Konfidenzintervalle, der optimalen sowie der Menge aller gültigen Parameter auf. Die Parametersensitivitäten werden ebenfalls untersucht.

In Kapitel 5 leiten wir einen mengenwertigen Zustandsbeobachter her, betrachten das Erreichbarkeitsproblem, und schlagen Modell-basierte Ansätze zur Detektion von Ausreißern vor. Die Ergebnisse ermöglichen zudem eine Unsicherheitsanalyse.

In Kapitel 6 betrachten wir die Planung von Experimenten zur optimalen Parameterschätzung. Dazu zeigen wir fundamentale Grenzen des Möglichen auf, leiten ein notwendig und hinreichendes Kriterium zur Identifizierbarkeit her, und ermitteln optimale Experimente unter Einbezug von Messungenauigkeiten.

In Kapitel 7 wenden wir schließlich die hier entwickelten Methoden der Hypothesenfalsifizierung, Schätzung, Analyse, und der Experimentenplanung auf ein reales Beispiel an. Wir untersuchen das Zellwachstum und den Stoffwechsel einer menschlichen Zelllinie. Die Ergebnisse liefern neue Einsichten in die Stoffwechselfvorgänge und Wachstumslimitierung dieser Zellen.

1. Introduction

The question of origin, evolution, and functioning of life has always fascinated mankind and influenced the way we think about and perceive ourselves. In 1665, Robert Hooke described the structure of cork, by using a microscope, as being composed of what he called cells. This discovery led to one of the fundamental axioms of modern biology, namely, that cells are the basic units of life. While the minute structures within these cells are generally colorless and transparent, it was the development of cell staining techniques, which made possible the impressive advances in our understanding of cells and organisms. Among the recent developments in this field of study, the combination of fluorescence and transfection techniques - especially green fluorescent protein - revolutionized the study of cells, i.e. to analyze, track, and quantify biological molecules *in vivo*. With the automation of these techniques advancing towards high throughput ones, it became possible to gather large amounts of data about genes, proteins and the interactions of both - pivotal for the understanding of the plenitude of cellular functions and their regulation.

Yet, many biological and in particular intra-cellular processes are often difficult to study by experimentation alone. They are influenced and regulated by a variety of factors, which can not be kept constant under experimental conditions. Besides, certain signals or components of a studied system, such as cytokines or transcription factors, may trigger or regulate several cellular functions simultaneously. Therefore, the study of cellular functions, their mechanisms and interplay requires, besides experimentation, a complementary methodological framework: an abstracted, integrative description in terms of a mathematical model. With simplicity and conciseness (law of parsimony) as guiding principles, modeling promises to structure and integrate the available knowledge and data about the process. Besides it provides an insight into the most important features and the relevant mechanisms, and possibly designs therapies or control strategies which influence the process under study in the desired way.

Nowadays, numerous modeling approaches are available for the variety of cellular functions and systems. Each approach hereby reflects upon the prior knowledge available about the process as well as the modeling purpose. Cellular decisions such as the chemotaxis of bacteria for example, can be described by stochastic models in many cases, whereas metabolic pathways, cell growth, and signal transduction are frequently being described by ordinary deterministic differential equations. More complex gene regulatory processes are often represented by Boolean or Petri networks. In any case, modeling has to be understood as an objective-oriented and iterative approach, de-

picted in Fig. 1.1. Initially it consists in formulating - potentially competing - model hypotheses about the inner mechanisms and interactions regarding the components of the considered process. In a next step, these hypotheses have to be evaluated by taking into account the available (experimental) data, which might also imply to reject hypotheses that do not represent the observed behavior. Thus, a preliminary model is obtained, which uses the experimental data to estimate unknown parameters. This model can then be analyzed, e.g. to verify qualitative features of the model. Accordingly, new experiments have to be designed which reveal the most important information about the ongoing process, which particularly consists in rejecting the remaining hypotheses and to refine or learn more about the unknown parameters.

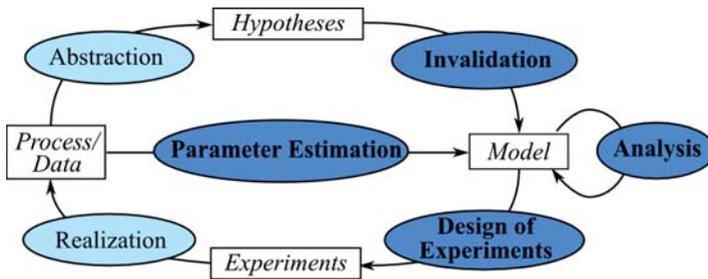


Figure 1.1.: Iterative modeling scheme and the major topics addressed in this thesis.

To complete the thus structured modeling process, an appropriate and consistent methodological framework is required to address the systems theoretical issues and problems, namely model falsification, estimation, analysis, and the design of the individual experiments. Such a framework, in turn, needs to be accurately adjusted to the model class in question and the particular challenges of biological systems. When focusing on the biological processes, described by ordinary differential equations (ODE), the most challenging characteristics are nonlinearities, noise, and the lack of information. Unlike technical systems, measurement uncertainty can be very large, and this is due not only to measurement imperfection, but also to the inherent heterogeneity of biological systems. Therefore, from a systems engineering perspective, errors and uncertainties cannot be neglected because valuable information would be ignored, nor can they be considered homogeneously distributed mostly. These facts motivate us to reconsider the description of uncertainties and to develop coherent and conclusive methods based on previous considerations within a membership setting in order to address these challenges.

1.1. Research topics

We further review and discuss the considered topics.

Hypotheses testing

Modeling within system biology is in often complicated because of limited information about the inner mechanisms of a considered process. For instance, when modeling biochemical reaction networks, it remains unclear how and if at all components interact, and the reaction kinetics is often unknown. This typically amounts in a number of rivaling hypotheses with unknown parameters. Modeling hence often necessitates, prior to estimation and analysis, a selection criterion as well as testing methods to distinguish concurring hypotheses.

In terms of existing concepts, validation or falsification approaches are considered for this purpose. The former consists in selecting the ‘most appropriate’ hypothesis, the latter in rejecting those hypotheses which do not meet certain criteria. One general observation is that models with a larger number of parameters are more flexible and fit the data better than models with a smaller number of parameters [Jaqaman and Danuser, 2006]. An example for validation approaches is the *Bayesian information criterion*. Here, a score - based on *maximum likelihood* and the number of parameters - is assigned to each hypothesis. Among these hypotheses the one with the minimum score is the most suitable one, see e.g. Jaqaman and Danuser [2006]. One problem, however, is that validation criteria are inherently subjective, see e.g. Oreskes and Belitz [2001], Smith and Doyle [1992]. Furthermore, as pointed out by Anderson and Papachristodoulou [2009], it is impossible to validate a hypothesis, since this would require an infinite number of experiments. There are various stochastic approaches for falsification. One example is the *Neyman-Pearson Lemma*. It is based on minimizing the probability of false acceptance [Lehmann, 2009], and for this purpose the maximum likelihood ratio can be considered. In addition, the *F-test* can be used to clarify if the introduction of extra parameters, and hence a more complicated hypothesis, is justified, see e.g. Jaqaman and Danuser [2006]. An overview over hypothesis testing approaches for systems biology can for example be found in Kremling et al. [2004].

Alternatives to stochastic approaches have been considered e.g. in Anderson and Papachristodoulou [2009], Prajna [2006], using *barrier certificates* for model invalidation. These barrier certificates, namely functions of state, parameter, and time, allow to separate possible model trajectories from measurement data. Locating them allows invalidating a model conclusively; however, deriving these certificates is non-trivial and often practically impossible. Another certificate-based approach has been considered for polynomial stationary systems in Kuepfer et al. [2007]. It consists in deriving *infeasibility certificates* by employing a suitable *semidefinite relaxation* [Parrilo, 2003].

The approach has been extended recently to dynamical systems in Borchers et al. [2009a,b,c], Hasenauer et al. [2010b], Rumschinski et al. [2010a].

Parameter and state estimation

Mathematical models allow us, by numerical simulation, to make, for instance, predictions about the outcome of an experiment. Given the available efficient ODE solvers, such simulations can be easily obtained if the initial conditions and the model parameters are known and defined. To obtain those parameters which characterize the process however, observations have to be converted into information, which defines the estimation problem, also known as the *inverse problem* [Tarantola, 2005].

Inverse problems are in general difficult to solve because they typically admit multiple solutions and are often *ill-posed*. Following the definition by J. Hadamard [Tikhonov et al., 1977], a *well-posed* estimation problem should include the following properties: an existing solution which is unique and depends ‘smoothly’ on the data. Evidently, these criteria are not met when considering nonlinear systems, particularly with regard to sparse and uncertain biological measurements; as measurement is not only an observed value, but an acquired *state of information* [Tarantola, 2005]. The information about the (un-)certainty of a measurement is crucial to evaluate the influence of the uncertainties as well as the precision of the estimated parameters (or states).

The classical notion of inverse problems is based on describing the uncertainty (i.e. the state of information) of the available data by *probability density distributions*, i.e. a stochastic setting. One particular approach to the inverse problem is to infer the optimal variables (parameter) from the available information with respect to a certain optimality criterion (e.g. Gauss ‘the most advantageous values’, Laplace ‘the most plausible values’ [Lehmann, 2009]). The most widely used criterion is the ordinary least squares, it leads to a regression problem. If the uncertainties are normally distributed, the ordinary least squares method is equivalent to the maximum likelihood, see e.g. Shao [2003]. For *generalized linear models* with normally distributed measurement errors, it can be shown that the maximum likelihood is well-defined (see e.g. Ljung [1998]).

Due to conceptual simplicity, the ordinary least squares approach is being applied the most frequently to estimate the optimal parameters of dynamical systems. Here, estimation consists in solving a constrained optimization problem. If the model is linear time-invariant and if the uncertainties are (unbiased and) normally distributed, see e.g. [Ljung, 1998], the estimation problem is well-posed. To obtain the optimal parameter values, gradient methods, e.g. the *direction of steepest ascent*, can be employed, see e.g. Boyd and Vandenberghe [2004], Marquardt [1963]. If, however, the model is nonlinear (in parameters or states), or if the data is not normally distributed, the optimization problem may admit multiple and non-unique optima. The problems can be partially

addressed considering stochastic global optimization approaches [Mendes and Kell, 1998, Moles et al., 2003, Robert and Casella, 2004, Rodriguez-Fernandez et al., 2006], such as *evolutionary algorithms* [Kikuchi et al., 2003], *multiple-shooting* approaches [Balsa-Canto et al., 2008, Peifer and Timmer, 2007], *clustering* approaches [Rinnooy-Kan and Timmer, 1987], or *simulated annealing* methods [Kirkpatrick et al., 1983].

The more general interpretation of the inverse problem consists in transforming the a priori into a posteriori *probability density distribution* [Tarantola, 2005]. This allows for the accuracy of the parameters to be determined, e.g. estimating the *confidence intervals* or limits, see e.g. Faller et al. [2003], Swameye et al. [2003], Toni et al. [2009]. To this end, several approaches based on the *Fisher information matrix* have been proposed to obtain the (symmetric) lower bound of a parameter’s variance, see e.g. Banga et al. [2002], Ljung [1998]. However, as noted by Kremling et al. [2004], this lower bound would be reached only if the model equations were linear within the parameters, which is seldom the case with biological systems. To estimate the precision of the parameters, besides some simple cases, a resolution in terms of samples of the probability density distribution [Tarantola, 2005] is considered. To this end, re-sampling techniques, e.g. *bootstrap*, *jackknife*, and *Monte-Carlo testing* [Efron, 1979, Joshi et al., 2006, Robert and Casella, 2004] have been elaborated. However, important solutions may be missed because these methods due to the stochastic setting. This is also why these methods typically require a large number of samples, growing exponentially with the number of unknown parameters and initial conditions, to obtain a certain confidence in the results.

Besides the stochastic setting, a *set-membership* approach can be considered to address inverse problems. This approach is based on unknown, but bounded uncertainties and disturbances, i.e. bounded error description. Then, the inverse problem consists in deriving, or approximating, the a posteriori solution set, e.g. the set of parameters or states which are consistent with the bounded uncertainties. In the following, we provide an overview, not intended to be complete, of methods which have been established within the bounded error setting.

The set-membership approach for linear systems has been pioneered in Schweppe [1968], Witsenhausen [1968], where a set-valued outer-approximate is obtained by linear programming (see also Bai et al. [1999], Milanese and Vicino [1991] and the references therein). To this end, ellipsoids [Fogel and Huang, 1982, Schweppe, 1968, 1973], orthotopes [Milanese and Belforte, 1982], and zonotopes [Mo and Norton, 1990, Walter and Piet-Lahanier, 1989] have been considered so far. For nonlinear systems, these algorithms are typically not directly applicable. To manage this problem, *interval analysis* and *constraint propagation* methods for parameter estimation have been developed, see e.g. Jaulin et al. [2001] and references therein. Set-membership approaches for nonlinear *system identification* have been established, e.g. for Wiener [Cerone and Regruto, 2006] and Hammerstein [Cerone and Regruto, 2003] models, and Aubin et al. [2002], Bemporad et al. [2005], Figueroa et al. [2008], Milanese and No-

vara [2004], Sznaier [2009]. A barrier-certificate approach has been considered [Prajna, 2006] to discard parameter sets. An alternative certificate-based approach for parameter estimation of stationary systems has been presented in Kuepfer et al. [2007]. The approach consists in deriving infeasibility-certificates by employing a suitable semi-definite relaxation [Parrilo, 2003]. Using a bi-sectioning algorithm, the set of feasible parameters can be outer-approximated. Such a relaxation approach has, for example, been extended to steady state sensitivity analysis [Waldherr et al., 2008], steady state analysis of (bio-)chemical processes [Hasenauer et al., 2010a], and linear dynamic systems [Cerone et al., 2010]. However, considering only stationary systems (as in Hasenauer et al. [2010a], Kuepfer et al. [2007], Waldherr et al. [2008]) is in general not sufficient to analyse dynamical systems and estimate their parameters, see e.g. Bullinger et al. [2008], Farina et al. [2006]. In this thesis, we extend this approach to dynamical systems, following Borchers et al. [2009a,b,c], Hasenauer et al. [2010b], Rumschinski et al. [2010a].

State estimation for dynamical systems has been addressed in a stochastic setup, where the process and measurement noises are assumed to be Gaussian [Raïssi et al., 2010]. *Kalman filtering* and *particle filtering* are two representative methods, see Lendek et al. [2006]. Under the assumption of linear state space systems, and that both state transition and the measurement noise is unbiased and normally distributed, the Kalman filtering is shown to be an optimal estimator in the mean least squares sense [Lendek et al., 2006]. The Kalman filtering has been extended to the nonlinear case with non-Gaussian noise, e.g. the extended Kalman filtering (see e.g. Anderson and Moore [1979]), or the unscented Kalman filtering, see e.g. Wan and Van Der Merwe [2000]. However, particularly if the parameters of the system are not exactly known, or the measurement uncertainties are non-homogeneously distributed, the Kalman filtering might fail. To overcome these problems, particle filtering, also known as sequential Monte Carlo methods, can be considered. A survey about particle filters is given e.g. in Crisan and Doucet [2002]. These methods use many random samples (i.e. particles) to represent the posterior probability distributions, which are then propagated over time [Crisan and Doucet, 2002]. Although this approach is generally applicable to nonlinear systems and non-homogeneous uncertainties, there exists no general rule how to choose ‘representative’ samples, and the required number of samples grow exponentially with the number of variables.

As for state estimation in membership settings, three main approaches have been considered (compare the review given in Raïssi et al. [2010]): A prediction/correction mechanism as e.g. proposed in Jaulin [2002], Kieffer and Walter [2004], although the approach is limited, due to the *wrapping effect*, to small (measurement and parametric) uncertainties [Raïssi et al., 2012]. The second approach (e.g. Kieffer and Walter [2006]) is again a prediction/correction approach which employs however the Müller Theorem [Müller, 1927] for the prediction step, and third closed loop *interval observers*, e.g. Bernard and Gouzé [2004], Gouzé et al. [2000], Mazenc and Bernard [2011], Moisan

et al. [2009]. Interval observers have in parts been applied to biological processes, e.g. uncertain bioreactors as in Moisan et al. [2009].

Uncertainty and outlier analysis

This analysis deals with investigating the uncertainty in the model output that is generated from the one in initial conditions or parameters [Marino et al., 2008]. Such investigation is motivated because measurements for biological processes are in general uncertain and lead to uncertain parameter estimates. Moreover, one feature of those processes is their robustness regarding perturbations [Eißing et al., 2005, Kitano, 2004, Stelling et al., 2004], which means that they show a particular behavior even under uncertain conditions and disturbances. Hence, a corresponding mathematical model should also describe such qualitative property [Morohashi et al., 2002], which in turn requires an uncertainty analysis.

To investigate the influence of uncertainties, *sensitivity analysis* is a frequently considered tool to evaluate the individual contribution of variations of certain variables onto the behavior of the system. Sensitivity analysis approaches are classified depending on whether small (local sensitivity analysis) or large perturbations are investigated (global sensitivity analysis). In addition, the influence of uncertainties can be evaluated regarding the systems dynamics (dynamical sensitivity analysis, see e.g. Saltelli et al. [2008], or with respect to steady states, see e.g. Mönnigmann et al. [2007], Waldherr et al. [2008]), which is also relevant for robust model synthesis as e.g. in Marquardt and Mönnigmann [2005]. A frequently used tool to study robustness is hereby bifurcation analysis, particularly to investigate the influence of uncertainties regarding a qualitative change in system behavior, as for example oscillation, multi-stability or even chaotic behavior, see for example Bagheri et al. [2007], Bates and Cosentino [2011], Conradi et al. [2007], Waldherr et al. [2011].

Outlier analysis, in turn, deals with the problem of detecting - and if appropriate - removing anomalous observations. It is the primary step towards obtaining estimates and coherent analysis [Ben-Gal, 2005] because outliers may contain valuable information or lead to falsely rejecting hypotheses or biased parameter estimates. Therefore it is crucial to identify outliers prior to modeling and analysis (refer to Liu et al. [2004], Williams et al. [2002]).

Outliers mainly arise due to faults, such as changes in system behavior, fraudulent behavior, human error, instrument error or simply through natural deviations in populations [Hodge and Austin, 2004]. There are several approaches to detect those outliers, each involving the challenge of defining the outlier, which again depends on the method used for detection. Existing outlier detection methods can be classified according to whether or not an (error) model is utilized, i.e. parametric (supervised or model-based) and non-parametric (non-supervised) respectively, see e.g. Hodge and Austin [2004] and Ben-Gal [2005] for a comprehensive survey. Non-parametric meth-

ods typically deal with large data sets and independent data. Model-based outlier detection methods are frequently used to detect outliers in time-series data (dependent data), as yet though only within a stochastic context. They can be further classified into model-specific or model generic approaches [Ben-Gal et al., 2003]. Examples of model-specific approaches for dependent data are *cumulative sum filters* or *moving average filters*. They typically require knowledge about the distribution of the data. In contrast, model-generic approaches are based on estimating the underlying model, i.e. they are parameter-dependent.

A general observation when considering data sets with multiple outliers is, that, following Ben-Gal [2005], they are possibly subject to *masking* and *swamping* effects. An intuitive, though not mathematically rigorous, understanding of these phenomena is given by Acuna and Rodriguez [2004]: An outlier masks another one, if the second is an outlier by itself, though not in the presence of the first outlier. Thus, only after the deletion of the first one, the second outlier emerges as such. Vice versa, an outlier is said to swamp a second observation, if the latter is an outlier only due to the first one. In other words, after the deletion of the first outlier the second observation becomes a non-outlying observation. Hence, masking and swamping effects may complicate the detection of outliers and have to be employed cautiously, e.g. to avoid neglecting important information.

Design of experiments

The design of experiments is an important link between modeling and experimentation. It addresses a priori how to perturb the process under study, as well as which states have to be observed to identify and reveal the most important features of a system, e.g. to invalidate a particular hypothesis or to learn the unknown parameters. To obtain best possible estimates concerning the parameters of a mathematical model, experiments have to be performed and measurements have to be taken. Experiments however are generally (resource and time) expensive, and poorly planned ones might only provide little information. It is therefore important to properly investigate the conditions required to identify parameters at all, and to design experiments with a maximum information gain, while explicitly taking uncertainties into consideration.

Fisher [1935] initiated the study of a priori experimental design, with the idea of “deciding what *patterns of factors combination*¹ will best reveal the properties of the systems response, and how this response is influenced by the factors” [Franceschini and Macchietto, 2008]. He focused on obtaining the most important information to reveal an input-output relationship in the presence of variations, in general of a stochastic nature, which is known nowadays as *black box* experimental design (see e.g. Franceschini and Macchietto [2008] for applications and a more comprehensive review). The black box approaches however are inappropriate for dynamical systems

¹i.e. inputs and initial conditions

with constrained outputs, as they do not take into account the available (or at least partially available) information about the system’s structure.

Therefore, experimental design approaches had to be extended to explicitly enclose knowledge of the considered system’s structure [Franceschini and Macchietto, 2008]. This is termed model-based experimental design. Early approaches admitted only steady state systems, including linear and nonlinear models, e.g. to study reaction kinetics as in Box and Lucas [1959]. The extension to dynamical systems has been slow [Shirt et al., 1994] and primarily been considered in a stochastic context. For example, one objective for experimental design, which has been considered, is to minimize either the variance (uncertainty of the estimates [Shirt et al., 1994]) or the bias [Ljung, 1998] of the transfer function. Various other optimality criteria have been pursued, all based on the *Fisher information matrix* (FIM). Here, the parameter uncertainty can be appropriately distinguished within the FIM “due to the asymptotic normality of parameter estimators and the *Cramer-Rao bound*” [Pronzato, 2008]. Optimality criteria are found by minimizing the (expected) variance of the unknown parameters, e.g. functionals of the invariants of the FIM. In this context, a widely used criterion is the *D-optimal* design which aims at maximizing the determinant of the FIM and thereby minimising the parameter variances. Alternatively, designs such as A-optimality, E-optimality, etc. have been considered (see e.g. [Boyd and Vandenberghe, 2004, p. 384–392] for a compact overview).

However, the proposed frequency domain and Fisher-information matrix based approaches all rely on the true system parameters, or at least on an appropriate and accurate a priori guess of the nominal system parameters. Hence, the quality of experiments designed using these standard techniques can be adversely affected by poor starting values of the parameters [Asprey and Macchietto, 2002]; such information however is in many cases simply not available. Thus, design methods that are insensitive to these starting values are required [Asprey and Macchietto, 2002]. This issue has been recognized in literature, and some approaches to the so called *robust experimental design* have been taken into account, as for example the *sequential design* (e.g. Walter and Pronzato [1997], Wynn [1970]), *Bayesian approaches* (see e.g. Chaloner and Verdinelli [1995] for a review), or *minimax design* (see e.g. Rojas et al. [2007] for an overview and references). However, apart from standard cases (linear systems, *white noise*), there has been little study on robust experimental design for engineering problems [Rojas et al., 2007], see also the survey presented in Hjalmarsson [2005]. In membership setting only few approaches have been made so far towards the robust design of experiments. Norton [1987] proposed a number of general guidelines, and Belforte et al. [1987] described an orthotopic approximation approach. Pronzato and Walter [1990] considered to use experimental design for linear regression models, by choosing as design policy a volume criterion which compares to the classic D-optimal design in the stochastic setting. Novara [2007] considered experimental design for nonlinear system identification, and recently, Marvel and Williams [2012]

set-membership experimental design has been considered in the context of biological systems.

With this overview of existing approaches for systems engineering, we outline next the particular research challenges when considering biological systems.

1.2. Challenges

Estimation, analysis, and design of experiments are well known issues in control engineering and it seems easy to apply available methods for technical systems to biological ones. However, their theoretical and practical applicability is severely limited for the following reasons: Non-linearity, the scope of existing methods, and the available data's characteristics. In the first place, biological processes are in general nonlinear dynamical systems within parameters, states, and stimuli. Secondly, available methods almost exclusively focus on obtaining optimal parameter/state values. However, the main features of biological systems are robustness and heterogeneity, therefore focusing on the nominal system's behavior alone is critical. Yet, one leading assumption of many available (stochastic) methods is normally distributed noise. This assumption is required, implicitly or explicitly, to assure unbiased results and convergence of methods, but is however often not given. Thirdly, a primary challenge considering biological processes is that measurement data of is typically uncertain, sparse, and possibly erroneous, i.e. uncertainties can be very large, non-homogeneous, and the data may be corrupted by outliers. This uncertainty extends to initial conditions and the model parameters, which are often completely unknown. Furthermore, lack of information about the underlying mechanisms or reactions frequently results in concurring model hypotheses.

Keeping these requirements in mind, the following contributions are made in this thesis.

1.3. Contribution

In this thesis, we elaborate a set-membership framework for falsification, estimation, analysis, and design of experiments for polynomial dynamic systems. The methods are in particular suited for modeling and analyzing biological processes, because uncertainties are taken explicitly into account using a bounded error description. Furthermore, we can integrate disturbances, a priori and qualitative data, which is of particular relevance considering that available data for biological processes is typically sparse. The presented methods provide a robust perspective yielding guaranteed results of in principle desired precision, and are build on convex optimization and set-membership approximation techniques. Convex optimization is a well established method and efficient solvers for convex optimization problems are available. The setting of the

proposed framework is flexible, and can thus be applied to a variety of processes and research questions, as for example to identify the main influencing factors of cell growth.

In addition to a falsification approach for hypothesis testing, we address the estimation of parameter confidence intervals, sets, and optimal parameter values. An interval observer is developed to reconstruct missing state values. Furthermore, we elaborate several analysis methods. Reachability analysis, for example, allows to study the evolution of the system under parametric uncertainties, as well as to analyze the influence of those uncertainties on the system dynamics. As shown, such information is very useful to direct model changes and allows us to detect outliers within a model-generic setting. Focusing on systems that are linear in the parameters, we propose an experimental design approach to obtain a minimum volume parameter set in worst case. The predominant idea of this approach is to consider a number of independent one-step experiments. We furthermore show that the necessary number of experiments (and observations) is equivalent to the number of unknown parameters, and also derive a sufficient criterion. Complementary, we investigate fundamental limits of experimental design. The methods are illustrated by several examples.

Finally, the proposed methods for falsification, estimation, analysis, and design of experiments are applied to a comprehensive real world application, a cell growth process of a novel human cell line. Particularly, we identify qualitatively different growth phases, and determine the main influencing factors of cell growth and metabolism for the available experiments.

1.4. Outline

This work is structured as follows:

Chapter 2 introduces the problem of mathematical modeling within biological processes as well as it outlines the bounded error description of commonly available data. Particularly, we describe a priori knowledge, empirical or measurement data, and structural constraints. This description is considered throughout the thesis.

Chapter 3 focuses on the deduction of the set-membership framework intended for estimation and analysis. First, the relevant premises and concepts are introduced and the problems of estimation and analysis within the bounded-error notion are formulated. In a second step, we describe the convexifying relaxation technique and provide infeasibility and dual certificates. These certificates form the basis of set-membership estimation techniques and branch-and-bound optimization and will be briefly outlined. We furthermore present complexity reduction techniques to implement the trade-off between the accuracy of estimation results and computational complexity. The main results have been presented in parts in Borchers et al. [2009a,b,c, 2012], Rumschinski et al. [2010a].

In **Chapter 4**, the framework obtained in Chapter 3 is applied to model invalidation and parameter estimation. We derive a coherent invalidity criterion and address the estimation of parameter uncertainty intervals and sets. Furthermore, we elaborate the estimation of optimal parameter values. The main results have been presented in parts in Borchers et al. [2009c], Rumschinski et al. [2010a].

In **Chapter 5**, the same framework is applied to analyze model and data. More precisely, we apply robust reachability analysis to examine uncertainties and propose a model-dependent outlier detection approach. The results have been presented in parts in Borchers et al. [2012, 2013].

Chapter 6 deals with the design of experiments in terms of parameter estimation. We derive necessary as well as sufficient conditions for experiments and observations so that the parameters can be clearly identified. We then turn to the design of optimal experiments, considering a volume and worst-case setting under additive disturbances. The results presented here are based on the works Borchers and Findeisen [2011] and Borchers et al. [2011b].

Chapter 7 presents a genuine application of the proposed framework, the cell growth and basic metabolism of a human cell line. The example demonstrates the applicability of the proposed method besides providing new insights into the underlying mechanisms of cell growth. This chapter is based on the work presented in Borchers et al. [2013].

Chapter 8 closes the work with a brief discussion and outlook.

2. Modeling and Data of Biosystems

The diversity of biological processes gives rise to a variety of modeling approaches suited to capture the essential features of a particular system. The scope of possible models ranges from stochastic and deterministic descriptions with concentrated or distributed parameters, toward cell ensemble models, Boolean networks, Markov chains, etc.. In this work, we focus on biological processes which can be described by ordinary differential equations. This class has drawn much attention in recent years, among other reasons because they are accessible to many system theoretical tools, and they cover e.g. biochemical reaction networks, and thus include metabolic and many important signal transduction systems.

A pivotal step for testing and estimation of such models is an appropriate description of the available information, i.e. all qualitative and quantitative knowledge about the process. This includes classical time-course measurements and its uncertainty, as well as a priori knowledge, e.g. about state restrictions or the initial parameters, and qualitative data, e.g. certain compounds admit a monotonic or oscillatory dynamics. For testing and estimation, it is of course advantageous to explicitly include as much of such information as possible.

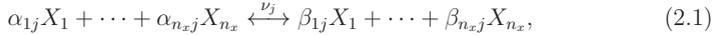
Next, we provide a background of modeling of such processes, and describe common data and its characteristics. Section 2.1 briefly introduces the modeling of deterministic biological processes, leading to ordinary differential equation systems and algebraic constraints. Section 2.2 asserts the most frequent available data and provides a suitable description of uncertainties.

2.1. Modeling biological processes

A variety of biological processes such as metabolic activities, signal transduction and gene regulation processes can be approximated by a common modeling framework, i.e. a reaction network. It is build on two main components, the compounds (species, states), i.e. quantifiable entities X_1, \dots, X_{n_x} , e.g. proteins, metabolites, RNA etc., and the interactions (reactions) among the compounds. The compounds are typically considered in terms of concentrations regarding a specified compartment (volume), e.g. $x_i = \frac{X_i}{V_i}$, $i \in [1 : n_x]$, where V_i denotes the respective volume of the compartment. We denote the collection of compounds by $x = \{x_1, \dots, x_{n_x}\}^T \in \mathbb{R}^{n_x}$. It is important to note that if these compounds uniformly distribute within the compartment by diffusion, and if the number of the compounds is large or the reactions are sufficiently fast,

spatial and stochastic effects can be neglected. In the following, we will neglect spacial effects, and exclusively consider deterministic process models.

The interactions generally describe the transformation of some specie into others; the most common forms of interactions are (biochemical) reactions, transport, and degradation processes. An interaction can be denoted in the form:



where α_{ij} and β_{ij} , $i \in [1 : n_x]$, define the stoichiometric relations of the j -th reaction, and ν_j denotes the reaction rate.

The reaction rates, and thereby the reaction kinetics, can be modeled in various ways, see e.g. Cornish-Bowden [2004], Klipp et al. [2005] for comprehensive overview. A frequently considered approach, in particular for elementary biochemical reactions, is the law of mass action, which derives from first principles, and where the reaction rates are proportional to the substrate concentrations, e.g.

$$\nu_j(t) \doteq p_j^+ \prod_{i=1}^{n_x} x_i^{\alpha_{ij}}(t) - p_j^- \prod_{i=1}^{n_x} x_i^{\beta_{ij}}(t). \quad (2.2)$$

Here, p_j^+ and p_j^- define the forward and backward reaction constants, and α_{ij}, β_{ij} , $i \in [1 : n_x]$, $j \in [1 : n_\eta]$ define the stoichiometric factors of the j -th reaction.

In addition, there exist various phenomenological kinetics to describe the effect of enzymes, e.g. limitation, saturation, or inhibition, which typically derive from limiting cases of mass action kintetics. Examples are the Monod, Hill, and Michaelis-Menten kinetics. Exemplary, the saturation effect of an enzyme can be described by the Monod equation (see e.g. Zeng and Deckwer [1995]), given by

$$\nu_j(t) = \nu_{j,max} \frac{x_i(t)}{K_j + x_i(t)}, \quad (2.3)$$

where $\nu_{j,max}$ denotes the maximum reaction rate, and K_j denotes the Monod constant.

Balancing the compounds considering the interactions, the process' dynamics can be described by

$$\dot{x}(t) = S\nu(t), \quad (2.4)$$

where $\nu(t) = (\nu_1(t), \dots, \nu_{n_\nu}(t))^T \in \mathbb{R}^{n_\nu}$ denotes the vector collecting all reactions of the considered network, and $S \in \mathbb{R}^{n_\nu \times n_x}$ denotes the stoichiometric matrix (see e.g. Horn and Jackson [1972]), constructed from the factors α_{ij} and β_{ij} with

$$S_{ij} = \beta_{ij} - \alpha_{ij}, \quad i \in [1 : n_x], \quad j \in [1 : n_\nu]. \quad (2.5)$$

We furthermore can include inputs $u(t) \in \mathbb{R}^{n_u}$ to the system description, to model external stimuli or changing environmental conditions of the process. Moreover, we

can include systemic disturbances, i.e. process noise, denoted by $w(t) \in \mathbb{R}^{n_w}$, to the system description.

Such (systemic) disturbances are typically unknown and not constant during a process, and can be used to approximate environmental influences (e.g. temperature, pH, etc.) besides modelling errors.

The overall dynamics can be summarized by the system of ordinary differential equations

$$\dot{x}_i(t) = f_i(x(t), p, u(t), w(t)), \quad i \in [1 : n_x], \quad (2.6)$$

where $p \in \mathbb{R}^{n_p}$ collects all rate constants (i.e. $p_j^+, p_j^-, \nu_{j,max}, K_j$) for all $j \in [1 : n_\nu]$. In the remainder, we consider the functions $f_i(\cdot) : \mathbb{R}^{n_x} \times \mathbb{R}^{n_p} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_w} \rightarrow \mathbb{R}$ to be rational, which is satisfied for above mentioned kinetics.

We are aiming to verify the validity of a given model based on physical measurements or to estimate unknown parameters. These objective are challenging, since often not all compounds can be measured, or only aggregated information might be available. We therefore distinguish between the system state variables $x(t) \in \mathbb{R}^{n_x}$, representing the components of the considered process, and the measured entities $y(t) \in \mathbb{R}^{n_y}$ of the process. Because a system output can be aggregated information of several compounds, it is represented for the sake of generality by a nonlinear function of the form:

$$y_i(t) = g_i(x(t), p, u(t), w(t)), \quad i \in [1 : n_y], \quad (2.7)$$

where n_y defines the number of outputs, and for the sake of generality $g_i(\cdot) : \mathbb{R}^{n_x} \times \mathbb{R}^{n_p} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_w} \rightarrow \mathbb{R}$ a polynomial/rational function in the variables. Typically, the outputs deduce from the systems states, i.e. $y(t) = Cx(t)$ with $C \in \mathbb{R}^{n_y \times n_x}$ a known matrix.

For many biological processes, the principle of mass conservation applies. This is, the entities of a closed system do not change over time, although the entities can be transformed into others. Biochemical reaction networks are often described as closed systems, e.g. when no inputs or degradation processes are considered. Mathematically, mass conservation can be expressed by vectors $\gamma \in \mathbb{R}^{n_x}$ which satisfy

$$\gamma^T \dot{x}(t) = \gamma^T S\nu(t) = 0. \quad (2.8)$$

Such conservation relations allow to reduce the dimension of the system, i.e. by reducing the system's order, see for details e.g. Heinrich and Schuster [1996], which can be required for analysis purposes, for example to avoid numerical problems as in Waldherr [2009]. Here, we take the conservation relations explicitly into account. By integrating Equation (2.8), the conservation relations can be conveniently expressed in terms of *algebraic constraints* of the form

$$\gamma_i^T x(t) = c_i, \quad i \in [1 : n_c], \quad (2.9)$$

where $\gamma_i \in \mathbb{N}^{n_x}$ denotes an integer vector satisfying Equation (2.9), and c_i for all $i \in [1 : n_c]$ a constant parameter. By including c_i for all $i \in [1 : n_c]$ into the parameter vector p , we summarize algebraic constraints for the sake of generality by

$$0 = h_i(x(t), p, u(t), w(t)), \quad i \in [1 : n_c], \quad (2.10)$$

where $h_i(\cdot) : \mathbb{R}^{n_x} \times \mathbb{R}^{n_p} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_w} \rightarrow \mathbb{R}$ a rational function.

Example 2.1

We consider the reaction mechanisms proposed by Henri [1902] to build a product (P) from a substrate (S) via an enzyme (E) and an enzyme-substrate complex (C):



By balancing, the simple reaction network can be described by:

$$\begin{aligned} [\dot{S}] &= -\nu_1 + \nu_2 \\ [\dot{E}] &= -\nu_1 + \nu_2 + \nu_3 \\ [\dot{C}] &= +\nu_1 - \nu_2 - \nu_3 \\ [\dot{P}] &= + \nu_3, \end{aligned}$$

where $[X]$ denotes the concentration of compound $X \in \{S, E, C, P\}$, and ν_i , $i = 1, 2, 3$ the reaction rates. By assuming the law of mass action, the reaction rates are:

$$\nu_1 = p_1[S][E], \quad \nu_2 = p_2[C], \quad \nu_3 = p_3[C],$$

where p_1, p_2, p_3 are the reaction constants.

For the considered reaction system, the following two conservation relations are found: $[\dot{E}] + [\dot{C}] = 0$, and $[\dot{S}] + [\dot{C}] + [\dot{P}] = 0$, and hence the algebraic constraints

$$h_1 = [E] + [C] - c_1, \quad h_2 = [S] + [C] + [P] - c_2$$

are derived, with c_1 and c_2 denoting the integration constants.

If we furthermore consider that only the concentration of the substrate and the product are measurable, the system outputs are given by:

$$y_1 = [S], \quad y_2 = [P].$$

Next, we turn on describing the available information for testing such models and estimating the unknown parameters.

2.2. Data and uncertainty description

Information about the process is required to decide if a model hypothesis has to be rejected or to infer unknown model parameters. Besides a priori and qualitative knowledge, e.g. conservation relations and monotonic behavior, such information is generated from experiments with the studied process.

One of the most important features of measurements of biological processes however is its uncertainty. Of course, “all physical measurements are subject to uncertainties” [Tarantola, 2005], the situation for biological processes is however particularly challenging. A typical situation is depicted in Fig. 2.1. Errors of biological data are typically non-homogeneous, outliers may corrupt the data, and the magnitude of uncertainty is very often significant, e.g. due to superposition of inherent and observational noise. A measurement therefore can not be understood only as an observed value, rather as an acquired ‘state of information’ [Tarantola, 2005]. This information about the (un-)certainties is crucial e.g. to evaluate the precision of the estimated parameters.

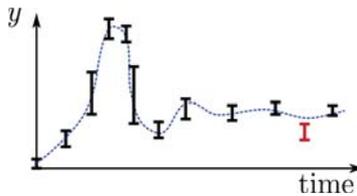


Figure 2.1.: Illustration of the characteristic challenges of biological measurement data. Uncertainties are frequently non-homogeneous, the data may be corrupted by outliers, and errors are significant.

The ‘state of information’ of the observations can be represented by probability density functions. There are however practical and methodological issues concerning this notion of errors. Firstly, the probability density functions of the observations in general have to be inferred from statistical and calibration procedures of the hardware sensing devices [Raissi et al., 2010]. To this end, the experiments may have to be repeated several times. While this is a standard procedure for determining the imprecision of measurement devices (observational errors), e.g. by calibration procedures, this may be very resource intense when considering experiments with the actual process. Then, only few repetitions may be realizable, which may be insufficient for a reliable determination of the probability density functions. Secondly, the information about the probability densities of the observations have then to be converted into information about the process, e.g. to infer the *a posteriori probability distributions*. This is in general a difficult task considering the nonlinear nature of the model and non-homogeneity of the errors.

For these reasons, we pursue an alternative approach to describe the ‘state of information’ by using the notion of bounded errors, i.e. we consider the uncertainties to be unknown, but set-bounded¹. Importantly, the bounded error description can be derived from available probability density functions if this information is available. This case is considered in Chapter 7. Otherwise, realistic uncertainty bounds can also be derived if less information is available, e.g. from a few repetitions.

Next, we turn on describing commonly available data and its associated uncertainty. We address more complicated scenarios such as outliers in the data separately in Section 5.3.

2.2.1. Measurement data

Uncertainty in measurement data can be introduced by the observation, may be due to disturbances, or may originate from the process itself. Typically, these three factors superpose. *Observational errors* result from limitations of the precision of the utilized measurement devices. Such errors are frequently accessible via a statistical analysis, and are described later on in detail. *Disturbances* in turn reflect upon limitations of the proposed model, and denote for example factors which influence the process although they are not explicitly modeled. This may be because the underlying mechanisms are not well known, or simply to keep the complexity of the model moderate. We explicitly include disturbances, which are considered unknown, but bounded. Furthermore, one key characteristic of biological processes, in comparison to many technical systems, is *variability*, which can be, in particular for intra-cellular processes, “so great that we rarely worry about measurement error” [Quinn and Keough, 2002]. Variability describes the phenomena that cells of the same type exhibit individuality and differences in behavior when supposed to similar conditions; and it has been shown for a wide variety of cellular processes in different cell types ranging from bacterial cells [Swain et al., 2002] to complex mammalian cells [Eissing et al., 2004, Ramsey et al., 2006, Weinberger et al., 2005]. This phenomena has drawn much attention in recent years, see e.g. Borchers [2007], Colman-Lerner et al. [2005], Elowitz et al. [2002], Hayot and Jayaprakash [2006], Levsky et al. [2002], Mantzaris [2005], Raser and O’Shea [2005], Schliemann et al. [2011]).

From a modeling perspective, such heterogeneity is important to recognize for two reasons. First, many biochemical measurement techniques such as Western Blotting and Electrophoretic Mobility Shift Assays (EMSA) provide observations of an average response combining many cells for analysis (bulk cell analysis), all of which might contribute in a different way to the overall observed dynamics. Single cell all-or-none or oscillatory behavior can be masked in the average response, and using average data for (estimation of parameters of) single-cell models can be misleading, and vice

¹This does not imply an uniform distribution with bounded support. We rather do not assume anything about the distribution in the set.

versa. Second, when we compare and utilize different experimental data to infer the model parameters or to discriminate model hypotheses. If inherent variability is not appropriately taken into account, e.g. by introducing additional uncertainty or by allowing variations in initial conditions and parameters, hypotheses may be falsely rejected.

Experiments and measurements

We refer to an *experiment* to a set of instructions which is performed with the process under study for the purpose to obtain measurements. Particularly, the instructions consists in an initial condition $x_0 \in X_0$ and a known and well defined inputs² $u(t) \in U(t)$, applied continuously without delays. Note that we address the design of experiments in Chapter 6.

By performing an experiment, denoted for shorthand by $Exp(x_0, u(t))$, with the process, measurements are taken at $\{t_0, t_1, \dots, t_M\}$. The respective observations are denoted by

$$\tilde{y}_i(t_j), \quad i \in [1 : n_y], \quad t_j \in [t_0 : t_M]. \quad (2.12)$$

Analogously, a state observation is denoted $\tilde{x}_i(t_j)$. These observations represent a state of information, i.e. they are uncertain to some extend. The uncertainty may be the result of the observation process, i.e. due to measurement imperfection, or caused by disturbances acting on the process. We next focus on the former case, and describe how to model observational uncertainty.

Note that in general, we allow that not all system states are measured (*permanently incomplete measurements*), and that not at every time instance $t_j \in [t_0 : t_M]$ a measurement is available (*casually incomplete measurements*). Also, the measurements can be correlated. However, for the sake of simplicity, we focus next on the frequent case that the measurements are independent from one another, i.e. component-wise interval bounds instead of more general set-valued bounds are derived.

Observational uncertainty

The observational errors are related to the measuring process itself, in particular resulting from imperfection of the measurement devises. Depending on the knowledge of the devises and there characteristics, the observational error can be modeled appropriately.

Absolute uncertainty is used to describe homogeneous uncertainty of the data, e.g. a possible bias of the measurements due to calibration errors. We model absolute

²which can be uncertain itself

uncertainty by an additive disturbance $\eta_a \in \mathbb{R}_+$, superposed onto the observation values. Then, the true measurement $y_i(t_j)$ is located in the interval

$$y_i(t_j) \in [\tilde{y}_i(t_j) - \eta_a, \tilde{y}_i(t_j) + \eta_a],$$

Such absolute uncertainty can for instance be motivated from a statistical analysis of the procedural errors. e.g. if one can show that uncertainties are homogeneously distributed according to the F-test, see e.g. Funk et al. [2007]. Then, it is reasonable to consider the standard deviation of the procedure $\eta_a = \kappa\sigma_i$, e.g. the κ -sigma confidence interval, as uncertainty bounds.

Relative uncertainty is a particular non-homogeneous error, where uncertainty growth with the values of the observation. This type of uncertainty is modeled by a relative error $0 \leq \eta_r \leq 1$. The respective bounding interval is given by

$$y_i(t_j) \in [(1 - \eta_r)\tilde{y}_i(t_j), (1 + \eta_r)\tilde{y}_i(t_j)].$$

Such a choice can be motivated, if statistical analysis of validation assays shows non-homogeneous uncertainty distribution with relative procedural standard deviation r_i , i.e. the variation coefficient (see e.g. Funk et al. [2007]). We then set $\eta_r = r_i$.

Limit of detection We furthermore have to take into account that some compounds may only be detectable above a certain threshold. We denote the *minimal detectable value* $\underline{\eta}_i$ as the lowest level at which a compound can be detected. The detection threshold is taken into account by

$$\tilde{y}_i(t_j) \leq \underline{\eta}_i \Rightarrow \underline{y}_i(t_j) = 0. \tag{2.13}$$

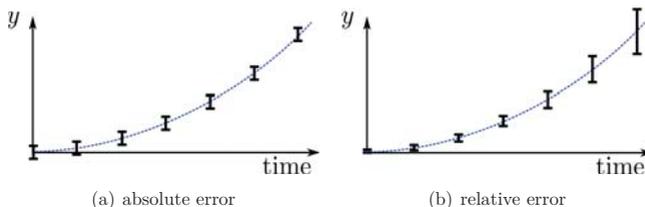


Figure 2.2.: Illustration of the absolute (homogeneous) and relative (non-homogeneous) measurement uncertainty.

The above descriptions of observational uncertainty are of course special cases, which nevertheless are often sufficient in practice, compare also the application example in

Chapter 7. In other cases however, these regular uncertainty distributions described above may not be applicable, and then the uncertainty description may have to be adapted to the particular case, e.g. by evaluating the concentration-dependent deviations of the calibration function with respect to the validation assays.

Furthermore, a special case is if the probability densities of the observational errors are known, or can be inferred because sufficient repetitions of an experiment are available. For completeness, we consider a number R of repetitions of an experiment $Exp(x_0, u)$ have been performed, and we have obtained the (independent) measurements

$$\tilde{y}_i^{(l)}(t_j), \quad l \in [1 : R], \quad t_j \in [t_0 : t_M], \quad i \in [1 : n_y]. \quad (2.14)$$

If R is sufficiently large, we can infer an approximate probability density distribution for the measured outputs. Exemplary, if the measurements can be assumed independent and normally distributed, i.e.

$$y_i(t_j) \sim N(\mu_i(t_j), \sigma_i^2(t_j)), \quad (2.15)$$

an estimate of the mean $\hat{\mu}_i(t_j)$ and standard deviation $\hat{\sigma}_i^2(t_j)$ is given by

$$\hat{\mu}_i(t_j) = \frac{1}{n_r} \sum_{l=1}^{n_r} \tilde{y}_i^{(l)}(t_j), \quad (2.16)$$

$$\hat{\sigma}_i^2(t_j) = \frac{1}{n_r - 1} \sum_{l=1}^{n_r} (\hat{\mu}_i(t_j) - \tilde{y}_i^{(l)}(t_j))^2. \quad (2.17)$$

Based on the (so derived) probability densities, the uncertainty bounds can be chosen according to the sigma-confidence levels. Exemplary, the uncertainty bound for $y_i(t_j)$ for the normal distributed observations $\tilde{y}_{i,l}(t_j)$, the confidence-level based uncertainty bound is given by

$$y_i(t_j) \in [\hat{\mu}_i(t_j) - \kappa \hat{\sigma}_i(t_j), \hat{\mu}_i(t_j) + \kappa \hat{\sigma}_i(t_j)], \quad (2.18)$$

where $\kappa = \{1, 2, 3\}$ define the $\{68.27\%, 95.45\%, 99.73\%\}$ confidence level. As suggested by Fogel and Huang [1982], the confidence-level based uncertainty description can be used to obtain *confidence set estimators*.

Inherent Variability

Besides observational uncertainty, the studied process may contain an inherent source of uncertainty, e.g. heterogeneity, for example if data is obtained by Western Plots or EMSA, or more generally by aggregation of independent experiments. This case is accommodated by considering the observations are not single values $\tilde{y}_i(t_j)$ but already set valued, i.e. $[\tilde{\underline{y}}_i(t_j), \tilde{\overline{y}}_i(t_j)]$. In this case, one proceeds with superimposing the

observational errors onto the lower and upper measurement value, $\tilde{y}_i(t_j)$ and $\bar{\tilde{y}}_i(t_j)$ respectively, and by deriving worst-case uncertainty bounds.

Summary uncertainty description The observational uncertainties can be evaluated from validation assays of the utilized measurement devices. Commonly used descriptions are the absolute and relative uncertainty, as well as the limit of detection. These descriptions can be extended to the case information about inherent variability is available. Furthermore, a set-valued uncertainty description of the observations can be derived from known probability density distributions of the errors by considering the κ -sigma confidence intervals. The uncertainty description is summarized by bounding sets for each observation by

$$D_{meas} : \begin{cases} u(t) \in U(t) \doteq \{u \in \mathbb{R}^{n_u} : A_{u(t)}u \leq a_{u(t)}\}, & t \in [t_0, t_N], \\ x(t_j) \in X(t_j) \doteq \{x \in \mathbb{R}^{n_x} : A_{x(t_j)}x \leq a_{x(t_j)}\} & t_j \in [t_0 : t_M], \\ y(t_j) \in Y(t_j) \doteq \{y \in \mathbb{R}^{n_y} : A_{y(t_j)}y \leq a_{y(t_j)}\} & t_j \in [t_0 : t_M]. \end{cases} \quad (2.19)$$

Note that very frequently the sets are simple component-wise intervals. A formal description of the bounding sets in terms of polytopes is provided in the Appendix B.

2.2.2. A priori data

Data independently available of actual measurements is termed a priori knowledge or data. Such knowledge might derive from first principles, or stem from further observations.

The system's states $x(t)$ (and the systems outputs $y(t)$) can typically be constrained from first principles, e.g. by considering symmetry properties or conservation relations (energy, mass, entropy), see e.g. Ederer and Gilles [2007]. Furthermore, disturbances are assumed unknown, but bounded in magnitude, i.e. $w(t) \in \Omega \subset \mathbb{R}^{n_w}$. Finally, a priori information might be available for the model parameters, e.g. an initial parameter domain $p \in P \subset \mathbb{R}^{n_p}$. Such bounds often derive from first principles, or can be based on literature values; if available, the bounding set can also be deduced from a priori probability distributions.

In summary, we model a priori data for the system variables by (compact and convex) bounding sets

$$D_{prior} : \begin{cases} P \doteq \{p \in \mathbb{R}^{n_p} : A_p p \leq a_p\}, \\ X \doteq \{x \in \mathbb{R}^{n_x} : A_x x \leq a_x\} \\ U \doteq \{u \in \mathbb{R}^{n_u} : A_u u \leq a_u\} \\ Y \doteq \{y \in \mathbb{R}^{n_y} : A_y y \leq a_y\} \\ \Omega \doteq \{w \in \mathbb{R}^{n_w} : A_w w \leq a_w\} \end{cases} \quad (2.20)$$

see Appendix B for further details.

2.2.3. Structural constraints

Additional information, besides measurements and a priori knowledge, might be available in terms of dependencies of variables or about some a priori known qualitative features the dynamics of the system. Exemplary, metabolites are often measurable and physical entities such as chemical compounds, i.e. non-negative $x_i(t) \geq 0$ for some $i \in [1 : n_x]$. Furthermore, some compounds might be known to obey a *monotone* behavior, e.g. when considering exponential cell growth. A monotone non-decreasing dynamics can be expressed by an inequality of the form $\dot{x}_i(t) \geq 0$, for some $i \in [1 : n_x]$, and $-\dot{x}_i \geq 0$ a monotonic non-increasing dynamic respectively. More generally, structural information in form of barrier functions might be available, including positivity and monotonicity. In summary, we include additional constraints which can be expressed by inequalities of the form

$$D_{str} : \left\{ q_i(p, \dot{x}(t), x(t), u(t), y(t)) \leq 0, \quad i \in [1 : n_q], \right. \quad (2.21)$$

where n_q denotes the number of constraints motivated above. Note that we here consider only $q_i(\cdot)$ linear functions of the system variables (see Appendix B).

2.3. Summary

Many biological processes can be described by means of ordinary differential equations with rational structure. The models equations derive from balancing, by considering a set of relevant compounds $x(t) \in \mathbb{R}^{n_x}$ and their interactions. By specifying the reactions kinetics, reaction parameters $p \in \mathbb{R}^{n_p}$ are introduced.

Inputs $u(t) \in \mathbb{R}^{n_u}$, disturbances $w(t) \in \mathbb{R}^{n_w}$, and outputs $y(t) \in \mathbb{R}^{n_y}$ can be introduced to model external stimuli, perturbations, and the measurable entities; conservation relations can be described by additional algebraic equations. The dynamical (continuous-time) model is then summarized by

$$M^c : \begin{cases} \dot{x}_i(t) = f_i(x(t), p, u(t), w(t)) & i \in [1 : n_x] \\ y_i(t) = g_i(x(t), p, u(t), w(t)) & i \in [1 : n_y] \\ 0 = h_i(x(t), p, u(t), w(t)) & i \in [1 : n_c] \end{cases}, \quad (2.22)$$

for $t_0 \leq t \leq t_N$ denoting the time window of interest.

Data for biological systems is subjected to various sources of uncertainty. In contrast to classical approaches for estimation, we model uncertainties using the notion of bounded errors. The uncertainty description can be based on knowledge about the measurement devices, or can be derived from a priori probability distributions if available. To this end, we denote a priori data D_{prior} (2.20) as the information available before experimentation. Measurement data and the associated observational uncertainty, may be incomplete and sparse, and is expressed by state and output bounding

sets for the time instances t_0, t_1, \dots, t_M where observations are available. The measurement data is summarized by D_{meas} (2.19). We also include additional knowledge e.g. about correlation of model variables or known qualitative features which can be expressed by inequality constraints, summarized by D_{str} (2.21).

With this preparations, we can now turn on the formulation of the estimation and analysis problems, regarding the introduced dynamical models and utilizing all available data.

3. Set-Membership Framework for Falsification and Estimation

One of the most important issues when modeling biological processes is to determine or to refine the models variables given the available data. For example, the estimation of the model parameters is pivotal for analysis and experimental design, and has fundamental impacts on the performance of model-based controllers. Prior to this, it is important to decide whether a model can reproduce an observed behavior at all, otherwise the model has to be rejected. As outlined in Chapter 2, data can be given, besides measurements from experimentation, as a priori knowledge or qualitative information, and the data is typically subjected to uncertainty such as measurement noise. By describing data uncertainty in terms of bounded errors, we can consider a set-membership setting for the hypothesis falsification and estimation problems. Particularly, the parameter/state estimation problem then translates into determining the *set of feasible parameters/states* of the model that respects the available data. For falsification it is sufficient to show that this set is empty.

In general, the set of feasible parameters or states can be very complicated, and for most applications outer-approximations are sufficient. Obtaining (simply-shaped) outer-approximations of given sets defines the set-membership problem. Parameter estimation has been addressed in this setting already for linear systems considering ellipsoids (e.g. Fogel and Huang [1982], Schweppe [1968, 1973]), orthotopes (e.g. Milanese and Belforte [1982]), zonotopes (e.g. Mo and Norton [1990], Walter and Piet-Lahanier [1989]), or general fixed shape approximations utilizing homothety [Borchers et al., 2011b]. For nonlinear systems, these algorithms are not generally applicable, because the underlying optimization problems are typically non-convex. To overcome this issue, interval analysis and constraint propagation methods for parameter estimation have been developed, see e.g. Jaulin et al. [2001] and references therein. Set-membership approaches for nonlinear *system identification* have been established, e.g. for Wiener [Cerone and Regruto, 2006] and Hammerstein [Cerone and Regruto, 2003] models, and Aubin et al. [2002], Bemporad et al. [2005], Figueroa et al. [2008], Milanese and Novara [2004], Sznaier [2009]. A barrier-certificate set-membership approach has been considered [Prajna, 2006] to discard parameter sets.

An alternative certificate-based approach for parameter estimation of stationary systems has been presented in Kuepfer et al. [2007]. The approach consists in deriving infeasibility-certificates by employing a suitable semi-definite relaxation [Parrilo,

2003]. Using a bi-sectioning algorithm, the set of feasible parameters can be outer-approximated. Such a relaxation approach has, for example, been extended to steady state sensitivity analysis [Waldherr et al., 2008], steady state analysis of (bio-)chemical processes [Hasenauer et al., 2010a], and linear dynamic systems [Cerone et al., 2010]. However, considering only stationary systems (as in Hasenauer et al. [2010a], Kuepfer et al. [2007], Waldherr et al. [2008]) is in general not sufficient to analyse dynamical systems and estimate their parameters, see e.g. Bullinger et al. [2008], Farina et al. [2006]. In this thesis, we extend this approach to dynamical systems, following Borchers et al. [2009a,b,c], Hasenauer et al. [2010b], Rumschinski et al. [2010a].

In this chapter, we extend the approach of Kuepfer et al. [2007] to nonlinear dynamical systems. The proposed approach builds on polynomial relaxations (see e.g. Kojima [2002], Lasserre [2001], Lasserre et al. [2008], Lovasz and Schrijver [1991], Parrilo [2003], Sherali and Adams [1990]), and employs convex optimization techniques to address the set-membership problem efficiently. Next, we formulate in Section 3.1 the falsification and estimation problems in the set-membership setting in terms of solutions to particular non-convex quadratic optimization/feasibility problems. In Section 3.2 we describe the procedure to relax the non-convex problem into a convex one, and subsequently derive infeasibility and dual certificates. These certificates form the basis for falsification, estimation, and branch-and-bound optimization as outlined in Section 3.3. In Section 3.4, we provide some computational notes how to reduce computational complexity. We finally provide some insights in Section 3.5. Parts of this chapter are based on Rumschinski et al. [2010a] and Borchers et al. [2009a,b,c, 2012].

3.1. The feasibility, estimation, and optimization problems

A preliminary step of the following set-membership falsification and estimation framework consists in deriving an discrete time system, which is close by its properties to the continuous ODE systems (2.22). In general, this is achieved by discretization, i.e. choosing an appropriate discretization scheme and sampling. The required steps were discussed in detail in Rumschinski [2012], Rumschinski et al. [2010b], and are outlined in the Appendix A. We remark that the sampling steps have to be chosen sufficiently small such that possible discretization errors are negligible, refer also Letellier et al. [2004], Rumschinski et al. [2012] for a comprehensive discussion of possible numerical stability issues. The resulting difference equation system with $k \in [1 : N]$ considered in the remainder is summarized by:

$$M : \begin{cases} f_i^k(x_k, x_{k-1}, p, u_{k-1}, w_{k-1}) = 0, & i \in [1 : n_x] \\ g_i^k(y_k, x_k, p, u_{k-1}, w_{k-1}) = 0, & i \in [1 : n_y] \\ h_i^k(y_{k-1}, x_k, x_{k-1}, p, u_{k-1}, w_{k-1}) = 0, & i \in [1 : n_c] \end{cases} \quad (3.1)$$

For the sake of simplicity of notation, we introduce the vector of variables $z \in \mathbb{R}^{n_z}$, which collects all (discrete) system variables for $k \in [1:N]$:

$$z \doteq (p_1, \dots, p_{n_p}, x_0, \dots, x_N, u_0, \dots, u_{N-1}, y_0, \dots, y_N, w_0, \dots, w_{N-1}).$$

By construction, we have $n_z = n_p + N(2 + n_x + n_u + n_y + n_w)$.

We furthermore denote the collection of available data, i.e. a priori, measurement, and structural data, informally by

$$D \doteq D_{\text{prior}} \cap D_{\text{meas}} \cap D_{\text{str}}. \quad (3.2)$$

Formally, this collection can be seen as the intersection of the respective bounding sets Z_{prior} , Z_{meas} , Z_{str} , see for details Appendix B. The overall data D (3.2) can then be summarized for simplicity of presentation by the set

$$Z = \{z \in \mathbb{R}^{n_z} : A_z z \leq a_z\}.$$

With these preparation, we can now focus on the formulation of the considered problems in membership setting. To this end, we first derive the set of consistent solutions as follows:

Proposition 1 (Solution set)

All solutions of the dynamical model (3.1), which are consistent with the data (3.2), belong to the set $\mathcal{Z} \subset \mathbb{R}^{n_z}$ with:

$$\mathcal{Z} \doteq \left\{ z \in Z : \begin{array}{l} f_i^k(z) = 0 \quad k \in [1:N], i \in [1:n_x] \\ g_i^k(z) = 0 \quad k \in [1:N], i \in [1:n_y] \\ h_i^k(z) = 0 \quad k \in [1:N], i \in [1:n_c] \end{array} \right\} \quad (3.3)$$

where $f^k(z)$, $g^k(z)$ and $h^k(z)$ are respectively (3.1) with an appropriate choice of z -components.

Proof. The proof is simply by construction. Let us choose a vector $\xi \in \mathbb{R}^{n_z}$, with $\xi \in Z$ so that ξ is consistent with the data D (3.2) (Z , see Appendix B.4). Only if we have $f_i^k(\xi) = 0$ for all $k \in [1:N]$ and for all $i \in [1:n_x]$, and respectively $g_i^k(\xi) = 0$ for all $k \in [1:N]$ and for all $i \in [1:n_y]$, and $h_i^k(\xi) = 0$ for all $k \in [1:N]$ and for all $i \in [1:n_c]$, then ξ is, by definition, a solution of M (3.1). \square

In other words, the set of solutions \mathcal{Z} contains all solutions of M (3.1) consistent with D (3.2), denoted hereafter as *strictly feasible solutions*. Note, by construction, $\mathcal{Z} \subseteq Z$.

Feasibility The general feasibility problem can now be formulated in terms of the solution set as follows:

Problem 1 (Feasibility)

Prove that Model (3.1) is consistent (not consistent) with the Data (3.2). That is, prove that there exists (no) solutions of the model respecting the data, i.e. \mathcal{Z} is non-empty (empty).

As we will outline later on, the feasibility problem templates several applications, for example model invalidation, reachability analysis, and outlier detection.

Estimation The goal of an estimation problem is to describe and characterize the solution set \mathcal{Z} , if \mathcal{Z} is non-empty, as good and as efficient as possible. More particular, one is typically only interested an estimate of *some* components of z , denoted hereafter by $s \in \mathbb{R}^{n_s}$ ($n_s \leq n_z$). Exemplary, in case of parameter estimation, we are interested in the n_p components of z corresponding to the parameters, i.e. $s = p$. In general, we select some components s ; this can be seen as an n_s -dimensional axis-parallel subspace of \mathbb{R}^{n_z} , formalized by the projection map $f_s(\cdot) : \mathbb{R}^{n_z} \rightarrow \mathbb{R}^{n_s}$. The general estimation problem can thus be formulated as:

Problem 2 (Estimation)

Given a selection $s \in \mathbb{R}^{n_s}$ of system variables of interest. Provide the set bounding all feasible solutions in s of M (3.1) respecting D (3.2), i.e. provide an estimate of the set $\mathcal{S} \doteq f_s(\mathcal{Z})$.

Note that in principle any desired selection of variables can be chosen, thus possibly defining estimation problems involving a combination of state, parameter, input, and output variables. Note also that we are only interested in guaranteed estimates/approximations in the remainder of this thesis, i.e. to find membership sets, refer to Section 3.3 for details.

Optimal estimation Often, not only a set, but also an optimal value for the variables of interest is required, referred to as the optimal estimation problem. The problem of optimal estimation consists of finding the optimal values $z^* \in \mathbb{R}^{n_z}$, or a selection $s^* \doteq f_s(z^*)$, minimizing some (polynomial) objective function $c(z)$. The respective optimum is denoted by $c(z^*)$. In case of optimal parameter estimation, frequently one considers a data fitting objective function measuring the distance of the output trajectory from the measurements, e.g. c_{slq} provided in (4.3). The problem of finding the optimal value is formulated as follows:

Problem 3 (Optimal estimation)

Determine the (global) optimum $c(z^)$ and the optimal values z^* of the objective function $c(z)$. That is, find a solution of the polynomial optimization problem:*

$$\begin{cases} \min_{z \in \mathbb{R}^{n_z}} c(z) \text{ s.t.} \\ z \in \mathcal{Z}. \end{cases} \quad (3.4)$$

The polynomial optimization problem (3.4) forms the basis of the following considerations in this chapter and applications considered in the following chapters. Note that by construction, the argument of above optimization problem is also a strictly feasible solution, and the optimum may not be unique.

3.2. Reformulation and relaxation

Due to set valued states and parameters and the (implicit) polynomial system equations, the solution set \mathcal{Z} (3.3) is non-convex, and might be composed of disconnected regions. Furthermore, (3.4) might admit several local optima, or the solutions might be set-valued. Showing that no strictly feasible solution exists, characterizing the solution set, or finding a global optimum, can therefore be very challenging.

We approach the feasibility and estimation problems via convexification, i.e. the non-convex problems are relaxed into convex ones, which are close by their properties to the original problems. In particular, the relaxation process is conservative: any strictly feasible solution remains feasible for the relaxed problem. The converse does not hold in general, as relaxations typically introduce ‘spurious’ solutions, that are feasible for the relaxed problem but not strictly feasible.

Some convex relaxations have been developed for polynomial optimization problems, in particular the Reformulation-Linearization Technique (e.g. Sherali and Adams [1990]) and semidefinite programming (SDP) relaxations (see e.g. Lasserre [2001], Lovasz and Schrijver [1991], Parrilo [2003]). SDP relaxations have become very popular due to their theoretical and practical properties. Semidefinite programming can be regarded as a generalization of linear programming over semi-definite cones, and efficient software packages for SDP optimization are nowadays available. SDPs can be solved with any desired precision in polynomial time by interior-point methods [Nesterov and Nemirovski, 1994]. Furthermore, the optimum of a fairly general class of polynomial problems with compact feasible region can be approximated with arbitrary precision by a finite sequence of SDPs [Lasserre, 2001]. Several SDP solvers with state-of-the-art implementations of the primal-dual interior-point algorithm are freely available, as e.g. SeDuMi [Sturm, 1999] and SDPT3 [Tütüncü et al., 2003], and can be directly used in Matlab using Yalmip [Lofberg, 2004] or GloptiPoly [Henrion and Lasserre, 2003] as interfaces. The ADMIT toolbox [Streif et al., 2012], as considered here, is build upon Yalmip.

In the following, we describe our framework using the SDP relaxation following Parrilo [2003], but we remark that it is independent of the specific relaxation employed. Indeed, relaxations to linear problems is numerically advantageous if problems with a large number of variables are considered. We discuss the use of weaker relaxations in Section 3.4, and refer to Kojima [2002] for a hierarchy of convex relaxations for semi-algebraic problems that are applicable to our framework.

We focus next on the optimal estimation problem, i.e. we aim to obtain a lower bound on the objective value. The feasibility problem immediately derives from the optimal estimation problem by considering a zero objective function. Lower bounds on particular objective value are furthermore required for approximating the solution set, as outlined in Section 3.1.

3.2.1. Reformulation of POP(Z)

Reconsider the polynomial optimization problem (3.4):

$$\text{POP}(Z) : \begin{cases} \min_{z \in \mathbb{R}^{n_z}} c(z) \text{ s.t.} \\ f_i^k(z) = 0 & k \in [1:N], l \in [1:n_x] \\ g_i^k(z) = 0 & k \in [1:N], l \in [1:n_y] \\ h_i^k(z) = 0 & k \in [1:N], l \in [1:n_h] \\ z \in Z. \end{cases} \quad (3.5)$$

As a first step, we reformulate $\text{POP}(Z)$ in terms of an equivalent quadratic optimization problem using *quadrification* (see e.g. Sherali and Tuncbilek [1997]). To this end, let \mathbb{S}^n be the set of real symmetric $n \times n$ matrices, and

$$\langle A, B \rangle = \sum_{i,j} a_{ij} b_{ij}$$

denote the usual Frobenius product. Quadrification [Sherali and Tuncbilek, 1997] consists in deriving a monomial vector ξ for which

$$c(z) = \langle C, \xi \xi^T \rangle, \quad f_i^k(z) = \langle F_i^k, \xi \xi^T \rangle, \quad g_i^k(z) = \langle G_i^k, \xi \xi^T \rangle, \quad h_i^k(z) = \langle H_i^k, \xi \xi^T \rangle,$$

for appropriate matrices $C, F_i^k, G_i^k, H_i^k \in \mathbb{S}^{n_\xi}$. Each *monomial* (of ξ) of degree two or more is thus represented by the product of two other monomials of lower degree. This is explicitly expressed by constraints of the form

$$\langle D_i, \xi \xi^T \rangle = 0 \quad i \in [1:n_d],$$

which enforce the inter-dependence between higher and lower degree monomials in ξ , for appropriate matrices $D_i \in \mathbb{S}^{n_\xi}$. As a technical requirement, we ask without loss of generality, that $\xi_1 = 1$. The next n_z components of the vector ξ are all the components of z , i.e.,

$$(\xi_2, \dots, \xi_{n_z+1}) = z.$$

Note we have $\xi \xi^T e_1 = \xi$, with $e_1 = (1, 0, \dots, 0)^T \in \mathbb{R}^{n_\xi}$. Note also that the quadrification is not unique. One can typically find various monomials defining the same polynomial. This degree of freedom can be exploited to decrease conservatism.

The polytopic constraints Z bound, by construction, the z equivalent components of ξ . The remaining components, which are monomials of degree two or higher, can be bounded directly from the data Z , particularly considering interval arithmetic, for details see Appendix C. The resulting bounding constraints for all the components of ξ are expressed by

$$A_\xi \xi \xi^T e_1 = A_\xi \xi \leq a_\xi.$$

The reformulation by quadrification is summarized by the quadratic program:

$$\text{QOP}(Z) : \begin{cases} \min_{\xi \in \mathbb{R}^{n_\xi}} \langle C, \xi \xi^T \rangle \text{ s.t.} \\ \langle F_i^k, \xi \xi^T \rangle = 0 & k \in [1:N], i \in [1:n_x] \\ \langle G_i^k, \xi \xi^T \rangle = 0 & k \in [1:N], i \in [1:n_y] \\ \langle H_i^k, \xi \xi^T \rangle = 0 & k \in [1:N], i \in [1:n_h] \\ \langle D_i, \xi \xi^T \rangle = 0 & i \in [1:n_d] \\ A_\xi \xi \xi^T e_1 \leq a_\xi \\ \xi_1 = 1. \end{cases} \quad (3.6)$$

Advantageously, the quadrification procedure is always possible for a system with the considered structure, though it is not uniquely determined. In contrast to the moment matrix approach [Lasserre, 2001], a simple quadrification yields a smaller representation which is less tight but computationally easier to solve [Kojima, 2002], and hence pursued in the remainder. It is important to note that the moments relaxation can be considered within the here proposed framework by choosing an appropriate monomial vector ξ and considering further regularizations, for details see Lasserre [2001].

3.2.2. Semidefinite programming relaxation

The resulting QOP(Z) (3.6) is, still, non-convex. To overcome this problem, we propose to relax 3.6. Specifically, (3.6) can be casted as a linear matrix optimization problem by replacing the product matrix $\xi \xi^T$ with a symmetric variable matrix $\Xi \in \mathbb{S}^{n_\xi}$, with the additional non-convex constraint $\text{rank}(\Xi) = 1$, see Ramana [1994] for further details. This problem can then be relaxed into a semidefinite problem by replacing the rank constraint with the convex positive semidefinite constraint $\Xi \succeq 0^1$. The resulting

¹Note that other relaxations can be considered as discussed in Section 3.4, e.g. a relaxation of QOP(Z) into a second order cone problem by imposing $\Xi \in \mathcal{K}_2$ (instead $\text{rank}(\Xi) = 1$) or into a linear program imposing the non-negativity constraint $\Xi \geq 0$, see e.g. Kojima and Tuncel [2001].

convex semidefinite program is thus:

$$\text{SDP}(Z) : \begin{cases} \min_{\Xi \in \mathbb{S}^{n_\xi}} \langle C, \Xi \rangle \text{ s.t.} \\ \langle F_i^k, \Xi \rangle = 0 & k \in [1:N], i \in [1:n_x] \\ \langle G_i^k, \Xi \rangle = 0 & k \in [1:N], i \in [1:n_y] \\ \langle H_i^k, \Xi \rangle = 0 & k \in [1:N], i \in [1:n_h] \\ \langle D_i, \Xi \rangle = 0 & i \in [1:n_d] \\ A_\xi \Xi e_1 \leq a_\xi \\ \Xi_{11} = 1 \\ \Xi \succeq 0. \end{cases} \quad (3.7)$$

For shorthand of notation, we denote in the sequel by $z_{\text{SDP}} = (\Xi_{1,2}, \dots, \Xi_{1,n_z+1})$ the projection of the respective elements of the matrix $\Xi \in \mathbb{S}^{n_\xi}$ onto \mathbb{R}^{n_z} , formalized by the (coordinate erasing) projection map $f_z(\cdot) : \mathbb{R}^{n_\xi \times n_\xi} \rightarrow \mathbb{R}^{n_z}$. Furthermore, we denote by \mathcal{Z}_{SDP} the set of all vectors $f_z(\Xi)$ with Ξ a feasible solution of $\text{SDP}(Z)$, i.e. the projection set of all feasible solutions of $\text{SDP}(Z)$ onto z , for details see Appendix C.2.

Since the relaxation is conservative, the following key relation between $\text{POP}(Z)$ (3.4) and $\text{SDP}(Z)$ (3.7) is as follows:

Theorem 1 (Relaxation)

If the $\text{SDP}(Z)$ is infeasible, then the $\text{POP}(Z)$ is infeasible. If both problems are feasible, then the optimum of $\text{SDP}(Z)$ is a lower bound for the minimum of $\text{POP}(Z)$. All strictly feasible solutions of $\text{POP}(Z)$, i.e. \mathcal{Z} , are feasible solutions of $\text{SDP}(Z)$, i.e. $\mathcal{Z} \subseteq \mathcal{Z}_{\text{SDP}}$.

Proof. For any $z \in \mathcal{Z}$ there is by construction a vector $\xi \in \mathbb{R}^{n_\xi}$ which is feasible for the $\text{QOP}(Z)$ such that $\langle C, \xi \xi^T \rangle = c(z)$, and hence a rank-one matrix $\Xi \doteq \xi \xi^T$ which is feasible for the $\text{SDP}(Z)$ such that $\langle C, \Xi \rangle = c(z)$ and $z_{\text{SDP}} = z$. This shows that $\mathcal{Z}_{\text{SDP}} \supseteq \mathcal{Z}$, and directly implies Theorem 1. \square

Note also that, conversely, if $\Xi \in \mathbb{S}^{n_\xi}$ is rank-one and feasible for $\text{SDP}(Z)$, then $z_{\text{SDP}} \in \mathcal{Z}$ and $c(z_{\text{SDP}}) = \langle C, \Xi \rangle$. In other words, there is a bijection among strictly feasible solutions and rank-one feasible matrices.

The $\text{SDP}(Z)$ can be relaxed into a second order cone problem or into a linear program, where the numerical efficiency over SDP is the major advantage. On the other hand, adding valid (polynomial) constraints can yield tighter convex relaxations, though this may affect the computational performance. For example, the constraints $A_\xi \Xi e_1 \leq a_\xi$ can be augmented to $A_\Xi \Xi \leq a_\Xi$, thus binding all the elements of Ξ . Moreover, constraints $A_\Xi \Xi A_\Xi^T \leq a_\Xi$ can be considered, refer to Appendix C.2.

Relaxation tightness is in general very difficult to assess, although for some classes of non-convex quadratic optimization problems exact solutions are obtained by SDP or SOCP relaxations [Kim and Kojima, 2003]. A comprehensive study of the effect of SDP and further strengthening constraints can be found in Kojima [2002], and a comparison of SDP and RLT relaxations can be found in Anstreicher [2009].

Theorem 1 requires a guarantee of infeasibility of $\text{SDP}(Z)$, or a guarantee of optimality of a given solution. Such a guarantee can be provided by dual certificates as described next.

3.2.3. Dual certificate

Any feasible solution for the *dual problem* of a SDP provides by weak duality a lower bound to the SDP optimum, see e.g. Boyd and Vandenberghe [2004]. Therefore, dual unboundedness provides an (easy to check) certificate of primal infeasibility. Moreover, if strong duality applies, then the optimum of the dual and of the primal coincide, which provides a certificate of optimality.

Several SDP duals have been proposed (see Ramana et al. [1997] for a discussion of SDP duals). For simplicity, we here consider the Lagrangean dual $\text{SDP}^*(Z)$, which is itself an SDP and for which strong duality holds under constraint qualification conditions [Nesterov and Nemirovski, 1994]. To obtain the Lagrangian dual, non-negative Lagrange multipliers (dual variables) are introduced to include the constraints to the objective function. The dual problem then consists in maximizing the augmented objective function with respect the dual variables. We denote for shorthand of notation the Lagrangian dual by $\text{SDP}^*(Z)$.

Using duality, we can derive guaranteed bounds of the objective function as follows:

Theorem 2 (Dual bound)

A feasible solution of $\text{SDP}^(Z)$ gives a lower bound on the optimum of $\text{SDP}(Z)$, and hence on the optimum of $\text{POP}(Z)$. If $\text{SDP}^*(Z)$ is unbounded, then $\text{POP}(Z)$ is infeasible.*

Proof. Theorem 2 is a direct consequence of Lagrangean weak duality and Theorem 1.

We denote by Ξ^* an optimal solution of (the dual problem) $\text{SDP}^*(Z)$; by weak duality, we have first that Ξ^* provides a lower bound on the objective value for the $\text{SDP}(Z)$, i.e.

$$\langle C, \Xi^* \rangle \leq \langle C, \Xi \rangle, \forall \Xi \in \mathcal{R}_{\text{SDP}},$$

with \mathcal{R}_{SDP} the set of feasible solutions Ξ of $\text{SDP}(Z)$, see Appendix C.2. By denoting $z^* = f_z(\Xi^*)$, following Theorem 1, the lower bound of $\text{POP}(Z)$ is bounded from below by

$$c(z^*) \leq c(z), \forall z \in \mathcal{Z}.$$

Furthermore, if $c(z^*) \rightarrow \infty$, i.e. if it becomes unbounded, we have by weak duality that $\text{SDP}(Z)$ is infeasible, and hence that the $\text{POP}(Z)$ is infeasible according to Theorem 1.

□

Theorem 2 therefore provides a sufficient criterion for infeasibility of $\text{POP}(Z)$. This provides a way to solve Problem 1 using the dual problem. This will be used for model invalidation (Section 4.1), reachability analysis (Section 5.2), and outlier detection

(Section 5.3). The dual bounds themselves are utilized for estimation (Problem 2) and optimization (Problem 3) as shown in the next section.

3.3. Estimation

In this section, we consider the estimation problems, i.e. how to obtain an outer-estimate of the solution set \mathcal{Z} and $f_s(\mathcal{Z})$, and how to obtain an optimal estimate.

An overview of the considered estimation techniques is provided in Fig. 3.1. The most efficient approach for outer-estimation is determining the uncertainty intervals for the unknown variables using dual certificates, outlined hereafter in Section 3.3.1. In Section 3.3.2, we present an approach by which fixed shape membership sets are obtained using homothety. To assess the solution set in more detail, a partitioning approach is presented in Section 3.3.3. Finally, in Section 3.3.4, we propose a branch-and-bound approach for optimization purposes.

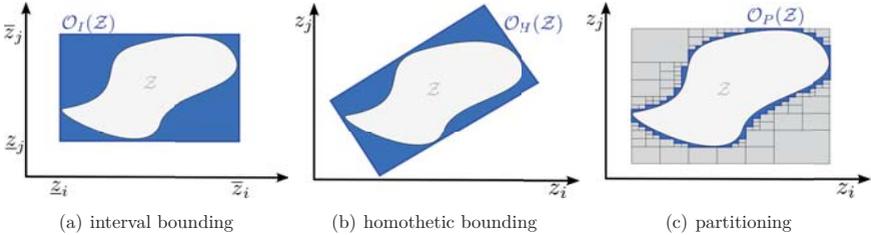


Figure 3.1.: Illustration of the basic estimation techniques.

3.3.1. Interval bounding

To obtain an as good as possible estimate of a (unknown) variable s_i , we choose for POP(\mathcal{Z}) the objective function $c(z) = s_i$. According to Theorem 2, the solution of the respective dual SDP*(\mathcal{Z})

$$\langle C_i, \Xi^* \rangle \doteq \underline{s}_i \quad (3.8)$$

provides a lower bound for s_i . Analogously, by choosing $c(z) = -s_i$, the respective dual solution

$$-\langle C_i, \Xi^* \rangle \doteq \bar{s}_i, \quad (3.9)$$

i.e. an upper bound of the respective variable s_i . The lower and upper bounds define the (a posteriori) uncertainty interval $[\underline{s}_i, \bar{s}_i]^2$.

²The uncertainty interval can be regarded as an outer-approximation of the n-sigma confidence interval, if the measurement uncertainties are modeled by their n-sigma confidence limits.

Proposition 2 (Uncertainty interval)

The interval $[\underline{s}_i, \bar{s}_i]$ bounds all strictly feasible solutions of s_i , and is found by 2 dual bounds of SDP(Z).

Proof immediately follows from construction and Theorem 2.

The length of the uncertainty interval, which is simply given by

$$\ell(s_i) \doteq \bar{s}_i - \underline{s}_i \quad (3.10)$$

can be considered as a measure for the sensitivity of the variable, as outlined in the next chapter.

If a number of unknown variables $s \in \mathbb{R}^{n_s}$ has to be estimated, the uncertainty intervals can be obtained independently from one another. We denote the overall solution set (given by the collection of n_s uncertainty intervals) as bounding orthotope

$$\mathcal{O}_I(\mathcal{S}) \doteq \mathcal{O}_I(f_s(\mathcal{Z}_{\text{SDP}})) = [\underline{s}_1, \bar{s}_1] \times \dots \times [\underline{s}_{n_s}, \bar{s}_{n_s}].$$

The properties of the bounding orthotope are summarized as follows:

Proposition 3 (Bounding orthotope)

The orthotope $\mathcal{O}_I(\mathcal{S})$ contains all strictly feasible solutions of s , and is derived from $2n_s$ dual bounds of SDP(Z).

Proof immediately follows from Proposition 2. The volume of the orthotope is simply given by the product of the intervals length.

Remark 1 (Iterations) *The quality of the relaxation \mathcal{Z}_{SDP} and thereby of the uncertainty intervals strongly depends on the (initial) bounding sets Z . Tightening the bound of one variable (e.g. by estimating the respective uncertainty interval) can propagate to the other variables. This enables us to refine the outer-estimate by updating the ‘initial’ bounds. This can be iterated until a fixed point is attained, or some precision threshold is reached.*

Interval bounding is the most efficient way to obtain an estimate of the unknown variables, i.e. the analytical complexity growth linearly with the number of unknown variables. Interval bounding is therefore used e.g. if several (time-variant) variables have to be estimated, e.g. for state estimation. Interval bounding often precedes a more detailed analysis of the solution set. However, interval bounding does not provide information about correlations of variables, nor the shape of the solution set. To this end, we next consider more general bounding sets for analysis purposes.

3.3.2. Homothetic bounding

We now consider obtaining more general bounding sets for the solution set \mathcal{Z} and $f_s(\mathcal{Z})$. To this end, we consider a framework for fixed-shape membership estimation

based on homothety, see e.g. Borchers et al. [2011b], Raković and Fiacchini [2008]. In the following, we consider families of homothetic sets defined as follows:

Definition 1 (Homothetic sets)

Sets $A \subset \mathbb{R}^n$ and $B \subset \mathbb{R}^n$ are called (positively) homothetic if $A = b + \alpha B$ for some $b \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}_+$.

Hereby, $+$ denotes the sum of sets. In other words, the sets A and B are homothetic if the one set can be obtained by appropriately orienting (shifting) and scaling the other set.

A family of homothetic sets with the same shape $H \subseteq \mathbb{R}^{n_h}$ is described by:

$$\mathcal{H}(H) \doteq \{h + \alpha H, c \in \mathbb{R}^{n_h}, \alpha \in \mathbb{R}_+\}. \quad (3.11)$$

Hereby, $h \in \mathbb{R}^{n_h}$ is an orientation vector, and $\alpha \in \mathbb{R}_+$ a scalar representing the ‘size’ of the set H . The set $H \subseteq \mathbb{R}^{n_h}$ denotes the basic shape, is designed off-line, and can in principle be an arbitrary non-empty compact, convex set.

We focus in the remained on an outer-bounding map, $\mathcal{O}_{\mathcal{H}}(\cdot) : \mathbb{R}^{n_h} \rightarrow \mathbb{R}^{n_h}$, defined by:

$$\mathcal{O}_H(X) \doteq \arg \inf_H \{\gamma(X, H) : H \in \mathcal{H} \text{ and } X \subseteq H\}, \quad (3.12)$$

where $\gamma(\cdot, \cdot) : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ is a selection criterion of the homothetic bounding set. In other words, we search for the member H of \mathcal{H} , which optimizes the selection criterion γ such that the set $X \in H$. Exemplary, if we chose as selection criterion the scaling factor α , we aim to determine the ‘smallest’ member of the family \mathcal{H} in which the basic shape X still fits.

Proposition 4 (Homothetic bounding)

Given a basic shape $H \subset \mathbb{R}^{n_s}$, a selection s of variables of interest, and the solution set $f_s(\mathcal{Z}_{\text{SDP}})$. The homothetic outer-bounding set of $f_s(\mathcal{Z}_{\text{SDP}})$ is given by $\mathcal{O}_H(S) \doteq \mathcal{O}_H(f_s(\mathcal{Z}_{\text{SDP}}))$ (3.12).

Proof. Proof follows immediately from definition of the outer-bounding map (3.12) and Thm. 2. □

Exemplary, we consider the basic shape to be an irreducible polytopical set $H = \{h \in \mathbb{R}^{n_s} : A_h h \leq a_h\}$ with known matrix-vector pair $(A_h \in \mathbb{R}^{n_h \times n_s}, a_h \in \mathbb{R}^{n_h})$. Then we have:

$$\begin{aligned} (h^*, \alpha^*) = & \arg \min_{h, \alpha} \{\alpha^2\} \\ \text{s.t.} & \begin{pmatrix} A_h & a_h \\ -A_h & -a_h \end{pmatrix} \cdot \begin{pmatrix} h \\ \alpha \end{pmatrix} \leq \begin{pmatrix} A_h h_0 + a_h \alpha_0 \\ -\bar{h} \end{pmatrix}, \end{aligned} \quad (3.13)$$

where (h_0, α_0) defines the initial basic shape, and with $\bar{h} = (\bar{h}_1, \bar{h}_2, \dots, \bar{h}_{n_h})^T$ obtained by

$$\begin{aligned} \bar{h}_j &= \max_s \{A_h(j)s\} \\ \text{s.t.} \quad & s \in f_s(\mathcal{Z}_{\text{SDP}}), \end{aligned}$$

with $A_h(j)$ denoting the j -th row of A_h . Then, h^* and α^* are the orientation vector and the scaling factor respectively of the homothetic family of polytopes \mathcal{H} . The volume of such a polytope, as required later on, can only be evaluated implicitly, e.g. following Lawrence [1991], Sheynin and Tuzikov [2001].

The homothetic approach complements and generalizes the idea of using fixed shapes such as polytopes or ellipsoids for outer-bounding purposes. The choice of the basic shape H might depend on a particular application or on some quality criterion.

In general, low computational effort along with simple basic shapes, e.g. ellipsoids, however to obtain a good approximation quality, more complicated basic shapes have to be considered. Overall, the computational costs grow linearly with the number of constraints defining the basic shape n_h . In case the basic shape is a simple cube, we have $n_h = 2n_s$, and the computational costs equals to the interval bounding (see Section 3.3.1).

Thus, homothetic bounding may be advantageous in case knowledge about the shape of the solution set is available, e.g. obtained from a principal component analysis, see e.g. Jolliffe [2002]. However, neither homothety nor interval bounding e.g. allow to verify whether the solution set is composed of disconnected regions. To this end, a partitioning approach is considered next.

3.3.3. Partitioning

Interval bounding and homothety provide bounding sets of the desired variables with relatively low computational demands, in particular interval bounding. Though, using these approaches, only convex approximations of the solution sets can be obtained. Sometimes however, a more detailed analysis of the solution set is required. To obtain estimates of desired accuracy, we can consider a partitioning approach, e.g. bisectioning [Kuepfer et al., 2007]. In turn, this approach allows to study disconnected solution sets and possibly nonlinear correlations of variables.

To this end, we consider a partition of the initial bounding set $S = f_s(Z)$ (e.g. the initial parameter set) into a number of smaller subsets, i.e. $Q_j \subseteq S$, $j \in [1:q]$. Each of these partitions can then be analyzed separately for infeasibility, whereas infeasible partitions are discarded.

For partitioning, we consider the recursive bisectioning algorithm 1 up to some desired volumetric resolution ε . We denote the union of partitions Q_j for which $\text{SDP}^*(Z \cap Q_j)$ is not infeasible, i.e. $c_j \doteq \langle C, \Xi \rangle$ of $\text{SDP}^*(Z \cap Q_j)$ is bounded, by

$$\mathcal{O}_P(\mathcal{S}) \doteq \bigcup_{\substack{c_j < \infty, \\ 1 \leq j \leq q}} Q_j. \quad (3.14)$$

The union of feasible partitions $\mathcal{O}_P(\mathcal{S})$ provide an outer-estimate of the solution set as follows:

Proposition 5 (Partition)

$\mathcal{O}_P(\mathcal{S})$ (3.14) bounds the solution set $f_s(\mathcal{Z}) \subseteq \mathcal{O}_P(\mathcal{S})$, and is derived by $2^{n_s/\varepsilon}$ dual bounds of $SDP(\mathcal{Z})$.

Proof immediately follows from construction of $\mathcal{O}_P(\mathcal{S})$ and Theorem 1.

To obtain the desired partition estimate, the following recursive bisection algorithm can be used:

Algorithm 1 (Bisectioning(Q, ε))

1. If $SDP^*(Z \cap Q)$ is unbounded then exit and return \emptyset
2. If $\|Q\| \leq \varepsilon$ then exit and return Q
3. Partition Q into Q_1 and Q_2
4. Set $Q'_1 \doteq \text{Bisectioning}(Q_1)$
5. Set $Q'_2 \doteq \text{Bisectioning}(Q_2)$
6. Return $Q'_1 \cup Q'_2$

Depending on the particular problem at hand, several exploration strategies can be considered to increase performance (e.g. using multisection instead bisectioning). Nevertheless, the computational costs of the partitioning approach grow exponentially with the number of variables n_s as well as with the threshold ε . Although the algorithm can be easily and efficiently paralleled, this limits the applicability of the partitioning approach.

3.3.4. Branch-and-bound optimization

For the purpose of optimization, e.g. to obtain optimal parameter values, an optimality criterion is required. A classical choice for the objective function is e.g. the sum of least squares, i.e. to minimize the deviation of the model outputs and the observations.

Theorem 2 immediately provides a lower bound for the objective value and the respective argument. Though, due to relaxation errors, the lower bound may be pessimistic. To overcome this situation, a branch-and-bound approach can be considered.

To this end, we slightly modify Algorithm 1. Instead only testing for feasibility, we evaluate for each partition the lower bound of c_{slq} (if not unbounded), and assign the value obtained to the (feasible) partition Q_j . Thus, we obtain a ‘heat map’ of the solution set.

Together with strictly feasible solutions derived e.g. by numerical (Monte Carlo) methods or, when applicable, by analytic methods, a branch-and-bound optimization scheme is devised as follows. Given the objective value of a strictly feasible solution, all the partitions with lower bound larger than this one are discarded. The location and the value of the global optimum is thus narrowed, until a desired precision threshold is attained.

3.4. Computational notes

With the proposed approaches, we showed that the estimation problems translates into solving a number of SDPs of the form (3.7). This number in turn mainly depends on the considered estimation approach, e.g. two SDPs per variable for interval bounding. Advantageously, these SDPs are independent from one another, and hence parallelization can be considered.

However, the tractable size of a single SDP problem, given by the number of variables and constraints, is limited in practice for various reasons, e.g. memory demands. We next discuss some strategies which allow us to overcome such limitations.

3.4.1. Relaxation hierarchy

Convex optimization problems with a large number of constraints and variables can be treated efficiently only if their specific structure can be exploited [Mittelmann, 2003], e.g. sparsity or symmetry properties. For biochemical reaction networks, this however remains a challenging task, refer for a discussion to Rumschinski [2012], Rumschinski et al. [2012].

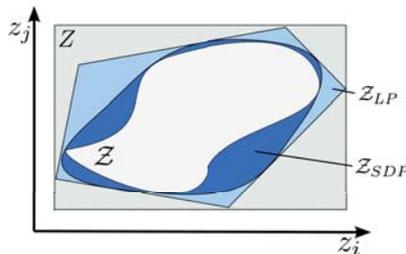


Figure 3.2.: Illustration of the hierarchy of convex relaxations.

A possibility to reduce the computational demands is by considering different relaxations chosen among a relaxation hierarchy (see Fig. 3.2 for an illustration). This comes as second order cone programming (SOCP) and linear programming (LP) allow handling much larger number of constraints and variables as semidefinite programming [Kim and Kojima, 2003, Kojima, 2002]. The SOCP and LP relaxations are obtained by replacing in (3.7) the positive semidefinite constraint $\Xi \succeq 0$ by the second order cone constraint $\Xi \in \mathcal{K}_2$ and the non-negativity constraint $\Xi \geq 0$ respectively. This yields less tighter but efficient dual bounds. A comparison of LP and SDP relaxations for polynomial problems can be found in Anstreicher [2009].

From experience with the current implementation [Streif et al., 2012], to obtain an infeasibility or dual certificate using semidefinite programming is in the order of fifty variables, linear programming allows to consider more than three hundred variables. Latter approximately corresponds to a dynamical system with ten state variables, each represented by twenty samples, and approx. twenty parameters.

3.4.2. Decomposition & integration of multiple experiments

Often, several input-output experiments are available for estimation and analysis, or different control objectives have to be met. Also, the experiments or control objectives might be correlated (e.g. share the same system parameters, the same initial state, etc.), and it is desirable to make use of such an interdependence.

For this, let D_1, D_2, \dots denote the available data sets, and let s denote the variables of interest. Each data set D_i relates to the solution set $\mathcal{S}_i = f_s(\mathcal{Z}_i)$, and the estimate $\mathcal{O}_I(\mathcal{S}_i)$, $\mathcal{O}_{\mathcal{H}}(\mathcal{S}_i)$, or $\mathcal{O}_P(\mathcal{S}_i)$ as obtained e.g. via Proposition 3. An overall estimate consistent with all available data sets D_i , is given by:

Proposition 6 (Multiple data sets)

Any enclosure of the solution sets \mathcal{S}_i or of their intersection $\bigcap_{j=1}^{n_j} \mathcal{S}_i$ encloses the solution set S .

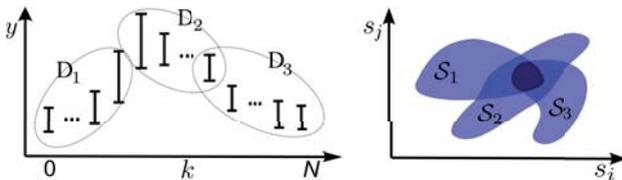


Figure 3.3.: Decomposition of the time window into smaller sequences that can be analyzed separately (left). The corresponding sub-estimates are intersected to obtain the desired estimate (right).

Hence, the integration of multiple data sets for estimation is straightforward, and obtained by intersection. Moreover, Proposition 6 can be applied for decomposition

purposes. To this end, let $0 \leq k \leq N$ be the time window to which the estimation problem is confined. Each solution set, associated with the estimation problem defined upon a smaller horizon, encloses by construction the overall solution set. Moreover, any intersection of solution sets of smaller problems enclose the overall solution set according to Proposition 6.

This decomposition approach can be generalized by reformulating $\text{SDP}(Z)$ into an equivalent problem with block-angular constraint matrix, whose structure can be exploited by Lagrangean relaxation (see e.g. Lemaréchal [2001], Sivaramakrishnan [2010]). Any decomposition strategy can be considered in principle, and a simple practical option is to consider a sliding time window of fixed size.

3.5. Summary

In this chapter, we provided a solution approach for the feasibility, estimation, and optimization problems. The framework applies to polynomial dynamical systems, and allows us to integrate uncertain and incomplete a priori, measurement, and structural data. The solution approach employs a convexifying relaxation, which transform the non-convex feasibility, estimation, and optimization problems into convex ones. Via duality, a sufficient criterion for infeasibility is provided, i.e. a solution to the feasibility problem 1. This will be used e.g. for model invalidation in the following chapter.

To answer the estimation problem 2, set-membership techniques were considered. The most efficient approach for outer-estimation consists in determining the uncertainty intervals for the unknown variables using dual certificates. These certificates were furthermore employed to obtain general fixed-shape membership sets using homothety. In addition, a partitioning approach based on infeasibility certificates and a bisection algorithm has been presented. Latter approach has been adapted to a branch-and-bound algorithm for answering the optimization problem 3.

A particular emphasis is put on balancing computational demands and precision. While in principle arbitrary precise results can be obtained, e.g. considering the Moment relaxation [Lasserre, 2001], in practice computational limitations have to be considered. We therefore elaborate on strategies to balance accuracy and computational costs, while still providing guaranteed results. A first and powerful option consists in choosing the class of dual certificates according to a relaxation hierarchy. This allows to choose, e.g., between semidefinite programming relaxations, yielding tightest although most expensive results, and more efficient linear programming relaxations. A second option regards the estimation technique employed. For estimation problems with many unknown variables (and constraints), interval bounding is recommended because it is the most efficient approach to obtain estimates. The most accurate estimates, which are often required for an detailed analysis e.g. about correlations of the variables, are obtained using partitioning. Additionally, general fixed shape sets can

be considered for outer-bounding using homothety. In particular by choosing simple, e.g. ellipsoids, or more complicated basic shapes, approximation quality and computational effort can be balanced. A third option to balance effort and precision is by decomposition respectively the integration of several interdependent experiments.

A limit of the framework is that it can not provide inner approximations, however, this can be overcome (when possible) by complementing the proposed approach with methods yielding regular solutions. With these methods at hand, we can next apply the framework to answer particular estimation and analysis problems.

4. Model Invalidation and Parameter Estimation

Modeling in systems biology is frequently complicated by limited knowledge about the studied process resulting in several competing hypotheses. Deciding on validity or invalidity of candidate hypotheses represents a particular challenge here, because in most situations the available data is uncertain and the parameters are unknown. Though, the selection of suitable hypotheses is only a first step toward the comprehensive quantitative description of a process. To this end, it is important to investigate how uncertainty affects the attainable precision of the parameter estimates, besides determining the optimal parameter values.

In Section 4.1, we consider a falsification approach for hypothesis testing based on the infeasibility certificates derived in the previous chapter. Thereafter, we focus on the parameter estimation problem in Section 4.2 by considering the outlined set-membership techniques. Parts of this chapter, particularly the carnitine example, are based on Rumschinski et al. [2010a].

4.1. Model invalidation

Limited knowledge about the process, e.g. the underlying biochemical reactions and mechanisms, often results in competing hypotheses. For instance, when modeling a biochemical reaction network, the specific type of an interaction may be unknown, e.g. if mass action or Hill kinetics has to be considered. Modeling hence often requires, prior to estimation of the unknown parameters, to evaluate and if possible reject some of the hypotheses by taking all available data into account, which is known as *model invalidation*.

By now, very different approaches exist for model invalidation. An overview on model invalidation approaches for systems biology can for example be found in Kremling et al. [2004]. Exemplary, the issue can be approached from an analytic perspective, i.e. to investigate the conditions such that two (or more) hypotheses can be distinguished. Schnell et al. [2006] proposed an approach which analyzes the *structural distinguishability*, i.e. if a transformation exists so as to transfer a system (hypothesis) into the other one. Whenever such a transformation exists, models can not be structurally distinguished. However, the proposed approach only provides a partial answer to the distinguishability problem, as the approach does not provide means to

distinguish model alternatives based on available data: even if two hypotheses are analytically distinguishable, they might not be regarding the measurements.

Therefore, a criterion is required which allows to distinguish model alternatives using the available data. Intuitively, such a criterion might consist in choosing the ‘most appropriate’ hypothesis, i.e. a validation approach. An example is the Bayesian information criterion. Here, a score, based on *maximum likelihood* and the number of parameters, is assigned to each hypothesis. Among these hypotheses the one with the minimum score is the most suitable one, see e.g. Jaqaman and Danuser [2006], which however does not mean that this hypothesis actually allows to describe the observed behavior. A time domain validation approach has been suggested in Poolla et al. [1994]. A main criticism hereby is that validation criteria are inherently subjective [Oreskes and Belitz, 2001, Smith and Doyle, 1992].

An alternative approach consists in *falsification*, i.e. to reject a model hypothesis if it does not meet a certain criterion, e.g. if it is not *consistent* with the available data inclusively the observed behavior. One example for a stochastic falsification approach is the Neyman-Pearson Lemma. It is based on minimizing the probability of false acceptance [Lehmann, 2009], and for this purpose the maximum likelihood ratio can be considered as criterion. One general observation is that models with a larger number of parameters are more flexible and fit the data ‘better’ than models with a smaller number of parameters [Jaqaman and Danuser, 2006]. In addition, the F-test can be used to clarify if the introduction of extra parameters, i.e. a more complicated hypothesis, is justified, see e.g. Jaqaman and Danuser [2006]. Several alternatives to stochastic approaches have been considered so far, which are based on showing *inconsistency* of the model with the available data. For example, Anderson and Papachristodoulou [2009] considered a worst-case approach using real algebraic geometry and semidefinite programming. Prajna [2006] used barrier certificates to show inconsistency, and Kuepfer et al. [2007] derived infeasibility certificates using semidefinite relaxations applicable to *stationary* systems.

Here, we pursue a similar approach as in Kuepfer et al. [2007] for model invalidation considering dynamic systems and uncertain data. To this end, consider a model hypothesis M (3.1) and the available data D (3.2), and note that the parameters are considered unknown or bounded (initial parameter region) if such information is available. Model falsification then consists in proving that the model and the data are *inconsistent*; i.e., there exists not a single trajectory (solution) of the model that ‘touches’ the data. In this case, we can conclude that the model is invalid (with respect to the initial parameter bounds if available). This can be formally stated by the following infeasibility formulation of model invalidation:

Corollary 1 (Model invalidation)

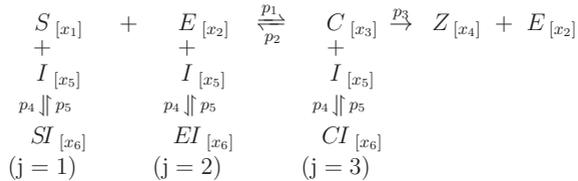
If the dual problem $SDP^(Z)$ is unbounded, then model (3.1) is inconsistent with the data (3.2), and therefore considered invalid.*

Proof is an immediate consequence of Theorem 2.

Remark 2 *The proposed rejection criterion is rigorous only under the premise that the uncertainty of the data is appropriately described. Particularly, we do not consider here measurement outliers in the data set. Because this is a strong assumption which may not be guaranteed in practice, we relax this assumption later on in Section 5.3, so as to admit some, non-consecutive outliers in the data set.*

Example 4.1: inhibition mechanisms

We consider three possible inhibition mechanisms of a Michaelis-Menten reaction [Cornish-Bowden, 2004], depicted in the following enzymatic reaction scheme:



Here, the substrate S is catalyzed by an enzyme E via the complex C to form a product Z . Three possible inhibition mechanisms [Cornish-Bowden, 2004] may interfere the reaction. An inhibitor I may (reversibly) bind either the substrate ('substrate inhibition', $j = 1$)¹, the enzyme ('competitive inhibition', $j = 2$), or the complex ('uncompetitive inhibition', $j = 3$).

We assume the law mass action, and consider the Euler backward discretization (see Appendix A.2) with time step size $h = 0.05$, which is sufficiently small in the considered setup². This way, we derive a discrete-time model for each of the three possible inhibition mechanisms, given by:

$$\begin{aligned}
 x_1^+ &= x_1 - \nu_1 + \nu_2 + \delta_{j1}(-\nu_4(j) + \nu_5) \\
 x_2^+ &= x_2 - \nu_1 + \nu_2 + \nu_3 + \delta_{j2}(-\nu_4(j) + \nu_5) \\
 x_3^+ &= x_3 + \nu_1 - \nu_2 - \nu_3 + \delta_{j3}(-\nu_4(j) + \nu_5) \\
 x_4^+ &= x_4 + \nu_3 \\
 x_5^+ &= x_5 - \nu_4(j) + \nu_5 \\
 x_6^+ &= x_6 + \nu_4(j) - \nu_5,
 \end{aligned}$$

where $x_i = x_i[k]$ the current and $x_i^+ = x_i[k+1]$ the future state, the Kronecker δ_{ji} with $j \in [1 : 3]$, and $\nu_1 = hp_1x_1x_2$, $\nu_2 = hp_2x_3$, $\nu_3 = hp_3x_3$, $\nu_4(j) = hp_4x_jx_5$, $\nu_5 = hp_5x_6$. In

¹this mechanism can be considered as an additional pathway which utilizes the substrate

²Note that choosing an appropriate step size is in general difficult, e.g. because the discretization error varies with the (unknown) parameters.

addition, mass conservation implies the relations $x_1+x_3+x_4+x_6 = p_6$, $x_2+x_3+x_6 = p_7$, and $x_5+x_6 = p_8$, where p_i are non-negative constants. The step size is $h = 0.05$.

We next demonstrate the ability of our approach to discriminate the three candidate mechanisms, and furthermore to evaluate the measurement error margins that still lead to rejection. For this, we take one inhibition mechanism as reference, generate measurements, and invalidate the remaining two hypotheses. For the sake of completeness, we check each possible combination of reference and hypothesis.

Reference model We consider either inhibition scheme ($j = \{1, 2, 3\}$) with reference parameters $p = (5, 1, 1, 5, 1, 1, 0.1, 0.1)$ and initial condition $x[0] = (1, 0.1, 0, 0.1, 0, 0)$, resulting in the reference trajectories $x^{(j)}[k]$, $k \in [0:N]^3$. We generate the uncertain measurements with the relative error (η_r) as given in Table 4.1:

$$x_i[k] \in \left[(1 - \eta_r)x^{(j)}[k], (1 + \eta_r)x^{(j)}[k] \right].$$

Candidate models We consider a reference inhibition mechanism to be chosen, and the remaining two mechanisms are tested for invalidity. We fix $p_6 = 1$, $p_7 = 0.1$, $p_8 = 0.1$, and the remaining five parameters are considered non-negative, and unknown to the extend of two orders of magnitudes

$$P = \{p \in \mathbb{R}^5 : p_i \in [0.1, 10], i = 1, \dots, 5\}. \quad (4.1)$$

Results The necessary computations were conducted using the ADMIT toolbox [Streif et al., 2012] on a standard 2.4 GHz Intel desktop with 4 GB RAM. The computation time to obtain the required infeasibility certificates varied with number of measurements considered, i.e. between in average approximately 0.1 s for two measurements and in average approximately 0.5 s for ten measurements.

Table 4.1 shows the estimated maximum error margins η_r of the measurement data which still lead to rejection of the other two hypotheses. In case of uncompetitive inhibition, two measurements are insufficient for invalidation. The results show however that the error margins (η_r) are increasing with the number of available measurements. For example, in case of substrate inhibition and considering ten measurements, error margins of 42% and 45% still allows to reject the hypotheses uncompetitive and competitive inhibition mechanism respectively.

4.2. Parameter estimation

For modeling, the selection of suitable hypotheses, e.g. by applying above invalidity criterion, is only a first step toward comprehensive description of the process. Particularly, the models which can not be rejected need to be further analyzed. To this

³We here consider that all states are measured, though this is not necessary in general

Table 4.1.: Model invalidation (inhibition mechanisms). Shown are the maximum relative error margins (η_r) allowing to reject the respective hypotheses.

Reference	Candidate	# of Measurements (k)			
		2	3	5	10
competitive inhibition	substrate inhibition	0.5%	1.0%	2.0%	3.0%
	uncompetitive inhibition	4.5%	6.0%	12%	13%
uncompetitive inhibition	substrate inhibition	—	8.5%	14%	15%
	competitive inhibition	—	11%	23%	27%
substrate inhibition	competitive inhibition	20%	28%	39%	45%
	uncompetitive inhibition	23%	32%	40%	42%

end, it is necessary to determine the unknown model parameters, and to evaluate the effects of measurement uncertainty on the parameter estimates.

We here address parameter estimation from two perspectives. Firstly, we show how to obtain the set of consistent parameters, i.e. all parameters for which the model admits a solutions consistent with the data. To this end, the set-membership techniques outlined in Section 3.3. This way, we obtain also information about the sensitivity and possible correlations of the parameters. Secondly, we address the problem of finding optimal parameter values regarding the sum of least squares (4.3).

The setup for parameter estimation is as follows: we consider a model M (3.1), and the available data D (3.2), consisting of a priori knowledge, measurements with an appropriate uncertainty description, and possibly structural data. The n_p parameters are unknown.

We denote for simplicity of notation the desired solution set by $\mathcal{P} \doteq f_p(\mathcal{Z}) \subset \mathbb{R}^{n_p}$. The solution set can be approximated following any of the proposed set-membership estimation methods, in particular interval bounding, homothety, and partitioning as outlined in Section 3.3. We start with the uncertainty intervals, i.e. the orthotopic projections of the solution set \mathcal{P} . They can be obtained - for each parameter - following Proposition 2, summarized as follows:

Corollary 2 (Parameter uncertainty interval)

The parameter uncertainty interval $[p_i, \bar{p}_i]$ is derived from 2 dual bounds (Theorem 2).

Thus, overall $2n_p$ dual bounds are required to obtain an orthotopic parameter estimate, i.e. the computational complexity growth linearly with the number of parameters.

Remark 3 *Under the condition that measurement uncertainties are described by their n -sigma confidence intervals, the obtained uncertainty intervals can be interpreted as the outer-approximation of the n -sigma parameter confidence intervals.*

The uncertainty intervals provide valuable information about the precision and the sensitivity of the parameters as shown next.

Parameter sensitivity Because any parameter value outside the interval leads to invalidity of the model by construction, the larger the interval, the less important is a parameter variation regarding invalidity; vice versa, a parameter is sensitive, if already small variations leads to rejection of the model (hypothesis). To measure this ‘sensitivity’ of the parameters, we evaluate the largest possible variation of a parameter p_i which does not lead to rejection; the sensitivity is derived from the interval bounds of the parameters, given by

$$\xi = \xi(p_i) \doteq \sqrt{\frac{\bar{p}_i}{\underline{p}_i}}. \quad (4.2)$$

By definition, we have $0 \leq \xi \leq 1$. The closer the sensitivity coefficient ξ of a parameter p_i is to 1, the more sensitive is the parameter ($\xi = 1$ means that already a small variation of the parameter leads to rejection of the model). Sensitive parameters have sensitivity coefficients between $0.5 \leq \xi \leq 1$, i.e. less than a 4-fold variation of the nominal parameter is possible. Values between $0.1 \leq \xi \leq 0.5$ indicate less sensitive, and $\xi \leq 0.1$ insensitive parameters (i.e. more than 100-fold variation is possible).

Parameter insensitivity may result from over-parametrization of the model, i.e. the available information is not sufficient to determine all degrees of freedom (unknown parameters). In these cases, it may be advantageous to perform a correlation analysis as shown next.

Parameter correlation analysis An efficient approach to study parameter correlations is by using homothety as outlined in Proposition 4. Here, differently oriented basic shapes $H \in \mathbb{R}^{n_s}$ can be used, e.g. by comparison of the volume, to obtain an idea how the parameters are correlated. Though, this approach is limited to convex basic shapes, and hence nonlinear correlation may not be suitably explored. To overcome this limitation, parameter correlations can be studied by partitioning, following Proposition 5:

Corollary 3 (Partitioning)

The set of consistent parameters \mathcal{P} is bounded by the partition estimate $\mathcal{O}_P(\mathcal{P})$, derivable by $2^{n_p/\varepsilon}$ dual bounds.

Because of the exponential complexity of the partitioning algorithm, typically only a subset of parameters is analyzed, e.g. to answer how a pair $s = \{p_i, p_j\}$ or a triple $s = \{p_i, p_j, p_k\}$ of parameters are correlated among another. The respective partition estimate then corresponds to the projection of the feasible solution set $f_p(\mathcal{Z})$ onto the desired subspace. The analysis can be combined e.g. with a principal component analysis [Le Roux and Rouanet, 2004].

Optimal parameter estimation For optimal parameter estimation, we consider as objective function the sum of least squares, i.e. to minimize the deviation of the model outputs and the observations. The weighted sum of least squares objective function is given by

$$c_{slq} \doteq \sum_{\substack{1 \leq i \leq n_y \\ 1 \leq j \leq M}} \frac{1}{a_i} (y_i(t_j) - \tilde{y}_i(t_j))^2, \quad (4.3)$$

where a_i denotes the weighting factors and $y_i(t_j)/\tilde{y}_i(t_j)$ the model output/observation y_i at t_j respectively.

Theorem 2 immediately provides a lower bound for the objective value of the sum of least squares. Though, due to relaxation errors, the lower bound may be pessimistic. To overcome this situation, a branch-and-bound approach is considered. To this end, we slightly modify Algorithm 1. Instead only testing for feasibility, we evaluate for each partition the lower bound of c_{slq} (if not unbounded), and assign the value obtained to the (feasible) partition Q_j . Thus, we obtain a ‘heat map’ of the solution set.

Together with strictly feasible solutions derived e.g. by numerical (Monte Carlo) methods or, when applicable, by analytic methods, a branch-and-bound optimization scheme is devised as follows. Given the objective value of a strictly feasible solution, all the partitions with lower bound larger than this one are discarded. The location and the value of the global optimum is thus narrowed, until a desired precision threshold is attained.

Example 4.2: competitive inhibition

Setup We consider again the competitive inhibition ($j = 2$) from the previous example in Section 4.1 with the reference parameter and initial condition $p^* = (10, 1, 1, 10, 1, 1, 0.1, 0.1)$ and $x^*[0] = (1, 0.1, 0, 0.1, 0, 0)$ respectively, resulting in the reference sequence $x_i^*[k]$, $k \in [0:40]$, $i \in [1:6]$. We generate artificial measurements by considering the relative error $\eta_r = 0.1$ and additional (randomly generated) absolute measurement errors $|\eta_a[k]| \leq 0.05$ for $k \in [0:40]$. The uncertain and incomplete measurements used in the following are described by the intervals

$$x_i[k] \in \left[(1 - \eta_r)(x_i^*[k] - \eta_a[k]), (1 + \eta_r)(x_i^*[k] + \eta_a[k]) \right] \quad (4.4)$$

for $i \in \{1, 2, 4, 5\}$ ⁴ and $k \in [0:40]$.

All eight parameters are unknown to the extend of four orders of magnitudes:

$$P = \{p \in \mathbb{R}^8 : p_i \in [0.01, 100], i = 1, \dots, 8\}. \quad (4.5)$$

Results We estimate the parameter uncertainty intervals according to Cor. 2 using the ADMIT toolbox [Streif et al., 2012]. The results are shown in Figure 4.1 (left).

⁴we consider that x_3 and x_6 cannot be measured

We found that all parameters are sensitive, with sensitivity coefficients ranging from 0.71–0.98.

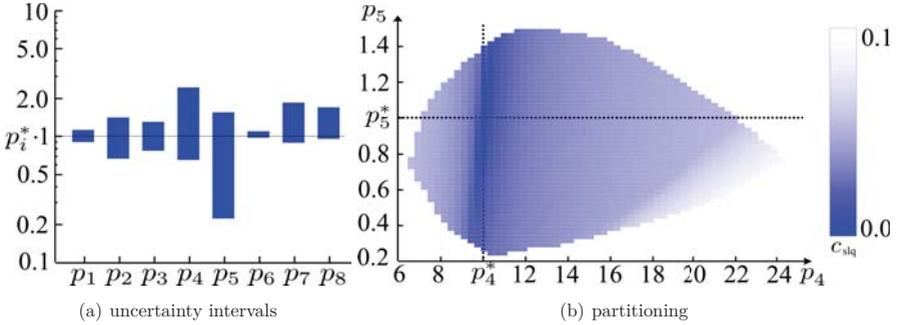


Figure 4.1: Parameter estimation (competitive inhibition). 4.1(a): Estimated parameter uncertainty intervals. 4.1(b): Partitioning and heat map of the consistent parameter set $\{p_4, p_5\}$. Dashed lines indicate the reference parameters.

We furthermore analyze the correlation of the parameters p_4 and p_5 in more detail, and determine the optimal parameter values regarding sum of least squares

$$\sum_{\substack{i \in I, \\ 0 \leq k \leq 40}} (x_i[k] - x_i^*[k])^2 \quad (4.6)$$

where $I = \{1, 2, 4, 5\}$. To this end, we apply Corollary 3 and the bisection algorithm (Alg. 1) to obtain the partitioning estimate of the pair $\{p_4, p_5\}$. The result is depicted in Fig. 4.1 (right).

We furthermore associate to each (feasible) partition the respective lower bound of the objective, using blue-scale heat map for the objective values from blue (partitions with least sum of squares) to white (infeasible partitions). As it can be seen from Fig. 4.1 (right), no clear correlation of the parameters p_4 and p_5 is apparent. The partition with the minimum sum of least squares coincides with the one containing the reference values of both parameters as expected.

Example 4.3: carnitine shuttle

We next consider the carnitine shuttle mechanism, which is a well known intra-cellular transport system for fatty acids. This example is outlined in detail in Rumschinski et al. [2010a], and taken from there to demonstrate the influence of uncertainty, sparsity and incompleteness of measurements, as well as a priori knowledge on the quality of the parameter estimates.

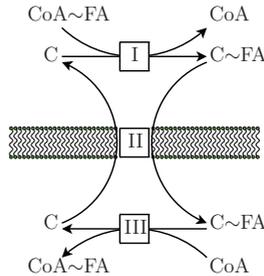


Figure 4.2.: Scheme of the carnitine shuttling system. Activated fatty acid ($\text{CoA}\sim\text{FA}$) are transferred to carnitine (C) via carnitine-acyltransferase [I] at the cytoplasm. The carnitine-fatty acid complex ($\text{C}\sim\text{FA}$) is then shuttled via a so called antiporter [II] into the mitochondria in exchange for a free carnitine. There, a mitochondrial isoform of the carnitine-acyltransferase [III] reactivates via Coenzyme A (CoA) the fatty acids. The activated fatty acid inside the mitochondria is a precursor for β -oxidation. Note that reactions [I] and [III] are reversible.

The carnitine shuttle is a step of mitochondrial β -oxidation, and is an important mechanism for fat catabolism. The considered reaction scheme (see Figure 4.2) is adapted from Bremer [1983], and models a specific transport system at the inner mitochondrial membrane involving fatty acids (FA), carnitine (C) and Coenzyme A (CoA). An activated fatty acid ($\text{CoA}\sim\text{FA}$) is transferred to carnitine (C) via carnitine-acyltransferase at the cytoplasm (reaction I). The carnitine-fatty acid complex ($\text{C}\sim\text{FA}$) is then shuttled via a so called antiporter into the mitochondria in exchange for a free carnitine (reaction II). There, a mitochondrial isoform of the carnitine-acyltransferase reactivates via Coenzyme A (CoA) the fatty acids (reaction III). The activated fatty acid inside the mitochondria is a precursor for β -oxidation. Note that reactions I and III are reversible.

Table 4.2.: Description of the variables for the carnitine shuttle model.

*

Symbol	Specie
x_1	$\text{CoA}\sim\text{FA}$ (Cy)
x_2	Carnitine (Cy)
x_3	$\text{CoA}\sim\text{FA}$ (Mi)
x_4	Carnitine (Mi)

By considering mass action kinetics and taking into account the conservation moieties, se e.g. Bremer [1983], the dynamic of the shuttle system can be expressed by the

following set of ordinary differential equations

$$\begin{aligned}\dot{x}_1 &= -p_2x_1x_2 + p_3(C_0 - x_2) + p_1u \\ \dot{x}_2 &= -p_2x_1x_2 + p_3(C_0 - x_2) + p_4(C_0 - x_2)x_4 \\ \dot{x}_3 &= -p_2x_3x_4 + p_3(C_0^m - x_4) - p_5x_3 \\ \dot{x}_4 &= -p_2x_3x_4 + p_3(C_0^m - x_4) - p_4(C_0 - x_2)x_4,\end{aligned}$$

where the variables $x_1 \dots x_4$ correspond to the participating compounds (as described in Table 4.2), the parameters $p_1 \dots p_5$ denote the (unknown) constant reaction rates, C_0 and C_0^m represent the initial concentrations of carnitine respectively outside and inside the mitochondria, and the input u is regarded as a binary function corresponding to active ($u = 1$) and inactive ($u = 0$, fat starvation) β -oxidation.

Applying Euler forward discretization (Appendix A.1, the difference equations for the above continuous-time model are given by

$$\begin{aligned}x_1^\dagger &= x_1 + h[-p_2x_1x_2 + p_3(C_0 - x_2) + p_1u] \\ x_2^\dagger &= x_2 + h[-p_2x_1x_2 + p_3(C_0 - x_2) + p_4(C_0 - x_2)x_4] \\ x_3^\dagger &= x_3 + h[-p_2x_3x_4 + p_3(C_0^m - x_4) - p_5x_3] \\ x_4^\dagger &= x_4 + h[-p_2x_3x_4 + p_3(C_0^m - x_4) - p_4(C_0 - x_2)x_4],\end{aligned}$$

where h is the discretization time-step, $x_i = x_i[k]$ the current and $x_i^\dagger = x_i[k + 1]$ the future state, and $u = u[k]$ the current input.

Scenarios and Setup

The discrete-time model has been simulated with time step $h = 5$ seconds using the reference parametrization p^* and initial condition as in Table 4.3, with values chosen from the literature [Bieber, 1988, Lysiak et al., 1986]. To test the robustness of the approach and study the influence of measurement quality and availability on the resulting estimates, we compare several experimental scenarios derived from the above simulation. Each scenario is obtained as a combination of the following options, as summarized in Table 4.4.

- A priori knowledge.** Two prior knowledge cases are considered, denoted by 3-PAR and 5-PAR. In the former case, parameters p_1 and p_5 are known with relative bounds $[0.95, 1.05]$, while the remaining parameters p_2, p_3, p_4 are unknown. In the latter case, all five parameters are unknown. For the unknown parameters we assume as initial bounds the relative bounds $[0.3, 3]$. C_0 and C_0^m are treated in the difference equations as constants, with values as in Table 4.3. Here relative bounds $[lb, ub]$ for a parameter p_i mean $p_i \in [lb \cdot p_i^*, ub \cdot p_i^*]$.

Table 4.3.: Reference values for parameters and initial conditions for the simulation of the carnitine shuttle model. The initial conditions are given by $x_1 = 0$, $x_2 = C_0$, $x_3 = 0$ and $x_4 = C_0^m$.

Symbol	Value	Unit
p_1^*	5.00e-4	$\mu M s^{-1}$
p_2^*	1.03e-1	$\mu(Ms)^{-1}$
p_3^*	2.36e-2	s^{-1}
p_4^*	1.85e-2	$\mu(Ms)^{-1}$
p_5^*	2.50e-2	s^{-1}
C_0	0.33	μM
C_0^m	1.00	μM

Table 4.4.: Summary of the measurement and knowledge options. Each parameter estimate scenario is obtained by selecting a value for each of the options.

Scenario Type	Options			
prior knowledge	3-PAR		5-PAR	
measurement density	DENSE		SPARSE	
measurement error	ERR-1%	ERR-2%	ERR-4%	
measured concentrations	ALL	NOT-X3	NOT-X4	NOT-X3-X4

- **Measurement density.** We consider two measurement density options, denoted DENSE and SPARSE. The former consists of two sequences of 15 consecutive measurements, taken in the transient phase ($k = 0, \dots, 14$) and in the equilibrium phase ($k = 300, \dots, 314$) respectively. The latter consists of two sequences of only five measurements, taken in the transient phase ($k = 0, 3, 5, 10, 14$) and in the equilibrium phase ($k = 300, 303, 305, 310, 314$) respectively.
- **Measurement errors.** To analyze the influence of measurement errors, we consider the three options ERR-1%, ERR-2%, and ERR-4%, with respectively 1%, 2% and 4% relative error (see Grube et al. [2005], Okamura et al. [2006] for examples of practical measurement errors compatible with our setup).
- **Measured concentrations.** The influence of incomplete measurements is also investigated. We consider four different cases, denoted ALL, NOT-X3, NOT-X4, and

NOT-X3-X4, where respectively all concentrations, all concentrations but x_3 , all concentrations but x_4 , and only concentrations x_1 and x_2 are measured. This choice reflects the fact that the inner mitochondrial concentrations x_3 and x_4 are more difficult to measure with simple techniques.

For each of the resulting $2^2 \times 3 \times 4 = 48$ different experimental scenarios, the parameter uncertainty intervals are estimated.

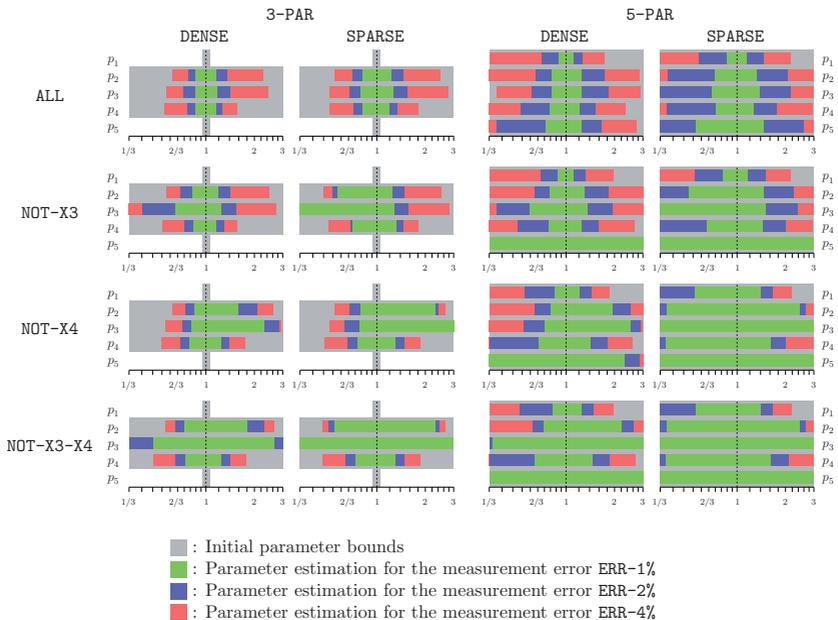


Figure 4.3.: Parameter estimation results for the carnitine shuttle example. In the first group of tests the parameters p_1 and p_5 are known (relative bounds fixed to $[0.95, 1.05]$), and the three remaining ones are unknown (initial relative bounds $[0.3, 3]$). In the second group all five parameters are unknown. The rows of the table report the parameter estimation results for different state-measurement scenarios.

The relative bounds resulting from parameter estimation are summarized in Figure 4.3 for all the considered scenarios. The figure is structured in a table-like fashion, with groups of experimental scenarios arranged from highest information (top-left) to lowest information (bottom-right). In each group, the bounds for the three error-measurement options are reported as nested intervals, using different colors.

The results clearly indicate that measurement error has a substantial impact on the estimates. With measurement error ERR-1%, the unknown parameters can be

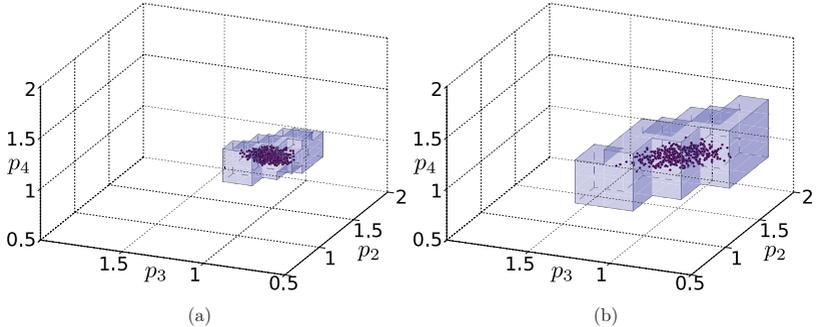


Figure 4.4.: Consistent parameter estimate $\mathcal{O}_P(\mathcal{P})$ for the scenario(3-PAR, DENSE, NOT-X3, ERR-1%), left, and (3-PAR, DENSE, NOT-X3, ERR-2%), right. The dots show consistent parameterizations obtained by Monte Carlo tests. The coordinate axes show values relative to the reference parameter p^* .

narrowed with sufficient precision for most scenarios. Conversely, with measurement error ERR-4%, reasonable estimates can only be obtained for the 3-PAR case, where the additional prior knowledge compensates for the larger uncertainty.

As for the influence of incomplete measurements, while clearly the best results are obtained when all species are measured (ALL), some improvements can still be obtained with incomplete measurements, in particular for the case NOT-X3. Note however that the bounds on parameter p_5 cannot be improved when x_3 is not measured (cases NOT-X3 and NOT-X3-X4), as p_5 only appears in the difference equation of x_3 . Considering the 3-PAR case, it is also interesting to note that the cases NOT-X3 and NOT-X4 have opposite effects on the estimates, improving more the upper and the lower parameter bounds respectively. As a remark, we noted in our tests that uncertainties with respect to x_2 (the carnitine-FA complex) have overall the largest impact on the quality of the parameters estimates.

Comparing the SPARSE and DENSE scenarios, it can be seen that very similar results are obtained when prior knowledge is available (3-PAR). As it can be expected, the impact of measurement errors is in general more noticeable for the SPARSE cases.

The bounds in Fig. 4.3 are the single-parameter projections of the actual bounding sets obtained with Alg 1. These sets, which provide additional information on the correlation among the parameters, are depicted for the scenarios (3-PAR, DENSE, NOT-X3, ERR-1%) and (3-PAR, DENSE, NOT-X3, ERR-2%) in Fig. 4.4(a) and 4.4(b) respectively. To indicate the estimate quality, some consistent parameterizations derived by Monte Carlo simulations are also plotted.

Note that this is a qualitative comparison, as the probability of finding a consistent parametrization is not uniform. Conversely, our approach guarantees that outside of the indicated sets there is *no* consistent parametrization. Considering larger initial bounds for the unknown parameters does not affect the results for the 15 scenarios

out of 48 in which all bounds are strictly improved. In the other cases, the parameter estimates will depend on how the feasible region extends outside of the area we considered.

4.3. Summary and discussion

Modeling in systems biology is often complicated by limited knowledge about the process. This frequently results in competing hypotheses, whereas the parameters might be completely unknown. Therefore, a reliable criterion to evaluate and to reject the hypotheses is required regarding the available data. We proposed a model selection criterion based on falsification, that is if the model hypothesis and the data are inconsistent, respectively the relaxed problem does not admit a solution, i.e. an infeasibility certificate can be provided. Hereby, the model parameters and, possibly, the initial conditions, do not need to be known. The falsification is guaranteed, i.e. a ‘valid’ model can this way never be rejected; however, as ‘spurious’ solutions might have been introduced by the relaxation, an invalid model might be considered as valid. This can for instance occur if only few data are available, e.g. no a priori data for the model parameters is available. Though, in this case, by estimating the parameters and iteratively updating the parameter estimates, the relaxation is tightened, and a certificate of infeasibility and hence model invalidity might result after refining the parameters.

We furthermore applied the set-membership framework to parameter estimation in this chapter. We take into account that not all states are measured, as it is frequently the case for the transient phase of biological experiments. Moreover, it is possible to include uncertain but set-bounded measurable inputs and disturbances. In case the (measurement) data uncertainty bounding sets correspond to (n-sigma) confidence-levels of the a priori uncertainty distributions, the obtained uncertainty intervals can be interpreted as the (n-sigma) parameter confidence intervals; this way, a comparison with established re-sampling techniques can be considered.

To investigate how data uncertainty influences the parameter estimates, and hence to evaluate the precision and sensitivity of the parameters, the uncertainty intervals are determined. Here, the evaluation of a single uncertainty interval requires to solve two semidefinite problems. Although the estimation should be iterated, as tightening of one variable might propagate to other variables, the number of computations, i.e. solving the relaxed problem, linearly depend on the number of parameters, and can therefore considered for systems with many unknown parameters.

Obtaining the optimal parameters, e.g. regarding a least squares objective, is typically required for (model-based) applications, e.g. control synthesis and prediction purposes. The optimal parameters can be determined by partitioning of the (initial)

parameter space and evaluating the lower bound of the objective value for each partition.

The achievable results will however depend on the problem at hand. If for instance only few measurements with large uncertainty are given, a successful result will rely on the available prior knowledge. Also, limited estimability with respect to measurement and parameter uncertainties is an intrinsic limit when dealing with guaranteed bounds. In such a case, it can be advantageous to study possible correlations among the parameters, e.g. using the proposed homothetic approach or via partitioning.

5. State Estimation, Reachability, and Outlier Analysis

In the previous chapter, we considered model invalidation and parameter estimation given uncertain measurements. The proposed methods yield rigorous results in case measurement uncertainties are appropriately described. Otherwise, in particular if uncertainties are underestimated, it is possible that a valid model is falsely rejected. For this reason, *outliers* in measurement data constitute a particular challenge to the considered framework, and it is thus important to identify outliers prior to hypothesis testing and parameter estimation.

Furthermore, we aim to investigate how uncertainty in initial conditions and parameters propagates to the model outputs. This is of particular relevance because investigating only the nominal system behavior (e.g. regarding fixed parameters and initial condition) does not provide insight into qualitative features such as the robustness of the model in general.

In this chapter, we analyze how uncertainties propagate to the model outputs and detect outliers in the measurements by estimating the model states under parametric uncertainty. Therefore, we first consider state estimation and provide an interval observer in Section 5.1. Then, we analyze reachability properties of the systems and show how to obtain reachable sets in Section 5.2. In Section 5.3, we finally approach the issue of detecting outliers in dependent data.

5.1. State estimation

State estimation denotes the problem of reconstructing unknown or missing state values, e.g. intra-cellular concentrations that cannot be assessed directly. For many processes including biological ones, very often not all states can be measured (i.e. *permanently* missing values), or the measurement process might have been interrupted, leading to *casually* missing values. In both cases, it might be required to reconstruct the missing state values. State estimation can be categorized depending on whether past, current, or future states are desired, which are respectively known as retrodiction, online estimation, and prediction van der Heijden et al. [2004].

The classical setup for state estimation is stochastic, where the process and measurement noises are assumed to be Gaussian [Raïssi et al., 2010]. Kalman filtering and particle filtering are two representative methods, see Lendek et al. [2006]. Former works,

under the assumption of linear state space systems, and that both state transition and the measurement noise is unbiased and normally distributed, and is then shown to be an optimal estimator in the mean least squares sense [Lendek et al., 2006]. The Kalman filtering has been extended to the nonlinear case with non-Gaussian noise, e.g. by considering the linearized system around the current estimate (extended Kalman filtering Anderson and Moore [1979]), or by considering the unscented Kalman filter, see e.g. Wan and Van Der Merwe [2000]. However, particularly if the parameters of the system are not exactly known, or the measurement uncertainties are nonhomogeneously distributed, the Kalman filter might fail. To overcome these problems, particle filtering, also known as sequential Monte Carlo methods, can be considered. A survey about particle filters is given e.g. in Crisan and Doucet [2002]. These methods use many random samples (i.e. particles) to represent the posterior probability distributions, which are then propagated over time [Crisan and Doucet, 2002]. Although this approach is generally applicable to nonlinear systems and nonhomogeneous noise, there exists no general rule how to choose ‘representative’ samples, and the required number of samples grow exponentially with the number of variables.

As for state estimation in membership settings, three main approaches have been considered so far (compare also the review given in Raïssi et al. [2010]): A prediction/correction mechanism as e.g. proposed in Jaulin [2002]. Applicability of this approach however is limited, due to the wrapping effect, to small (measurement and parametric) uncertainties [Raïssi et al., 2012]. The second approach (e.g. Kieffer and Walter [2006]) is again a prediction/correction approach, whereas the Müller Theorem [Müller, 1927] is applied for the prediction step, and third closed loop interval observers Bernard and Gouzé [2004], Mazenc and Bernard [2011], Moisan et al. [2009], which have in parts also been applied to bioengineering processes. These interval observers have been devised initially for exponentially stable linear systems, where certain robustness guarantees can be provided, see e.g. Gouzé et al. [2000] and Mazenc and Bernard [2011]. Recently, the approach has been extended to some classes of nonlinear systems considering e.g. partial exact discretization [Raïssi et al., 2012].

As setup for state estimation in our setting, we consider a model M (3.1) with unknown parameters $p \in P$, and summarize the available data, consisting in particular of a priori knowledge, measurements, an structural data by D (3.2).

We confine the state estimation problem to the time window $[\alpha : \beta]$, with $\alpha \leq 0 \leq N \leq \beta$. We aim to provide an interval estimate of the (missing) state variables $x_k \in \mathbb{R}^{n_x}$ for a $k \in [\alpha : \beta]$. We denote the desired consistent state set by $\mathcal{X}_k = f_{x_k}(\mathcal{Z})$, with $f_{x_k} : \mathbb{R}^{n_z} \rightarrow \mathbb{R}^{n_x}$ the respective projection map. The result follows from Theorem 2 and Proposition 2:

Corollary 4 (State uncertainty intervals)

The state uncertainty intervals $[\underline{x}_k, \bar{x}_k]$ are found by $2n_x$ dual bounds.

By appropriately choosing α and β , the retrodiction, online, and prediction problems are addressed.

Remark 4 *The applicability of the interval observer Cor. 4 for on-line (real time) estimation will depend on the sampling time, the considered time horizon, and on the precision demands.*

Remark 5 *To investigate the correlation of the state variables, or to locate steady states (as in Hasenauer et al. [2009, 2010a]), homothety (Prop. 4) or the partitioning approach (Prop. 5) can be considered analog to parameter estimation, see Section 4.2.*

The proposed interval observer Cor. 4 is applied next to the competitive inhibition mechanism.

Example 5.1: competitive inhibition

We again consider the competitive inhibition ($j = 2$) from Example 2 (Section 4.1), where we considered that x_3 and x_6 are not measured (permanently missing state values). Taking the obtained parameter uncertainty intervals and the uncertain measurements of x_1, x_2, x_4, x_5 (4.4) into account, we can re-evaluate the measured states (for instance x_1) and reconstruct the missing states (e.g. x_3), i.e. $x_1[k]$ and $x_3[k]$, $k = \{1, \dots, 40\}$ using Cor. 4. The results are shown in Fig. 5.1.

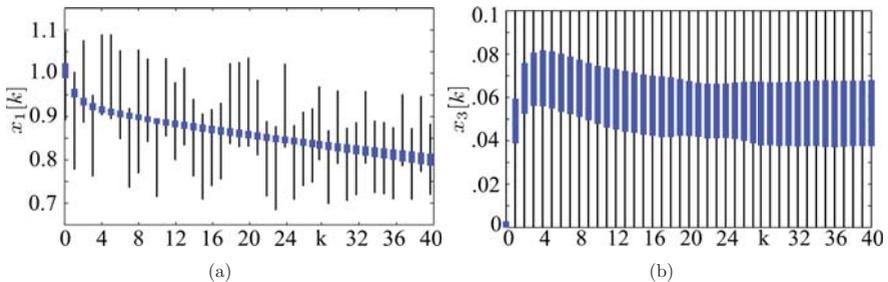


Figure 5.1.: Interval observer. Reconstruction of the state values $x_1[k]$ (5.1(a)) and $x_3[k]$ (5.1(b)), $k = \{1, \dots, 40\}$ for the competitive inhibition mechanism. Depicted are the prior (black) and the estimated (blue) state uncertainty intervals.

5.2. Reachability analysis

While state estimation as shown in the previous section is concerned with the problem of reconstructing missing state values considering measurement data, reachability is

concerned with the question if certain states regions can be reached from other ones, considering parametric uncertainties and disturbances. Reachability is an important property of dynamical systems with direct implications on various modeling and control problems [Rumchev and Adeane, 2004]. The reachable sets, i.e. the states reachable by the model trajectories, provide useful information to assess the capacity of a proposed model, to analyze the influence of uncertainties on the system dynamics, and to determine control possibilities, effectiveness and performance of control laws, in particular when considering disturbances or parametric uncertainty. Reachability-based approaches have also been considered already for model verification in the context of biological systems, see e.g. Dang et al. [2011], Yang and Lin [2010], i.e. to verify whether a certain model admits a desired dynamical behaviour.

Reachability under disturbances or parametric uncertainty has been studied intensively for linear systems, in particular positive time-invariant and time-variant systems [Rakovic et al., 2006, Sontag, 1998, Vincent and Grantham, 1997]. However for nonlinear systems only partial results are available. There are two basic types of reachable sets, depending on whether an initial or a final condition is specified [Mitchell and Tomlin, 2003].

Given a state space domain $X \subseteq \mathbb{R}^{n_x}$, the *weak* reachability problem can be defined as follows: Given an initial set $X_0 \subset X$ and an target set $X_N \subset X$, and the input domain $u_k \in U$, $k \in [1 : N - 1]$; show that there exist points in X_0 and points in X_N that are *connected* by trajectories of the system M (3.1), see e.g. Maler [2004], Mitchell and Tomlin [2003].

To formalize the setup, consider a model of the form (3.1). The available data includes a priori data, i.e. a parameter region $P \subset \mathbb{R}^{n_p}$, a state space $X \subset \mathbb{R}^{n_x}$, an input domain $u_k \in U \subset \mathbb{R}^{n_u}$, $k \in [0 : N - 1]$, and an initial and target set $X_0 \subseteq X$, $X_N \subseteq X$ respectively. If required, bounded disturbances can be included. The available data including the design constraints X_0 , X_N , is summarized by D (3.2).

Analogously to model invalidation, a sufficient criterion for non-reachability is as follows:

Corollary 5 (Reachability)

If the optimum of the dual problem $SDP^(Z)$ is unbounded, then no point in X_N is reachable from any point in X_0 in N steps.*

Furthermore, the reachable sets are the feasible future state sets considering a given input domain with respect to a given initial set. Estimating these sets can be recasted in terms of the state estimation problem as outlined in the previous section; the difference consists in the available data, i.e. to estimate the reachable sets no measurement data is used but an initial set. Without loss of generality, we confine the reachable set estimation problem to the time window $[\alpha : \beta]$, with $\alpha \leq 0 \leq N \leq \beta$. We consider the state $x_k \in \mathbb{R}^{n_x}$ of interest, i.e. aim to determine the respective reachable set, i.e. $s = \{x_k, k \in [\alpha : \beta]\}$, and denote the desired reachable set state set $\mathcal{X}_k = f_{x_k}(Z)$,

with $f_{x_k} : \mathbb{R}^{n_z} \rightarrow \mathbb{R}^{n_z}$ the respective projection map. The respective state uncertainty intervals $\mathcal{O}_I(\mathcal{X}_k)$ are obtained by Cor. 4. For a more detailed analysis, the reachable sets can be approximated considering homothety $\mathcal{O}_{\mathcal{H}}(\mathcal{X}_k)$ or via partitioning $\mathcal{O}_P(\mathcal{X}_k)$ following Prop. 4 and Prop. 5 respectively.

Example 5.2: Van der Pol oscillator

Consider the nonlinear Van der Pol oscillator, which can e.g. be used to describe the action potentials of neurons [FitzHugh, 1955]:

$$\begin{aligned} x_1^+ &= x_2, \\ x_2^+ &= -p_1 \cdot x_1 - p_2 \cdot (x_1^2 - p_3) \cdot x_2 + u, \end{aligned} \quad (5.1)$$

where p_i are the system parameters, $x_i = x_i[k]$, $u = u[k]$ and $x_i^+ = x_i[k+1]$ the current state, input, and future state. Let us assume the uncertain parameters

$$P = \{p_1 \in [0.08, 0.12], p_2 = 1.1, p_3 = 1\}, \quad (5.2)$$

and the (rectangular) initial set

$$X_0 = \{[0.47, 0.53] \times [0.48, 0.52]\}.$$

We estimate the reachable sets \mathcal{X}_k ($1 \leq k \leq 4$) considering two feasible input domains

$$\begin{aligned} U(1) &= \{u_k \in [0.00, 0.10], k \in [1 : 4]\}, \\ U(2) &= \{u_k \in [0.20, 0.30], k \in [1 : 4]\}. \end{aligned}$$

The states are constrained, for all times, to the state constraint set $X = \{(x_1 \in [0, 1], x_2 \in [0, 1])\}$.

Fig. 5.2 shows the estimated reachable sets $\mathcal{O}_P(\mathcal{X}_k)$, $k = [1 : 4]$ for the respective input domains, compared with Monte Carlo tests (spots). Note that the remaining state-regions are certified non-reachable.

5.2.1. Open loop control

An immediate application following the reachability considerations is the estimation of admissible inputs, i.e. the finite-time open loop control problem to steer the system from X_0 to X_N .

If, for a given system, there exists points in X_0 which are connected by state trajectories to some points in X_N , then there exist (at least one) input sequence u_0, u_1, \dots, u_{N-1} to steer the system (3.1) from a point in X_0 to a point in X_N . In this case, we can estimate the feasible inputs $\mathcal{U} \doteq f_u(\mathcal{Z})$, $f_u : \mathbb{R}^{n_z} \rightarrow \mathbb{R}^{n_u}$ which allow for this transition. The result immediately follows from Prop. 3/Prop. 5:

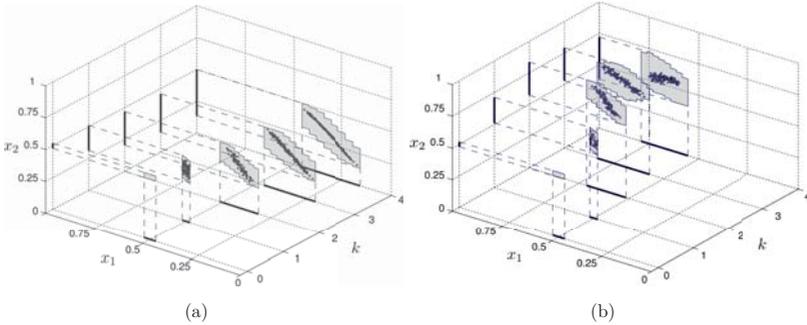


Figure 5.2.: Reachable sets $\mathcal{O}_P(\mathcal{X}_k)$, $k = [1 : 4]$ and Monte Carlo tests for the Van der Pol oscillator. (a) Reachable sets for input domain $U(1)$, (b) for $U(2)$.

Corollary 6 (Admissible inputs)

The admissible inputs \mathcal{U} are bounded by $\mathcal{O}_I(\mathcal{U})$ or $\mathcal{O}_P(\mathcal{U})$, derivable by $2Nn_u$ or $2^{Nn_u/\varepsilon}$ dual bounds respectively.

Hereby, an optimal control problem can be considered by defining optimal values of the target set, and to consider branch-and-bound scheme (Section 3.3.4) to obtain an optimal input sequence.

Example 5.3: Open loop control

Reconsider the Van der Pol oscillator (5.1) with uncertain parameters (5.2). The inputs are constrained to the domain $U = [0.00, 0.30]$. As initial and target sets we choose the rectangles $X_0 = \{[0.47, 0.53] \times [0.48, 0.52]\}$ and $X_N = \{[0.75, 0.80] \times [0.40, 0.45]\}$, respectively.

We have that no point in X_N is reachable from any point in X_0 for $N = 1, 2$. For $N = 3$, the reachability problem is feasible, and the corresponding inputs are estimated according to Corollary 6. The resulting set of admissible inputs $\mathcal{O}_P(\mathcal{U})$ is depicted in Fig. 5.3. Tightness of the result is demonstrated by comparing this set with strictly feasible solutions obtained by Monte Carlo tests, see Fig. 5.3.

We now turn our attention to another topic, considering if measurements obtained are physically possible or not. For this we will exploit reachability and state estimation concepts.

5.3. Outlier analysis

Outlier analysis deals with the problem of detecting, and if appropriate removing, anomalous observations in data, and is a primary step towards obtaining estimates and coherent analysis [Ben-Gal, 2005]. As pointed out by Ben-Gal [2005], outliers may carry valuable information, e.g. about variability of the process, although they may

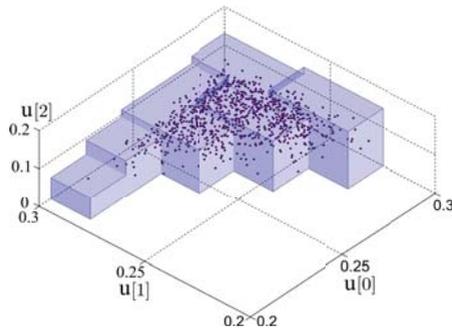


Figure 5.3.: Van der Pol oscillator. Partition estimate of the admissible inputs $u[k]$ for $k = 0, 1, 2$ to steer the oscillator from X_0 to X_N . Monte Carlo tests yield strictly feasible inputs (dots).

conversely lead to model misspecification, biased parameter estimation and incorrect results. Therefore it is important to identify them prior to modeling and analysis [Liu et al., 2004, Williams et al., 2002].

Even so an exact definition of an outlier often depends on hidden assumptions regarding the data structure and the applied detection method [Ben-Gal, 2005], some general definitions have been proposed. Hawkins [1980] defines an outlier as “an observation that deviates so much from other observations as to arouse suspicion that it was generated by a different mechanism”, and Barnett and Lewis [1994] “an outlying observation, or outlier, is one that appears to deviate markedly from other members of the sample in which it occurs”. Furthermore, Johnson and Wichern [2001] defines an outlier as “an observation in a data set which appears to be inconsistent with the remainder of that set of data”. For more case-specific definitions, see e.g. Ben-Gal [2005].

Outliers often arise due to faults, changes in system behavior, fraudulent behavior, human error, instrument error or simply through natural deviations in populations [Hodge and Austin, 2004]. There are many ways how to detect outliers, depending on the specific applications. Existing outlier detection methods can be classified according to whether a (error) model is utilized or not, i.e. parametric (model-based) and non-parametric (non model-based) respectively, see e.g. Hodge and Austin [2004] and Ben-Gal [2005] for a comprehensive survey. While latter typically deal with large data sets and independent data, model-based outlier detection methods are commonly used for detecting outliers in time-series data (dependent data). Model-based methods can be further classified into *model-specific* or *model-generic* approaches [Ben-Gal et al., 2003]. While model-specific approaches rely on a given model to perform the outlier

analysis, model-generic approaches provide the flexibility to estimate the underlying model parameters.

A general observation when considering data sets with *multiple outliers* are *masking* and *swamping effects* Ben-Gal [2005]. An intuitive, though not mathematically rigorous understanding of these phenomena is given by Acuna and Rodriguez [2004]. An outlier *masks* another outlier, if the second outlier is an outlier by itself, but not in the presence of the first outlier. Thus, only after the deletion of the first outlier, the second outlier emerges as one. Conversely, an outlier is said to *swamp* a second observation, if the latter is an outlier only due to the first one. In other words, after the deletion of the first outlier the second observation becomes a non-outlying observation. Masking and swamping effects hence may complicate the detection of outliers, and have to be considered carefully, e.g. to avoid neglecting important information.

We here focus on model-generic outlier detection strategies for time-series data considering nonlinear and uncertain systems. To this end, we first devise a combinatorial outlier detection approach to evaluate outlier hypotheses using infeasibility certificates. Furthermore, we consider a second model-generic outlier detection procedure which is based on reachability analysis.

Setup

We consider a model M of the form (3.1), and the available data D (3.2), including a priori and structural knowledge as well as measurement data. Regarding the measurements, we focus for simplicity on the single output case ($n_y = 1$), and remark that the extension to multiple outputs is straightforward. We assume that the time instances, at which measurements for the output are available, correspond to the indexes $k \in [0 : N - 1]$ (Ass. 1). The uncertainty of the measurements is described by intervals, i.e.

$$D_{meas} = \{y_k \in [\underline{y}_k, \bar{y}_k]_{meas}, k \in [0 : N - 1]\},$$

where $[\underline{y}_k, \bar{y}_k]_{meas}$ ($k = 0, \dots, N - 1$) denote the measurement uncertainty intervals. Analogously, we denote the a priori uncertainty intervals by $[\underline{y}_k, \bar{y}_k]_{prior}$ ($k = 0, \dots, N - 1$).

We start from the observation that the model and the data are inconsistent according to Corollary 1. Instead immediately rejecting the model, we now consider the possibility that the data contains *some* outliers, i.e. the uncertainty of some measurements is not appropriately described¹. In this regard, we define an outlier as follows:

Definition 2 (Outlier)

An outlier denotes a measurement which is not appropriately described by the proposed uncertainty description, i.e. the actual values are not covered.

¹This case can for instance occur by natural probability if the confidence intervals of the measurements probability density distribution are considered as measurement uncertainty intervals.

The number of outliers in D_{meas} is denoted by n_o , which is of course unknown in the beginning of the detection process (and can be considered small in comparison with N). When we say an outlier y_k for a $k \in [0:N-1]$ is *discarded*, we replace the measurement uncertainty interval with the a priori interval $[\underline{y}_k, \bar{y}_k]_{prior}$. We denote this case with $D_{meas} \setminus y_k$.

Finally, we require a measure of the distance of two sets later on, in particular the minimum distance of two intervals:

Definition 3 (Minimum distance)

We define the minimum distance between two non-empty compact sets $A \in \mathbb{R}^n$, $B \in \mathbb{R}^n$ by the minimum distances between any two of their respective points, i.e.

$$d_{min}(A, B) = \min\{\|a - b\| : a \in A, b \in B\}.$$

Note that this distance is zero if the two sets overlap.

With the preparations above, we can now focus on the outlier detection problem.

5.3.1. Combinatorial outlier detection approach

An approach to detect possible outliers in the data set D_{meas} consists in formulating outlier hypotheses, i.e. to select possible outlier candidates and discard them, and subsequently to perform an inconsistency test utilizing Corollary 1. To this end, we consider first the simple case where a single outlier is suspected.

Single outlier

We assume that a single outlier is present in the measurement data D_{meas} , i.e. $n_o = 1$. Then, the combinatorial approach consists in systematically *discarding* a single measurement y_k (starting from $k = 0$) and perform the consistency test following Corollary 1 with M and $D_{meas} \setminus y_k$. This leads to the following result:

Proposition 7 (Single outlier case)

A single outlier is detected by at most N infeasibility certificates according to Corollary 1.

The respective outlier, y_k , can be further analyzed, i.e. the distance of the outlier from the reachable state can be estimated. To this end, we perform a state estimation (see Section 5.1) with M and $D \setminus y_k$. Then, according to Corollary 4, we obtain the a posteriori uncertainty interval $[\underline{y}_k, \bar{y}_k]_{post}$. The distance according to Definition 3 of the obtained interval and the measurement interval is given by

$$d_{min}([\underline{y}_k, \bar{y}_k]_{meas}, [\underline{y}_k, \bar{y}_k]_{post}).$$

Remark 6 *Note that, even in this most simple case, there exists the possibility that there exists no unique solution to the outlier problem, i.e. there might exist alternative,*

possibly exclusive, outliers. This is independent from the proposed method, rather an intrinsic issue of the (in this case ill-posed) outlier detection problem².

Multiple outliers

The proposed outlier detection approach extends to the multiple outlier case. To this end, we discard n_o measurements, not necessarily consecutive, from the data set analogously to the single outlier case, and perform the consistency test following Corollary 1. This leads to the result:

Proposition 8 (Combinatorial outlier detection)

To detect n_o outliers within N measurements, at most

$$\binom{N}{n_o} = \frac{N!}{n_o!(N - n_o)!} \quad (5.3)$$

infeasibility certificates according to Corollary 1 are required.

Proof. The number of possible arrangements of n_o outliers in N measurements is equivalent to the classical combinatorial problem, i.e. n_o -combinations (zero elements) in a sequence of otherwise one elements of length N . Hence, the number of possible combinations is the given by the binomial coefficient (5.3). \square

Note that in practice the number of outliers is unknown. Therefore, a strategy in this case consist in increasing successively the number of suspected outliers starting from a single one. This ensures to obtain a minimal number of outliers, with the advantage that masking and swamping are avoided, although uniqueness can not be guaranteed in general, refer Remark 6. The so obtained outliers can be analyzed as outlined in Section 5.3.1.

The disadvantage of this approach is that the number of combinations and therefore the number of required evaluations according to Corollary 1 increases with the number of the measurements N and with the number of suspected outliers n_o ; particularly if N is large, the complexity of the proposed approach is approximately N^{n_o} , i.e. exponential in the number of suspected outliers. Hence, the combinatorial outlier detection approach is in general not suited for large data sets (with possibly many outliers), given that no particular outlier hypotheses can be formulated and a comprehensive combinatorial search has to be considered. This is because the outliers have to be detected instantaneously (all at once).

For larger data sets, a sequential approach may be advantageous, e.g. to detect the most extreme outliers first. To find the most extreme outliers, a possibility consists in ‘relaxing’ the uncertainty description of the measurements, i.e. to introduce an additional pessimism. For example, the 2-sigma confidence intervals of the measurement probability density function can be considered as measurement uncertainty intervals

²This case can occur e.g. if N is very small, refer the following example.

instead the 1-sigma confidence interval. This obviously will lead to a fewer number of outliers in most cases, which in turn facilitates (combinatorial) detection. To formalize this sequential approach, we introduce a *tolerance* as follows.

Definition 4 (Tolerance)

The tolerance $\epsilon \geq 0$ admits the following properties:

- If $\epsilon = 0$, then $D_{meas}^\epsilon = D_{meas}$.
- For any $0 < \epsilon$, we have $D_{meas} \subset D_{meas}^\epsilon$.
- For any $\epsilon_2 < \epsilon_1$, we have $D_{meas}^{\epsilon_2} \subset D_{meas}^{\epsilon_1}$.

Exemplary, a tolerance can be modeled by an additional absolute or relative error, imposed onto the measurement data according to (2.2.1) or (2.2.1) respectively.

A sequential (model-generic) outlier detection procedure can be stated as follows:

1. Consider a tolerance ϵ and determine D_{meas}^ϵ .
2. Perform the combinatorial outlier detection using D_{meas}^ϵ (Prop. 8) and *discard* detected outliers.
3. Decrease the tolerance ϵ and repeat until $\epsilon = 0$ or a desired threshold is achieved.

The challenge hereby consists in choosing an appropriate initial tolerance, such that only few (at best only one, i.e. the most extreme) outliers remain, which are easier to detect. Note that, in general, the tolerance can be considered as a ‘weighting’ of uncertainty; if no additional knowledge is available, it is reasonable to consider a tolerance which acts equally onto all measurements. Conversely, knowledge might be available that certain outputs (or specific measurements) might be more prone to errors than others; then, the tolerance may be adapted accordingly. Exemplary, initial conditions or steady state measurements may be known with certainty, and then it is reasonable to not relax the respective uncertainty description.

Example 5.3: linear and quadratic regression

As an illustrative example, we consider a regression problem given the six input/measurements pairs (k, y_k) with $(0, 0.8 \pm 0.25)$, $(1, 2 \pm 0.25)$, $(2, 2.2 \pm 0.25)$, $(3, 4.1 \pm 0.25)$, $(4, 4.85 \pm 0.25)$, $(5, 5.2 \pm 0.25)$ subjected to an absolute measurement error ($\eta_a = 0.5$).

We first turn on a linear regression given by

$$y_k = a_1 k + a_0,$$

where $a_1 \in \mathbb{R}$ and $a_0 \in \mathbb{R}$ are the unknown parameters.

We found, regarding linear regression, at least two outliers according to Prop. 5.3, i.e. at $k = 2$ and $k = 5$. Regarding the remaining measurements, the parameters are estimated according to Cor. 3, see Fig.5.4(a), and the reachable states are depicted in Fig.5.4(b).

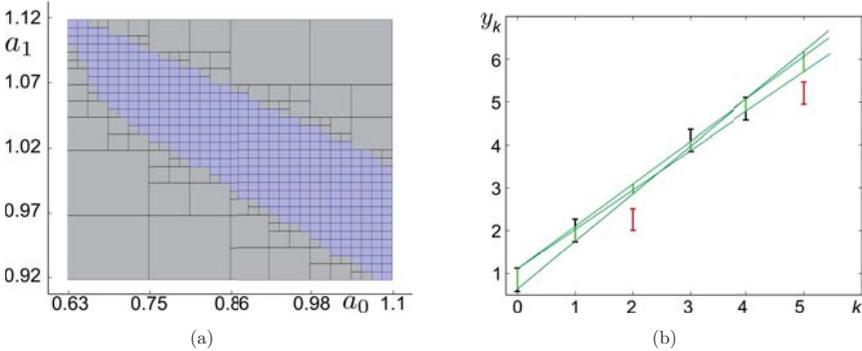


Figure 5.4.: Model-generic outlier detection (linear regression model). 5.4(a): Consistent parameters. 5.4(b): Reachable sets and the detected outliers.

We furthermore consider the quadratic regression given by

$$y_k = a_2 k^2 + a_1 k + a_0,$$

where now $a_2 \in \mathbb{R}$, $a_1 \in \mathbb{R}$, $a_0 \in \mathbb{R}$ are the unknown parameters. For this regression model, we found one outlier at $k = 2$. With the remaining measurements, the parameters are estimated according to Cor. 3, see Fig.5.5(a), and the reachable states are depicted in Fig.5.5(b).

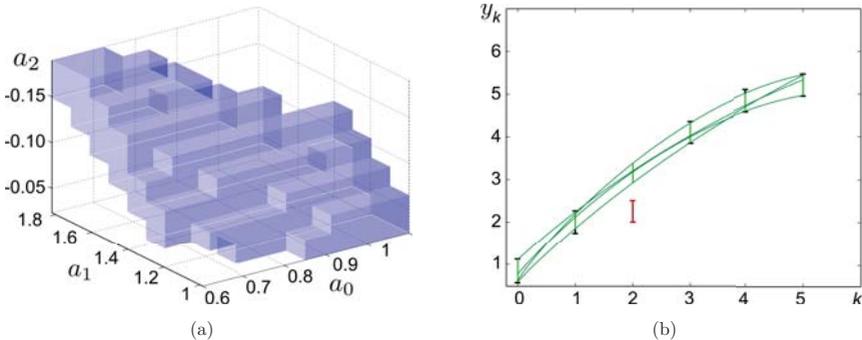


Figure 5.5.: Model-generic outlier detection. Quadratic regression model. 5.5(a): Consistent parameters. 5.5(b): Reachable sets and the detected outlier.

5.3.2. Reachability-based outlier detection

The difficulty to find appropriate tolerances motivates us to explore alternative approaches for outlier detection. Complementary to the combinatorial detection approach, reachability analysis can be used for detecting outliers as follows. We start from the observation that the model M (3.1) and the data D (3.2) are inconsistent according to Corollary 1. Subsequently, we follow the procedure given below:

1. Consider a tolerance ϵ and determine D_{meas}^ϵ . The tolerance has to be chosen sufficiently large, i.e. such that the model is not invalid according to Cor. 1.
2. Estimate the reachable sets $[\underline{y}_k, \bar{y}_k]_{post}$, for all $k \in [0 : N - 1]$ regarding the data D^ϵ (Cor. 4).
3. Compare the reachable sets and the measurements considering the minimum distance (Def. 3): Each measurement with $d_{min}([\underline{y}_k, \bar{y}_k]_{meas}, [\underline{y}_k, \bar{y}_k]_{post}) > 0$ is an outlier and discarded.
4. Decrease the tolerance ϵ and repeat the procedure until $\epsilon = 0$ or a desired threshold is achieved.

The main advantages of this approach are, first, that finding a sufficiently large tolerance is not difficult. Secondly, by computing the reachable sets, consecutive outliers and thus possible model misspecification can be detected. Thirdly, the approach allows to refine the parameters even if not all outliers are detected.

We provide a comprehensive example for outlier detection based on reachability in Chapter 7.

5.4. Summary

In this section, we applied the set-membership framework to state estimation, reachability analysis, and the detection of outliers. Missing state values, e.g. concentrations of intra-cellular metabolites, can be reconstructed by an interval observer. To evaluate e.g. the influence of parametric uncertainties and disturbances onto the systems dynamic, a reachability analysis can be considered. This way, an insight into robustness properties of the considered system can be obtained.

The here considered methods including model invalidation and parameter estimation are based on a bounded error description of the measurement uncertainties. This uncertainty description however is in general prone to outliers, which impose a particular challenge to the framework. Outlying observation might lead to reject a valid model (false positive), or biased estimation results. Their detection is thus crucial for the applicability of set-membership methods to actual process data.

We therefore considered the detection of outliers within a model-generic setting. Particularly, we presented a combinatorial approach, based on a combinatorial evaluation of outlier hypotheses using infeasibility certificates. The approach is suitable only if the number of outliers is small. For larger data sets with possibly several outliers, we proposed a different approach based on a reachability analysis. By estimating the reachable states regarding a more conservative uncertainty description, the most extreme outliers can be detected first, and the pessimism can be reduced sequentially.

An important consequence of the model-generic setting is that the detection of outliers depends on the model used for detection. In other words, the outliers detected with respect to one model may deviate with the outliers detected with respect to another model.

6. Design of Experiments

The design of experiments is an important link between modeling and experimentation, to address a priori how to perturb the process under study, as well as which states to be observed, so as to learn and reveal the most important features of the system, e.g. to learn the unknown model parameters. To this end, experiments have to be performed to obtain measurements of the process which can be used for estimation subsequently. However, poorly planned or designed experiments may only yield little information, resulting in poor parameter estimates.

In this chapter, we address the problem of designing optimal experiments in membership setting for the purpose of parameter estimation, focusing on polynomial systems which are linear in the parameters. After reviewing existing concepts in Section 6.1, we derive necessary and sufficient conditions such that the unknown parameters can be identified at all (Section 6.3). We then take uncertainties explicitly into account, and investigate fundamental limitations of experimental design in Section 6.4. Based on these considerations, we propose a robust optimal experimental design method in Section 6.5. Several examples illustrate the proposed approach. This chapter is based in parts on the works Borchers and Findeisen [2011], Borchers et al. [2011b].

6.1. Review

Obtaining as good as possible estimates of the parameters of a mathematical model describing a dynamic process is an ubiquitous problem and required for purposes such as model selection, prediction, or control synthesis.

Fisher [1935] initiated the study of a priori experimental design, with the idea of “deciding what patterns of factors combinations (inputs) will best reveal the properties of the systems response, and how this response is influenced by the factors” [Franceschini and Macchietto, 2008]. He focused on obtaining the most important information to reveal an input-output relationship in the presence of variations of stochastic nature, which is known nowadays as *black box* experimental design (see e.g. Franceschini and Macchietto [2008] for applications and a more comprehensive review). The black box approaches however are inappropriate for dynamical systems with constrained outputs, as they do not take into account the available (or at least partially available) information about the system’s structure.

Therefore, experimental design approaches had to be extended to explicitly enclose knowledge of the considered system’s structure [Franceschini and Macchietto, 2008].

This is termed model-based experimental design. Early approaches admitted only steady state systems, including linear and nonlinear models, e.g. to study reaction kinetics as in Box and Lucas [1959]. The extension to dynamical systems has been slow [Shirt et al., 1994] and primarily been considered in a stochastic context. For example, one objective for experimental design, which has been considered, is to minimize either the variance (uncertainty of the estimates [Shirt et al., 1994]) or the bias [Ljung, 1998] of the transfer function. Various other optimality criteria have been pursued, all based on the *Fisher information matrix* (FIM). Here, the parameter uncertainty can be appropriately distinguished within the FIM “due to the asymptotic normality of parameter estimators and the *Cramer-Rao bound*” [Pronzato, 2008]. Optimality criteria are found by minimizing the (expected) variance of the unknown parameters, e.g. functionals of the invariants of the FIM. In this context, a widely used criterion is the *D-optimal* design which aims at maximizing the determinant of the FIM and thereby minimising the parameter variances. Alternatively, designs such as A-optimality, E-optimality, etc. have been considered (see e.g. [Boyd and Vandenberghe, 2004, p. 384–392] for a compact overview).

However, the proposed frequency domain and Fisher-information matrix based approaches all rely on the true system parameters, or at least on an appropriate and accurate a priori guess of the nominal system parameters. Hence, the quality of experiments designed using these standard techniques can be adversely affected by poor starting values of the parameters [Asprey and Macchietto, 2002]; such information however is in many cases simply not available. Thus, design methods that are insensitive to these starting values are required [Asprey and Macchietto, 2002]. This issue has been recognized in literature, and some approaches to the so called *robust experimental design* have been taken into account, as for example the *sequential design* (e.g. Walter and Pronzato [1997], Wynn [1970]), *Bayesian approaches* (see e.g. Chaloner and Verdinelli [1995] for a review), or *minimax design* (see e.g. Rojas et al. [2007] for an overview and references). However, apart from standard cases (linear systems, *white noise*), there has been little study on robust experimental design for engineering problems [Rojas et al., 2007], see also the survey presented in Hjalmarsson [2005]. In membership setting only few approaches have been made so far towards the robust design of experiments. Norton [1987] proposed a number of general guidelines, and Belforte et al. [1987] described an orthotopic approximation approach. Pronzato and Walter [1990] considered to use experimental design for linear regression models, by choosing as design policy a volume criterion which compares to the classic D-optimal design in stochastic settings. Novara [2007] considered experimental design for non-linear system identification, and recently, Marvel and Williams [2012] set-membership experimental design has been considered in the context of biological systems.

6.2. Setup

The setup for experimental design considered in the following is less general than the setting for estimation and analysis, see Chapter 3. The most important differences are we consider polynomial systems that are *linear in the parameters*¹, subject to *additive* disturbances, and that *all states can be measured*. Although we can avoid some of these simplifications as shown in the examples later on, for now, the systems have the following form:

$$x_i^+ = \sum_{j=1}^{n_p} f_{ij}(x, u)p_j + w_i = F_i(x, u)p + w_i, \quad i = 1, 2, \dots, n_x \quad (6.1)$$

where $x \in \mathbb{R}^{n_x}$, $u \in \mathbb{R}^{n_u}$ and $w = (w_1, w_2, \dots, w_{n_x})^T \in \mathbb{R}^{n_x}$ are the current state, control and the unknown disturbance respectively, x_i^+ is the successor state of x_i ; $f_{ij}(x, u) : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \rightarrow \mathbb{R}$ for all $i \in \{1, \dots, n_x\}$ and $j \in \{1, \dots, n_p\}$ are polynomial functions (in the states and inputs), and $F_i(x, u) \in \mathbb{R}^{n_p}$ for all $i \in \{1, \dots, n_x\}$. $p \in \mathbb{R}^{n_p}$ denote the vector of unknown parameters. Note that typically $n_p \gg n_x$, and that some $f_{ij}(x, u) = 0$ independent of the choice of x and u by construction.

We consider in the remainder that the parameters $p \in \mathbb{R}^{n_p}$ are (completely) unknown, but constant. The disturbances w_i , $i = 1, \dots, n_x$, are not constant, and known only to be bounded, e.g.

$$w_i \in [-\bar{w}_i, \bar{w}_i], \quad (6.2)$$

with known \bar{w}_i for all $i = 1, \dots, n_x$. The *design variables* consists in the initial state and the inputs, which can be chosen from the domains

$$x \in X_0 = [\underline{x}_i, \bar{x}_i]^{n_x}, \quad u \in U = [\underline{u}_i, \bar{u}_i]^{n_u}, \quad (6.3)$$

with known $\underline{x}_i, \bar{x}_i$ for all $i = 1, \dots, n_x$, and known $\underline{u}_i, \bar{u}_i$ for all $i = 1, \dots, n_u$.

Remark 7 *The generalization to a comprehensive polytopic setting for the disturbances, initial conditions and inputs is possible, although not considered here for simplicity of presentation.*

A pivotal idea of the proposed design approach consists in designing and choosing a number of *one-step experiments* and observations respectively, to provide as good as possible parameter estimates.

We denote an *one-step experiment* by $Exp(x, u)$, where $x \in X_0$ and $u \in U$ (6.3). By performing such an one-step experiment, a state measurement (an observation) $z_i \doteq x_i^+$ can be taken.

¹as e.g. resulting from mass action kinetics

Remark 8 One measurement z_i is related to one experiment $Exp(x, u)$, and hence each experiment involves a choice which state is measured. The measurement furthermore depends on the current disturbance, i.e. $z_i = F_i(x, u)p + w_i$, where w_i (6.2) not constant. An important consequence is that the same experiment (e.g. a repetition) in general does not lead to the same measurement, because w_i can take any value from the interval $[-\bar{w}_i, \bar{w}_i]$. However, the probability distributions of the disturbances are unknown, and thus we focus on the worst case throughout this chapter.

Consider n experiments $Exp(x^{(l)}, u^{(l)})$, and respectively n measurements $z_{i(l)}^{(l)}$, with $l = 1, \dots, n$ and $i(l)$ from the indexset $i(l) \in [1 : n_x]$, i.e. the choices which states will be measured. Then, we have with

$$\mathbf{z} = \begin{pmatrix} z_{i(1)}^{(1)} \\ \vdots \\ z_{i(n)}^{(n)} \end{pmatrix} \in \mathbb{R}^n, \quad A = \begin{pmatrix} F_{i(1)}(x^{(1)}, u^{(1)}) \\ \vdots \\ F_{i(n)}(x^{(n)}, u^{(n)}) \end{pmatrix} \in \mathbb{R}^{n \times n_p}, \quad \mathbf{w} = \begin{pmatrix} w_{i(1)}^{(1)} \\ \vdots \\ w_{i(n)}^{(n)} \end{pmatrix} \in \mathbb{R}^n, \quad (6.4)$$

$$\mathbf{z} = Ap + \mathbf{w}.$$

For shorthand, we denote the disturbance set by $\mathbf{w} \in \Omega$, constructed from (6.2).

For ease of notation, we call the matrix A the *design matrix*. Thus, the design matrix depends the number of experiments, on the choice which states will be measured, and the initial conditions and inputs for all considered experiments. For simplicity of presentation, we say that the possible design matrices A (with n experiments/measurements) belong to the family of design matrices \mathcal{A} , defined as

$$\mathcal{A} \doteq \left\{ \begin{pmatrix} F_{i(1)}(x^{(1)}, u^{(1)}) \\ \vdots \\ F_{i(n)}(x^{(n)}, u^{(n)}) \end{pmatrix} \in \mathbb{R}^{n \times n_p} : x^{(l)} \in X_0, u^{(l)} \in U, l = 1, \dots, n, i(l) \in [1 : n_x] \right\}. \quad (6.5)$$

We can next define the *consistent parameter set* depending on the design matrix:

Definition 5

The set of consistent parameters Θ induced by n experiments is given by

$$\Theta = \{p \in \mathbb{R}^{n_p} : \mathbf{z} = Ap + \mathbf{w}, \mathbf{z} \in \mathbb{R}^n, A \in \mathcal{A}, \mathbf{w} \in \Omega\}. \quad (6.6)$$

Remark 9 It is very important to note that the consistent parameter set Θ can be interpreted as a ‘family’ of sets. As an anticipation of the following considerations, the ‘members’ of the consistent parameter set (i.e. the possible a posteriori parameter sets) have a certain ‘shape’ (which is determined by the disturbance set and the design matrix), an a certain ‘orientation’ in \mathbb{R}^{n_p} (which is determined by the actual measurements \mathbf{z}). Robust experimental design consists in choosing the design matrix from \mathcal{A} such that the ‘shape’ of Θ has minimum volume, which corresponds to minimizing the

volume of the a posteriori parameter set in worst case (Section 6.5). Beforehand, we investigate how ‘small’ Θ can be made at all in this setting (Section 6.4).

The consistent parameter set can be interpreted geometrically as a set with a certain orientation in \mathbb{R}^{n_p} and a certain size/shape, i.e.

Proposition 9

The set of consistent parameters can be described by set addition

$$\Theta = \Theta_z + \Theta_\Omega, \tag{6.7}$$

where

$$\Theta_z \doteq \{\theta_z \in \mathbb{R}^{n_p} : z = A\theta_z, A \in \mathcal{A}\}, \tag{6.8}$$

and

$$\Theta_\Omega \doteq \{\eta \in \mathbb{R}^{n_p} : 0 = A\eta + w, w \in \Omega, A \in \mathcal{A}\}. \tag{6.9}$$

Proof. By (6.7) and (6.8), we have respectively $A\Theta_z \in \{z\}$ and $A\Theta_\Omega \in -\Omega$. For the set $\Theta_z + \Theta_\Omega$ it holds $A(\Theta_z + \Theta_\Omega) \in \{z\} - \Omega$, i.e.

$$A(\Theta_z + \Theta_\Omega) + \Omega \in \{z\} \equiv \{A(\theta_z + \eta) + w = z, \theta_z \in \Theta_z, \eta \in \Theta_\Omega, w \in \Omega\},$$

respectively

$$\Theta_z + \Theta_\Omega = \{\theta_z + \eta \in \mathbb{R}^{n_p} : A(\theta_z + \eta) + w = z, w \in \Omega\}.$$

With $p \doteq \theta_z + \eta$, $\theta_z \in \Theta_z$, $\eta \in \Theta_\Omega$, we have $\Theta_z + \Theta_\Omega = \Theta$ according to Def. 5. □

Particularly, we interpret the sets Θ_z as the orientation of the set Θ_Ω . Note that both sets depend on the design matrix themselves.

We show next that under a certain condition onto the choice of the design matrix, Θ_z becomes a singleton set, i.e. a vector, in worst case. Then, the shape of the set of consistent parameters only depends on the shape of Θ_Ω , which we investigate thereafter.

6.3. Parameter identifiability

We first turn on the question how to obtain a *compact* parameter set. Remind that the parameters are completely unknown in the beginning. The necessary and sufficient condition for obtaining a compact consistent parameter set is the following:

Proposition 10

The consistent parameter set Θ is compact if and only if the design matrix A has full row rank, i.e. $\text{rank}(A) = n_p$.

Proof. Proof immediately follows from Prop. 9 the and compactness of Ω . □

In consequence, n_p measurements (experiments) are required to bound all model parameters. Note also the choice of full row rank design matrices implicitly involves an appropriate choice which states are measured.

Remark 10 *It is important to note, under mild assumptions on the domains X_0 and U , that a choice of n_p experiments such that A has full rank is always possible (for polynomial systems). This is because the determinant of A are polynomials in the states and the inputs, which are chosen from the compact domains X_0 and U .*

By choosing full rank design matrix, we achieve that Θ_z becomes a vector, denoted hereafter by $\theta_z \in \mathbb{R}^{n_p}$. However, the vector θ_z does not correspond to the ‘true’ parameters in general, with the exception if no disturbances were acting (noise free case), i.e. $\Theta_\Omega = \{0\}$. In this case, full row rank of the design matrix is necessary and sufficient for *parameter identifiability*, i.e. that the model parameters can be determined *uniquely*. This result immediately follows from Proposition 10 and Def. 5.

Remark 11 *Note that parameter identifiability here is motivated from a practical point of view, i.e. the necessary and sufficient conditions such that an unique parameter estimate is actually obtained, differing from structural identifiability issue, compare e.g. Bellman and Åström [1970], Cobelli and Distefano [1980]]. Furthermore, due to the worst case setting, the results are globally valid, compare e.g. Audoly et al. [2001].*

With these theoretical results at hand, we can discuss how to obtain a full row rank design matrix in practice.

A full row rank design matrix can be chosen in practice in two stages. First, an appropriate choice of the states which have to be measured, has to be made. In a next step, the initial conditions and inputs are chosen. To this end, the main idea is to treat the unknown parameters separately for each system equation. We rewrite, for all $i = 1, \dots, n_x$, the system equations into

$$x_i^+ = F_i'(x, u)\mathbf{p}_i + w_i,$$

where now $\mathbf{p}_i \in \mathbb{R}^{n_{p_i}}$ denotes the parameters appearing in the i -th state equation, and $F_i'(x, u)$ containing only the non-zero elements f_{ij} .

Following Prop. 10, we require (necessary condition) n_{p_i} state measurements $z_i^{(l)}$, $l = 1, \dots, n_{p_i}$, (respectively n_{p_i} experiments $Exp(x^{(l)}, u^{(l)})$) for obtaining a compact consistent parameter set of the now n_{p_i} parameters. The respective (i -th) design matrix is then given by

$$A_i = \begin{pmatrix} F_i'(x^{(1)}, u^{(1)}) \\ \vdots \\ F_i'(x^{(n_{p_i})}, u^{(n_{p_i})}) \end{pmatrix} \in \mathbb{R}^{n_{p_i} \times n_{p_i}}.$$

where $x^{(l)} \in X_0$ and $u^{(l)} \in U$ for $l = 1, \dots, n_{p_i}$.

In a next step, the inputs and initial conditions have to be chosen appropriately, which is however not difficult, compare Remark 10. Note also that the matrices A_i for $i = 1, \dots, n_x$ are smaller in general, which furthermore eases the choices.

Example 6.1: noise-free system

Consider the polynomial (noise-free) system

$$x_i^+ = p_{i1}x_1 + p_{i2}x_1^2 + p_{i3}x_2 + p_{i4}x_2^2 + p_{i5}x_1x_2 + p_{i6}u.$$

With $\mathbf{p}_i \doteq (p_{i1}, \dots, p_{i6})^T$, $i = \{1, 2\}$, and $z_i = x_i^+$ denoting the respective observations, we can write

$$z_i = F_i(x, u)\mathbf{p}_i = \begin{pmatrix} x_1 & x_1^2 & x_2 & x_2^2 & x_1x_2 & u \end{pmatrix} \mathbf{p}_i.$$

Following Prop. 10, twelve experiments/measurements are necessary to identify the unknown parameters (six measurements per state, denoted by $\mathbf{z}_1 \in \mathbb{R}^6$ and $\mathbf{z}_2 \in \mathbb{R}^6$ respectively). Exemplary, we consider the following six experiments $Exp(x_1, x_2, u)$:

$$Exp(x_1, 0, 0), Exp(0, x_2, 0), Exp\left(\frac{1}{2}x_1, 0, 0\right), Exp\left(0, \frac{1}{2}x_2, 0\right), Exp(x_1, x_2, 0), Exp(0, 0, u).$$

The corresponding design matrices ($i = \{1, 2\}$) are

$$A_i = \begin{pmatrix} x_1 & (x_1)^2 & 0 & 0 & 0 & 0 \\ \frac{1}{2}x_1 & (\frac{1}{2}x_1)^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & x_2 & (x_2)^2 & 0 & 0 \\ 0 & 0 & \frac{1}{2}x_2 & (\frac{1}{2}x_2)^2 & 0 & 0 \\ x_1 & (x_1)^2 & x_2 & (x_2)^2 & x_1x_2 & 0 \\ 0 & 0 & 0 & 0 & 0 & u \end{pmatrix}.$$

It is then easy to see that for $x_1 \neq 0$, $x_2 \neq 0$, and $u \neq 0$, the design matrices have full rank. The parameters are

$$\mathbf{p}_i = (A_i)^{-1}\mathbf{z}_i, \quad i = \{1, 2\}.$$

Note that the choice of the experiments is of course not unique, and the remaining degrees of freedom can be utilized for designing worst case optimal experiments as shown later on.

6.4. Limits of design

We have shown in the previous section that n_p measurements/experiments are necessary to obtain a compact estimate of the parameters. We furthermore showed that the

consistent parameter set can be considered, following Prop. 9, as the set Θ_Ω which is oriented in \mathbb{R}^{n_p} by the set Θ_z . Next, we consider the issue how ‘small’ Θ can be made at all in worst case², e.g. by considering infinitely many experiments.

To this end, reconsider the system equations given by

$$x_i^+ = \sum_{j=1}^{n_p} f_{ij}(x, u) p_{ij} + w_i, \quad i = 1, \dots, n_x.$$

We focus on the case that Θ_z is a singleton set (the vector θ_z), i.e. we choose at minimum n_p experiments with full row rank design matrix³. For simplicity of notation, we denote the i -th component of the system by the superscript index (i) .

By Proposition 9, we have given the consistent parameter set $\Theta^{(i)} \subset \mathbb{R}^{n_p}$ (for all $i = \{1, \dots, n_x\}$) by

$$\Theta^{(i)} = \theta_z^{(i)} + \Theta_\Omega^{(i)},$$

where

$$\begin{aligned} \theta_z^{(i)} &= (\theta_1^{(i)} \dots \theta_{n_p}^{(i)})^T \in \mathbb{R}^{n_p}, \\ \Theta_\Omega^{(i)} &= \{(\eta_{i1}, \dots, \eta_{in_p})^T \in \mathbb{R}^{n_p} : \sum_{j=1}^{n_p} f_{ij}(x, u) \eta_{ij} \in [-\bar{w}_i, \bar{w}_i]\}. \end{aligned}$$

By linearity of η_{ij} , it is easy to see that we have for all $j = 1, \dots, n_p$ and for all $i = 1, \dots, n_x$:

$$\eta_{ij} \in [-\bar{\eta}_{ij}, \bar{\eta}_{ij}] \doteq \frac{[-\bar{w}_i, \bar{w}_i]}{|\bar{f}_{ij}|}, \quad |\bar{f}_{ij}| \doteq \max_{x \in X_0, u \in U} |f_{ij}(x, u)|, \quad (6.10)$$

if $|\bar{f}_{ij}| \neq 0$. Otherwise, i.e. if $|\bar{f}_{ij}| = 0$, the parameter p_{ij} is *not identifiable* via the state measurement x_i^+ , and we have $\eta_{ij} \in (-\infty, \infty)$ ⁴.

In other words, considering X_0 and U to be a given compact sets, the uncertainty η_{ij} (i.e. the uncertainty of the desired parameters p_{ij}) can not be decreased below the interval given in Eq. (6.10), i.e. this uncertainty remains independent of the experiments. Moreover, again due to linearity of the η_{ij} ,

$$\Theta_\Omega = \text{ConH}(\{\pm \bar{\eta}_{ij}\}, \quad i = \{1, \dots, n_x\}, \quad j = \{1, \dots, n_p\}),$$

i.e. Θ_Ω is the convex hull of the interval limits of all η_{ij} (Eq. (6.10)), see Fig. 6.1. Thus, the precision which can be obtained at all by experiments a priori is limited by $\Theta_\Omega \subseteq \Theta$. We denote hereafter Θ_Ω as the *residual set*. Geometrically, the residual set is a rhombus in two dimensions, in three dimensions an octahedron. More generally,

²This is, we do not make any assumptions on the likelihood of measurements

³This is necessary to obtain a compact parameter estimate, and hence generality is not affected.

⁴i.e. the parameter p_{ij} remains completely unknown in this case

by construction, Θ_Ω is the dual body of the n_p dimensional hyper-rectangle defined by the uncertainty intervals Eq. (6.10).

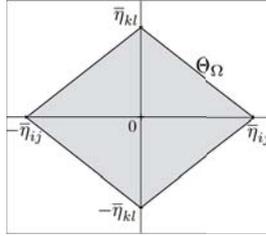


Figure 6.1.: Illustration of the residual set.

Allowing as many experiments as desired, the optimal experimental design translates into finding those experiments such that the uncertainty of the consistent parameter set is exactly the residual set. As shown in next example, the number of experiments to achieve this can be finite, or may be impossible in other cases.

Example 6.2: linear systems

We consider the (input-free) linear system equation

$$x_i^+ = \sum_{j=1}^{n_x} x_j p_{ij} + w_i, \quad (6.11)$$

where $w_i \in [-\bar{w}_i, \bar{w}_i]$ denotes the unknown, but bounded disturbance, and x_j/p_{ij} denote the initial conditions/completely unknown parameters for $i = \{1, \dots, n_x\}$ and $j = \{1, \dots, n_x\}$. The overall number of unknown parameters is thus n_x^2 . The initial conditions can be chosen from the intervals $x_j \in [\underline{x}_j, \bar{x}_j]$, with known \underline{x}_j and \bar{x}_j for all $j = 1, \dots, n_x$. We will investigate next how the initial condition domain influences the best possible estimate in the worst case setting.

Particularly, we consider two different initial condition domains, a symmetric and a non-negative⁵ one. To obtain the best estimate of the parameters p_{ij} for $j = \{1, \dots, n_x\}$, we consider the following *augmented basis* $V = \{v^{(1)}, \dots, v^{(q)}\}$ of experiments obtained as follows:

- i) $v^{(k)} = \begin{pmatrix} v_1^{(k)} \\ \vdots \\ v_{n_x}^{(k)} \end{pmatrix}$, $v_j^{(k)} = \{\underline{x}_j, \bar{x}_j\}$, $\forall j \in [1 : n_x]$ and $k \in [1 : q]$

⁵Non-negative initial condition domains are typical for biological systems, because concentrations are in general non-negative.

- ii) any combination of n_p vectors $v^{(k)} \in V$, i.e. $\{v^{(k(1))}, \dots, v^{(k(n_x))}\}$ with $k(l) \in [1 : q]$, is linearly independent.

Each vector of the basis V correspond to a particular experiment, e.g. $v^{(k)}$ to the experiment $Exp(x)$ with $x = (v_1^{(k)}, \dots, v_{n_x}^{(k)})$. To obtain finally the optimal parameter estimate p_{ij} , $j = \{1, \dots, n_x\}$, we apply the experiments for $k = \{1, \dots, q\}$ and measure the state x_i^+ . The resulting consistent parameter sets are depicted in Fig. 6.2, where $\bar{p}_{ij} = \theta_j^{(i)} + \frac{\bar{w}_i}{|\bar{x}_j|}$, $\underline{p}_{ij} = \theta_j^{(i)} - \frac{\bar{w}_i}{|\bar{x}_j|}$, and $\theta_z = (\theta_j^{(i)}, \theta_l^{(k)})^T$ for all $i, j, k, l \in \{1, \dots, n_x\}$.

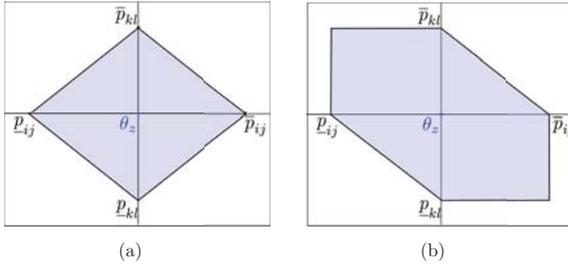


Figure 6.2.: Consistent parameter set Θ for the linear system (6.11) depending on the initial condition domain X_0 . 6.2(a): symmetric initial condition domain. 6.2(b): non-negative initial condition domain.

In case we a symmetric initial condition domain is considered, i.e. $\underline{x}_j = -\bar{x}_j$ for all $i \in [1 : n_x]$, we have $q = 2^{n_x-1}$. These q experiments and the respective $n_x \cdot q$ measurements (q for each state) according to the basis V yield the best possible consistent parameter set (in worst case), shown in Fig. 6.2(a). Under these conditions, the residual set is actually obtained.

In case a non-negative domain of initial conditions is considered, i.e. $\underline{x}_i = 0$, we have that $q = 2^{n_p} - 1$ experiments provide the best possible estimate, shown in Fig. 6.2(b). However, in this case, the consistent parameter set is larger than in the previous case. Note that in both cases, a finite number of experiments is sufficient to obtain the best possible estimate. This is however not in general the case.

6.5. Optimal experimental design

We have shown that the design matrix must have full row rank to obtain a compact set estimate of the parameters Θ . To this end, n_p experiments/measurements are required. To obtain furthermore the best possible estimate, considered in the previous section, in general more than n_p experiments are required, possibly infinitely many experiments. Because this may in practice not be realizable, we next consider the issue of obtaining the n_p experiments which provide a *maximum of information* for parameter estimation, thereafter denoted as the optimal experimental design problem.

The design of optimal experiments first requires a choice of an appropriate design criterion, i.e. what defines the maximum of information. Our motivation is to design the experiments which minimize the uncertainty in worst case. In other words, we aim to design experiments which provide a volume-minimum⁶ set estimate Θ .

To formalize this, reconsider Prop. 9, where now the possible design matrices \mathcal{A} have n_p rows. As design criterion, we choose the volume of Θ , denoted by $Vol(\Theta)$, where $Vol : \mathbb{R}^{n_p} \rightarrow \mathbb{R}_+$ is the volume map, see e.g. Schneider [1993]. We denote by $|\cdot|$ the absolute values.

Proposition 11 (Robust volume-optimal DOE)

The volume-optimal robust design criterion translates into maximizing the determinant of the (full rank) design matrix, i.e.

$$\min_{A \in \mathcal{A}} |Vol(\Theta)| \equiv \max_{A \in \mathcal{A}} |Det(A)|.$$

Proof. By Prop. 9, we have that $|Vol(\Theta)| = |Vol(\Theta_z + \Theta_\Omega)|$. For full rank design matrices (A^{-1} exists) Θ_z is a singleton set (compare Prop. 10), and we have $|Vol(\Theta)| = |Vol(\Theta_\Omega)| = |Vol(A^{-1}\Omega)|$. With $Vol(\Omega)$ known and nonzero, we have $|Vol(A^{-1}\Omega)| = |Det(A^{-1})| \cdot |Vol(\Omega)|$. Therefore, we have

$$\min_{A \in \mathcal{A}} |Vol(\Theta)| = \min_{A \in \mathcal{A}} |Det(A^{-1})| \cdot |Vol(\Omega)| = \max_{A \in \mathcal{A}} |Det(A)| \cdot |Vol(\Omega)|.$$

□

We next show how to determine maximum determinant design matrices. To this end, we exploit the special structure of the design matrix and compactness of the design variables. It is easy to see that each entry of A , $a(j, k)$, is bounded, and the bound is the same for all elements within a column, i.e.

$$0 \leq a(j, k) \leq \bar{a}_k, j = 1, \dots, n_x, k = 1, \dots, n_p, \tag{6.12}$$

where $\bar{a}_k = \prod_l \bar{x}_l \prod_m \bar{u}_m$ for some $l \in [1 : n_x]$ and $m \in [1 : n_u]$.

We can then state, exemplary for *non-negative domains* X_0 and U , the following result:

Corollary 7

The maximum determinant of A is constrained by

$$\max_{A \in \mathcal{A}} |Det(A)| \leq \prod_{k=1}^{n_p} \bar{a}_k |Det(H)|, \tag{6.13}$$

where H is a $n_p \times n_p$ $\{0, 1\}$ binary matrix with maximum determinant $|Det(H)|$. Equality holds in particular for linear systems.

⁶We focus on a volumetric criterion, although different criterias can be considered, e.g. a norm criterion.

Proof. The proof follows immediately from construction of A , (6.12), and multilinearity of determinant. \square

For $n = 1, 2, 3, \dots$ the (absolute) maximum determinants of $\{0,1\}$ binary matrices are 1, 1, 2, 3, 5, 9, 32, 56, 144, 320, 1458, 3645, 9477, \dots ⁷. In general, the maximum determinant matrices are not unique, e.g. any row or column permutation does not change the absolute value of the determinant. This leads to additional degrees of freedom when considering optimal design, which can be used to choose experiments which are easier to implement in practice. Exemplary, the first six Hadamard $\{0,1\}$ matrices are the square matrices from top left [Osborn, 2002]:

$$\begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 \end{pmatrix}.$$

Note also, exemplary for the case $n = 2$,

$$|\text{Det} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}| = |\text{Det} \begin{pmatrix} 1 & 0 \\ * & 1 \end{pmatrix}| = |\text{Det} \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}|, \quad (6.14)$$

where $*$ denotes an arbitrary choice in the interval $[0, 1]$. This degree of freedom can be exploited for choosing between optimal design matrices so as to ease implementation as shown in the illustrative Examples 6.2 and 6.3.

It is possible to consider the symmetric domains $X_0 = [-\bar{x}_i, \bar{x}_i]^{n_x}$ and $U = [-\bar{u}_i, \bar{u}_i]^{n_u}$. This leads to the $\{-1,1\}$ Hadamard maximum determinant problem, see e.g. Curtis and Kincaid [2006], Neubauer and Radcliffe [1997]. Respective matrices and determinant values (currently up to the order 28x28) can be found at Sloane [2002] (Sequence A003433).

In conclusion, for linear systems the Hadamard matrix based design is always possible under mild assumptions on the domains of the design variables. For polynomial systems, the case is more complicated, because the design matrix may be additionally constrained. In this case, Cor. 7 immediately provides an upper bound of the optimum. However, the matrix determinant problem needs to be solved for the specific case, e.g. by using matrix determinant algorithms under constraints such as presented in Vandenberghe et al. [1998].

⁷see Sloane [2002], Sequence A003432

Example 6.3: linear system

Consider the linear system

$$\begin{pmatrix} x_1^+ \\ x_2^+ \end{pmatrix} = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} p_{13} \\ p_{23} \end{pmatrix} u + \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}, \quad (6.15)$$

where $x_i \in [0, \bar{x}_i]$ for $i = \{1, 2\}$ and $u \in [0, \bar{u}]$ are non-negative domains, the disturbance $w_i \in [-\bar{w}_i, \bar{w}_i]$ for $i = \{1, 2\}$, p_{ij} the unknown parameters.

According to Proposition 10, six observations (three per state) are necessary to estimate the unknown parameters. The design matrices for $i = \{1, 2\}$ are given by

$$\mathcal{A}_i = \begin{pmatrix} x_1^{(1)} & x_2^{(1)} & u^{(1)} \\ x_1^{(2)} & x_2^{(2)} & u^{(2)} \\ x_1^{(3)} & x_2^{(3)} & u^{(3)} \end{pmatrix},$$

with $x_i^{(j)} \in [0, \bar{x}_i]$ and $u^{(j)} \in [0, \bar{u}]$ for $i = \{1, 2\}$ and $j = \{1, 2, 3\}$. The maximum determinant design matrices are given by 3×3 Hadamard matrices according to Cor. 7, e.g.

$$\max_{A_i \in \mathcal{A}_i} |Det(A_i)| = \bar{x}_1 \bar{x}_2 \bar{u} |Det(H_3)| = \bar{x}_1 \bar{x}_2 \bar{u} |Det \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}|.$$

The respective optimal three experiments $Exp(x_1, x_2, u)$ are

$$Exp(0, \bar{x}_2, \bar{u}), \quad Exp(\bar{x}_1, 0, \bar{u}), \quad Exp(\bar{x}_1, \bar{x}_2, 0).$$

Then, we have the optimal design considering n_p experiments according Cor. 7, with

$$Det(A) = \bar{x}_1 \bar{x}_2 \bar{u} \cdot 2.$$

It is important to note that additional experiments (considering the augmented basis V in Sec. 6.4, in total 7 experiments with the four remaining: $Exp(\bar{x}_1, \bar{x}_2, \bar{u})$, $Exp(\bar{x}_1, 0, 0)$, $Exp(0, \bar{x}_2, 0)$, $Exp(0, 0, \bar{u})$) allow to decrease the determinant toward the fundamental limit, compare Fig. 6.2(b).

As illustrated by the previous example, the maximum determinant of the design matrix along with the maximum amplitudes of initial conditions and inputs, i.e. that only ‘strong’ stimulations are relevant. Otherwise, only poor parameter estimates (large uncertainty) can be obtained.

The situation for polynomial systems is more intricate, because, in contrast to the linear case, the elements of the design matrix A and respectively may be not independent, and hence it is not always possible to directly implement the Hadamard design as shown in the following example.

Example 6.4: nonlinear system

Consider the nonlinear scalar system

$$x^+ = p_1 x + p_2 x^2 + w,$$

with $x \in [0, \bar{x}]$ and $w \in [-\bar{w}, \bar{w}]$. According to Prop. 10, two experiments are necessary to identify the two parameters p_1 and p_2 ; the design matrix is

$$A = \begin{pmatrix} x^{(1)} & (x^{(1)})^2 \\ x^{(2)} & (x^{(2)})^2 \end{pmatrix}.$$

Because the entries of the design matrix are not independent from one another, we can not implement the Hadamard design. In this case however, we can determine the maximum determinant by

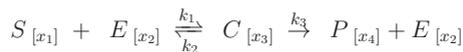
$$\max_{\substack{x^{(i)} \in X_0 \\ i=1,2}} |Det(A)| = \max |(x^{(1)}(x^{(2)})^2 - (x^{(1)})^2 x^{(2)})| = |x^{(1)} x^{(2)} ((x^{(2)}) - x^{(1)})|.$$

It is easy to see that the maximum is given by $x^{(1)} = \bar{x}$ and $x^{(2)} = \frac{1}{2}\bar{x}$, or vice versa. The maximum determinant of the A is then

$$\max_{\substack{x^{(i)} \in X_0 \\ i=1,2}} |Det(A)| = \frac{1}{4} \bar{x}^3.$$

Example 6.5: Michaelis-Menten

In the following example, we consider the design of optimal experiments for a frequently occurring reaction in systems biology, the Michaelis-Menten reaction:



The enzyme-substrate mechanism can be expressed using mass action kinetics and the Euler-forward discretization with step size h by

$$\begin{aligned} x_1^+ &= x_1 + h(-k_1 x_1 x_2 + k_2 x_3) + w_1 \\ x_2^+ &= x_2 + h(-k_1 x_1 x_2 + (k_2 - k_3) x_3) + w_2 \\ x_3^+ &= x_3 + h(+k_1 x_1 x_2 - (k_2 - k_3) x_3) + w_3 \\ x_4^+ &= x_4 + h(+k_3 x_3) + w_4, \end{aligned}$$

where x_i and x_i^+ denotes the initial and future state respectively, and the disturbances $w_i \in [-\bar{w}_i, \bar{w}_i]$. k_1 , k_2 , and k_3 denote the unknown reaction parameters.

We aim to determine the optimal experiments, denoted by $Exp(x_1, x_2, x_3, x_4)$, for identifying the three unknown parameters. We can choose non-negative initial conditions with:

$$0 \leq x_1 \leq \bar{S}, \quad 0 \leq x_2 \leq \bar{E}, \quad 0 \leq x_3 \leq \bar{C}, \quad 0 \leq x_4 \leq \bar{P}.$$

Let us furthermore consider that we cannot measure neither the enzyme's x_2 nor the complex' x_3 concentration.

According to Prop. 10, three measurements are necessary to identify the model parameters. Thus, we choose three experiments pairs $Exp(x_1^{(l)}, x_2^{(l)}, x_3^{(l)}, x_4^{(l)})$, $l \in \{1, 2, 3\}$, measuring two times the substrate concentration ($z_1^{(i)}$, $i = \{1, 2\}$) and one time the product concentration (z_4). We have

$$\begin{pmatrix} z_1^{(1)} \\ z_1^{(2)} \\ z_4 \end{pmatrix} = \begin{pmatrix} x_1^{(1)} \\ x_1^{(2)} \\ x_4^{(3)} \end{pmatrix} + \begin{pmatrix} -x_2^{(1)}x_1^{(1)} & x_3^{(1)} & 0 \\ -x_2^{(2)}x_1^{(2)} & x_3^{(2)} & 0 \\ 0 & 0 & x_3^{(3)} \end{pmatrix} \cdot \begin{pmatrix} hk_1 \\ hk_2 \\ hk_3 \end{pmatrix} + \begin{pmatrix} w_1^{(1)} \\ w_1^{(2)} \\ w_4 \end{pmatrix}.$$

By Prop. 11, the optimal design consists in maximizing the determinant:

$$\max |Det \begin{pmatrix} -x_2^{(1)}x_1^{(1)} & x_3^{(1)} & 0 \\ -x_2^{(2)}x_1^{(2)} & x_3^{(2)} & 0 \\ 0 & 0 & x_3^{(3)} \end{pmatrix}| = \bar{C} \cdot \max |Det \begin{pmatrix} -x_2^{(1)}x_1^{(1)} & x_3^{(1)} \\ -x_2^{(2)}x_1^{(2)} & x_3^{(2)} \end{pmatrix}|,$$

i.e. the first experiment is $Exp(*, *, \bar{C}, *)$, where '*' denotes an arbitrary initial concentration.

To determine the remaining two experiments, we can apply here the Hadamard design. The maximum 2×2 determinant corresponds e.g. to the Hadamard matrix

$$H_2 = \begin{pmatrix} 1 & 0 \\ * & 1 \end{pmatrix}.$$

This optimum is achieved by the experiments

$$Exp(\bar{S}, \bar{E}, 0, *), Exp(*, *, \bar{C}, *),$$

which yield the parameters:

$$\begin{aligned} k_1 &\in \frac{-z_1^{(1)} + \bar{S}}{h\bar{S}\bar{E}} + \frac{[-\bar{w}_1, \bar{w}_1]}{h\bar{S}\bar{E}}, \\ k_2 &\in \frac{-z_1^{(2)} + \bar{S}}{h\bar{C}} + \frac{[-\bar{w}_1, \bar{w}_1]}{h\bar{C}}, \\ k_3 &\in \frac{z_4 - P}{h\bar{C}} + \frac{[-\bar{w}_4, \bar{w}_4]}{h\bar{C}}, \end{aligned}$$

where P is the initial product concentration. The example shows while the optimal experimental design problem is non unique, the degrees of freedom can be utilized to facilitate realizations of the proposed experiments, and thereby to select among optimal experiments which, for instance, are easier to perform with the real process. In the present case, we can first of all choose which states are measured. Among the possible combinations, only some allow to provide a compact estimate of all the parameters, e.g. (x_1, x_4) . In the case (x_2, x_3) for instance however, k_2 and k_3 cannot be estimated. If this choice is made, the Hadamard design provides additional degrees of freedom, because the volume-optimal experiments are not unique.

In some cases, these degrees of freedom can be used to design (few) multi-step experiments with optimal properties. For the Michaelis-Menten reaction, we can perform the experiment $Exp(\overline{S}, \overline{E}, 0, *)$, until the steady state is reached (batch experiment). In the transient phase, the complex concentration will reach maximum levels, before the steady state is reached. Hence, the transient dynamic of the single experiment provides optimal information for parameter estimation, in particular at the beginning of the experiment (to infer k_1), and after the complex reached its maximum concentration (for inferring k_2 and k_3). Note also that this integration in comes to the price that more than three observations have to be made, since the timing when the complex concentration is maximum depends on the unknown parameters.

6.6. Summary and conclusions

We considered the design of robust optimal experiments for the purpose of parameter estimation and regarding this we studied the limits of the possible designs. Focusing on polynomial systems which are linear in the parameters, we pursued a worst case membership setting to design a number of one-step experiments so as to obtain a maximum of information for parameter estimation.

Within this setting, we showed that the number of experiments necessary to guarantee a compact set estimate of the beforehand completely unknown parameters is equivalent to the number of unknown parameters. For sufficiency, the design matrix, which is constructed from the experiments, must have full row rank. The compact set estimate of the parameters however can not be made as small as desired, even when considering an infinite number of experiments. We showed that the residual set defines a lower bound for the achievable results.

Having these fundamental limitations in mind, we finally considered the robust optimal experimental design problem considering the minimal number of experiments. As optimality criterion, we considered the worst-case volume of the consistent parameter set. We showed that the optimal design then translates into maximizing the determinant of the design matrix. For linear systems, this optimal design is always realizable already under mild assumptions regarding the possible input and initial

conditions. For polynomial systems, the Hadamard design may not always be directly implementable, because the elements of the design matrix may not be independent, nevertheless it gives an lower bound on the achievable optimum. In such cases, the determinant maximization can e.g. be solved using matrix determinant algorithms under constraints, e.g. Vandenberghe et al. [1998].

The proposed approach provides some general insights into the problem of learning the unknown system parameters. The initial conditions are of particular importance for design of experiments; if the initial conditions can be manipulated, then the unknown parameters can be bounded. Furthermore, strong stimulations (inputs as well as initial conditions) are optimal to learn as good as possible the unknown parameters. The limitations of the single-step approach are if some initial conditions can not be manipulated. Then, it may not be possible to obtain a set-valued estimate at all. A second limitation is if some states can not be measured; then, the parameters associated with the respective system equation can not be estimated. To overcome these limitations, a multi-step approach can be considered as in Borchers and Findeisen [2011].

7. Application Example: Cell Growth in Batch

In this chapter, we present a comprehensive application of the proposed falsification, estimation, analysis, and experimental design methods outlined in the Chapters 2–6 considering a cell growth process of AGE1.HN cells [Niklas et al., 2011]. This chapter is based on the work presented in Borchers et al. [2013], a research collaboration with S. Freund, A. Rath, and Udo Reichl from the Max-Planck Institute of Complex Dynamical Systems, Magdeburg. The experiments were conducted by A. Rath, Max-Planck Institute of Complex Dynamical Systems, Magdeburg, and the assay validation was conducted and evaluated by Susann Freund and coworkers, in particular Verena Lohr. For further details, on the cell cultivation and the experimental procedures, refer to Borchers et al. [2013].

7.1. Introduction

Production of bio-pharmaceuticals, for instance in mammalian cell culture, are frequently described by unstructured and segregated models. Although the compartmental structure of cells and the underlying metabolic pathways are not taken into account explicitly, these models do provide a sound mechanistic description of the considered process, which is required e.g. for model-based experimental design, optimization purposes, or controller synthesis. A main advantage of these models is they can be tailored to particular growth phases and process conditions. This is important since cell growth and product formation is known to depend on a variety of factors, e.g. the availability of substrates, inhibitors, or changes in the cultivation conditions (e.g. oxygen, temperature, pH [Tziampazis and Sambanis, 1994, Yu et al., 2011]). However, within a particular experimental setting, only some of these factors actually contribute to the observed cell dynamics. To obtain a concise model of the process, it is necessary to identify the main influencing factors of cell growth and basic metabolism and to distinguish apparent growth phases.

Due to noisy and erroneous experimental data, unknown kinetic parameters, and the large number of combinations of influencing factors, this issue however has only been addressed by a limited number of studies, so far. We here consider set-membership methods for falsification, estimation, analysis, and design of experiments to identify different growth phases and factors influencing cell growth and metabolism for a mam-

malian suspension cell line. Particularly, we investigate growth of AGE1.HN cells using the data obtained from two batch culture experiments, in bioreactor and shaker flask in serum-free medium. Besides cell concentrations, the uptake of glucose and glutamine as well as the release of ammonia and lactate were measured. The uncertainty of the experimental data is described in terms of bounded errors based on an assay validation performed in a separate study (data not shown).

We first distinguish apparent cell growth phases based on an outer-approximation of the specific growth rate as a function of time considering the observed increase of viable cell concentration. This way, we can show that cell growth in batch culture for the considered suspension cell line is divided into two main growth phases. The first phase is characterized by a maximum and constant specific growth rate. This phase is described consistently by a relatively simple segregated model including the main metabolites and the dynamics of viable and dead cells. The second phase however is more intricate, characterized by a declining specific growth rate until growth completely ceases. We demonstrated via falsification and analysis that glucose limitation and the pH of the medium are the governing mechanisms for the decline of the specific growth rate in both cultivation systems. An extended model is provided which describes the observed dynamics of cell growth and main metabolites, and the corresponding kinetic parameters as well as their confidence intervals are estimated. The study is complemented by an uncertainty and outlier analysis. Overall, we demonstrate that the proposed models to describe the complete time course of the experiments were in good accord with the observations.

The structure of this example is as follows: We first describe in Section 7.2 the cell growth process and measurement uncertainties obtained from an assay validation. In Section 7.3 we distinguish qualitatively different growth phases for AGE1.HN cells in our batch experiments. Subsequently, the growth phases are analyzed in detail in the Sections 7.4 and 7.5. Thereafter, we discuss briefly the design of experiments in Section 7.6, and finally provide some insights in the computational procedures in Section 7.7.

7.2. Data and process description

Growth of mammalian cells is known to be dependent on various factors, essentially on the availability of the substrates glucose (Glc) and glutamine (Gln). As a by-product of Glc and Gln consumption, lactate (Lac) and ammonia (Amn) are released. Basic properties of cell growth have been described in various publications for hybridoma [Batt and Kompala, 1989, Pörtner and Schäfer, 1996], myeloma [Frahm et al., 2003] and CHO cells considering unstructured models, refer also [Bailey and Ollis, 1986, Haag et al., 2005], and metabolic shifts have been investigated for AGE1.HN cells using metabolic flux analysis [Niklas et al., 2011].

In general, substrate and by-product yield factors as well as specific growth rates strongly depend on the cell line, the used medium, and the process strategy (batch or continuous). In the present case, two batch culture experiments have been performed, and the observations of the extra-cellular metabolites and cell concentrations are denoted by

$$\tilde{x}_i(t_j), \quad i \in \{1, \dots, n_x\}, \quad j \in \{0, 1, \dots, N\}. \quad (7.1)$$

Measurement uncertainty in the present can be inferred by assay validation, see e.g. Johnson and Wichern [2001], Wilkinson [1961]. First, we evaluate if the variances, for each extra-cellular metabolite and the cell concentrations, are homogeneously distributed or not. Subsequently, the standard deviation or the relative standard deviation of the method respectively is used to determine the respective 1-sigma confidence intervals; they are used thereafter as uncertainty bounds for the measurements. To this end, we used the F-test, i.e. comparing the variances at the lower and the upper measurement range (obtained by 8-fold measurements at the respective concentrations). We can then describe measurement uncertainty as follows:

Homogeneous (absolute) errors In case variances are homogeneously distributed (according to the F-test), we consider the standard deviation of the method σ_i regarding a calibration function of first order (two degrees of freedom) to derive the 1-sigma confidence intervals, see e.g. Funk et al. [2007]. The 1-sigma confidence interval is given by $x_i(t_j) \in [\underline{x}_i(t_j), \bar{x}_i(t_j)]$, where

$$\begin{aligned} \underline{x}_i(t_j) &= \tilde{x}_i(t_j) - \sigma_i, \\ \bar{x}_i(t_j) &= \tilde{x}_i(t_j) + \sigma_i. \end{aligned} \quad (7.2)$$

Non-homogeneous (relative) errors In case the variances are non-homogeneously distributed (according to the F-test), we consider the relative standard deviation of the method r_i (variation coefficient), see Funk et al. [2007] for details. The confidence intervals are then described by $x_i(t_j) \in [\underline{x}_i(t_j), \bar{x}_i(t_j)]$, where

$$\begin{aligned} \underline{x}_i(t_j) &= \tilde{x}_i(t_j)(1 - r_i/100), \\ \bar{x}_i(t_j) &= \tilde{x}_i(t_j)(1 + r_i/100). \end{aligned} \quad (7.3)$$

We furthermore have to take into account that the compounds are only detectable above a certain threshold. We denote the **limit of detection (LOD)** η_i as the lowest level at which a compound concentration can be detected. The detection threshold is taken into account by

$$\tilde{x}_i(t_j) \leq \eta_i \Rightarrow \underline{x}_i(t_j) = 0. \quad (7.4)$$

A summary of the measurement errors obtained by assay validation (data not shown) is provided in Tab. 7.1.

Table 7.1.: Statistical analysis of the measurement errors by validation assay.

	Amn	Glc	Gln*	Lac	X_d*	X_v*
	[mM]	[mM]	[mM]	[mM]	[10 ⁶ <i>cells</i> / <i>ml</i>]	[10 ⁶ <i>cells</i> / <i>ml</i>]
LOD (η_i)	0.30	3.91	0.82	2.98	0.00	0.00
SD of the method (σ_i)	0.03	0.39	(0.08)	0.30	(0.02)	(0.02)
% SD of the method (r_i)	(2.1%)	(1.9%)	5.9%	(1.7%)	6.2%	6.2%
monotonic behavior	↗	↘	↘	↗	↗	↗

*non-homogeneous variance. LOD: limit of detection. SD: standard deviation. % SD: relative standard deviation.

7.3. Identification of growth phases

Cell growth in batch culture typically follows certain growth phases. Initially, sufficient substrates for cell growth are available, while metabolic by-product concentrations are low. In this situation, cells have ideal conditions to grow, where the specific growth rate is at maximum. Subsequently, particularly when substrates are nearly depleted and by-product concentrations rise, the specific growth rate declines until growth completely ceases. This defines a second growth phase. Finally, because no substrates are available any more and by-product concentrations are high, total cell number decreases.

Before analyzing the cell growth dynamics and the main metabolites in detail, we aim to identify the cell growth phases based on the dynamics of the viable cell concentration. A simple mechanistic description for the dynamics of the viable cells is given by

$$\dot{X}_v = (\mu - K_d)X_v, \quad (7.5)$$

where μ the unknown specific growth rate, K_d the specific cell death rate, and X_v denotes the concentration of viable cells. For now, we consider for simplicity the specific cell death rate fixed to the reference value ($K_d = 0.003 \text{ h}^{-1}$, data not shown). To identify the growth phases for both experiments, we can consider the following reverse engineering approach. We treat $\mu = \mu(t)$ as an unknown and time-variant parameter, and determine the values of $\mu(t)$ which are consistent with the data (i.e. X_v) and the simple model (7.5), i.e. we estimate the lower and upper 1-sigma confidence intervals of $\mu(t)$ at each time sample. The results are depicted in Fig. 7.1.

The time-dependent specific growth rates are used subsequently to distinguish qualitatively different phases of growth of AGE1.HN cells. We thus characterize the first phase by assuming the specific growth rate μ_{\max} to be constant, i.e.

$$\mu(t) = \mu_{\max} = \text{const.} \quad (7.6)$$

Such a constant maximum specific growth rate corresponds to an exponential growth dynamics. This first phase starts at the beginning of the experiments ($t = 0$ h), and terminates at that time point when the specific growth rate can not be considered constant any more, compare Fig. 7.1. The phase lasts in the bioreactor for a maximum of 125 h, and in the shaker for a maximum of 128 h. After the phase of maximum growth, the specific growth rate decreases until growth completely ceases, as shown in Fig. 7.1. The second phase terminates when no cell growth is observed any more, i.e. when $\mu(t) = 0$. We determined therefore the second growth phase by

$$0 \leq \mu(t) < \mu_{\max}. \quad (7.7)$$

For both experiments, cell growth is observed for a maximum of 180 h. The final phase is characterized by a declining cell concentration, i.e.

$$\mu(t) < 0, \quad (7.8)$$

observed for $t \geq 180$ h.

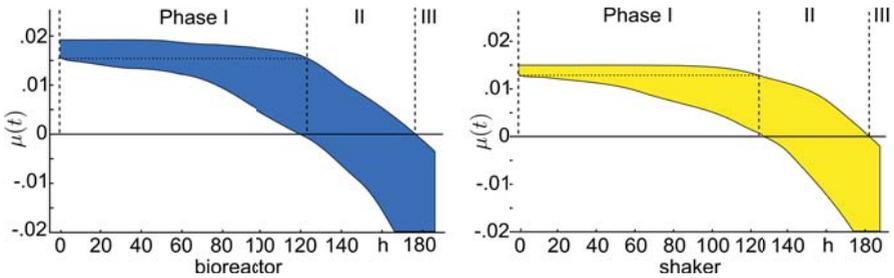


Figure 7.1.: Specific growth rate $\mu(t)$ and growth phases. Depicted is the 1-sigma confidence interval of the specific growth rate for the bioreactor (left) and the shaker (right) experiment. Phase I: exponential cell growth. Phase II: decreasing cell growth. Phase III: declining cell concentration.

The identification of the growth phases so far is based on the dynamics of the viable cell concentration alone. In the following, we aim to investigate the first two indicated growth phases more comprehensively by taking the dynamics of the metabolites into account.

7.4. Phase of exponential cell growth

We first investigate the exponential growth phase. To this end, we consider a mechanistic description of the uptake of glucose (Glc) and glutamine (Gln), the release of lactate (Lac) and ammonia (Amn), as well as the dynamics of dead (X_d) and viable

cells (X_v), following [Bailey and Ollis, 1986] and references therein:

$$\begin{aligned}
 d/dt \text{ Amn} &= \frac{\mu_{\max}}{Y_{X_v/\text{Amn}}} X_v + K_{\text{deg}} \text{Gln} \\
 d/dt \text{ Glc} &= -\frac{\mu_{\max}}{Y_{X_v/\text{Glc}}} X_v \\
 d/dt \text{ Gln} &= -\frac{\mu_{\max}}{Y_{X_v/\text{Gln}}} X_v - K_{\text{deg}} \text{Gln} \\
 d/dt \text{ Lac} &= \frac{\mu_{\max}}{Y_{X_v/\text{Lac}}} X_v \\
 d/dt X_d &= K_d X_v - K_{\text{lys}} X_d \\
 d/dt X_v &= (\mu_{\max} - K_d) X_v.
 \end{aligned} \tag{7.9}$$

Model (7.9) describes cell growth under ideal conditions. It includes the uptake of Glc and Gln, the release of Lac and Amn, and the lysis of dead cells. In addition, the spontaneous degradation of Gln to Amn is taken into account (see e.g. Bailey and Ollis [1986] and references therein). Note that this basic model does not include feedbacks, i.e. the specific growth rate μ_{\max} does not depend on the concentration of substrates or released products. Note also that the simple model (7.9) is only valid for non-negative concentrations and for low levels of accumulated by-products.

Parameter estimation and sensitivities Besides the values of the parameters K_{deg} and K_{lys} , which are known from previous experiments (data not shown, see Tab. 7.2), the parameters of the model (7.9) were unknown. To estimate the four yield factors, the death rate K_d , and the specific growth rate μ_{\max} , we consider the available data in Phase I.

Remark 12 *Parameter and state estimation does not depend on a guess neither for the initial parameters nor the initial conditions. Instead, the range of initial parameters covers several orders of magnitudes, compare Tab. 7.2, and also the initial conditions were uncertain.*

Subsequently, we determine the 1-sigma (68.3%) parameter confidence intervals and evaluated their sensitivity according to Eq. (4.2). The results are shown in Fig. 7.2 and Tab. 7.2. Results show that all the unknown parameters are sensitive. Conversely, this means that the experimental data contains sufficient information for identification of the unknown parameters. The maximum specific growth rate μ_{\max} is the most sensitive parameter ($\xi \approx 0.9$); the sensitivities ξ of the yield factors range from 0.6–0.9.

In a next step, we estimate the optimal parameter values regarding the least squares criterion (4.3) by using the proposed branch-and-bound scheme, refer Section 3.3.4. The optimization results are depicted in Fig. 7.2. Note that the confidence intervals are not symmetric regarding the optimal parameter values, which results from non-homogeneous errors and nonlinearity of the estimation problem.

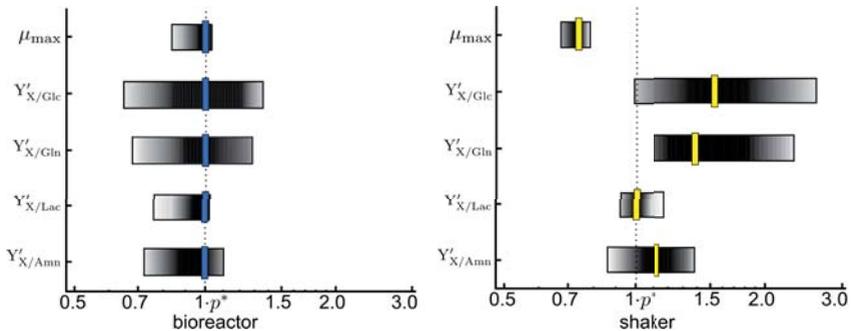


Figure 7.2.: Optimal parameter estimation using branch-and-bound. Depicted are the parameter confidence intervals (logarithmic scale, normalized), and the optimal parameters (vertical bars) regarding the sum of least squares.

Comparing both setups, the maximum specific growth rate is found to be larger in the bioreactor than in the shaker flask. In conclusion, the bioreactor provided more suitable growth conditions for AGE1.HN cells. Furthermore, the yield factors for the substrates, $Y'_{X/Glc}$ and $Y'_{X/Gln}$, are significantly lower in the in the bioreactor, i.e. the substrates are utilized more efficiently in the bioreactor than in the shaker to form viable cells.

Uncertainty and outlier analysis. To evaluate the effect of uncertain parameters and to detect outliers, we estimate the reachable states of Model (7.9) regarding the determined parameter confidence intervals. The results are depicted in Fig. 7.3. Results show that the model is rather robust with respect to parametric variations as expected, because the variations did not lead to significant or qualitatively different behavior.

Furthermore, by direct comparison of the reachable states with the measurement data, outliers were detected, see Fig. 7.3. Besides some lactate measurements from the shaker flask, we detected only few and isolated outliers. These isolated outliers can probably be explained from sampling or sample preparation errors, as well as the fact that we only considered the 1-sigma confidence limits of the parameters. Subsequently, we remove the outliers from the data set.

On the other hand, consecutive outliers as found for lactate in the shaker flask (see Fig. 7.3, right), can neither be explained by sampling nor sample preparation errors nor by statistics. Consecutive outliers typically indicate a model mismatch, i.e. a significant deviation of considered kinetics, e.g. additional metabolic pathways such as pyruvate pathway.

In summary, both parameter and uncertainty analysis support the proposed model. Only isolated outliers have been detected, besides lactate dynamics in the shaker flask.

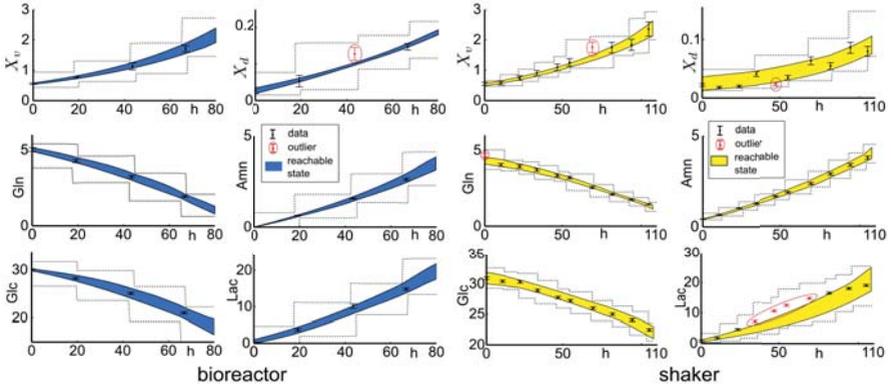


Figure 7.3.: Uncertainty analysis and outlier detection (exponential growth phase). Reachable state sets are shaded, outliers indicated by circles.

The model parameters are all sensitive, and the uncertainty analysis demonstrate robustness of the proposed model with respect to parametric uncertainties.

7.5. Phase of decreasing cell growth

We next consider the decrease in the specific growth rate with progressing time. In particular, we aim to provide a concise model which describes consistently the observed dynamics for $0 \leq t \leq 180$ h, i.e. covering the complete time course of both experiments.

To this end, it is necessary to modify the structure of the basic model (7.9), because this model is based on the simplifying assumption that substrates were (indefinitely) available and by-product concentrations were low, which is no longer the case toward the end of the experiments. To describe a substrate uptake kinetics, we use the Monod equation (see e.g. Zeng and Deckwer [1995], and below Eq. (7.11)). Substrate uptake kinetics also affects the production of Amn and Lac, because Amn is primarily produced from Gln (see Wahl et al. [2008]), and Lac from Glc, see Neermann and Wagner [1996]. Therefore, the production of Amn and Lac directly depends on the availability of the Gln and Glc, which had to be taken into account. The extended

Table 7.2.: Summary of parameters corresponding to the bioreactor and shaker flask experiment.

par.	unit	references	bioreactor			shaker		
			$[p_i, \bar{p}_i]$	opt	ξ	$[p_i, \bar{p}_i]$	opt	ξ
μ_{max}	1/h	2e-2-1.3e-1	[1.54,1.91]e-2	1.90e-2	0.90	[1.26,1.48]e-2	1.44e-2	0.92
$Y'_{X/Glc}$	$\frac{10^9}{mmol}$	6e-2-1.7	[0.93,1.93]e-1	1.44e-1	0.69	[1.41,3.75]e-1	2.21e-1	0.61
$Y'_{X/Gln}$	$\frac{10^9}{mmol}$	3e-2-1.6	[3.31,6.23]e-1	4.89e-1	0.73	[5.37,11.4]e-1	6.50e-1	0.69
$Y'_{X/Lac}$	$\frac{10^9}{mmol}$	7e-2-2.5e-1	[6.20,8.28]e-2	8.22e-2	0.87	[7.58,9.60]e-2	8.22e-2	0.89
$Y'_{X/Amn}$	$\frac{10^9}{mmol}$	5.0e-1-2.0	[3.98,6.03]e-1	5.48e-1	0.81	[4.69,7.55]e-1	5.78e-1	0.79
K_d	1/h	2.8e-4-3e-1	[1.66,3.45]e-3	2.66e-3	0.69	[5.09,12.4]e-4	7.20e-4	0.64
K_{Glc}	mM	1.5e-1-1.0	[0.89,2.43]	1.21	0.61	—*	—	—
K_{Gln}	mM	6e-2-8.0e-1	[0.01,1.35]	0.55	0.11	[0.13,1.52]	0.49	0.29
K_{pH}	pH	—	—	— ²	—	[0.51,4.91]	3.01	0.32
K_{Amn}	mM	1.0-2.0e1	—	—*	—	[5.16,15.8]	7.21	0.57
K_{Lac}	mM	8.0-1.4e2	—	—*	—	[27.7,72.9]	54.4	0.62
K_{deg}	1/h	1.5e-3 ¹	—	—	—	—	—	—
K_{lys}	1/h	1.0e-2 ¹	—	—	—	—	—	—
N	1/h	7.15	—	—	—	—	—	—

Literature values taken from [Cruz et al., 1999, Doyle and Griffiths, 1998, Goergen et al., 1993, Haggstrom, 2000, Meier et al., 1999, Ozturk and Palsson, 1990, Portner and Schafer, 1996]. p_i and \bar{p}_i denote the lower and upper limit of the 1-sigma parameter confidence interval. ξ denotes the sensitivity coefficient (Eq. (4.2)). ¹unpublished data. ²pH constant. *insensitive parameter.

model considered in the remainder is given by:

$$\begin{aligned}
d/dt \text{ Amn} &= \frac{\mu_{max}}{Y'_{X/Amn}} \frac{Gln}{Gln+K_{Gln}} X_v + K_{deg} \text{ Gln} \\
d/dt \text{ Glc} &= -\frac{\mu_{max}}{Y'_{X/Glc}} \frac{Glc}{Glc+K_{Glc}} X_v \\
d/dt \text{ Gln} &= -\frac{\mu_{max}}{Y'_{X/Gln}} \frac{Gln}{Gln+K_{Gln}} X_v - K_{deg} \text{ Gln} \\
d/dt \text{ Lac} &= \frac{\mu_{max}}{Y'_{X/Lac}} \frac{Glc}{Glc+K_{Glc}} X_v \\
d/dt X_d &= K_d X_v - K_{lys} X_d \\
d/dt X_v &= (\mu - K_d) X_v.
\end{aligned} \tag{7.10}$$

Furthermore, we have to identify the factors that explain the declining specific growth rate. In particular, we assume that the decline of the specific growth rate results from negative feedbacks, e.g. substrate depletion, by-product side effects, or

the pH of the medium in the shaker flask. To evaluate which of these factors actually contribute to the observed dynamics, we extend the model as described below.

First, we consider that the specific growth rate may be limited by either of the substrates Glc or Gln, e.g. Bailey et al. [1996]:

$$\mu = \mu_{\max} \frac{S}{S + K_s}, \quad (7.11)$$

where S denotes the substrates concentration, and K_s the (unknown) Monod constant. Second, accumulation of by-products may influence cell growth, i.e. Amn or Lac [Bailey et al., 1996]. Such an influence can be described by a non-competitive inhibition mechanism by

$$\mu = \mu_{\max} \frac{K_I}{I + K_I}, \quad (7.12)$$

where I is the by-product (inhibitor) concentration, and K_I the respective (unknown) inhibition constant. Third, for bacteria and hybridoma, the influence of the pH-value on cell growth has been reported by McQueen and Bailey [1990] and Eagle [1973], Ozturk et al. [1992]. Based on their studies, the influence of the pH on cell growth can be described qualitatively by a parabola

$$\mu = \mu_{\max} \cdot (K_{pH}(-pH^2 + 2NpH) - N^2) + 1) = \mu_{\max} \cdot \eta_{pH}, \quad (7.13)$$

where $N = 7.15$ (vertex) denotes the pH value where the specific growth rate is at its maximum, and K_{pH} an unknown parameter. Notice that all proposed feedback hypotheses contain besides μ_{\max} one unknown parameter. The single factor hypotheses for the specific growth rate are summarized in Tab. 7.3, and were analyzed hereafter.

Table 7.3.: Specific growth rates hypotheses.

factor	hypothesis	bioreactor	shaker
Glc	$\mu = \mu_{\max} \cdot \frac{Glc}{Glc + K_{Glc}}$	+ ^{3,4}	-
Gln	$\mu = \mu_{\max} \cdot \frac{Gln}{Gln + K_{Gln}}$	- ¹	+
Lac	$\mu = \mu_{\max} \cdot \frac{K_{Lac}}{Lac + K_{Lac}}$	-	+
Amn	$\mu = \mu_{\max} \cdot \frac{K_{Amn}}{Amn + K_{Amn}}$	- ²	+
pH	$\mu = \mu_{\max} \cdot \eta_{pH}$	-	+ ⁴

+ valid, - invalid hypothesis. ¹see Fig. 7.4A, ²see Fig. 7.4B, ³see Fig. 7.4C, ⁴compare Fig. 7.5.

Evaluating the feedback hypotheses. For evaluation, we choose again a reverse engineering approach. We already estimated the specific growth rates $\mu(t)$, depicted in Fig. 7.1, which reflects the ‘observed’ cell growth dynamics. In addition, we determine the 1-sigma confidence limits for the specific growth rates according to the hypotheses listed in Tab. 7.3. To this end, we consider the 1-sigma confidence interval μ_{max} as determined before and constrain the remaining unknown parameter to the range of the reported literature values, compare Tab. 7.2. Thus, we obtain the ‘hypothetical’ specific growth rates, which we can compare with the ‘observed’ specific growth rate for falsification purposes as shown in Fig. 7.4. Exemplary, Fig. 7.4A and Fig. 7.4B show the results for Gln-limitation and Amn-inhibition in the bioreactor, respectively. Because the ‘observed’ and the ‘hypothetical’ specific growth rates in both cases do not overlap at any time, we conclude that neither Gln-limitation nor Amn-inhibition alone explained the observed growth dynamics. On the contrary, Glc-limitation, see Fig. 7.4C, may be a valid hypothesis.

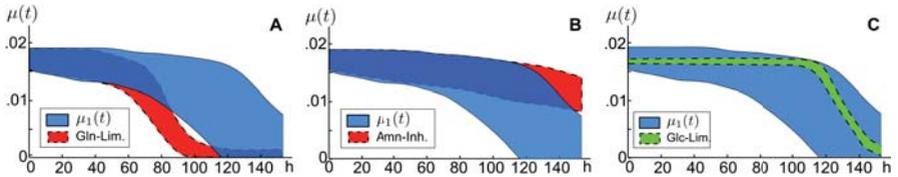


Figure 7.4.: Evaluation of feedback hypotheses. Comparison of the ‘observed’ (bioreactor) and three ‘hypothetical’ specific growth rates: Gln-limitation (A, falsified), Amn-inhibition (B, falsified), and Glc-limitation (C, validated).

The results are summarized in Tab. 7.3. Results show that Glc is essential for cell growth in the bioreactor. In contrast, Gln limitation does not affect growth of AGE1.HN cells. Furthermore, the by-products Amn and Lac do not affect cell growth within the observed concentration ranges significantly (considering physiologically meaningful inhibition constants).

The situation in the shaker flask is different, because Glc is available until the end of the experiment, i.e. Glc is not responsible for the decrease of the specific growth rate here. Instead, the decrease may be either explained by by-product inhibition, the proposed pH-dependency, or Gln-limitation. Hence, without additional knowledge, the results appear to be non-conclusive for the shaker flask. However, since we showed that cell growth is not affected by Gln-limitation in the bioreactor, we can rule out this hypothesis for the shaker flask. Furthermore, since the observed concentration ranges of Amn and Lac are comparable in the bioreactor and in the shaker flask (both slightly lower in the shaker), we can rule out Amn nor Lac inhibition too. For the shaker, only the pH dependency hypothesis remains. This is plausible, because it is known that the pH value decreases due to the release of the acid Lac; we thus

conclude that the decrease of the specific growth rate in the shaker is the result of the acidification of the medium by Lac. For further analysis, we determined the unknown parameters and the corresponding confidence intervals for Glc-limitation (K_{Glc}) and pH-dependency (K_{pH}), see Tab. 7.2. The parameters were found sensitive and in accord with the literature values. Finally, for Glc-limitation (bioreactor) and pH-dependence (shaker flask), we performed an uncertainty and outlier analysis as described before, see Fig. 7.5; this analysis demonstrated robustness against parametric variations, and only few (non-consecutive) additional outliers.

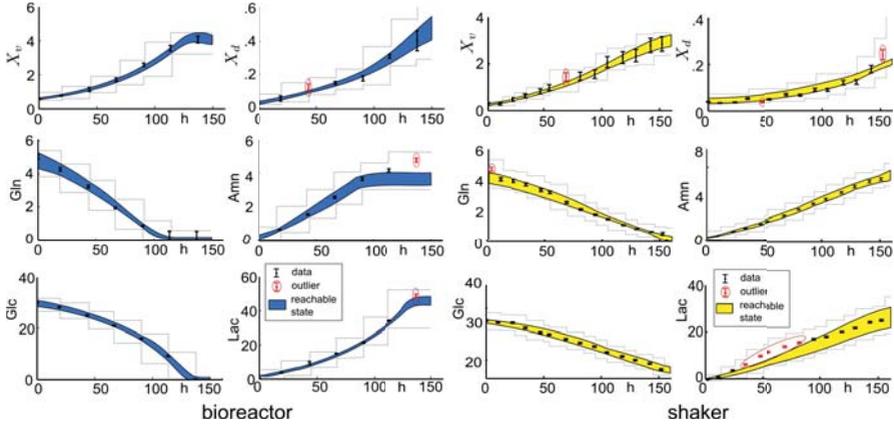


Figure 7.5.: Uncertainty analysis and outlier detection. Reachable state sets are shaded, outliers indicated by circles. Bioreactor: Glc-limitation, Shaker: pH-dependency.

7.6. Design of experiments

We finally consider the design of experiments for the exponential growth phase for the purpose of parameter estimation. To this end, reconsider the Model (7.9), where we rename for shorthand of notation the states as follows: x_1 and x_4 denote the ammonia and lactate concentration, x_2 and x_3 the glucose and glutamine concentration, and x_5 and x_6 the concentration of dead and viable cells respectively. Note that the Model (7.9) is nonlinear in the parameters, but we can consider a (bijjective) transformation $g : \mathbb{R}^8 \rightarrow \mathbb{R}^8$ of the model parameters, given by: $p_1 = \frac{1}{Y_{Amn}}\mu_{max}$, $p_2 = \frac{1}{Y_{Glc}}\mu_{max}$, $p_3 = \frac{1}{Y_{Gln}}\mu_{max}$, $p_4 = \frac{1}{Y_{Lac}}\mu_{max}$, $p_5 = K_d$, $p_6 = k_{deg}$, $p_7 = K_{lys}$, $p_8 = (\mu_{max} - K_d)$. The inverse transformation $g^{-1} : \mathbb{R}^8 \rightarrow \mathbb{R}^8$ exists for $p_1 \neq 0$, $p_2 \neq 0$, $p_3 \neq 0$, and $p_4 \neq 0$, given by: $\mu_{max} = p_8 + p_5$, $Y_{Amn} = \frac{p_8 + p_5}{p_1}$, $Y_{Glc} = \frac{p_8 + p_5}{p_2}$, $Y_{Gln} = \frac{p_8 + p_5}{p_3}$, $Y_{Lac} = \frac{p_8 + p_5}{p_4}$, $K_d = p_5$, $k_{deg} = p_6$, $K_{lys} = p_7$.

We furthermore consider the Euler forward discretization (Appendix A.1) with time step h , and introduce disturbances for each state to model the measurement uncertainties; thus, we obtain the difference equation system considered in the remainder:

$$\begin{aligned}
 x_1^+ &= x_1 + h(p_1x_6 + p_6x_3) + w_1 \\
 x_2^+ &= x_2 + h(-p_2x_6) + w_2 \\
 x_3^+ &= x_3 + h(-p_3x_6 - p_6x_3) + w_3 \\
 x_4^+ &= x_4 + h(p_4x_6) + w_4 \\
 x_5^+ &= x_5 + h(p_5x_6 - p_7x_5) + w_5 \\
 x_6^+ &= x_6 + h(p_8x_6) + w_6,
 \end{aligned} \tag{7.14}$$

where x_i and $x_i^+ = z_i$ denote the initial and future state concentrations respectively, and p_j , $j = \{1, \dots, 8\}$, the unknown parameters. We assume that the disturbances $w_i \in [-\bar{w}_i, \bar{w}_i]$ are unknown but bounded. The possible initial conditions are non-negative, i.e. $x_i \in [0, \bar{x}_i]$, $i = \{1, \dots, 6\}$. An experiment is denoted by $Exp(x_1, x_2, x_3, x_4, x_5, x_6)$.

We aim to design the optimal experiments to identify the unknown model parameters. To this end, according to Prop. 10, eight experiments/observations are required. Because each parameter is only appearing once, for each state equation, separate design matrices can be considered:

We have the one dimensional design matrices $A_2 = (x_6)$, $A_4 = (x_6)$, and $A_6 = (x_6)$, and hence consider the experiment $Exp(*, *, *, *, *, *, \bar{x}_6)$ (and measure z_2 , z_4 , z_6 , i.e. Glc, Lac, and the viable cell concentration X_v after the time h).

The design matrix

$$A_5 = \begin{pmatrix} -x_5^{(1)} & x_6^{(1)} \\ -x_5^{(2)} & x_6^{(2)} \end{pmatrix},$$

attains optimal determinant for e.g.

$$\max |Det(A_5)| = \bar{x}_5 \bar{x}_6 \cdot \begin{pmatrix} 1 & 0 \\ * & 1 \end{pmatrix}.$$

Thus, we choose the two experiments $Exp(*, *, *, *, \bar{x}_5, 0)$ and $Exp(*, *, *, *, *, \bar{x}_6)$ (and measure $z_5^{(1)}$ and $z_5^{(2)}$, i.e. the dead cell concentrations X_d respectively).

The remaining design matrices A_1 and A_3 , we have

$$A_1 = A_3 = \begin{pmatrix} x_3^{(1)} & x_6^{(1)} \\ x_3^{(2)} & x_6^{(2)} \end{pmatrix},$$

and choose the same maximum Hadamard matrix as before, corresponding to the remaining experiments $Exp(*, *, \bar{x}_3, *, *, 0)$ and $Exp(*, *, *, *, *, \bar{x}_6)$ (and measure $z_1^{(1)}$, $z_1^{(2)}$ and $z_3^{(1)}$, $z_3^{(2)}$, i.e. Amn and Gln concentrations respectively).

The optimal experiments are summarized as follows: We chose an experiment of the form $Exp(*, *, *, *, *, \bar{x}_6)$, and measure all state concentrations. Furthermore, we chose $Exp(*, *, *, *, \bar{x}_5, 0)$, and measure the dead cell concentration X_d . Last, we perform the experiment $Exp(*, *, \bar{x}_3, *, *, 0)$, and measure Amn concentration if $w_1 \leq w_3$, otherwise Gln concentration. This choice is possible because the parameter p_6 can be identified using either of the concentrations.

A multi-step experiment based on the above results can be considered as follows. We start the batch process with the medium and no viable cells $x_6(0) = 0$, particularly with a high level of glutamine (i.e. $x_3(0) = \bar{x}_3$). We then add a small portion of viable cells after some time T , i.e. $x_6(T) = \underline{x}_6$. The cells will grow until the medium is depleted. In this experiment setting, the most important information will be obtained at the beginning of the process, after inserting the portion of viable cells, and shortly before the medium will be depleted.

As for the second growth phase, a design of experiments should be considered by which the possible influencing factors are investigated one by one. This avoids superposition of several influencing factors, and hence more precise parameter estimates can be expected. For evaluating by-product influences, a pulse administration during the exponential growth phase will be advantageous. This pulse should be strong enough to decrease the influence of uncertainties, within biologically meaningful limits. Similarly, the influence of pH should be studied explicitly this way.

7.7. Conclusions

We investigated the growth and basic metabolism of AGE1.HN cells by using set-based methods and batch experiments performed in two commonly used environments. To this end, we described the uncertainty of the measurements by their 1-sigma confidence intervals obtained from a validation assay.

By using the set membership methods for falsification and estimation, we identified two qualitatively different growth phases. In both experiments, the first phase was characterized by a constant maximum specific growth rate corresponding to exponential cell growth. We demonstrated that this phase could be described very well by a relatively simple model including the main metabolites as well as dynamics of viable and dead cells. We estimated the optimal model parameter and the 1-sigma confidence intervals, and showed that the parameters are sensitive.

Using the outlier detection approach based on reachability analysis, besides lactate dynamics for the shaker flask experiment, only few and isolated outliers were detected. By comparing the results for both experiments, we showed that the bioreactor provided more suitable growth conditions than the shaker.

The second phase was characterized by a declining specific growth rate. To describe the observed dynamics for the complete time course of both experiments, we extended

the previous model including substrate limitations, and identified the factors which lead to the decrease of the specific growth rate using a falsification setting. Thus, we demonstrated that the governing mechanism for this was glucose limitation in the bioreactor, and the decrease of the pH value due to the release of lactate in the shaker. Only few additional outliers were detected.

Overall, we showed that the proposed dynamical models were in good accord with the experimental data, and we demonstrated that the set membership methods are valuable tools for modeling and analyzing cell growth dynamics.

8. Conclusions

8.1. Discussion and conclusions

In this thesis, we presented a set-membership framework for falsification, estimation, analysis, and design of experiments for polynomial dynamic systems. The proposed methods are based on the notion of bounded errors, and employ concepts from linear algebra and convex optimization. Key to falsification and estimation are the semidefinite and linear programming relaxations, which make it possible to use efficient set-membership techniques such as orthothopic bounding and homothety to obtain set-valued estimates of the parameters, states, or inputs. For the design of optimal experiments for the purpose of parameter estimation, a minimax strategy is considered, i.e. a worst-case and minimum volume setting. The key joint features of the methods are they apply to polynomial dynamic systems, they yield a robust perspective because uncertainties are taken explicitly into account, and they provide guaranteed and conclusive results. For these reasons, the proposed set-membership methods provides the means to address some of the challenges encountered when modeling biological processes.

The methods take as its first premise that uncertainty of the available data and disturbances can be described by bounded sets, for example that the actual value of a measurement is *unknown, but bounded*. This approach is motivated from the facts, firstly, that the probability density distributions, as e.g. required for hypothesis testing and estimation in the classical stochastic setting, are frequently complicated, very difficult to obtain, or not available at all considering biological systems. For example, time-series measurements in biology are typically *not* independently and homogeneously distributed, and determining the probability density distributions may require many repetitions of the experiment, which however is often not practicable; moreover, a priori parameter density distributions, as e.g. required for state estimation and outlier analysis, are frequently completely unknown. Secondly, even if the probability density distributions are known for a particular case, determining the unknown probability density distributions, e.g. of the parameters, can be very challenging or computationally demanding for biological systems due to the nonlinear model equations and non-homogeneity of the measurement's distributions. On the other hand, a bounded error description of the uncertainties can be easily derived from known probability density distributions, e.g. by considering the n-sigma confidence intervals of the measurement's probability density distributions.

Though, complications can arise from the bounded error description in regard to real world applications. If the uncertainties are underestimated, a model may be wrongly rejected, or the estimates may be biased. On the other hand, if uncertainties are overestimated, the results may be too pessimistic or non-conclusive. In particular outlying observations, or outliers, are a challenge to the proposed framework which requires specific measures. We addressed this issue in Sec. 5.3, and proposed model-generic outlier detection approaches. By allowing for few, non-consecutive outliers in the measurements, we successfully investigated the cell growth process (Ch. 7).

Besides uncertain and possibly erroneous measurement data, lack of data is a primary challenge when modeling biological processes. It is thus, first of all, important to integrate all available data for falsification, estimation, and analysis. Advantageously, the proposed methods account for data alongside measurements, e.g. a priori knowledge derived from first principles, information about possible correlations of the model's variables, as well as qualitative behavior such as monotonic dynamics, can be used (Sec. 2.2).

Lack of knowledge about the underlying mechanisms or the type of reactions often results in a number of competing model hypotheses with unknown parameters. To this end, we proposed an invalidation criterion (Sec. 4.1), i.e. to reject the hypotheses that are inconsistent with the available data. The parameter estimation problem (Sec. 4.2) in turn is considered from several perspectives in this thesis. To investigate how data uncertainty propagates to the parameter estimates, and hence to evaluate the precision of the estimates, the uncertainty intervals can be determined using interval bounding. The uncertainty intervals furthermore provide a measure of the sensitivity of the parameter with respect to model rejection. Optimal parameter values, e.g. regarding the sum of least squares, can be determined via branch-and-bound strategy. To study possible correlations among the parameters, homothety or partitioning can be considered.

In addition, it is often not possible to measure all states of a process, e.g. the concentrations of intra-cellular metabolites, or the measurement process might have been casually interrupted and some state values are missing. To reconstruct the missing state values we proposed an interval observer (Sec. 5.1). Conversely, to evaluate the influence of parametric uncertainties and disturbances onto the system's dynamics, a reachability analysis can be considered (Sec. 5.2). The reachable sets provide valuable information about the dynamical features, e.g. robustness with respect to parametric uncertainty, of the studied system, which is required for a comprehensive uncertainty and outlier analysis (Sec. 5.3).

Complementary, we considered in Ch. 6 the design of robust optimal experiments and studied the limits of the possible designs so as to obtain as good as possible estimates of the unknown parameters. For polynomial systems which are linear in the parameters, e.g. biochemical reaction networks derived from law of mass action, we derived necessary and sufficient conditions to obtain a bounded set estimate in

worst case of the beforehand completely unknown parameters (Sec. 6.3) using one-step experiments. The worst case parameter set however can not be made as small as desired, even when considering an infinite number of experiments; particularly, the attainable precision is bounded from below by the residual set (Sec. 6.4). Having these fundamental limitations in mind, we showed that the robust optimal design translates into maximizing the determinant of the design matrix (Sec. 6.5). For linear systems, this optimal design corresponds to the Hadamard maximum determinant problem, and is always realizable under mild assumptions regarding the possible input and initial conditions only. For polynomial systems, the proposed design may not always be implementable, nevertheless it gives an lower bound on the achievable optimum. In such cases, the determinant maximization can e.g. be solved using matrix determinant algorithms under constraints. These results provide general guidelines for designing experiments for parameter estimation, e.g. that the initial conditions are particularly important. If the initial conditions can be manipulated, then the unknown parameters can be determined using one-step experiments, or more precisely, a compact set of the parameters can be obtained. Furthermore, *extreme* stimuli (inputs as well as initial conditions) are significant to learn as good as possible the unknown parameters, compensating for additive disturbances.

A limitation of the proposed framework, which has to be considered whenever applying these methods for particular cases, is the computational cost. Overall, from experience, the reasonable size of models that can be considered in this framework using semidefinite programming relaxation is in the order of fifty variables, and using linear programming relaxation approx. 300 variables. Latter corresponds e.g. to a typical biochemical reaction network based on mass action, with ten state variables, each represented by twenty samples, and approx. twenty parameters. This figure refers to the current implementation [Streif et al., 2012]. Besides, there is a trade-off between computational demands and precision. To find here appropriate compromises for specific cases, it is e.g. possible to choose a specific relaxation from those applicable to our framework (Sec. 3.2), by considering different estimation techniques (Sec. 3.3), or to decompose an estimation problem into a number of smaller problems by taking the time ordering explicitly into account (Sec. 3.4).

8.2. Outlook

In this thesis, we derived set-membership methods for falsification, estimation, analysis, and design of experiments within an *off-line* context. Extending the existing methods to an on-line and real time setting, e.g. by allowing the optimization problems to be updated whenever a new measurement is available (see e.g. Mattingley and Boyd [2010]), would expand the applicability and possibly improve performance of

these methods, besides making new applications possible such as recursive estimation, closed-loop control, and adaptive experimental design.

Further work should also focus on improving the performance of the proposed methods. To this end, it is possible to tailor the construction of the underlying optimization problems to specific system structures, e.g. to exploit the special structure of biochemical reaction network based on the law of mass action. Here, the Horn-Jackson-Feinberg scheme Feinberg [1987], may be considered e.g. following Bullinger et al. [2007]. Besides, exploiting certain sparsity and symmetry properties of these specific systems may be crucial to apply these methods to system with many variables.

An interesting future direction consist in integrating stability constraints into the estimation and analysis problems, e.g. to investigate regions of robust stability and convergence of uncertain systems. Stability constraints may derive from a bounded-input bounded-output stability concept as e.g. considered in Cerone and Regruto [2007], Cerone et al. [2011], or input-to-state stability for discrete-time nonlinear systems as in Jiang and Wang [2001]. In addition, such an approach would also benefit from inner approximations of the solution which are so far not available.

The proposed approach for the design of optimal experiments may be extended to more general setups, e.g. a comprehensive polytopic setting for the design variables and disturbances, or a multi-step setting as considered already in Borchers and Findeisen [2011] for linear systems. Also, alternative design criteria should be considered in future, such as design of experiments for purpose of model falsification. To this end, the graph-based approach presented in Borchers et al. [2011a] may be combined with the proposed estimation methods.

A. Discretization of ODE Systems

Discretization refers to the transformation of a continuous-time ordinary differential equation system into an equivalent system of difference equations. Such a procedure is often required when an analytical solution is difficult or impossible to derive, e.g. for numerical integration (see e.g. Davis and Rabinowitz [1975]) or the implementation of controllers.

For simplicity of presentation, we consider the scalar initial value problem for $t_0 \leq t \leq t_N$:

$$\dot{x}(t) = f(x(t), p, u(t)), \quad x(t_0) = x_0. \quad (\text{A.1})$$

Discretization allows us to derive an approximate solution to the definite integral

$$\int_{x_0}^{x(t_N)} dx = x(t_N) - x_0 = \int_0^{t_N} f(x(\tau), p, u(\tau)) d\tau. \quad (\text{A.2})$$

Numerous algorithms exist to approximate the definite integral, depending on the desired properties of the approximation, i.e. the desired accuracy. The fundamental concept here is to utilize the truncated Taylor-series expansion of the analytical function $f(x(t), p, u(t))$, while the truncation order defines the order of the approximation method. Here, we describe for simplicity first-order approximation schemes to derive an approximate discrete-time model. First order methods, particularly the Euler forward and Euler backward method, are sufficient for most of the cases, and the accuracy of the approximation can be influenced by the sampling size. Though, it is important to note that higher order discretization schemes, such as trapezoidal, midpoint, or generally Runge-Kutta methods, allow for an adaptive sampling (variable step size, e.g. Runge-Kutta 45) and hence accurate solutions for a sparse sampling, and can be considered if necessary. For a comprehensive overview of higher order discretization schemes and related numerical stability issues, see e.g. Davis and Rabinowitz [1975], Ralston and Rabinowitz [2001].

A.1. Forward Euler

The simplest numerical approximation scheme is the first-order forward Euler, which advantageously leads to an explicit representation of the discrete-time model. A disadvantage of this approach is that it generally requires small sampling intervals leading to many samples to avoid numerical stability issues (see e.g. Stuart and Humphries [1996] for a numerical study on dynamical systems).

The scheme derives from the truncation of the Taylor series considering only the first order derivative, i.e. the linear approximation (extrapolating the tangent at t_0) with

$$x(t_0 + h) \approx x(t_0) + h\dot{x}(t_0), \quad (\text{A.3})$$

where h denoted the step size, and we have

$$\dot{x}(t_0) \approx \frac{x(t_0 + h) - x(t_0)}{h}. \quad (\text{A.4})$$

Denoting $\dot{x}(t_0) = f(x(t_0), p, u(t_0))$, assuming $u(t)$ being constant within the time interval $t \in [t_0, t_0 + h]$ (zero-order-hold), and denoting $x_{k-1} = x(t_0)$, $u_{k-1} = u(t_0)$, $x_k = x(t_0 + h)$, we finally obtain the discretization scheme (approximate solution of A.2)

$$x_k = x_{k-1} + hf(x_{k-1}, p, u_{k-1}). \quad (\text{A.5})$$

To overcome numerical stability issues, the Backward Euler scheme can be considered alternatively.

A.2. Backward Euler

The backward Euler scheme derives similarly from extrapolating the tangent at $t_0 + h$, i.e.

$$x(t_0 + h) \approx x(t_0) + h\dot{x}(t_0 + h). \quad (\text{A.6})$$

With $\dot{x}(t_0 + h) = f(x(t_0 + h), p, u(t_0 + h))$, assuming $u(t)$ being constant within the time interval $t \in [t_0, t_0 + h]$ (zero-order-hold), and denoting $x_{k-1} = x(t_0)$, $u_k = u(t_0 + h)$, $x_k = x(t_0 + h)$, we finally obtain the discretization scheme (numerical solution)

$$x_k = x_{k-1} + hf(x_k, p, u_k). \quad (\text{A.7})$$

The backward Euler method hence yields an implicit difference equation, and has advantageous numerical stability properties in contrast to explicit difference equation obtained by the forward Euler.

Note both schemes allow to consider a variable step size h , which is relevant e.g. when considering stiff systems or to adapt the sampling to the measurements. It should also be noted that for the considered approach it is not significant whether an explicit or implicit scheme is considered, thus we prefer the implicit Euler due to its advantageous numerical properties, see also Rumschinski [2012], Rumschinski et al. [2010b, 2012].

A.3. Sampling

Finally, an approximate solution of the finite integral is obtained by defining an appropriate sampling of the time window, e.g. $\{t_0, t_1, \dots, t_N\}$, and denote the corresponding sample index by k with $k \in \{0, 1, \dots, N\}$. We furthermore consider the general case where measurements are available at $\{t_0, \dots, t_M\}$. In this case, the technical requirement is that for each t_j , $j \in [0 : M]$ there exists a corresponding index k_j , $j \in [0 : M]$ with $k_j \in [0 : N]$ for all $j \in [0 : M]$. For example, a simple choice consists in choosing $N = M$, and a variable step size $h = (h_0, \dots, h_{N-1})$ with

$$h_j = t_{j+1} - t_j, \quad \forall j \in [0 : N - 1]. \quad (\text{A.8})$$

Then, then time indexes $k = \{0, \dots, N\}$ correspond to time instances $\{t_0, \dots, t_N\}$.

If the considered time window is large, the system admits stiff dynamics, or to use adaptive sampling, it might be necessary to introduce more samples as measurements are available. Although the sampling can in principle be chosen arbitrarily, we assume in the remainder:

Assumption 1 (Appropriate sampling)

The sampling is chosen, such that for each time instance at which a measurement is available, there is a corresponding integer time index. Furthermore, the sampling steps are chosen sufficiently small such that discretization errors can be neglected.

Overall, the ODE system 2.22 is transformed into the difference equation system 3.1 by considering a (first order) approximation scheme and a sampling according to Assumption 1. The resulting difference equation system is summarized by:

$$M : \begin{cases} f_i^k(x_k, x_{k-1}, p, u_{k-1}, w_{k-1}) = 0, & i \in [1:n_x] \\ g_i^k(y_{k-1}, x_{k-1}, p, u_{k-1}, w_{k-1}) = 0, & i \in [1:n_y] \\ h_i^k(y_{k-1}, x_k, x_{k-1}, p, u_{k-1}, w_{k-1}) = 0, & i \in [1:n_c] \end{cases} \quad (\text{A.9})$$

Here, x_k, u_k, y_k, w_k denote respectively the system states, inputs, outputs, and disturbances at the (integer) time index $k \in [1 : N]$, and $p \in \mathbb{R}^{n_p}$ the constant parameters. Note that the functions $f_i^k(\cdot), g_i^k(\cdot), h_i^k(\cdot)$ may be implicit, and we assume they are polynomial in the system variables.

B. Data Description

To formalize the data in terms of polytopic sets, we require the following definition:

Definition 6 (Direct sum)

The direct sum of two matrices $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{o \times q}$ is defined by

$$A \oplus B = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \in \mathbb{R}^{m+o \times n+q}.$$

B.1. A priori data

A priori knowledge is modeled by polytopic sets bounding the possible variable's values. The a priori bounding sets of the parameters p , states $x(t)$, inputs $u(t)$, outputs $y(t)$, disturbances $w(t)$, with $t_0 \leq t \leq t_N$ are respectively given by:

$$D_{prior} : \begin{cases} P \doteq \{p \in \mathbb{R}^{n_p} : A_p p \leq a_p\}, \\ X \doteq \{x \in \mathbb{R}^{n_x} : A_x x(t) \leq a_x\} \\ U \doteq \{u \in \mathbb{R}^{n_u} : A_u u(t) \leq a_u\} \\ Y \doteq \{y \in \mathbb{R}^{n_y} : A_y y(t) \leq a_y\} \\ \Omega \doteq \{w \in \mathbb{R}^{n_w} : A_w w(t) \leq a_w\}, \end{cases} \quad (\text{B.1})$$

with known the matrix-vector pairs (A_p, a_p) , (A_x, a_x) , (A_u, a_u) , (A_y, a_y) , (A_w, a_w) of appropriate dimensions. For the sampled system with $k \in [0 : N]$ corresponding to $\{t_0, t_1, \dots, t_N\}$, we have respectively

$$D_{prior} : \begin{cases} p \in P, \\ x_k \in X & k \in [0 : N], \\ u_k \in U & k \in [0 : N - 1], \\ y_k \in Y & k \in [0 : N], \\ w_k \in \Omega & k \in [0 : N - 1]. \end{cases} \quad (\text{B.2})$$

For simplicity of notation, we collect in the remainder the variables induced by the sampling by

$$z \doteq (p_1, \dots, p_{n_p}, x_0, \dots, x_N, u_0, \dots, u_{N-1}, y_0, \dots, y_N, w_0, \dots, w_{N-1}),$$

where $z \in \mathbb{R}^{n_z}$, $n_z = n_p + N(2 + n_x + n_u + n_y + n_w)$.

The a priori data can be summarized for shorthand of notation by

$$Z_{\text{prior}} \doteq P \times \prod_{k=0}^N X \times \prod_{k=0}^{N-1} U \times \prod_{k=0}^N Y \times \prod_{k=0}^{N-1} \Omega,$$

i.e.,

$$Z_{\text{prior}} \doteq \{z \in \mathbb{R}^{n_z} : A_{\text{prior}} z \leq a_{\text{prior}}\}, \quad (\text{B.3})$$

where

$$A_{\text{prior}} = A_p \oplus A_x \oplus \dots \oplus A_w, \quad a_{\text{prior}} = (a_p, a_x, \dots, a_w)^T. \quad (\text{B.4})$$

Note that the block structure can be exploited to increase performance.

Example As a simple example, we consider the parameters $p \in \mathbb{R}^{n_p}$ to be unknown, but bounded with $0 \leq p_i \leq 1$, for all $i \in [1 : n_p]$. We have

$$P_i = \left\{ p_i \in \mathbb{R} : \begin{pmatrix} 1 \\ -1 \end{pmatrix} p_i \leq \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\},$$

i.e.

$$P_i = \{A_{p_i} p_i \leq a_{p_i}\}.$$

The overall bounding set is given by concatenation, i.e. the Cartesian product

$$P \doteq P_1 \times P_2 \times \dots \times P_{n_p} = \{p \in \mathbb{R}^{n_p} : A_p p \leq a_p\},$$

with

$$A_p = \bigoplus_{i=1}^{n_p} A_{p_i}, \quad p = \begin{pmatrix} p_1 \\ \vdots \\ p_{n_p} \end{pmatrix}, \quad a_p = \begin{pmatrix} a_{p_1} \\ \vdots \\ a_{p_{n_p}} \end{pmatrix}.$$

B.2. Measurement data

We consider the inputs known and well-defined for any time index $k \in [0 : N]$. Furthermore, the measurements are given at $\{t_0, t_1, \dots, t_M\}$, and the associated uncertainties are known polytopic sets, i.e.

$$D_{\text{meas}} : \begin{cases} u_k \in U_k \doteq \{u \in R^{n_u} : A_{u_k} u \leq a_{u_k}\}, & k \in [0 : N - 1], \\ x_k \in X(t_k) \doteq \{x \in R^{n_x} : A_{x_k} x \leq a_{x_k}\} & k \in [0 : M], \\ y_k \in Y(t_k) \doteq \{y \in R^{n_y} : A_{y_k} y \leq a_{y_k}\} & k \in [0 : M], \end{cases} \quad (\text{B.5})$$

with know matrix-vector pair (A_{u_k}, a_{u_k}) for all $k \in [0 : N - 1]$, and $(A_{x_k}, a_{x_k}), (A_{y_k}, a_{y_k})$ for all $k \in [0 : M]$.

The measurement data is summarized for shorthand by

$$Z_{\text{meas}} \doteq \{z \in \mathbb{R}^{n_z} : A_{\text{meas}} z \leq a_{\text{meas}}\}.$$

B.3. Structural data

The structural data defines interrelations of variables, defined for $t_0 \leq t \leq t_N$, i.e. $k = 0, \dots, N$. We describe structural data by a set of inequalities of the form

$$D_{str} : \left\{ q_i(x_k, x_{k-1}, p, u_k, y_k) \leq 0 \quad i \in [1 : n_q] \quad k \in [0 : N], \right. \quad (\text{B.6})$$

where n_q denotes the number of constraints. Assuming that q_i for all $i \in [1 : n_q]$ are linear equations, we can write the equations by

$$Z_{q_i} \doteq \{z \in \mathbb{R}^{n_z} : A_{q_i} z \leq a_{q_i}\}, \quad i \in [1 : n_q],$$

with known matrix-vector pairs (A_{q_i}, a_{q_i}) . We summarize for shorthand the structural data by

$$Z_{str} \doteq Z_{q_1} \times \dots \times Z_{q_{n_q}} = \{z \in \mathbb{R}^{n_z} : A_{str} z \leq a_{str}\},$$

where

$$A_{str} = \begin{pmatrix} A_{q_1} \\ A_{q_2} \\ \vdots \\ A_{q_{n_q}} \end{pmatrix}, \quad a_{str} = \begin{pmatrix} a_{q_1} \\ a_{q_2} \\ \vdots \\ a_{q_{n_q}} \end{pmatrix}.$$

B.4. Summary

Finally, for shorthand of notation, we can summarize the available data, denoted by $D = D_{prior} \cap D_{meas} \cap D_{str}$, by

$$Z \doteq \{z \in \mathbb{R}^{n_z} : A_z z \leq a_z\},$$

where

$$A_z = \begin{pmatrix} A_{prior} \\ A_{meas} \\ A_{str} \end{pmatrix}, \quad a_z = \begin{pmatrix} a_{prior} \\ a_{meas} \\ a_{str} \end{pmatrix}.$$

Note that (A_z, a_z) typically contains some redundant constraints, which can be detected and removed following e.g. Mattheiss [1973] to ease computations.

C. Relaxation

C.1. Bounding sets for monomials

To derive bounds for the higher order monomials, which define the last $n_\xi - n_z - 1$ components of n_ξ , we utilize simple interval arithmetic. Consider the monomial of degree two, e.g.

$$\xi_{n_z+1} = z_i z_j.$$

For each element of z we assume given interval bounds, i.e. $z_i \in [\underline{z}_i, \bar{z}_i]$ and $z_j \in [\underline{z}_j, \bar{z}_j]$. Note that if the interval bounds are not given explicitly, they be easily derived by linear optimization, i.e. via vertex enumeration. Since the components are assumed positive, we have the bounding set

$$\xi_{n_z+1} \in [\underline{z}_i \underline{z}_j, \bar{z}_i \bar{z}_j] \Leftrightarrow \{A_{\xi_{n_z+1}} \xi_{n_z+1} \leq a_{\xi_{n_z+1}}\},$$

with

$$A_{\xi_{n_z+1}}, a_{\xi_{n_z+1}} : \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} \bar{z}_i \bar{z}_j \\ -\underline{z}_i \underline{z}_j \end{pmatrix}.$$

Finally, the matrix-vector pair (A_ξ, a_ξ) is given by

$$(A_z \oplus A_{\xi_{n_z+1}} \oplus \dots \oplus A_{\xi_{n_\xi}}, a_z \oplus a_{\xi_{n_z+1}} \oplus \dots \oplus a_{\xi_{n_\xi}}).$$

Similarly, a bounding set for each component of Ξ is obtained, e.g. considering the relation $\Xi = \xi \xi^T$. We denote the respective matrix-vector pair by (A_Ξ, a_Ξ) .

C.2. Feasible solution set

The set of solutions feasible for $\text{SDP}(Z)$ (3.7) is given by

$$\mathcal{R}_{\text{SDP}} \doteq \left\{ \Xi \in \mathbb{R}^{n_\xi \times n_\xi} : \begin{array}{l} \langle F_i^k, \Xi \rangle = 0 \quad k \in [1:N], i \in [1:n_x], \\ \langle G_i^k, \Xi \rangle = 0 \quad k \in [1:N], i \in [1:n_y], \\ \langle H_i^k, \Xi \rangle = 0 \quad k \in [1:N], i \in [1:n_h], \\ \langle D_i, \Xi \rangle = 0 \quad i \in [1:n_d] \\ A_\xi \Xi e_1 \leq a_\xi, \\ \Xi_{11} = 1, \\ \Xi \succeq 0 \end{array} \right\}.$$

We denote the projection of the $(\Xi_{1,2}, \dots, \Xi_{1,n_z+1})$ elements of the matrix $\Xi \in \mathbb{S}^{n_\xi}$ onto \mathbb{R}^{n_z} by $f_z(\Xi)$, $\mathbb{R}^{n_\xi \times n_\xi} \rightarrow \mathbb{R}^{n_z}$. The set of feasible solutions can then be written as

$$\mathcal{Z}_{\text{SDP}} = f_z(R_{\text{SDP}}),$$

with $\mathcal{Z}_{\text{SDP}} \subseteq \mathcal{Z}$ according to Theorem 2.

Analogously, we denote by \mathcal{R}_{LP} the solution set of the LP relaxation,

$$\mathcal{R}_{\text{LP}} \doteq \left\{ \Xi \in \mathbb{R}^{n_\xi \times n_\xi} : \begin{cases} \langle F_i^k, \Xi \rangle = 0 & k \in [1:N], i \in [1:n_x], \\ \langle G_i^k, \Xi \rangle = 0 & k \in [1:N], i \in [1:n_y], \\ \langle H_i^k, \Xi \rangle = 0 & k \in [1:N], i \in [1:n_h], \\ \langle D_i, \Xi \rangle = 0 & i \in [1:n_d] \\ A_\xi \Xi e_1 \leq a_\xi, \\ \Xi_{11} = 1, \\ \Xi \geq 0 \end{cases} \right\}.$$

In this case, it is easy to see that the set of solutions \mathcal{R}_{LP} and the respective projection $\mathcal{Z}_{\text{LP}} = f_z(\mathcal{R}_{\text{LP}})$ are polytopic sets by construction.

Bibliography

- E. Acuna and C. Rodriguez. A meta analysis study of outlier detection methods in classification. *Technical Paper*, 2004.
- B.D.O. Anderson and J.B. Moore. *Optimal filtering*. Prentice-Hall, 1979.
- J. Anderson and A. Papachristodoulou. On validation and invalidation of biological models. *BMC Bioinformatics*, 10(1):132, 2009.
- M. Anstreicher. Semidefinite programming versus the reformulation-linearization technique for nonconvex quadratically constrained quadratic programming. *Journal of Global Optimization*, 43(2-3):471–484, 2009.
- Z. Artstein and S. V. Raković. Set invariance under output feedback: a set-dynamics approach. *International Journal of Systems Science*, 42(4):539–555, 2011.
- Z. Artstein and S.V. Raković. Feedback and invariance under uncertainty via set-iterates. *Automatica*, 44(2):520–525, 2008.
- S.P. Asprey and S. Macchietto. Designing robust optimal dynamic experiments. *Journal of Process Control*, 12(4):545–556, 2002.
- J.P. Aubin, J. Lygeros, M. Quincampoix, S. Sastry, and N. Seube. Impulse differential inclusions: A viability approach to hybrid systems. *IEEE Transactions on Automatic Control*, 47(1):2–20, 2002.
- S. Audoly, G. Bellu, L. D’Angio, M.P. Saccomani, and C. Cobelli. Global identifiability of nonlinear models of biological systems. *IEEE Transactions on Biomedical Engineering*, 48(1):55–65, 2001.
- N. Bagheri, J. Stelling, and F.J. Doyle. Quantitative performance metrics for robustness in circadian rhythms. *Bioinformatics*, 23(3):358–364, 2007.
- E.W. Bai, Y. Ye, and R. Tempo. Bounded error parameter estimation: A sequential analytic center approach. *IEEE Transactions on Automatic Control*, 44:1107–1117, 1999.
- J.E. Bailey and D.F. Ollis. *Biochemical engineering fundamentals*. McGraw-Hill Science, 2nd edition, 1986.
- J.E. Bailey, A. Sburlati, V. Hatzimanikatis, K. Lee, W.A. Renner, and P.S. Tsai. Inverse metabolic engineering: a strategy for directed genetic engineering of useful phenotypes. *Biotechnology and Bioengineering*, 52(1):109–121, 1996.
- E. Balsa-Canto, M. Pfeifer, J. R. Banga, J. Timmer, and C. Fleck. Hybrid optimization method with general switching strategy for parameter estimation. *BMC Systems*

- Biology*, 2:26, 2008.
- J.R. Banga, K.J. Versyck, and J.F. Van Impe. Computation of optimal identification experiments for nonlinear dynamic process models: a stochastic global optimization approach. *Industrial & Engineering Chemistry Research*, 41(10):2425–2430, 2002.
- V. Barnett and T. Lewis. *Outliers in Statistical Data*. John Wiley & Sons, 3rd edition, 1994.
- D.G. Bates and C. Cosentino. Validation and invalidation of systems biology models using robustness analysis. *IET Systems Biology*, 5(4):229–244, 2011.
- B.C. Batt and D.S. Kompala. A structured kinetic modeling framework for the dynamics of hybridoma growth and monoclonal antibody production in continuous suspension cultures. *Biotechnology and Bioengineering*, 34(4):515–531, 1989.
- G. Belforte, B. Bona, and S. Frediani. Optimal sampling schedule for parameter estimation of linear models with unknown but bounded measurement errors. *IEEE Transactions on Automatic Control*, 32(2):179–182, 1987.
- R. Bellman and K.J. Åström. On structural identifiability. *Mathematical Biosciences*, 7(3-4):329–339, 1970.
- A. Bemporad, A. Garulli, S. Paoletti, and A. Vicino. A bounded-error approach to piecewise affine system identification. *IEEE Transactions on Automatic Control*, 50(10):1567–80, 2005.
- I. Ben-Gal. Outlier detection. In O. Maimon and L. Rokach, editors, *Data Mining and Knowledge Discovery Handbook*. Springer US, 2005.
- I. Ben-Gal, G. Morag, and A. Shmilovici. Context-based statistical process control. *Technometrics*, 45(4):293–311, 2003.
- O. Bernard and J.L. Gouzé. Closed loop observers bundle for uncertain biotechnological models. *Journal of Process Control*, 14(7):765–774, 2004.
- L.L. Bieber. Carnitine. *Annual Review of Biochemistry*, 57:261–283, 1988.
- S. Borchers. Cell-to-cell variability in TNF-induced apoptosis signalling. *Diploma thesis, Universität Stuttgart, Germany*, 2007.
- S. Borchers and R. Findeisen. Design of experiments for guaranteed parameter estimation in membership setting. In *Proc. 50th IEEE Conference on Decision and Control (CDC)*, pages 2602–07, 2011.
- S. Borchers, P. Rumschinski, S. Bosio, R. Weismantel, and R. Findeisen. Model discrimination and parameter estimation via infeasibility certificates for dynamical biochemical reaction networks. In *Proc. International Conference on Mathematical Modelling (MATHMOD)*, pages 787–795, 2009a.
- S. Borchers, P. Rumschinski, S. Bosio, R. Weismantel, and R. Findeisen. Model invalidation and system identification of biochemical reaction networks. In *Proc. 16th IFAC Symposium on Identification and System Parameter Estimation (SYSID)*,

- pages 245–250, 2009b.
- S. Borchers, P. Rumschinski, S. Bosio, R. Weismantel, and R. Findeisen. A set-based framework for coherent model invalidation and parameter estimation of discrete time nonlinear systems. In *Proc. 48th IEEE Conference on Decision and Control (CDC)*, pages 6786–92, 2009c.
- S. Borchers, S. Bosio, R. Findeisen, U. Haus, P. Rumschinski, and R. Weismantel. Graph problems arising from parameter identification of discrete dynamical systems. *Mathematical Methods of Operations Research*, 73(3):381–400, 2011a.
- S. Borchers, S.V. Raković, and R. Findeisen. Set membership parameter estimation and design of experiments using homothety. In *Proc. 18th IFAC World Congress*, 2011b.
- S. Borchers, S. Bosio, P. Rumschinski, A. Savchenko, R. Weismantel, and R. Findeisen. Estimation and analysis for polynomial systems: A relaxation-based approach. 2012. In preparation.
- S. Borchers, S. Freund, S. Streif, U. Reichl, and R. Findeisen. Identification of growth phases and influencing factors in cultivations with AGE1.HN cells using set-based methods. *PLOS ONE*, 2013. Accepted.
- G.E.P. Box and H.L. Lucas. Design of experiments in non-linear situations. *Biometrika*, 46(1–2):77–90, 1959.
- S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.
- J. Bremer. Carnitine - metabolism and function. *Physiol. Rev.*, 63:1420–1480, 1983.
- E. Bullinger, R. Findeisen, D. Kalamatianos, and P. Wellstead. System and control theory furthers the understanding of biological signal transduction. In I. Queinnec, S. Tarbouriech, G. Garcia, and S. Niculescu, editors, *Biology and Control Theory: Current Challenges*, volume 357 of *Lecture Notes in Control and Information Sciences*, pages 123–135. Springer Berlin / Heidelberg, 2007.
- E. Bullinger, D. Fey, M. Farina, and R. Findeisen. Identifikation biochemischer Reaktionsnetzwerke: Ein beobachterbasierter Ansatz (identification of biochemical reaction networks: An observer based approach). *AT-Automatisierungstechnik*, 56(5): 269–279, 2008.
- V. Cerone and D. Regruto. Parameter bounds for discrete-time Hammerstein models with bounded output errors. *IEEE Transactions on Automatic Control*, 48(10): 1855–1860, 2003.
- V. Cerone and D. Regruto. Parameter bounds evaluation of Wiener models with noninvertible polynomial nonlinearities. *Automatica*, 42(10):1775–1781, 2006.
- V. Cerone and D. Regruto. Bounding the parameters of linear systems with input backlash. *IEEE Transactions on Automatic Control*, 52(3):531–536, 2007.

- V. Cerone, D. Piga, and D. Regruto. Set-membership EIV identification through LMI relaxation techniques. In *Proc. of the American Control Conference (ACC)*, pages 2158–63, 2010.
- V. Cerone, D. Piga, and D. Regruto. Enforcing stability constraints in set-membership identification of linear dynamic systems. *Automatica*, 2011.
- K. Chaloner and I. Verdinelli. Bayesian experimental design: a review. *Statistical Science*, 10(3):273–304, 1995.
- C. Cobelli and J.J. Distefano. Parameter and structural identifiability concepts and ambiguities: a critical review and analysis. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, 239(1):7–24, 1980.
- A. Colman-Lerner, A. Gordon, E. Serra, T. Chin, O. Resnekov, D. Endy, C.G. Pesce, and R. Brent. Regulated cell-to-cell variation in a cell-fate decision system. *Nature*, 437:699–706, 2005.
- C. Conradi, D. Flockerzi, J. Raisch, and J. Stelling. Subnetwork analysis reveals dynamic features of complex (bio-)chemical networks. *Proceedings of the National Academy of Sciences USA*, 104(49):19175–80, 2007.
- A. Cornish-Bowden. *Fundamentals of Enzyme Kinetics*. Portland Press, 3rd edition, 2004.
- D. Crisan and A. Doucet. A survey of convergence results on particle filtering methods for practitioners. *IEEE Transactions on Signal Processing*, 50(3):736–746, 2002.
- H. Cruz, J. Moreira, and M. Carrondo. Metabolic shifts by nutrient manipulation in continuous cultures of BHK cells. *Biotechnology and Bioengineering*, 66(2):104–113, 1999.
- F.E. Curtis and R.K. Kincaid. Determinant optimization on binary matrices. *American Journal of Mathematical and Management Sciences*, 26(1–2):33–70, 2006.
- T. Dang, C. Le Guernic, and O. Maler. Computing reachable states for nonlinear biological models. *Theoretical Computer Science*, 2011.
- P.J. Davis and P. Rabinowitz. *Methods of numerical integration*. Academic Press New York, 1975.
- A. Doyle and J.B. Griffiths. *Cell and tissue culture laboratory procedures in biotechnology*. John Wiley & Sons, 1998.
- H. Eagle. The effect of environmental pH on the growth of normal and malignant cells. *Journal of Cellular Physiology*, 82(1):1–8, 1973.
- M. Ederer and E.D. Gilles. Thermodynamically feasible kinetic models of reaction networks. *Biophysical Journal*, 92(6):1846–1857, 2007.
- B. Efron. Bootstrap methods: another look at the jackknife. *The Annals of Statistics*, 7(1):1–26, 1979.

- T. Eissing, H. Conzelmann, E.D. Gilles, F. Allgöwer, E. Bullinger, and P. Scheurich. Bistability analyses of a caspase activation model for receptor-induced apoptosis. *The Journal of Biological Chemistry*, 279(35):36892–7, 2004.
- T. Eißing, F. Allgöwer, and E. Bullinger. Robustness properties of apoptosis models with respect to parameter variations and intrinsic noise. *Systems Biology*, 152(4):221, 2005.
- M.B. Elowitz, A.J. Levine, E.D. Siggia, and P.S. Swain. Stochastic gene expression in a single cell. *Science*, 297(5584):1183–6, 2002.
- D. Faller, U. Klingmüller, and J. Timmer. Simulation methods for optimal experimental design in systems biology. *Simulation*, 79(12):717–725, 2003.
- M. Farina, R. Findeisen, E. Bullinger, S. Bittanti, F. Allgöwer, and P. Wellstead. Results towards identifiability properties of biochemical reaction networks. In *Proc. 45th IEEE Conference on Decision and Control (CDC)*, pages 2104–09, San Diego, USA, 2006.
- M. Feinberg. Chemical reaction network structure and the stability of complex isothermal reactors-I. the deficiency zero and deficiency one theorems. *Chemical Engineering Science*, 42(10):2229–68, 1987.
- JL Figueroa, SI Biagiola, and OE Agamennoni. An approach for identification of uncertain Wiener systems. *Mathematical and Computer Modelling*, 48(1-2):305–315, 2008.
- R.A. Fisher. *The design of experiments*. Oliver & Boyd, 1935.
- R. FitzHugh. Mathematical models of threshold phenomena in the nerve membrane. *Bulletin of Mathematical Biology*, 17:257–278, 1955.
- E. Fogel and Y.F. Huang. On the value of information in system identification - bounded noise case I. *Automatica*, 18(2):229–238, 1982.
- B. Frahm, P. Lane, H. Märkl, and R. Pörtner. Improvement of a mammalian cell culture process by adaptive, model-based dialysis fed-batch cultivation and suppression of apoptosis. *Bioprocess and Biosystems Engineering*, 26(1):1–10, 2003.
- G. Franceschini and S. Macchietto. Model-based design of experiments for parameter precision: State of the art. *Chemical Engineering Science*, 63(19):4846–72, 2008.
- W. Funk, V. Dammann, and G. Donnevert. *Quality assurance in analytical chemistry: applications in environmental, food, and materials analysis, biotechnology, and medical engineering*. John Wiley & Sons, 2007.
- J. Goergen, A. Marc, and J. Engasser. Determination of cell-lysis and death kinetics in continuous hybridoma cultures from the measurement of lactate-dehydrogenase release. *Cytotechnology*, 11(3):189–195, 1993.
- J.L. Gouzé, A. Rapaport, and M.Z. Hadj-Sadok. Interval observers for uncertain biological systems. *Ecological modelling*, 133(1-2):45–56, 2000.

- M. Grube, H. Meyer zu Schwabedissen, K. Draber, D. Präger, K.-U. Möritz, K. Linemann, C. Fusch, G. Jedlitschky, and H.K. Kroemer. Expression, localization, and function of the carnitine transporter OCTN2 (SLC22A5) in human placenta. *Drug metabolism and disposition*, 33(1):31–37, 2005.
- J. Haag, A. Wouwer, and M. Remy. A general model of reaction kinetics in biological systems. *Bioprocess and Biosystems Engineering*, 27(5):303–309, 2005.
- L. Häggström. *Cell metabolism, animal*. John Wiley & Sons, 2000.
- J. Hasenauer, P. Rumschinski, S. Waldherr, S. Borchers, F. Allgöwer, and R. Findeisen. Guaranteed steady-state bounds for uncertain chemical processes. In *Proc. of the International Symposium on Advances in Control of Chemical Processes (ADCHEM)*, pages 674–679, Istanbul, Turkey, 2009.
- J. Hasenauer, P. Rumschinski, S. Waldherr, S. Borchers, F. Allgöwer, and R. Findeisen. Guaranteed steady state bounds for uncertain (bio-) chemical processes using infeasibility certificates. *Journal of Process Control*, 20(9):1076–83, 2010a.
- J. Hasenauer, S. Waldherr, K. Wagner, and F. Allgöwer. Parameter identification, experimental design and model falsification for biological network models using semidefinite programming. *IET Systems Biology*, 4(2):119–130, 2010b.
- D. Hawkins. *Identification of outliers*. Chapman and Hall, London, 1980.
- F. Hayot and C. Jayaprakash. NF- κ B oscillations and cell-to-cell variability. *Journal of Theoretical Biology*, 240(4):583–591, 2006.
- R. Heinrich and S. Schuster. *The regulation of cellular systems*. Springer US, 1996.
- V. Henri. General theory of action of certain hydrolases. *Comptes Rendus Hebdomadaires des Seances de l Academie des Sciences (Paris)*, 135:916–919, 1902.
- D. Henrion and J.B. Lasserre. GloptiPoly: Global optimization over polynomials with MATLAB and SeDuMi. *ACM Transactions on Mathematical Software (TOMS)*, 29(2):165–194, 2003.
- H. Hjalmarsson. From experiment design to closed-loop control. *Automatica*, 41(3):393–438, 2005.
- V. Hodge and J. Austin. A survey of outlier detection methodologies. *Artificial Intelligence Review*, 22(2):85–126, 2004.
- F. Horn and R. Jackson. General mass action kinetics. *Archive for Rational Mechanics and Analysis*, 47:81–116, 1972.
- K. Jaqaman and G. Danuser. Linking data to models: data regression. *Nature Reviews Molecular Cell Biology*, 7(11):813–819, 2006.
- L. Jaulin. Nonlinear bounded-error state estimation of continuous-time systems. *Automatica*, 38(6):1079–1082, 2002.
- L. Jaulin, M. Kieffer, O. Didrit, and É. Walter. *Applied Interval Analysis*. Springer, London, 2001.

-
- Z.P. Jiang and Y. Wang. Input-to-state stability for discrete-time nonlinear systems. *Automatica*, 37(6):857–869, 2001.
- R.A. Johnson and D.W. Wichern. *Applied multivariate statistical analysis*. Prentice Hall, 5th edition, 2001.
- I.T. Jolliffe. *Principal component analysis*. Springer, 2nd edition, 2002.
- M. Joshi, A. Seidel-Morgenstern, and A. Kremling. Exploiting the bootstrap method for quantifying parameter confidence intervals in dynamical systems. *Metabolic Engineering*, 8(5):447–455, 2006.
- M. Kieffer and E. Walter. Guaranteed nonlinear state estimator for cooperative systems. *Numerical algorithms*, 37(1):187–198, 2004.
- M. Kieffer and E. Walter. Guaranteed nonlinear state estimation for continuous-time dynamical models from discrete-time measurements. In *Proc. 5th IFAC Symposium on Robust Control Design*. Citeseer, 2006.
- S. Kikuchi, D. Tominaga, M. Arita, K. Takahashi, and M. Tomita. Dynamic modeling of genetic networks using genetic algorithm and s-system. *Bioinformatics*, 19:643–650, 2003.
- S. Kim and M. Kojima. Exact solutions of some nonconvex quadratic optimization problems via SDP and SOCP relaxations. *Computational Optimization and Applications*, 26(2):143–154, 2003.
- S. Kirkpatrick, C. D. Gellatt, and M. P. Vecchi. Optimization by simulated annealing. *Science*, 220:671–680, 1983.
- H. Kitano. Biological robustness. *Nature Reviews Genetics*, 5(11):826–837, 2004.
- E. Klipp, R. Herwig, A. Kowald, C. Wierling, and H. Lehrach. *Systems Biology in Practice: Concepts, Implementation and Application*. Wiley-VCH, 2005.
- M. Kojima. A general framework for convex relaxation of polynomial optimization problems over cones. *Journal of Operations Research Society of Japan*, 46:125–144, 2002.
- M. Kojima and L. Tuncel. Second order cone programming relaxation of nonconvex quadratic optimization problems. *Optimization Methods and Software*, 15:201–224, 2001.
- M Krauss, S Schaller, S Borchers, R Findeisen, J Lippert, and L Kuepfer. Integrating cellular metabolism into a multiscale whole-body model. *PLoS Computational Biology*, 2012.
- A. Kremling, S. Fischer, K. Gadkar, F.J. Doyle, T. Sauter, E. Bullinger, F. Allgöwer, and E.D. Gilles. A benchmark for methods in reverse engineering and model discrimination: problem formulation and solutions. *Genome Research*, 14(9):1773–85, 2004.

- L. Kuepfer, U. Sauer, and P. A. Parrilo. Efficient classification of complete parameter regions based on semidefinite programming. *BMC Bioinformatics*, 8:12, 2007.
- J.B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization*, 11:796–817, 2001.
- J.B. Lasserre, D. Henrion, C. Prieur, and E. Trélat. Nonlinear optimal control via occupation measures and LMI-relaxations. *SIAM Journal on Control and Optimization*, 47:1643–66, 2008.
- J. Lawrence. Polytope volume computation. *Mathematics of Computation*, 57(195): 259–271, 1991.
- B. Le Roux and H. Rouanet. *Geometric data analysis: from correspondence analysis to structured data analysis*. Springer, 2004.
- E.L. Lehmann. Some history of optimality. *Lecture Notes, Monograph Series*, pages 11–17, 2009.
- C. Lemaréchal. Lagrangian relaxation. *Computational Combinatorial Optimization*, 2241:112–156, 2001.
- Lendek, R. Babuška, and B. De Schutter. State estimation under uncertainty: A survey. Technical Report 06-004, Center for Systems and Control, Delft University of Technology, 2006.
- C. Letellier, S. Elaydi, L.A. Aguirre, and A. Alaoui. Difference equations versus differential equations, a possible equivalence for the Rössler system? *Physica D*, 195: 29–49, 2004.
- J.M. Levsky, S.M. Shenoy, R.C. Pezo, and R.H. Singer. Single-cell gene expression profiling. *Science*, 297:836–840, 2002.
- H. Liu, S. Shah, and W. Jiang. On-line outlier detection and data cleaning. *Computers & Chemical Engineering*, 28(9):1635–47, 2004.
- L. Ljung. *System identification. Theory for the user*. Prentice Hall, 2nd edition, 1998.
- J. Lofberg. YALMIP: a toolbox for modeling and optimization in MATLAB. In *IEEE International Symposium on Computer Aided Control Systems Design*, pages 284–289, 2004.
- L. Lovasz and A. Schrijver. Cones of matrices and set functions, and 0-1 optimization. *SIAM Journal on Optimization*, 1:166–190, 1991.
- W. Lysiak, P.P. Toth, C.H. Suelter, and L.L. Bieber. Quantitation of the efflux of acylcarnitines from rat heart, brain and liver mitochondria. *Journal of Biological Chemistry*, 261:13698–703, 1986.
- O. Maler. Computing reachable sets: An introduction. 2004. unpublished.
- N.V. Mantzaris. Single-cell gene-switching networks and heterogeneous cell population phenotypes. *Computers and Chemical Engineering*, 29:631–643, 2005.

- S. Marino, I.B. Hogue, C.J. Ray, and D.E. Kirschner. A methodology for performing global uncertainty and sensitivity analysis in systems biology. *Journal of Theoretical Biology*, 254(1):178–196, 2008.
- D.W. Marquardt. An algorithm for least-squares of nonlinear parameters. *SIAM Journal of Applied Mathematics*, 11:431–441, 1963.
- W. Marquardt and M. Mönnigmann. Constructive nonlinear dynamics in process systems engineering. *Computers & Chemical Engineering*, 29(6):1265–75, 2005.
- SW Marvel and CM Williams. Set membership experimental design for biological systems. *BMC Systems Biology*, 6:21, 2012.
- T.H. Mattheiss. An algorithm for determining irrelevant constraints and all vertices in systems of linear inequalities. *Operations Research*, 21(1):247–260, 1973.
- J. Mattingley and S. Boyd. Automatic code generation for real-time convex optimization. In Y.C. Eldar and D.P. Palomar, editors, *Convex optimization in signal processing and communications*, pages 1–41. Cambridge University Press, 2010.
- F. Mazenc and O. Bernard. Interval observers for linear time-invariant systems with disturbances. *Automatica*, 47(1):140–147, 2011.
- A. McQueen and J.E. Bailey. Effect of ammonium ion and extracellular pH on Hybridoma cell-metabolism and antibody-production. *Biotechnology and Bioengineering*, 35(11):1067–1077, 1990.
- S. Meier, T. Hatton, and D. Wang. Cell death from bursting bubbles: Role of cell attachment to rising bubbles in sparged reactors. *Biotechnology and Bioengineering*, 62(4):468–478, 1999.
- P. Mendes and D.B. Kell. Non-linear optimization of biochemical pathways: Applications to metabolic engineering and parameter estimation. *Bioinformatics*, 14: 869–883, 1998.
- M. Milanese and G. Belforte. Estimation theory and uncertainty intervals evaluation in presence of unknown but bounded errors: Linear families of models and estimators. *IEEE Transactions on Automatic Control*, 27:408–414, 1982.
- M. Milanese and C. Novara. Set membership identification of nonlinear systems. *Automatica*, 40(6):957–975, 2004.
- M. Milanese and A. Vicino. Optimal estimation theory for dynamic systems with set membership uncertainty: An overview. *Automatica*, 27(6):997–1009, 1991.
- I.M. Mitchell and C.J. Tomlin. Overapproximating reachable sets by Hamilton-Jacobi projections. *Journal of Scientific Computing*, 19(1):323–346, 2003.
- H.D. Mittelmann. An independent benchmarking of SDP and SOCP solvers. *Mathematical Programming*, 95(2):407–430, 2003.
- S.H. Mo and J.P. Norton. Fast and robust algorithm to compute exact polytope parameter bounds. *Mathematics and Computers in Simulation*, 32(5-6):481–493,

- 1990.
- M. Moisan, O. Bernard, and J.L. Gouzé. Near optimal interval observers bundle for uncertain bioreactors. *Automatica*, 45(1):291–295, 2009.
- C.G. Moles, P. Mendes, and J.R. Banga. Parameter estimation in biochemical pathways: A comparison of global optimization methods. *Genome Research*, 13:2467–74, 2003.
- M. Mönnigmann, W. Marquardt, C. Bischof, T. Beelitz, B. Lang, and P. Willems. A hybrid approach for efficient robust design of dynamic systems. *SIAM Review*, 49(2):236–254, 2007.
- M. Morohashi, A.E. Winn, M.T. Borisuk, H. Bolouri, J. Doyle, and H. Kitano. Robustness as a measure of plausibility in models of biochemical networks. *Journal of Theoretical Biology*, 216(1):19–30, 2002.
- M. Müller. Über das Fundamentaltheorem in der Theorie der gewöhnlichen Differentialgleichungen. *Mathematische Zeitschrift*, 26(1):619–645, 1927.
- J. Neermann and R. Wagner. Comparative analysis of glucose and glutamine metabolism in transformed mammalian cell lines, insect and primary liver cells. *Journal of cellular physiology*, 166(1):152–169, 1996.
- Y. Nesterov and A. Nemirovski. *Interior-point polynomial algorithms in convex programming*. Society for Industrial Mathematics, 1994.
- M.G. Neubauer and A.J. Radcliffe. The maximum determinant of ± 1 matrices. *Linear Algebra and its Applications*, 257:289–306, 1997.
- J. Niklas, E. Schraeder, V. Sandig, T. Noll, and E. Heinzle. Quantitative characterization of metabolism and metabolic shifts during growth of the new human cell line AGE1.HN using time resolved metabolic flux analysis. *Bioprocess and Biosystems Engineering*, 34(5):533–545, 2011.
- J.P. Norton. Identification and application of bounded-parameter models. *Automatica*, 23(4):497–507, 1987.
- C. Novara. Experiment design in nonlinear set membership identification. In *Proc. of the American Control Conference (ACC)*, pages 1566–71, 2007.
- N. Okamura, S. Ohnishi, H. Shimaoka, R. Norikura, and H. Hasegawa. Involvement of recognition and interaction of carnitine transporter in the decrease of L-carnitine concentration induced by pivalic acid and valproic acid. *Pharmacological Research*, 23(8):1729–35, Aug 2006.
- N. Oreskes and K. Belitz. Philosophical issues in model assessment. *Model validation: Perspectives in hydrological science*, pages 23–41, 2001.
- J.H. Osborn. The Hadamard maximal determinant problem. Honours Thesis, University of Melbourne, 2002.

-
- S. Ozturk and B. Palsson. Chemical decomposition of glutamine in cell-culture media - effect of media type, pH, and serum concentration. *Biotechnology Progress*, 6(2): 121–128, 1990.
- S.S. Ozturk, M.R. Riley, and B.O. Palsson. Effects of ammonia and lactate on hybridoma growth, metabolism, and antibody production. *Biotechnology and Bioengineering*, 39(4):418–431, 1992.
- P.A. Parrilo. Semidefinite programming relaxations for semi-algebraic problems. *Mathematical Programming, Series B*, 96:293–320, 2003.
- M. Peifer and J. Timmer. Parameter estimation in ordinary differential equations for biochemical processes using the method of multiple shooting. *IET Systems Biology*, 1(2):78–88, 2007.
- K. Poolla, P. Khargonekar, A. Tikku, J. Krause, and K. Nagpal. A time-domain approach to model validation. *IEEE Transactions on Automatic Control*, 39:951–959, 1994.
- R. Pörtner and T. Schäfer. Modelling hybridoma cell growth and metabolism - a comparison of selected models and data. *Journal of Biotechnology*, 49(1-3):119–135, 1996.
- S. Prajna. Barrier certificates for nonlinear model validation. *Automatica*, 42:117–126, 2006.
- L. Pronzato. Optimal experimental design and some related control problems. *Automatica*, 44(2):303–325, 2008.
- L. Pronzato and E. Walter. Experiment design for bounded-error models. *Mathematics and Computers in Simulation*, 32(5):571–584, 1990.
- G.P. Quinn and M.J. Keough. *Experimental design and data analysis for biologists*. Cambridge University Press, 2002.
- T. Raïssi, G. Videau, and A. Zolghadri. Interval observer design for consistency checks of nonlinear continuous-time systems. *Automatica*, 46(3):518–527, 2010.
- T. Raïssi, D. Efimov, and A. Zolghadri. Interval state estimation for a class of nonlinear systems. *IEEE Transactions on Automatic Control*, 57(1):260–265, 2012.
- S.V. Raković and M. Fiacchini. Approximate reachability analysis for linear discrete time systems using homothety and invariance. In *Proc. 17th IFAC World Congress*, 2008.
- S.V. Rakovic, E.C. Kerrigan, D.Q. Mayne, and J. Lygeros. Reachability analysis of discrete-time systems with disturbances. *IEEE Transactions on Automatic Control*, 51(4):546–561, 2006.
- A. Ralston and P. Rabinowitz. *A first course in numerical analysis*. Courier Dover Pub, 2001.

- M.V. Ramana. *An algorithmic analysis of multiquadratic and semidefinite programming problems*. PhD thesis, John Hopkins University, 1994.
- M.V. Ramana, L. Tunçel, and H. Wolkowicz. Strong duality for semidefinite programming. *SIAM Journal on Optimization*, 7(3):641–662, 1997.
- S. Ramsey, A. Ozinsky, A. Clark, K.D. Smith, P. de Atauri, V. Thorsson, D. Orrell, and H. Bolouri. Transcriptional noise and cellular heterogeneity in mammalian macrophages. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 361:495–506, 2006.
- J.M. Raser and E.K. O’Shea. Noise in gene expression: origins, consequences, and control. *Science*, 309(5743):2010–3, 2005.
- A.H.G. Rinnooy-Kan and G.T. Timmer. Stochastic global optimization methods. part I: Clustering methods. *Mathematical Programming*, 39:27–56, 1987.
- C.P. Robert and G. Casella. *Monte Carlo Statistical Methods*. Springer, 2004.
- M. Rodriguez-Fernandez, P. Mendes, and J. Banga. A hybrid approach for efficient and robust parameter estimation in biochemical pathways. *Biosystems*, 83:248–265, 2006.
- C.R. Rojas, J.S. Welsh, G.C. Goodwin, and A. Feuer. Robust optimal experiment design for system identification. *Automatica*, 43(6):993–1008, 2007.
- V. Rumchev and J. Adeane. Reachability and controllability of time-variant discrete-time positive linear systems. *Control and Cybernetics*, 33(1):85–94, 2004.
- P. Rumschinski. *Dissertation*. PhD thesis, Otto von Guericke Universität Magdeburg, Germany, 2012.
- P. Rumschinski, S. Borchers, S. Bosio, R. Weismantel, and R. Findeisen. Set-based dynamical parameter estimation and model invalidation for biochemical reaction networks. *BMC Systems Biology*, 4:69, 2010a.
- P. Rumschinski, D. Shona-Laila, S. Borchers, and R. Findeisen. Influence of discretization errors on set-based parameter estimation. In *Proc. 49th IEEE Conference on Decision and Control (CDC)*, pages 296–301, 2010b.
- P. Rumschinski, D. Shona-Laila, and R. Findeisen. Discrete-to-continuous dynamics reconstruction for bilinear systems. In *Proc. 51th IEEE Conference on Decision and Control (CDC)*, pages 1–7, 2012.
- A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola. *Global sensitivity analysis: the primer*. John Wiley & Sons, 2008.
- M. Schliemann, E. Bullinger, S. Borchers, F. Allgöwer, R. Findeisen, and P. Scheurich. Heterogeneity reduces sensitivity of cell death for TNF-stimuli. *BMC Systems Biology*, 5:204, 2011.
- R. Schneider. *Convex bodies: the Brunn-Minkowski theory*. Cambridge University Press, 1993.

- S. Schnell, M.J. Chappell, N.D. Evans, and M.R. Roussel. The mechanism distinguishability problem in biochemical kinetics: The single-enzyme, single-substrate reaction as a case study. *Comptes Rendus Biologies*, 329:51–61, 2006.
- F.C. Scheppe. Recursive state estimation: unknown but bounded errors and system inputs. *IEEE Transactions on Automatic Control*, 13(1):22–28, 1968.
- F.C. Scheppe. *Uncertain Dynamic Systems*. Prentice Hall, 1973.
- J. Shao. *Mathematical Statistics*. Springer, New York, USA, 2003.
- H.D. Serali and W.P. Adams. A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems. *SIAM Journal of Discrete Mathematics*, 3:411–430, 1990.
- H.D. Serali and C.H. Tuncbilek. Comparison of two Reformulation-Linearization Technique based linear programming relaxations for polynomial programming problems. *SIAM Journal of Global Optimization*, 10(4):381–390, 1997.
- S.A. Sheynin and A.V. Tuzikov. Explicit formulae for polyhedra moments. *Pattern Recognition Letters*, 22(10):1103–1109, 2001.
- R.W. Shirt, T.J. Harris, and D.W. Bacon. Experimental design considerations for dynamic systems. *Industrial & engineering chemistry research*, 33(11):2656–2667, 1994.
- K.K. Sivaramakrishnan. A parallel interior point decomposition algorithm for block angular semidefinite programs. *Computational Optimization and Applications*, 46(1):1–29, 2010.
- NJA Sloane. The on-line encyclopedia of integer sequences, 2002. URL <http://oeis.org/>.
- R.S. Smith and J.C. Doyle. Model validation: A connection between robust control and identification. *IEEE Transactions on Automatic Control*, 37(7):942–952, 1992.
- E.D. Sontag. *Mathematical control theory: deterministic finite dimensional systems*. Springer, 1998.
- J. Stelling, U. Sauer, Z. Szallasi, F.J. Doyle III, and J. Doyle. Robustness of cellular functions. *Cell*, 118(6):675–685, 2004.
- S. Streif, A. Savchenko, P. Rumschinski, S. Borchers, and R. Findeisen. ADMIT: a toolbox for guaranteed model invalidation, estimation and qualitative-quantitative modeling. *Bioinformatics*, 28(9):1290–91, 2012.
- A.M. Stuart and A.R. Humphries. *Dynamical Systems and Numerical Analysis*. Cambridge University Press, 1996.
- J.F. Sturm. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization Methods and Software*, 11:625–653, 1999.
- P.S. Swain, M.B. Elowitz, and E.D. Siggia. Intrinsic and extrinsic contributions to stochasticity in gene expression. *The Proceedings of the National Academy of Sci-*

- ences USA*, 99(20):12795–800, 2002.
- I. Swameye, TG Müller, J. Timmer, O. Sandra, and U. Klingmüller. Identification of nucleocytoplasmic cycling as a remote sensor in cellular signaling by databased modeling. *Proceedings of the National Academy of Sciences USA*, 100(3):1028–33, 2003.
- M. Sznaier. Computational complexity analysis of set membership identification of Hammerstein and Wiener systems. *Automatica*, 45(3):701–705, 2009.
- A. Tarantola. *Inverse problem theory and methods for model parameter estimation*. Society for Industrial Mathematics, 2005.
- A.N. Tikhonov, V.I.A. Arsenin, and F. John. *Solutions of ill-posed problems*. Winston, 1977.
- T. Toni, D. Welch, N. Strelkowa, A. Ipsen, and M.P.H. Stumpf. Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *Journal of the Royal Society Interface*, 6(31):187–202, 2009.
- R.H. Tütüncü, K.C. Toh, and M.J. Todd. Solving semidefinite-quadratic-linear programs using SDPT3. *Mathematical Programming*, 95(2):189–217, 2003.
- E. Tziampazis and A. Sambanis. Modeling of cell culture processes. *Cytotechnology*, 14(3):191–204, 1994.
- F. van der Heijden, R.P.W. Duin, D. de Ridder, and D.M.J. Tax. *Classification, Parameter Estimation and State Estimation: An engineering approach using MATLAB*. John Wiley & Sons, 2004.
- L. Vandenberghe, S. Boyd, and S.P. Wu. Determinant maximization with linear matrix inequality constraints. *SIAM Journal on Matrix Analysis and Applications*, 19:499–533, 1998.
- T.L. Vincent and W.J. Grantham. *Nonlinear and optimal control systems*. John Wiley & Sons, 1997.
- A. Wahl, Y. Sidorenko, M. Dauner, Y. Genzel, and U. Reichl. Metabolic flux model for an anchorage-dependent MDCK cell line: Characteristic growth phases and minimum substrate consumption flux distribution. *Biotechnology and Bioengineering*, 101(1):135–152, 2008.
- S. Waldherr. *Uncertainty and robustness analysis of biochemical reaction networks via convex optimisation and robust control theory*. PhD thesis, Universität Stuttgart, Germany, 2009.
- S. Waldherr, R. Findeisen, and F. Allgöwer. Global sensitivity analysis of biochemical reaction networks via semidefinite programming. In *Proc. of the 17th IFAC World Congress*, pages 9701–06, 2008.
- S. Waldherr, J. Hasenauer, M. Doszczak, P. Scheurich, and F. Allgöwer. Global uncertainty analysis for a model of TNF-induced NF- κ B signalling. *Advances in the Theory of Control, Signals and Systems with Physical Modeling*, pages 365–377, 2011.

-
- E. Walter and H. Piet-Lahanier. Exact recursive polyhedral description of the feasible parameter set for bounded-error models. *IEEE Transactions on Automatic Control*, 34:911–915, 1989.
- E. Walter and L. Pronzato. *Identification of parametric models from experimental data. Communications and Control Engineering Series*. Springer, 1997.
- E.A. Wan and R. Van Der Merwe. The unscented Kalman filter for nonlinear estimation. In *IEEE Adaptive Systems for Signal Processing, Communications, and Control Symposium*, pages 153–158. IEEE, 2000.
- L.S. Weinberger, J.C. Burnett, J.E. Toettcher, A.P. Arkin, and D.V. Schaffer. Stochastic gene expression in a lentiviral positive-feedback loop: HIV-1 TAT fluctuations drive phenotypic diversity. *Cell*, 122:169–182, 2005.
- GN Wilkinson. Statistical estimations in enzyme kinetics. *Biochemical Journal*, 80(2): 324–332, 1961.
- G. Williams, R. Baxter, H. He, S. Hawkins, and L. Gu. A comparative study of RNN for outlier detection in data mining. In *Proc. IEEE International Conference on Data Mining (ICDM)*, pages 709–712. IEEE, 2002.
- H. Witsenhausen. Sets of possible states of linear systems given perturbed observations. *IEEE Transactions on Automatic Control*, 13(5):556–558, 1968.
- H.P. Wynn. The sequential generation of D-optimum experimental designs. *The Annals of Mathematical Statistics*, 41(5):1655–64, 1970.
- Y. Yang and H. Lin. Reachability analysis based model validation in systems biology. In *Proc. IEEE Conference on Cybernetics and Intelligent Systems (CIS)*, pages 14–19, 2010.
- M. Yu, Z. Hu, E. Pacis, N. Vijayasankaran, A. Shen, and F. Li. Understanding the intracellular effect of enhanced nutrient feeding toward high titer antibody production process. *Biotechnology and Bioengineering*, 108(5):1078–1088, 2011.
- A.P. Zeng and W.D. Deckwer. A kinetic model for substrate and energy consumption of microbial growth under substrate-sufficient conditions. *Biotechnology Progress*, 11(1):71–79, 1995.

