Verification of System Properties of Polynomial Systems using Discrete-time Approximations and Set-based Analysis

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

von
Philipp Rumschinski
geboren am 27. Oktober 1980 in Tübingen

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

Gutachter:
Prof. Dr.-Ing. Rolf Findeisen
Prof. Dr. rer. nat. Nicole Radde
Dr. Dina Shona Laila

eingereicht am 24. November 2014
Promotionskolloquium am 18. August 2015

To my parents and grandparents.

Acknowledgment

Foremost I want to express my gratitude towards my supervisor Prof. Findeisen for giving me the opportunity to pursue this work and for providing his guidance and advice during my time at the Otto-von-Guericke University, Magdeburg.

My gratitude extends to Dina Shona Laila and Stefan Streif, for their mentorship, the fruitful collaborations, numerous discussions, and in case of Stefan for the time we spent together outside of work.

I would also like to recognize my other collaborators with whom I was able to work on so many interesting problems and projects. In particular I want to mention, Steffen Borchers and Sandro Bosio with whom I began the work on set-based methods.

Besides I want to thank my colleagues at the Institute for Automation to make this time so much more enjoyable. I am especially grateful to Anton Savchenko, Benjamin Kern, Lisa Carius, Paolo Varutti, Pablo Zometa, Petar Andonov, and Timm Faulwasser. Other past and present group members that I have had the pleasure to work with or alongside of are Daniel Hast, Eric Bullinger, Friedrich von Haeseler, Janine Matschek, Jürgen Ihlow, Khalid J. Kazim Al-Chaabawi, Markus Kögel, Matthias Rausch, Michael Kopf, Michael Maiworm, Monica Schliemann-Bullinger, Nadine Rudolph, Patrick Pascheka, Reinhardt Klein, Saša V. Raković, Sergio Lucia, Solvey Maldonado, and Tobias Bäthge.

On a personal level, my strength and dedication to pursue a PhD mainly derives from my family and my wish to make them proud. I am extremely grateful for their love, support, and patience over the whole time span. Thank you so much!

Contents

Abstract

To satisfy the steadily increasing requirements on safety and quality of processes in industry and natural sciences, model-based system and control theoretical methods become increasingly important. To obtain a suitable model for control and analysis of processes is, however, a nontrivial task in general.

In this thesis, we present methods for verifying and analyzing dynamical models helpful in the derivation of suitable model candidates. In particular, the problems of proving model invalidity, estimating initial conditions, parameters and states, deriving reachable sets, and analyzing observability of continuous-time and discrete-time systems are considered.

The fundamental basis of the derived results is the reformulation of these model analysis and verification tasks as polynomial feasibility and polynomial optimization problems. This reformulation allows on the one hand that these problems can be solved efficiently with the help of semi-definite and linear relaxations and on the other hand that the considered system properties can be verified. Moreover, the reformulation allows not only the consideration of set-valued uncertainties in the initial conditions, the parameters, and inputs, but also the consideration of semi-quantitative and qualitative observations and safety requirements.

The analysis of uncertain continuous-time systems is mainly based on discrete-time approximations. For this reason, we study the necessary conditions such that results obtained for the discrete-time system can be transferred to the continuous-time system. For instance, under which conditions can we conclude from inconsistency of the discrete-time model with the available measurement data and safety requirements to inconsistency of the continuous-time model.

The proposed observability analysis of uncertain dynamical systems is based on the investigation of algebraic observability notions and the bounded L_2-norm of the output. We show that both conditions can be addressed by a converging hierarchy of semi-definite programs.

The derived results and methods are validated considering examples from systems biology, biotechnology, and academic examples.

Zusammenfassung

Um die ständig ansteigenden Anforderungen an die Sicherheit und Qualität von Prozessen in der Industrie erfüllen zu können, werden modell-basierte system- und regelungstheoretische Ansätze immer wichtiger. Die Konstruktion eines geeigneten Modells stellt jedoch in der Regel eine schwierige Aufgabe dar.

In dieser Arbeit werden daher Verfahren vorgestellt, die die Modellierung erleichtern. Insbesondere werden Fragestellungen zur Modellinvalidierung, der Schätzung von Anfangsbedingungen, Parametern und Zuständen, der Berechnung von Erreichbarkeitsmengen sowie der Beobachtbarkeitsanalyse für zeitdiskrete und zeitkontinuierliche Systeme näher betrachtet.

Die Grundlage der erreichten Ergebnisse bilden Reformulierungen dieser Fragestellungen in Form von polynomialen Optimierungs- und Machbarkeitsproblemen. Diese Reformulierungen können zudem mit Hilfe von semi-definiten und linearen Relaxierungen effizient gelöst und die untersuchten Systemeigenschaften garantiert nachgewiesen werden. Des Weiteren lassen sich sowohl mengenbasierte Unsicherheiten in den Anfangsbedingungen, den Parametern und Eingängen als auch semi-quantitative und qualitative Beobachtungen und Sicherheitsanforderungen berücksichtigen.

Die Betrachtung von unsicheren, zeitkontinuierlichen Systemen basiert vor allem auf der Analyse von zeitdiskreten Approximationen. Daher werden die notwendigen Bedingungen untersucht, so dass Eigenschaften, die für das zeitdiskrete Modell nachgewiesen werden können auch auf das zeitkontinuierliche übertragbar sind. So lässt sich beispielsweise die Frage, unter welchen Bedingungen von der Modellinkonsistenz der Approximation bezüglich Messungen und Sicherheitsanforderungen auf die Modellinkosistenz des zeitkontinuierlichen Systems geschlossen werden kann, beantworten.

Die Beobachtbarkeitsanalyse für unsichere zeitdiskrete und zeitkontinuierliche Systeme kann direkt mit Hilfe von algebraischen Beobachtbarkeitsbedingungen und der L_2-Norm des Ausgangs untersucht werden. Besagte Bedingungen lassen sich wiederum mit einer konvergierenden Hierarchie an semi-definiten Programmen prüfen.

Deutsche Kurzfassung

Einleitung

Die Grundlage für die Anwendung von modellbasierten regelungs- und systemtheoretischen Ansätzen ist immer die Erstellung eines mathematischen Modells, das den zu untersuchenden Prozess in ausreichender Genauigkeit beschreibt. Hierbei bezieht sich Genauigkeit zum einen auf die quantitative Beschreibung der Systemzustände, wie z. B. Temperatur, Druck, Neigewinkel, und zum anderen auf die Möglichkeit Vorhersagen über das qualitative Systemverhalten, wie z. B. Stabilität, zu liefern. Die Forderung nach Genauigkeit begründet sich durch die ständig ansteigenden Sicherheits- und Qualitätsansprüche in der Industrie und durch die benötigte Flexibilität industrieller Prozesse unter Wirtschaflichkeitsgesichtspunkten. Dies führt ebenfalls dazu, dass ein Modell immer gößere Arbeitsbereiche und Aufgabenbereiche abdecken muss. Jedoch nicht nur in der Industrie werden Modelle benötigt, sondern auch in der Forschung stellen mathematische Modelle oftmals die Basis für die Untersuchung von Prozessen dar. Insbesondere sei hier der Bereich der Systembiologie genannt, dessen Ziel es ist, das einheitliche Verständnis von biologischen Organismen zu erreichen. Hier werden z. B. modellbasierte Analysen dazu verwendet die Zusammenhänge von Enzymen, Botenstoffen, und Stimuli zu verstehen, die nach dem derzeitigen Stand der Wissenschaft bisher noch nicht erklärt werden können.

Die Erstellung eines mathematischen Modells wird allerdings durch verschiedene Faktoren erschwert. Reale Prozesse verhalten sich normalerweise nichtlinear, d. h. ein realer Prozess kann Phänomene wie mehrere isolierte Ruhelagen, Limitzyklen oder Chaos aufweisen. Solche Phänomene schränken die Anzahl verwendbarer Systemidentifikationsmethoden erheblich ein, da die Mehrzahl dieser Methoden auf linearen Modellen beruhen [99, 120]. Des Weiteren kommen heutzutage immer häufiger digitale Sensoren zum Einsatz, um Messungen an diskreten Zeitpunkten von einem Prozess vorzunehmen. Daher ist es im Allgemeinen einfacher ein zeitdiskretes Modell zu bestimmen als ein zeitkontinuierliches [120]. Reale Prozesse verhalten sich jedoch normalerweise zeitkontinuierlich. Dieser Gegensatz stellt einer der grundlegenden Motivationen der vorliegenden Arbeit dar, nämlich die kontinuierliche und die zeitdiskrete Welt in Bezug auf die Entwicklung von Methoden zur Verifikation von Systemeigenschaften zusammenzubringen, welche für die Erstellung genauer Modelle für die Industrie und Forschung benötigt werden.

Eine Weitere Schwierigkeit, die für die Entwicklung von Methoden zur Verifikation von Modellen betrachtet werden muss, ergibt sich aus den unausweichlichen Unsicher-

heiten bei der Modellierung eines Prozesses. Diese Unsicherheiten können unterschiedliche Ursprünge haben. Ein Prozess kann derart komplex sein, dass eine vollständige physikalische-chemische Modellierung nicht zweckmässig ist, was direkt zu strukturellen Unsicherheiten führt [170]. Dies ist insbesondere in der Modellierung von biologischen Systemen der Fall, da in den seltensten Fällen alle involvierten Enzyme und Gene bekannt sind [202]. Ein weiterer Ursprung ergibt sich aus den Messungen. Verwendete Messgeräte und Messverfahren sind normalerweise unvollkommen, und nicht alle interessierenden physikalischen Grössen können direkt gemessen werden. Eine weitere Quelle an Messunsicherheiten ist die möglicherweise eingeschränkte Wiederholbarkeit einer Messung. Dies führt dazu, dass der mögliche Wertebereich mit der dazugehörigen Wahrscheinlichkeitsverteilung nicht genau bestimmt werden kann [179].

In dieser Arbeit werden verschiedene Ansätze entwickelt, um die oben genannten Schwierigkeiten in der Modellierung und in der Analyse von Modellen zu bewältigen. Insbesondere wird hier mit Hilfe von mengenbasierten Verfahren die Parameterschätzung von zeitkontinuierlichen Modellen betrachtet. Diese auf Optimierungsverfahren basierenden Methoden erlauben es bestimmte Systemeigenschaften, wie z. B. die Konsistenz eines Modells bezüglich vorhandener Messungen, zu garantieren. Im Folgenden werden kurz die betrachteten Probleme, die möglichen Anwendungsgebiete, sowie die entwickelten Methoden skizziert.

Grundlagen und Illustration der behandelten Probleme

Es wird angenommen, dass sich die in dieser Arbeit betrachtenden Prozesse durch ein gewöhnliches Differentialgleichungssystem der Form ($t \in \mathbb{R}$)

$$\begin{aligned} \dot{x}(t) &= f(x(t), u(t), p), \quad x(0) = x_0, \\ y(t) &= h(x(t), u(t), p), \end{aligned}$$

beziehungsweise als Differenzengleichungssystem der Form ($k \in \mathbb{N} \cup \{0\}$)

$$\begin{aligned} x(k+1) &= f_D(x(k), u(k), p), \quad x(0) = x_0, \\ y(k) &= h_D(x(k), u(k), p), \end{aligned}$$

beschreiben lassen. Wobei $x \in \mathbb{R}^{n_x}$ den Zustandsvektor, $u \in \mathbb{R}^{n_u}$ den Eingang, $y \in \mathbb{R}^{n_y}$ den Ausgang darstellen. Der Parametervektor wird mit $p \in \mathbb{R}^{n_p}$ bezeichnet. Im Differenzengleichungssystem beschreibt $D > 0$ die Zeitschrittweite. Die rechten Seiten der obigen Systeme werden stets als polynomial angenommen. Des Weiteren wird angenommen, dass die Werte von x, u, y und p durch geschlossene, kompakte, semialgebraische Mengen $\mathcal{X}, \mathcal{U}, \mathcal{Y}, \mathcal{P}$ beschränkt sind. Diese Mengen repräsentieren die Modell-, Eingangs- und Messunsicherheiten, sowie mögliche Beschränkungen durch verbotene Zustände oder Limitationen des Eingangsignals. In späteren Abschnitten

werden zudem logische Verküpfungen eingeführt, die noch allgemeinere Unsicherheiten und Beschränkungen, wie z. B. eine qualitative Beschreibung von Systemanforderungen, erlauben. Ein Beispiel für eine solche Unsicherheit ist die Übertragung von wenn-dann Beobachtungen in der Systembiologie: wenn das Substrat A in ausreichender Menge vorliegt dann steigt die Konzentration von Produkt B kontinuierlich an. Für die Illustration der im Folgenden vorgestellten Fragen und Ergebnisse, ist die einfachere Beschreibung der Unsicherheit als Mengen jedoch ausreichend.

Messungen der Systemzustände und Ausgänge werden in dieser Arbeit ebenfalls als semi-algebraische Mengen der Form

$$\mathcal{X}_t := \{x : g_x(x) \geq 0\} \subseteq \mathbb{R}^{n_x} \text{ bzw. } \mathcal{Y}_t := \{y : g_y(y) \geq 0\} \subseteq \mathbb{R}^{n_y}$$

angenommen, wobei die Funktion $g_x(\cdot), g_y(\cdot)$ Polynome darstellen.

Diese allgemeine Systembeschreibung lässt es zu verschiedenste Probleme zu berücksichtigen. So werden nachfolgend unter anderem die folgenden Fragen betrachtet und entsprechende Lösungsverfahren entwickelt.

- **(Modellkonsistenz)** Wann kann garantiert werden, dass ein unsicheres Modell eine Messreihe bestehend aus Mengen exakt wiedergeben kann?

- **(Erreichbarkeit)** Wann kann sichergestellt werden, dass ein zeitkontinuierliches Modell, sowie dessen zeitdiskrete Approximation, die selbe Erreichbarkeitsmenge besitzen?

- **(Beobachtbarkeit)** Wann kann bei einem unsicheren System aus möglichen Werten des Ausgangs der Zustandsvektor rekonstruiert werden?

Die erste Frage stellt den wichtigsten Schritt im Modellbildungskreislauf dar, nämlich der Verifikation des Modells mit Hilfe von (experimentellen) Messdaten und Beobachtungen [58]. Nur wenn ein Modell in der Lage ist das vorhandene Wissen wiederzugeben war die Modellierung erfolgreich. Anwendungsbeispiele für die hier entwickelten Verfahren bei der Modellierung und Modellinvalidierung von biologischen Systemen lassen sich in den folgenden Arbeiten finden [33, 164, 170]. Ein weiteres Gebiet in dem Modellkonsistenz eine wichtige Rolle spielt, ist die modellbasierte Fehlerdiagnose in technischen Systemen, siehe z. B. [169, 174]. Hier werden die Messdaten dazu verwendet Fehler zu erkennen, um gegebenenfalls Gegenmaßnahmen zur Sicherstellung des korrekten Betriebs des Prozesses einleiten zu können.

Erreichbarkeit wird ähnlich wie die Modellkonsistenz in der Verifikation von Spezifikationen an einen Prozess eingesetzt. Dies ist besonders im Bereich der hybriden Systeme der Fall, also bei Systemen in denen zeitkontinuierliche und zeitdiskrete Vorgänge betrachtet werden müssen. Hierzu gehört die Regelung eines physikalischen Prozesses mit einem digitalen Regler [128, 174, 181]. Im Allgemeinen bezieht sich der Begriff

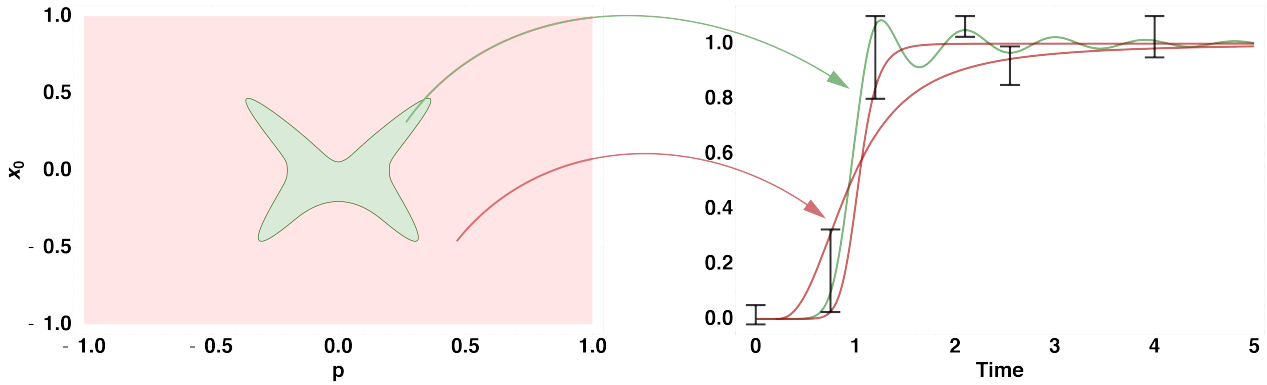

Abbildung I: Illustration der Modellkonsistenz eines Systems. Messdaten sind als schwarze Intervalle an diskreten Zeitpunkten dargestellt. Die grüne Menge korrespondiert zu Anfangsbedingungen x_0 und Parametern p die zu einer konsistenten Trajektorie $\phi(t)$ führen (ebenfalls grün dargestellt), während die rote Menge zu inkonsistenten Trajektorien korrespondiert.

der Erreichbarkeit als die Möglichkeit ein System von einer gegebenen Anfangsbedingung x_0 zu einem gewissen Endzustand $x(T)$ zu überführen. Daher ist Erreichbarkeit zusammen mit Stabilisierbarkeit auch in dem Begriff der Steuerbarkeit enthalten.

Die letzte der drei Fragen bildet häufig die Grundlage für die Anwendbarkeit einer Regelung in der Praxis. Normalerweise lassen sich nicht alle interessierenden Zustände eines Prozesses direkt messen. Der Wert dieser Grössen kann jedoch notwendig sein, um den Prozess regeln zu können [11, 67, 94]. In der Praxis muss infolgedessen ein sogenannter Zustandsschätzer eingesetzt werden, der aus den Messwerten die möglichen Werte der Zustände bestimmt, siehe z. B. [85, 186]. Beobachtbarkeit lässt sich in der Theorie als die Möglichkeit den Zustand eines Systems aus den Messwerten rekonstruieren zu können definieren [78]. Eine äquivalente Definition ergibt sich aus der Existenz eines geeigneten Zustandsschätzers [78].

Im Folgenden wird die in dieser Arbeit entwickelte Methodik an Hand der ersten Frage illustriert, siehe auch Abbildung I. Es wird hierzu angenommen, dass Messungen der Ausgänge $\mathcal{Y}_{t_i}$ für bestimmte Zeitpunkte $t_i \in \mathcal{T} := \{t_0, \ldots, t_k\}$, sowie ein bekanntes Eingangssignal $u(t) = u_s$ verfügbar sind. Die theoretische Fragestellung lautet in diesem Fall wie folgt:

Frage I (Modellkonsistenz)**.** *Existieren Anfangsbedingungen und Parameter, die zu einer Systemtrajektorie* $\phi(t) := \phi(x(t)|u_s, p, x_0)$ *führen, so dass der Ausgang* $y(t_i) = h(\phi(t_i), u_s, p)$ *zu allen Zeitpunkten* $t_i \in \mathcal{T}$ *innerhalb der korrespondierenden Menge* $\mathcal{Y}_{t_i}$ *liegt, d. h.* $y(t_i) \in \mathcal{Y}_{t_i},\ \forall t_i \in \mathcal{T}$*?*

Diese Frage kann beantwortet werden in dem diese als ein polynomiales Machbar-

keitsproblem der Form

$$\begin{aligned} &\text{finde } x_0,\ p \\ &\text{u.d.N. } \phi(t) \in \mathcal{X}, y(t) \in \mathcal{Y}, u(t) = u_s, \\ &\qquad p \in \mathcal{P}, y(t_i) \in \mathcal{Y}_{t_i}, t_i \in \mathcal{T}, \end{aligned} \tag{I}$$

umformuliert wird, wie z. B. in den Arbeiten [27, 164] gezeigt wurde. Frage I ist in diesem Fall äquivalent zu der Lösungsmenge des Machbarkeitsproblems. Wenn nun die Lösungsmenge leer ist, ist das betrachtete Modell nicht konsistent mit den Messungen. Das heißt, dass keine Anfangsbedingungen und Parameter existieren, die zu einer entsprechenden Systemtrajektorie führen, vergleiche hierzu auch die rote Menge und die roten Trajektorien in Abbildung I. Ist sie nicht leer, kann die Lösungmenge dazu genutzt werden die Bereiche in dem mögliche Werte der Variablen des Machbarkeitsproblems liegen, wie z. B. der Parameter, besser einzuschränken (vgl. mit der grünen Menge und der grünen Trajektorie in Abbildung I). Die Einschränkung der Mengen kann zudem wichtige Einblicke in das Systemverhalten liefern. So kann beispielsweise die Größe eines Parameterbereiches als Kennzeichen für die Robustheit eines Prozesses gegenüber Störungen herangezogen werden [202].

Jedoch ist zu beachten, dass die Lösungsmenge unter den vorgestellten Annahmen typischerweise nicht konvex ist. Dadurch ist die Bestimmung der Lösungsmenge im Allgemeinen nicht einfach möglich und daher auch nur wenige geeignete Methoden bekannt. Eine weitere Schwierigkeit ergibt sich, falls sich das System zeitkontinuierlich verhält. Im Regelfall kann die Lösungsmenge dann nur mit Hilfe von numerischer Integration des Differentialgleichungssystems bestimmt werden. Allerdings kann hierbei in der Regel keine Garantie gegeben werden, dass die so bestimmte Lösungsmenge die Modellkonsistenz des zeitkontinuierlichen Systems exakt wiedergibt [97]. Aus diesem Grund ist es möglich, dass beispielsweise ein inkonsistentes Modell fälschlicherweise als konsistent angenommen wird.

Die Grundlage zur Bestimmung der Lösungsmengen, die in dieser Arbeit verwendet wird, ist die Reformulation des Machbarkeitsproblems mit Hilfe der von Lasserre in [108] vorgestellten Methodik. Diese beruht auf der Äquivalenz von polynomialen Optimierungsproblemen bzw. Machbarkeitsproblemen mit dem sogenannten problem of moments, welches im Folgenden als Momentenproblem bezeichnet wird. Dieses Problem beschreibt die Existenz einer Wahrscheinlichkeitsverteilung auf dem von den Nebenbedingungen beschriebenen Gebiets, wobei der Support der Wahrscheinlichkeitsverteilung gleichbedeutend ist zu der zu bestimmenden Lösungsmenge. Da das Momentenproblem über den Raum der Wahrscheinlichkeitsverteilungen formuliert wird, ist dieses unendlichdimensional und daher oftmals nicht exakt lösbar. Daher ist es notwendig dieses mit Hilfe einer konvergierenden Hierarchie von semi-definiten Programmen zu approximieren. Jedes dieser Programme kann jedoch in polynomialer Zeit in Abhängigkeit von den Eingangsgrößen gelöst werden [23]. Der große Vorteil der ver-

wendeten Methodik, neben der effizienten Lösbarkeit, ist die kohärente Betrachtung der angesprochenen Probleme mit Hilfe eines einzigen Verfahrens, das es zudem erlaubt die geforderten Garantien zu liefern. Zum Beispiel ist es möglich eine Garantie zu geben, dass die Lösungsmenge leer ist oder auch dass die Lösungsmenge exakt bestimmt wurde.

Die Formulierung geeigneter Machbarkeitsprobleme zur Verifikation der oben genannten Systemeigenschaften, sowohl für zeitkontinuierliche als auch zeitdiskrete Systeme, und die Bestimmung der korrespondierenden Lösungsmengen bilden die Schwerpunkte dieser Arbeit. Nachfolgend werden die erreichten Ergebnisse näher vorgestellt.

Ergebnisse und Beiträge

Die im Rahmen dieser Arbeit entwickelten Beiträge lassen sich grob in zwei Bereiche unterteilen. Zum einen wurde die mengenbasierte Schätzung durch Machbarkeitsprobleme für zeitdiskrete und zeitkontinuierliche Systeme entwickelt. Zum anderen wurden mit Hilfe von Machbarkeitsproblemen neue Methoden entwickelt die Bebochbarkeit eines unsicheren Systems zu untersuchen. Die Beiträge zu beiden Bereichen werden im Folgenden weiter erläutert.

Mengenbasierte Schätzung

Das Ziel der mengenbasierten Schätzung ist die Bestimmung des Wertebereichs der Parameter und Zustände aus den durch Mengen beschriebenen Messungen. Das Hauptaugenmerk liegt hierbei auf der Garantie, dass der Wertebereich vollständig ist, d. h. keine Wahl an Parametern und Anfangsbedingungen außerhalb des Wertebereiches darf dazu führen, dass das Modell die Messungen wiedergibt. Im Rahmen dieser Arbeit wurde ein solches Schätzverfahren für zeitdiskrete Systeme entwickelt, welches in den Arbeiten [27, 164, 170] vorgestellt und in die Toolbox ADMIT [204] implementiert wurde. Die Grundlage dieses Verfahrens ist die Formulierung eines Machbarkeitsproblems der Form (I) und die Bestimmung der dazugehörigen Lösungsmenge. Diese Vorgehensweise ist jedoch nicht direkt auf zeitkontinuierliche Systeme anwendbar, da hier eine dynamische Nebenbedingung mitberücksichtigt werden muss.

Eine Möglichkeit zeitkontinuierliche Systeme dennoch mit diesem Verfahren zu betrachten, ist es eine zeitdiskrete Approximation zu verwenden. Allerdings konnte in [166] gezeigt werden, dass in diesem Fall die Aussagen bezüglich Modellkonsistenz im Allgemeinen von der Approximation nicht auf das zeitkontinuierliche System übertragbar sind. Dies liegt daran, dass bei der zeitdiskreten Approximation ein Diskretisierungsfehler gemacht wird, der dazu führt dass konsistente Parameter der Approximation nicht mit den konsistenten Parametern des zeitkontinuierlichen Systems übereinstimmen. Der Fokus wurde daher darauf gelegt eine Beziehung zwischen den konsistenten Parametern der beiden Systeme herzustellen, so dass von der Modellkonsistenz der

Approximation auf die Modellkonsistenz des zeitkontinuierlichen Systems geschlossen werden kann.

Zu diesem Zweck wurden zwei Verfahren entwickelt. Das erste Verfahren basiert auf der Abschätzung des gemachten Diskretisierungsfehlers. Hierzu wird zunächst eine Trainingsmenge mit Hilfe des zeitkontinuierlichen Modells erstellt. Das zeitdiskrete Modell wird dann um den Diskretisierungsfehler erweitert und anschließend wird der Diskretisierungsfehler anhand der Trainingsmenge bestimmt. Dies führt dazu, dass die gewünschte Beziehung für eine Teilmenge an Parametern erfüllt ist. Jedoch kann nicht garantiert werden, dass der in der mengenbasierten Schätzung bestimmte Wertebereich vollständig ist.

Um die Garantie der Vollständigkeit zu geben, wurde ein zweites Verfahren entwickelt, welches auf der Außenapproximation der Erreichbarkeitsmenge des kontinuierlichen Systems mit Hilfe einer zeitdiskreten Approximation basiert. Um sicherstellen zu können, dass die Beziehung der konsistenten Parametermengen erfüllt ist, wurde das Theorem von Picard-Lindelöf für die Existenz und Eindeutigkeit einer Lösung einer Differentialgleichung zur Konstruktion des Machbarkeitsproblems verwendet. Es kann dadurch garantiert werden, dass die geschätzten Paramter für die Approximation eine Übermenge der Parameter des zeitkontinuierlichen Systems sind. Daher ist es ebenfalls möglich von der Inkonsistenz der Approximation auch auf die Inkonsistenz des zeitkontinuierlichen Systems zu schließen.

Beobachtbarkeit von unsicheren Systemen

Während sich der erste Teil dieser Arbeit mit dem konkreten Schätzen von Parametern und Zuständen aus Messungen beschäftigt, wird im zweiten Teil die Beobachtbarkeit eines Systems untersucht. Der Begriff der Beobachtbarkeit bezieht sich hier auf die prinzipielle Möglichkeit die anfangs angenommen Parameter- und Zustandsbereiche mit Hilfe von Messungen zu verkleinern. Obwohl Beobachtbarkeit für die hier betrachteten polynomialen Systeme relativ gut verstanden ist, siehe z. B. [90, 91, 191], sind für unsichere Systeme praktisch keine anwendbaren Methoden bekannt.

Zu diesem Zweck wurden hier neue Methoden entwickelt, die ebenfalls auf der Konstruktion von geeigneten Machbarkeitsproblemen basieren. Hierzu wurden die algebraischen Beobachtbarkeitsbedingungen von [91] auf Systeme mit mengenbasierten Unsicherheiten erweitert. Zunächst wurde gezeigt, dass die algebraischen Bedingungen direkt durch ein Machbarkeitsproblem nachweisbar sind. Darauf aufbauend wurde eine Methodik entwickelt, um Unsicherheiten mitberücksichtigen zu können. Die Parameteridentifizierbarkeit kann ebenfalls direkt mit den hier vorgestellten Methoden betrachtet werden; es muss hierfür lediglich der Zustandsvektor, um die Parameter erweitert werden.

1 Introduction

1.1 Overview

Control and systems theoretical methods are employed to analyze and influence dynamical systems in a desired way throughout science and practice. The basis of successfully employing model-based methods is the derivation of a quantitative and predictive model. The modeling task is considerably simplified by the rapid development of digital computers and, hence, the possibility to use system identification and modeling tools, see e. g. [120, 135, 142]. While digital computers facilitate many new methods, this development also imposes certain challenges. In particular, practically all phenomena observed in nature behave continuously in time, however, such phenomena are typically analyzed with the help of digital computers by discrete-time (or numerical) approximations. This contradiction motivates this work and the connection of both "worlds" in the context of system identification and model analysis is studied.

The connection between continuous-time systems and discrete-time approximations becomes especially important if results obtained for one have to be transferred to the other. A notable research area concerned with this connection is sampled-data control systems, i. e. a continuous-time process that is controlled by a digital device, see e. g. [86, 106, 144]. Here, a controller is often devised for the discrete-time approximation, and the goal is to prove that this controller also stabilizes the continuous-time process. However, such transferable results are of importance in numerous other research areas as well. In particular, if guarantees have to be given, e. g. one is interested to verify that a continuous-time process operates within some safety margins, employing a discrete-time approximation in the corresponding analysis might result in wrong conclusions due to the discretization error, e. g. [166, 168]. To illustrate this problem consider the problem of model (in-)validation, i. e. the problem of proving the existence of an initial condition/parameterization of a model such that given measurement data are reproduced, see also Figure 1.1. If it can be shown that the discrete-time model is not able to reproduce the data, e. g. with the method presented in [27, 164], one cannot necessarily transfer this result to the continuous-time system, cf. Figure 1.2.

In this work, we are interested in providing conditions which guarantee when results are transferable from a discrete-time approximation to the continuous-time system, in particular, if the continuous-time and the discrete-time system are influenced by various types of uncertainties, e. g. in parameters and measurements. We do so for the rich class of polynomial systems, for which typical examples include mass action models in systems biology and mechatronic systems. Note that considering polynomial

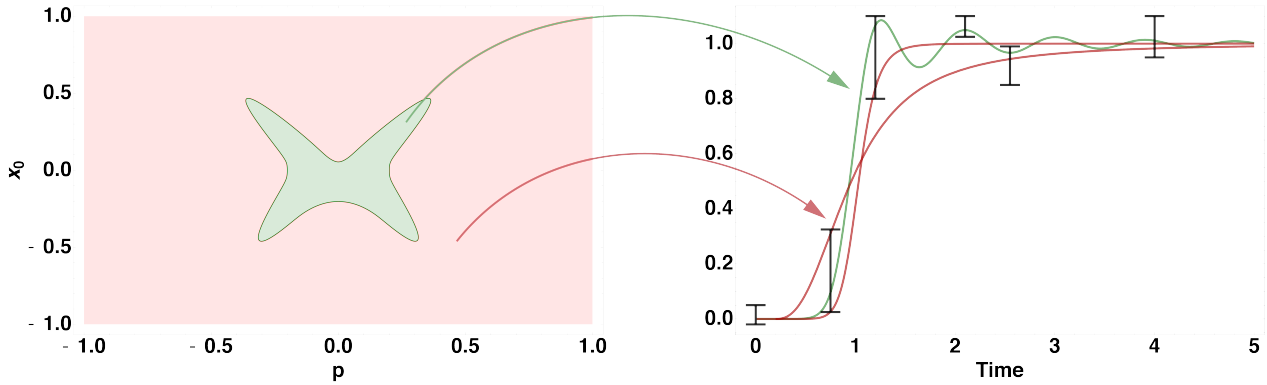

Figure 1.1: Illustration of the model (in-)validation problem. Measurement data is here depicted as black intervals. The green set corresponds to initial conditions and parameterizations of a dynamical system that lead to a trajectory that visits all measurements (green trajectory). The red set corresponds to initial conditions and parameterizations that do not lead to trajectories that visit all measurements (red trajectories). If the green set is empty the system is considered invalid (or inconsistent with the measurement data), otherwise it is considered as valid.

systems is not overly restrictive since every continuous nonlinear equation can be approximated to arbitrary precision by a polynomial, see e. g. [37].

For this purpose, a number of techniques are developed not only to analyze this connection, but also to analyze additional system properties of continuous-time as well as discrete-time systems. This is done by means of set-based considerations that lead to efficiently solvable (resp. computable) formulations. The central contributions are concerned with the following three problems.

- **(Model Consistency)** how to prove that an uncertain dynamical system is able to reproduce the available data and process insight, cf. Figure 1.1.
- **(Reachability)** find a (uncertain) discrete-time system that envelopes all possible trajectories of the (uncertain) continuous-time system, cf. Figure 1.3.
- **(Observability)** which conditions have to hold such that an uncertain (discrete-time or continuous-time) system can be observable.

The first two problems address the aforementioned problem of model (in-)validation as well as the reachability of dynamical systems. Model (in-)validation is a crucial step in the modeling process, namely the verification that a derived model is able to reproduce the available measurement data. Reachability is the corresponding generic system property, as a model can only reproduce the measurement data exactly if the data are contained in the reachable set. In general, reachability corresponds to the ability of a system to be transferred from a given initial state to another end state.

The third problem is concerned with the observability of a system. In other words, whether it is possible to derive the initial conditions of a system from made observations

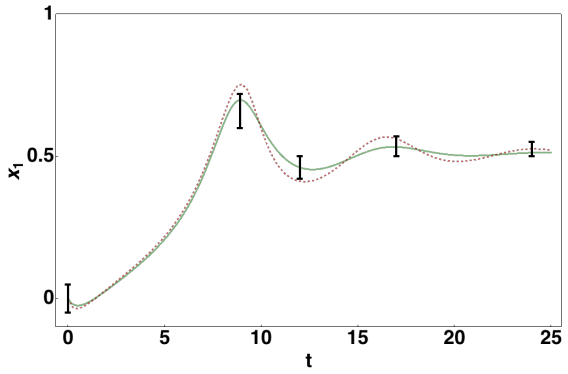

Figure 1.2: Illustration of the influence of the discretization error in model invalidation. The green trajectory depicts the continuous-time system that goes through all unknown-but-bounded measurements (black intervals) for a given initial condition and parameterization. The red (dashed) trajectory depicts the trajectory of the corresponding discrete-time approximation obtained by Euler discretization with time-step size of 0.5. To conclude from the fact that the discrete-time trajectory is not visiting all measurements, that the continuous-time is unable to reproduce the data is clearly wrong.

of the output or not. This system property is of particular interest in the case that not all states of a process can be measured directly. Therefore, it is the basis for employing controller design techniques that employ full state information.

In the following two sections, we give a more detailed description of these problems and a short review of related approaches in different application areas. Although, the developed ideas are applicable in many areas we restrict our attention mainly to set-based methods and examples from systems biology. It should also be noted that this introduction is kept rather short and more detailed discussions of available methods are given at the appropriate places throughout this work to minimize repetitions.

1.2 Set-based Estimation

1.2.1 Motivation and Existing Approaches

The increasing industrial requirements for product quality and safety, e. g. in pharmaceutical, medical and automotive industry, are nowadays often met by employing controls [30, 57]. Despite the availability of input-output based control strategies, model-based approaches often provide better performance and typically more system insight, e. g. through model analysis, can be gained [160]. A predictive model becomes especially important if not every state can be measured (for technical or economical reasons), but the information of a nonmeasured state is needed for the controller.

System identification or the extraction of a mathematical model from measurement data has, therefore, attracted a lot of attention in literature, see e. g. [58, 83, 119, 120,

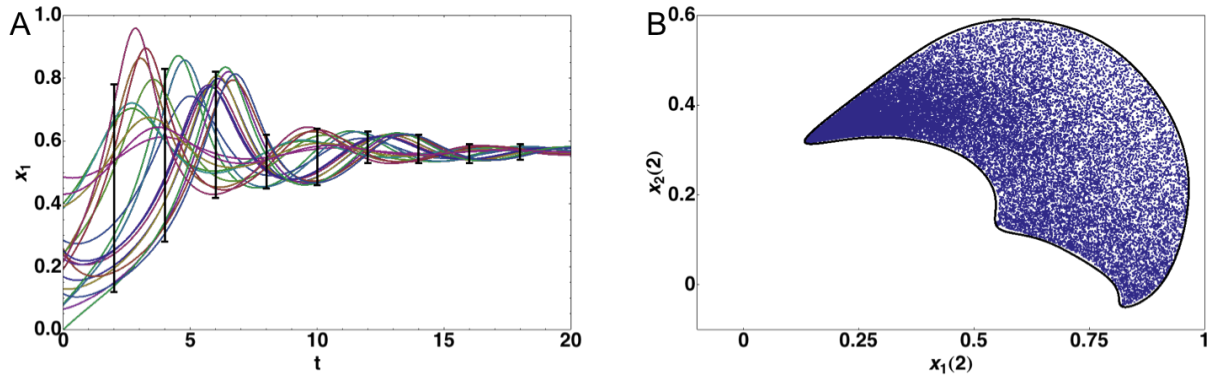

Figure 1.3: Illustration of a discrete-time enclosure. A: The figure depicts simulation data of a 2D continuous-time system and its discrete-time approximation. The continuous-time system and its discrete-time approximation have the same (or similar) reachable set, since at every sampling instance the possible values of the discrete-time approximation enclose the possible values of the continuous-time system. The trajectories depict the time evolution of x_1 of the continuous-time system for different initial conditions. Black intervals depict the possible values of the discrete-time approximation at different sampling times. B: The blue dots represent Monte Carlo samples of the continuous-time trajectories at $t = 2$. The black boundary corresponds to the possible values of the discrete-time system.

135] for an overview. Strangely, most system identification methods are concerned with the identification of discrete-time models, although, most practitioners prefer continuous-time models [120]. To not rely only on measurement data for identifying an accurate model the system structure is often inferred e. g. by first principles or system insight. In particular, this approach of building up a model from known interactions between species builds the basis of systems biology [99, 220].

If some prior knowledge on the structure is available, model identification is reduced to parameter estimation, e. g. [97, 116]. Applicable methods for parameter estimation depend significantly on the considered problem setup and the made assumptions. On the one hand parameter estimation methods can be classified depending on the description of the measurement data. Either the data are given as point values (possibly with additional statistical information) or as in this work they are given as sets. On the other hand one has to distinguish between methods for continuous-time systems and methods for discrete-time systems, in particular for set-based methods. The following paragraphs are therefore structured as follows. First, the differences of set-based and classical parameter estimation methods are discussed. Second, a general overview of set-based methods is provided and we end this section with a discussion of set-based methods for continuous-time systems.

Differences of Set-based and Classical Parameter Estimation

Classical parameter estimation is typically performed based on numerical optimiza-

tion [210]. Here, data are described by point values (possibly with statistical information) of the output and the goal is to minimize a chosen objective function that depends on these point values, while the optimization variables are the initial conditions and the parameters. A commonly chosen objective function is the distance of the system trajectory to the measurements at the corresponding time points. Although there are several reliable tools available for parameter estimation (see e. g. [130, 162]), parameter estimation is still in general a difficult task for nonlinear systems due to the involved nonconvexities in the optimization problem. Verifying additionally that the obtained model guarantees a certain safety requirement is even more involved and often depends on finding a global optimum [3] or some specific certificates [157].

In contrast to classical parameter estimation, measurement uncertainty is interpreted in set-based estimation as variables that belong to some bounded set (unknown-but-bounded variables) [219], instead of a variable that is subject to stochastic effects such as noise. The goal of set-based methods differs, therefore, notably from the classical methods as optimality becomes less important, see Figure 1.4 for an illustration. Instead of searching an optimal solution, one is interested in finding or bounding all possible parameterizations and initial conditions for which the system trajectories are consistent with the measurements, see e. g. [83, 164] and Figure 1.1. One major advantage of the set-based approach is, therefore, that a complete investigation of the parameter/initial conditions space can be performed. This provides a valuable complement to statistical information as it allows invalidating a model, in case no feasible initial condition/parameterization is found. This guaranteed invalidation further allows verifying if a system violates a safety requirement or, for example, to identify knockout targets in biological systems.

Clearly, the set-based viewpoint is not restricted to model invalidation and parameter estimation, but rather can be employed for all estimation and verification tasks as soon as a suitable model was derived. Examples include model analysis, control, state estimation, state prediction and related problems like fault diagnosis, outlier detection, uncertainty and robustness analysis, see e. g. [33, 101, 169, 174, 201, 202] and references therein.

Overview on Set-based Methods

Set-based considerations are often performed by interval analysis, see e. g. [83] for an overview. For this reason, variables are substituted by intervals and standard arithmetic operations are extended to intervals. By employing this one obtains a so called interval extension of the system dynamics. By checking which (sub-)intervals belong to the data an estimate of the variables can be derived, see e. g. [83, 139, 183].

In [27, 164, 170] a different set-based method to estimate the states and parameters (or more general variables of interest) of uncertain discrete-time systems was proposed. To do so, a nonlinear feasibility problem is constructed for which the projection of the solution space onto the variable space is equivalent to the considered estimation

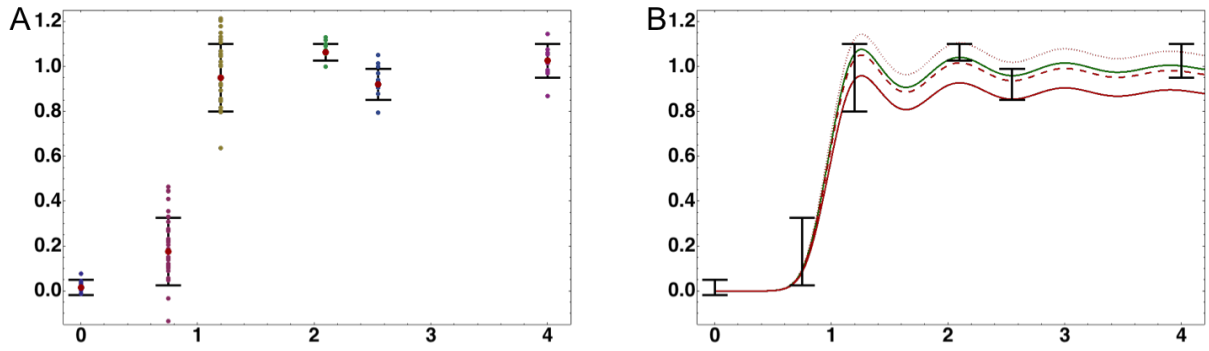

Figure 1.4: Comparison of set-based measurements and measurements with statistical information. Dots represent measurement data, large red dots correspond to the mean values of the data. Black intervals correspond to the mean value plus/minus the standard deviation. Trajectories correspond to optimal trajectories w.r.t. a least-squares objective function that penalizes the distance of the trajectory to the reference values at the measurement time points. The red dashed, the red doted, and the red trajectories correspond to optimal trajectories for which the references are chosen to be the mean values, the upper bounds of the intervals, or the lower bounds, respectively. The green trajectory corresponds to the optimal trajectory for which the reference values are chosen randomly to be either the lower or the upper bound of the intervals. Only the green trajectory goes through all black intervals.

problem. By testing where the solution space of the feasibility problem is empty the variable space can be classified into regions that are consistent with the measurements or safety requirements and regions that are not. This test is possible since the feasibility problem can be relaxed into an easier to solve convex semi-definite program, see e. g. [104, 108, 151, 182, 184] for general descriptions of relaxations of polynomial programs. Depending on the employed relaxation technique and relaxation order (higher degrees result in better approximations) one can derive different outer-approximations of the feasible set as depicted in Figure 1.5.

However, interval methods as well as the feasibility approach typically cannot consider continuous-time systems directly. This derives from the fact that the dynamical component has to be treated with special care such that the benefits of the set-based perspective are not lost.

Set-based Methods for Continuous-Time Systems

For continuous-time systems set-based approaches can be divided into the following two categories: reformulation of the continuous-time dynamics and approximation of the dynamics with a discrete-time system.

The possibly most straightforward reformulation that allows the consideration of continuous-time systems is the steady-state assumption, as employed e. g. in [71, 104].

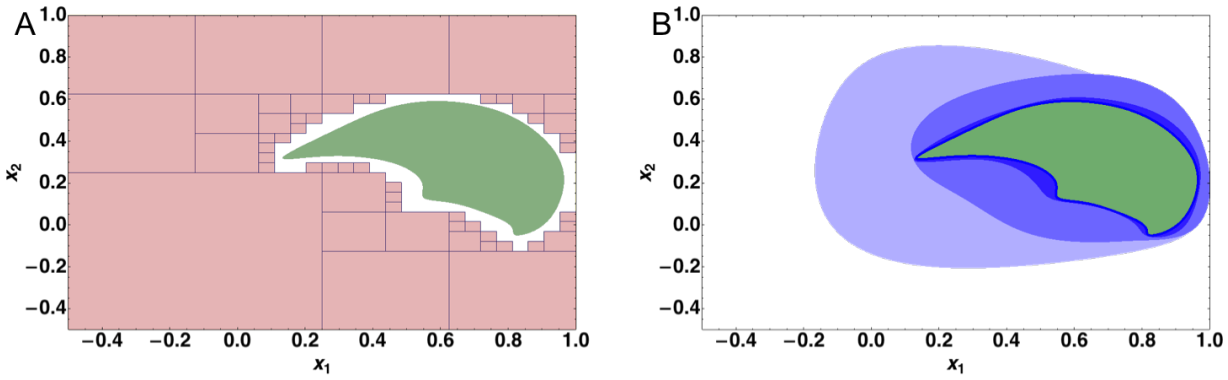

Figure 1.5: Illustration of approximating the solution space of a feasibility problem by a relaxation. The green set depicts the feasible set to be approximated. Red sets depict partitions for which it was shown that they do not contain feasible solutions [164]. Blue sets correspond to sets obtained by employing Lasserre's converging hierarchy of semi-definite programs [108], i. e. for every semi-definite program one obtains a valid outer-approximation that converges (in the Lebesgue measure) to the true feasible set for increasing relaxation order.

Clearly, this leads to a loss of information over the transient behavior of a system, but still can provide valuable insight e. g. on steady-state multiplicity or saddle-node bifurcations [71]. Another possibility to consider continuous-time systems, is to make additional assumptions. In [52] it was assumed that the derivatives of the states are also available as unknown-but-bounded measurements. The problem then reduces to a multi-stage steady-state investigation as the dynamics can basically be neglected and an approach similar to [27, 104] can be employed. A more elaborate approach is reformulating the dynamics in terms of so-called occupation measures [74, 112, 165, 176, 203]. These measures allow replacing the dynamics with linear (in-)equalities in an optimization problem over an infinite function space, which is then solved by a hierarchy of semi-definite programs [108]. This allows treating continuous-time systems directly, however, the relationship between continuous-time and discrete-time systems cannot be directly investigated which is one of the focuses of this work.

We investigate here how a continuous-time system can be analyzed with the help of a discrete-time approximation. The main difficulty that arises in this case is that results derived for the approximation do not necessarily hold for the continuous-time system. In particular, if a discrete-time approximation is inconsistent with the data it cannot be concluded that the continuous-time model is inconsistent as well. The main reason for this fact is the discretization error introduced by approximating the continuous-time system.

Discretization errors as they appear in solving initial value problems are a well studied topic in numerical mathematics, see e. g. [46, 198], and in the context of systems theory [68, 206, 207]. However, here only initial value problems with fixed parame-

terizations are considered and a direct application to set-based methods is in general difficult. Discretization errors also play an important role in the field of differential inclusions [10, 12, 54], and in particular the discretization of differential inclusions [48, 63, 171]. So for instance in [7] it was shown that the discrete-time approximation of a differential inclusion based on the Euler discretization only approximates the differential inclusion for vanishing time-steps. However, this area of research is typically only concerned with theoretical properties and as the proofs are often nonconstructive an application in system analysis and estimation is difficult.

One applicable approach dealing with this problem is based on interval arithmetics for deriving validated solutions to ordinary differential equations [38, 97, 103, 116, 123, 139, 140, 161, 216, 217]. One common feature of these works is that they try to minimize the discretization error by employing multi-step discretization schemes. In the feasibility approach multi-step schemes, however, lead to an unwanted increase of the employed relaxations.

1.2.2 Contributions

The main goal of this work is to construct and describe the relationship between a continuous-time model and its discrete-time approximation. The desired relationship is such that model consistency verified for the approximation also holds for the continuous-time system employing the aforementioned feasibility approach.

This work extends our earlier works [27, 164, 170] to be able to consider continuous-time systems. This extension is achieved by constructing a discrete-time system that encloses all trajectories of the continuous-time system, cf. Figure 1.3. In contrast to our previous works, the solution space of the involved feasibility problems is not approximated by semi-definite programs of fixed size. Instead the converging hierarchy proposed in [108] is employed, cf. Figure 1.5. This allows for the considered system class the derivation of more rigorous results as described next.

This work proposes two different methods to deal with the problem of verifying model consistency with the help of a discrete-time approximation. The first approach employs the classical idea of constructing a discrete-time enclosure of the reachable set of the continuous-time system. The main difference to classical approaches is here the possibility to give guaranteed convergence results. In particular, this means that it can be guaranteed that the derived envelope converges to the actual reachable set of the continuous-time model at all considered time points. Furthermore, the maximal time-step size that guarantees existence and uniqueness of solutions can be derived. This is possible by formulating the conditions of the Picard-Lindelöf theorem in terms of an infinite-dimensional linear program that is subsequently approximated by a converging hierarchy of semi-definite programs. Moreover, this approach of deriving a discrete-time enclosure guarantees that the model invalidation as well as the parameter estimation problem can be solved for the continuous-time system.

The second approach employs a nominal (possibly uncertain) continuous-time model that is used to generate a training set. The training set is then employed to estimate a set of errors needed such that the discrete-time model is not inconsistent with the training set. For this purpose, it is shown that the techniques developed in our earlier works can be directly employed. As a second step, an augmented Euler model is constructed that incorporates these error sets as time varying parameters. To conclude model consistency, the augmented model is then tested with respect to the available data. This allows the following statement. If the augmented model is found inconsistent, it can be concluded that the nominal system is also inconsistent.

1.3 Observability of Uncertain Polynomial Systems

1.3.1 Motivation and Existing Approaches

In control engineering, clearly stability of a system is the most prominent system property that one wants to verify. Many approaches are known to prove stability, from a theoretical perspective as well as from a computational perspective, see e. g. the standard textbooks [82, 96, 192]. In this work, we are not trying to contribute to this vast literature of stability results. Instead we aim at deriving computationally efficient methods for investigating the observability for uncertain polynomial systems. Namely the ability to reconstruct from given observations the initial state of a system.

This type of question is not only important from a theoretical point of view, but also from a practical one. In practice, it is typically not possible to employ a controller based solemnly on the system states as the states typically cannot be measured directly. Instead the states are derived from measurements of the output by an observer. The ability of deriving the states from measurements is then referred to as observability. Observability is not only a requirement for state controller design, but also for parameter estimation as parameters can only be estimated if the corresponding states can be reconstructed [190].

Observability as a system property is well studied for deterministic linear systems (systems with no uncertainties), see e. g. [29, 84, 138]. This is also true to some extend for deterministic nonlinear systems [67, 78, 79, 195]. However, observability for nonlinear systems is on the one hand more involved to define than in the linear case and on the other hand only few general approaches to investigate observability exist, e. g. [192]. For polynomial systems the situation is more favorable, see e. g. [15, 55, 72, 189, 190, 194] for general definitions and requirements, and [89–92, 94, 143, 212] for constructive algebraic conditions and algorithms. Similar algebraic conditions are derived for parameter identifiability in [14].

However, for uncertain polynomial systems this is not the case. Most results in this field are connected to differential inclusions and are mostly concerned with the existence than with the actual computation. So, for instance, a characterization of the

needed local injectivity condition for observability is given in [11] and the connection between this condition and the viability kernel of differential inclusions is proposed in [88]. Both approaches allow in principle to investigate the observability, however, an application is difficult as typically no constructive proofs are possible. This work aims at a computational solution to verify observability of an uncertain polynomial system. This is achieved by extending the algebraic conditions of Kawano and Othsuka [94] such that they can be solved by the methods derived before.

1.3.2 Contributions

This work proposes several extensions to the algebraic observability analysis of Kawano and Othsuka [90, 94]. These extensions are built on the theory of real algebraic geometry as detailed next.

At first it is demonstrated that the algebraic conditions for local-at-a-point observability can be tested efficiently with the help of semi-definite programs. Furthermore, we show that observability indeed induces a finite convergence property of the semi-definite program such that the solution of the semi-definite program is necessary and sufficient for observability for deterministic polynomial systems. It should be noted that the conditions derived in this work can be directly applied to parameter identifiability analysis by extending the states by the parameters appropriately.

It is also shown that determining whether an uncertain system is observable is considerably harder and the sufficiency of the deterministic case is lost. However, this work shows that the observability and unobservability can still be investigated. For this reason, an infinite-dimensional linear program is constructed that encodes the semi-algebraic extension of the algebraic conditions. This linear program is then solved via a converging hierarchy of relaxations [108]. It is furthermore shown that if at a certain relaxation order a rank condition holds the system is not observable. The obtained results show that observability can also be investigated for uncertain systems and with the methods developed (resp. employed) in other parts of this work. For continuous-time systems a similar approach based on the energy visible at the output is derived, which allows similar conclusions as the (semi-)algebraic approach.

1.4 Thesis Outline

- Chapter 2 briefly introduces the considered system classes and the used notations. Additionally, we define the system properties which are investigated in later chapters.

- Chapter 3 presents the concepts of model consistency for uncertain discrete-time and continuous-time systems with unknown-but-bounded measurements as well as qualitative information.

- Chapter 4 presents the set-based parameter estimation framework that is central for the following chapters. Here, it is shown that model invalidation and parameter estimation can be addressed by relaxing a polynomial feasibility problem. We discuss two relaxation approaches and illustrate the methods with several examples.
- Chapter 5 proposes a method for estimating the parameters of continuous-time models via their discrete-time approximations. This is done by extending the results of Chapter 4 such that a relationship between the discrete-time and the continuous-time model exists. This relationship allows us to conclude from model inconsistency of the approximation to the invalidity of the continuous-time model.
- Chapter 6 is concerned with the verification of observability of uncertain polynomial discrete-time and continuous-time systems with a clear focus on computational methods based on Chapters 4 and 5.
- Finally, Chapter 7 summarizes and concludes this thesis. It includes an outlook on possible research topics related to this contribution.

2 Considered System Classes and Properties

The main goal of this work is the development of methods for the estimation and analysis of uncertain polynomial systems. The main focus is put hereby on the unified treatment of continuous-time and discrete-time systems with unknown-but-bounded uncertainties. For the analysis we concentrate on three particular system properties, namely model consistency, reachability and observability. In the following, we provide a short overview over polynomial systems and define the considered system properties.

This chapter is structured as follows. First, the system class of polynomial systems is introduced for the continuous-time and the discrete-time case and the employed notation is fixed. Second, a short overview over the three system properties is given and relevant notions are defined. We conclude with a summary and an outline of following topics. Note that this chapter is kept rather concise to serve as a reference for the employed notation and defined notions in later chapters.

2.1 System Classes

2.1.1 Continuous-time Systems

In this work, we consider systems of the form

$$\begin{aligned} \dot{x}(t) &= f(x(t), u(t), p), \quad x(t_0) = x_0, \\ y(t) &= h(x(t), u(t), p), \end{aligned} \tag{2.1}$$

where $x(t) \in \mathbb{R}^{n_x}$ denotes the *states*, $u(t) \in \mathbb{R}^{n_u}$ the *input*, $y(t) \in \mathbb{R}^{n_y}$ the *output*. The system parameters are denoted with $p \in \mathbb{R}^{n_p}$. Throughout, this work we assume that the initial condition x_0 belongs to the set $\mathcal{X}_0 \subseteq \mathbb{R}^{n_x}$, p to the set $\mathcal{P} \subseteq \mathbb{R}^{n_p}$ and the input belongs to the compact set $\mathcal{U} \subset \mathbb{R}^{n_u}$ also called *admissible control set*, where $\mathcal{X}_0, \mathcal{P}, \mathcal{U}$ are semi-algebraic sets, i. e. the sets are defined by a finite number of polynomial inequalities. The vector field $f : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_p} \to \mathbb{R}^{n_x}$, the control function $u : \mathbb{R} \to \mathcal{U}$ and the output function $h : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_p} \to \mathbb{R}^{n_y}$ are assumed to be continuous polynomial functions that satisfy a *Lipschitz condition* on a time interval $[0, T] \subset \mathbb{R}$, cf. [68, 106]). Furthermore, we assume the input is piecewise-constant, i. e. it fulfills the zero-order hold condition [68]. For further reference the ring of polynomials with real coefficients in the variables x, u, p is denoted as $\mathbb{R}[x, u, p]$. In practice polynomial systems are often used to model biological or chemical reaction systems, see e. g. [164, 170] and the example sections of later chapters. Note that

the considered sets can typically be derived from measurements, process insight and conservation relations, also cf. Chapter 3.

We denote the *trajectories* (also *flow*) of (2.1) for an initial condition $x_0 \in \mathcal{X}_0$, a parameter $p \in \mathcal{P}$ and a control function $u(\cdot) \in \mathcal{U}$ with $\phi(x(t)|x_0, u(t), p)$. We assume that $\phi(x(t)|x_0, u(t), p)$ exists for all $t > t_0$, $x_0 \in \mathcal{X}_0, p \in \mathcal{P}$ and all $u(\cdot) \in \mathcal{U}$, i. e. ϕ is forward complete [4]. Note that this property results from the assumed Lipschitz continuity. For a more detailed discussion on the existence and uniqueness of solutions to initial value problems of form (2.1) see e. g. [54]. The corresponding output trajectories are denoted by $\phi_y(x(t)|x_0, u(t), p) := h\big(\phi(x(t)|x_0, u(t), p), u(t), p\big)$.

Next we introduce the considered class of discrete-time systems.

2.1.2 Discrete-time Systems

We define a parameterized discrete-time system of the form

$$\begin{aligned} x(k+1) &= f_D(x(k), u(k), p), \quad x(0) = x_0, \\ y(k) &= h_D(x(k), u(k), p). \end{aligned} \tag{2.2}$$

In some cases, we employ an implicit formulation, e.g. due to an implicit numerical integration of (2.1),

$$\begin{aligned} 0 &= f_D(x(k+1), x(k), u(k), p), \quad x(0) = x_0 \\ 0 &= h_D(y(k), x(k), u(k), p), \end{aligned} \tag{2.3}$$

where in (2.2) and (2.3) the variables $x(k), u(k), y(k), p$ denote the states, the input, the output and the parameters equivalent to the continuous-time system. We denote with $D > 0$ some positive *time-step* and $k \in \mathbb{N}$ denotes the *discrete-time index*. We assume that f_D, h_D are polynomials, i. e. $f_D, h_D \in \mathbb{R}[x, u, p]$, and that the variables $x_0, u(t), p$ are bounded by semi-algebraic sets. To distinguish variables of discrete-time systems and those of continuous-time systems (if necessary), we add a bar to the sets, i. e. $x_0 \in \bar{\mathcal{X}}_0 \subseteq \mathbb{R}^{n_x}, u(k) \in \bar{\mathcal{U}} \subset \mathbb{R}^{n_u}, p \in \bar{\mathcal{P}} \subseteq \mathbb{R}^{n_p}$.

Remark 1. *In most instances* (2.2) *and* (2.3) *will be considered as numerical approximations (or discretizations) of* (2.1)*. However, the results presented in Chapter 4 and large parts of Chapter 6 still hold for general discrete-time systems.*

As for the continuous-time system (2.1), we assume that f_D, h_D fulfill a Lipschitz condition depending on the time-step D, see [68]. We denote a state trajectory of (2.2) and (2.3) at $k \in \mathbb{N}$ with $\phi_D(x(k)|x_0, u(k), p) := f_D^k(x(k-1), u(k-1), p)$, where $f_D^k := f\big(x(k-1), u(k-1), p\big) = f_D \circ \cdots \circ f_D(x_0, u(0), p)$, and an output trajectory by $\phi_{D,y}(x(k)|x_0, u(k), p) := h_D \circ f_D^k(x(k-1), u(k-1), p)$. To be able to correlate ϕ_y of (2.1) and $\phi_{D,y}$ on a set of discrete-time points $\mathcal{T} := \{t_1, \ldots, t_n\}, t_i \in \mathbb{R}$, we assume throughout this work that for every D there exists a k such that $k \cdot D \equiv t_i$ for all

$t_i \in \mathcal{T}$ and we simply write $k \in \mathcal{T}$. Furthermore, we write in the following with slight abuse of notation $k \in [0, T]$ instead of the formally correct $k \in \{0, \dots, T\}, T \in \mathbb{N}$.

In the following, we introduce for continuous-time and discrete-time systems the considered system properties.

2.2 Considered System Properties

In this work, mainly three system properties are investigated, namely model consistency, reachability and observability, which are introduced in the subsequent sections.

2.2.1 Model Consistency

In this work, model consistency is the central system property for deriving a relationship between continuous-time systems and their discrete-time approximation. Therefore, we provide here only a basic definition and elaborate on this topic further in Chapter 3. In general, with model consistency one refers to the ability of a model to reproduce measurement data for certain initial conditions and parameterizations. If we assume that measurements are given as semi-algebraic sets $\mathcal{M}(t_i) \subseteq \mathbb{R}^{n_y}$, we can define model consistency for (2.1) and (2.2) as follows.

Definition 1 (Model consistency)**.** *Given output measurements $\mathcal{M}(t_i)$ at time-points $t_i \in \mathcal{T} := \{t_1, \dots, t_n\}, t_i \in \mathbb{R}$. Model (2.1) (resp. (2.2)) is said to be consistent with the measurements, if there exists $x_0 \in \mathcal{X}_0, u \in \mathcal{U}$ and $p \in \mathcal{P}$ such that $\phi_h(\phi(x(t_i)|x_0, u(t), p) \in \mathcal{M}(t_i)$ (resp. $\phi_{D,h}(x(k)|x_0, u(k), p) \in \mathcal{M}(t_i)$) for all $t_i \in \mathcal{T}$ (resp. $k \in \mathcal{T}$).* ∎

In Chapter 3, we consider more general data types than the measurement description employed here. This allows us to consider not only quantitative, but also semi-quantitative data like *if-then* observations. We further present methods to investigate model consistency in Chapter 4 and Chapter 5. Moreover, these methods allow deriving outer-approximations of the initial conditions, inputs and parameters that lead to the model being consistent. One important concept employed in the development of such methods for continuous-time systems is reachability as introduced next.

2.2.2 Reachability

Reachability can be seen as the generic system property that corresponds to model consistency. Indeed, a system can only be consistent with the available measurement data if the data are contained in the reachable set of the system. To derive methods for proving model consistency of a continuous-time system (2.1), we construct in Chapter 5 a discrete-time system that envelopes the reachable set of (2.1). Following [68, 69], we define the reachable set of (2.1) for a single initial condition as the set-valued mapping

$$\mathcal{R}(x_0) = \bigcup_{u(t)\in\mathcal{U},\ p\in\mathcal{P}} \{\phi(x(t)|x_0, u(t), p)\}. \tag{2.4}$$

If we are interested in the reachable set for all initial conditions x_0 in the set $\mathcal{X}_0$, the reachable set becomes

$$\mathcal{R} = \bigcup_{x_0\in\mathcal{X}_0} \mathcal{R}(x_0). \tag{2.5}$$

Equivalently, we define the reachable set of a discrete-time system (2.2) (resp. (2.3)) as

$$\bar{\mathcal{R}}(x_0) = \bigcup_{u_k\in\mathcal{U},\ p\in\mathcal{P}} \{\phi_D(x(k)|x_0, u(k), p)\}, \tag{2.6}$$

and

$$\bar{\mathcal{R}} = \bigcup_{x_0\in\mathcal{X}_0} \bar{\mathcal{R}}(x_0). \tag{2.7}$$

In Chapter 4, we provide methods for determining the sets (2.6) and (2.7), and methods for deriving the sets (2.4) and (2.5) are given in Chapter 5. The next considered system property corresponds to the ability to estimate the initial conditions from measurement data.

2.2.3 Observability

Observability corresponds to the possibility to derive the states of a system from given measurements of the output. This system property is of particular interest in the case that not all states can be measured directly in a process, but the state information is needed to derive a suitable controller. In this section, we define the notion of observability following the works [94, 95, 212]. For a more general treatment of observability of polynomial systems see [190].

Note that throughout this section, it is assumed that the parameter vector p is fixed. The extension to the uncertain case is given in Chapter 6. To be able to define observability, consider the notion of (finite-time) *distinguishability*.

Definition 2 (Distinguishability)**.** *A pair of initial conditions $\alpha, \beta \in \mathcal{X}_0 \subseteq \mathbb{R}^{n_x}$ are said to be distinguishable for (2.1) (resp. (2.2)) on the time interval $[0, T]$ if there exists a piecewise constant input $u(t)$ (resp. $u(k)$) such that $\phi_h(x(t)|\alpha, u(t), p) \neq \phi_h(x(t)|\beta, u(t), p)$ for some $t \in [0, T]$ (resp. $\phi_{D,y}(x(k)|\alpha, u(k), p) \neq \phi_{D,y}(x(k)|\beta, u(k), p)$). An initial condition α is said to be indistinguishable from $\beta \in \mathcal{X}_0$ if $\phi_y(x(t)|\alpha, u(t), p) = \phi_h(x(t)|\beta, u(t), p)$ holds for all $t \in [0, T]$ (analogously for $\phi_{D,y}$ for all $k \in [0, T]$). The set of all points $\beta \in \mathcal{X}_0$ that are indistinguishable from $\alpha \in \mathcal{X}_0$ is denoted by $\mathcal{O}(\alpha, \mathcal{X}_0)$.* ∎

For practical purposes, it is often sufficient that not all α can be distinguished from every other point in $\mathcal{X}_0$, but only from a neighborhood $\mathcal{N} \subset \mathcal{X}_0$ around α, i. e.

$\mathcal{O}(\alpha, \mathcal{X}_0) \cap \mathcal{N}(\alpha) = \{\alpha\}$. Note also that we consider here distinguishability on the semi-algebraic set $\mathcal{X}_0$ instead of the classical definition over $\mathbb{R}^{n_x}$. This derives from the fact that we consider uncertain systems and we, therefore, have to derive a semi-algebraic conditions for observability. Using the above definition we can define *local-at-a-point observability* as follows.

Definition 3 (Local-at-a-point Observability)**.** *A system is said to be locally-at-a-point observable for a finite-time interval* $[0, T]$*, if there exists a neighborhood* $\mathcal{N} \subseteq \mathcal{X}_0$ *of the initial condition* $\alpha \in \mathcal{X}_0$ *such that for all* $\beta \in \mathcal{N} \setminus \{\alpha\}$ *the pair* α, β *is distinguishable, i. e.* $\mathcal{O}(\alpha, \mathcal{N}) = \{\alpha\}$. ∎

Local observability of a system is defined accordingly, i. e. there exists a neighborhood for every $\alpha \in \mathcal{X}_0$ such that α can be uniquely determined from the input and output or for short:

Definition 4 (Local Observability)**.** *A system is said to be locally observable if for every initial condition* $\alpha \in \mathcal{X}_0$ *the system is locally-at-a-point observable.* ∎

Remark 2. *These observability notions are in some sense stronger than the classical observability notion [78] as it implies that local information is sufficient to determine the initial condition and, therefore, a limit on the time interval length is guaranteed.*

The final observability notion employed in this work is defined as follows.

Definition 5 (Global Observability)**.** *A system is said to be globally observable if all initial conditions* $\alpha, \beta \in \mathcal{X}_0$ *are distinguishable for the system.* ∎

The main idea of investigating observability for polynomial systems is to reformulate the introduced observability notions as algebraic sets, i. e. a set of finitely many polynomial equalities. This is done by means of polynomial ideals and their localizations as shown e. g. in [94]. However, the algebraic investigation of observability is only applicable for fixed parameterizations. For this reason, the method of Kawano and Ohtsuka [94] is extended in Chapter 6 to uncertain polynomial systems.

2.3 Summary

In this chapter, we have defined the considered system classes as well as the properties we investigate in the following chapters. In particular, we defined model consistency, reachability and observability for polynomial systems.

Model consistency, i. e. the ability of a model to reproduce some measurement data, is hereby of particular importance. On the one hand the concept of consistency is the basis for deriving the set-based estimation approach for discrete-time systems in Chapter 4. On the other hand, model consistency is employed in Chapter 5 to extend the set-based estimation approach to continuous-time systems. As this important concept was introduced here in its simplest form, we provide next a more detailed discussion and generalize the considered uncertainties to semi-quantitative data.

3 Uncertainties and Model Consistency

In this chapter, we define and briefly discuss the set-based uncertainties considered in this work. Furthermore, we extend the notion of model consistency introduced in the previous chapter from quantitative to semi-quantitative data. Semi-quantitative data can be employed to describe observations and requirements on a system in a more general setting [170]. For example, if a controller is supposed to lead to a limited amount of overshoot, it is often easier to formulate such a requirement in terms of semi-quantitative constraints instead of a fully quantitative description [22].

This chapter is structured as follows. We describe at first the main advantages of set-based uncertainties. Afterwards, we define set-based (or unknown-but-bounded) uncertainties and introduce the notion of model consistency for quantitative and semi-quantitative data. For this purpose, we introduce temporal logic constraints that allow for instance to transfer *if-then* conditions into a mathematical description, cf. Chapter 4.

3.1 Introduction

In the development of a mathematical model an inherent difficulty is the consideration of uncertainties [202]. Such uncertainties can range from external perturbations to limited structural knowledge and can have several sources. A major source of uncertainty derives directly from the available measurement data. For instance, the experimental data might be sparse, incomplete or the employed measurement techniques are indirect or have low accuracy and resolution [164, 202]. This results in large uncertainties of the absolute quantities of the measured output. Another source of uncertainty can be the formulation of requirements the model has to fulfill. In many cases such requirements are formulated as conditional or temporal statements and *if-then* conditions, instead of a quantitative description of the requirements [22, 170].

These uncertainties therefore pose the following challenges: first, how can uncertainties be described; second, how can uncertainties be incorporated in a model; third, how can uncertain models be analyzed. All three challenges are addressed in this chapter.

In this work, we focus on a set-based uncertainty description [179, 219] instead of a probabilistic description [152, 202]. This description is particularly useful in the case that measurements are of limited precision with low sampling frequencies, and only a low number of experimental replicates are available. In this case it is typically not possible to determine the actual probabilistic uncertainty distribution. Therefore, sets derived from standard deviations or worst-case approximations, can be helpful [202].

Furthermore, as the set-based uncertainty description describes all possible values a variable can take, it is particularly well suited for the formulation of safety and quality requirements as seen later in this chapter.

The analysis methods that we propose in this work are based on the notion of model consistency as introduced in this chapter. In the following, we define how data and safety requirements can be formulated and incorporated into a model. Afterwards, in Chapter 4 and Chapter 5 we propose a method to test a model for model consistency. Note that this chapter is based on [170].

3.2 Considered Uncertainties

In reality measurements are always subject to uncertainties. On the one hand such uncertainties might arise due to the accuracy or precision of the employed measurement device or on the other hand due to limitations such as e. g. limited repeatability [34, 158]. Uncertainties, however, do not only arise from measurement uncertainties, but also from unknown parameter values and initial condition. Especially this is true for biological systems [99, 220]. In this work, we employ the concept of *unknown-but-bounded* variables (or bounded-error description), that has been used e. g. [135, 136, 179, 219], which we define as follows.

Definition 6 (Unknown-but-bounded variable (**ubb**))**.** *A variable $m(t_1)$ is said to be unknown-but-bounded at time $t_1 \in \mathbb{R}$, if it lies in the semi-algebraic set $\mathcal{M}(t_1) = \{u \in \mathbb{R}^{n_u}, p \in \mathcal{P}, y \in \mathbb{R}^{n_y} : g(u(t_1), p, y(t_1)) \geq 0\}$, where $g : \mathbb{R}^{n_u} \times \mathbb{R}^{n_p} \times \mathbb{R}^{n_y} \to \mathbb{R}^{n_g}$ is a polynomial (vector) function for some $n_g \in \mathbb{N}$, and u, p and y are as in* (2.1) *(resp.* (2.2) *and* (2.3)*).* ∎

Function g in Definition 6 describes in the simplest case an interval given by a lower and an upper bound on a variable, for instance, $\mathcal{M}(t_1) = \{y \in \mathbb{R}^{n_y} : \underline{y} \leq y \leq \overline{y}\}$, where $\underline{y}, \overline{y} \in \mathbb{R}^{n_y}$ denote the lower and upper bound, respectively. Such bounds might be derived from the accuracy specifications of the measurement device, the noise level or from physical insight (concentrations have to be positive). However, the determination of bounds depends in general from the considered system.

One particular example, where such data is considered is constraint control problems. Either this data describe a limitation of a variable, e. g. input constraint, or a safety requirement like $x(t_1)$ is not allowed to take values outside of the set $\mathcal{M}(t_1)$

In this work, we consider the initial conditions, parameters as well as the output as unknown-but-bounded variables. This allows us to consider model uncertainties and quantitative measurement uncertainties. However, it does not allow for the consideration of data and requirements given in a semi-quantitative description. Therefore, we extend in the following section the considered uncertainty description. For simpler presentation, we do so in terms of the notion of model consistency.

3.3 Model Consistency

Model consistency refers to the ability of a model to reproduce measurements or the fulfillment of safety requirements. In the following, we denote with data either measurements or such requirements. The notion of model consistency depends on the description of the data, and we distinguish here between quantitative and semi-quantitative data as seen next.

3.3.1 Model Consistency for Quantitative Data

We denote the collection of **ubb** variables $m(t_i)$ at time points $t_i \in \mathbb{R}, i \in \mathcal{I}_m$ as

$$\mathcal{M} = \{\mathcal{M}(t_i), t_i \in \mathcal{T}\}, \tag{3.1}$$

where $\mathcal{T}$ denotes the collection of time points t_i, i. e. $\mathcal{T} = \{t_i, i \in \mathcal{I}_m\}$.

Consequently, if we have data of the output variable given by a collection of **ubb** variables, we can state the question of model consistency as:

Definition 7 (Model consistency)**.** *Given a collection of* ***ubb*** *measurements $\mathcal{M}$. Model* (2.1) *(resp.* (2.2) *and* (2.3)*) is said to be consistent with $\mathcal{M}$, if and only if there exists $x_0 \in \mathcal{X}_0, u \in \Omega$ and $p \in \mathcal{P}$ such that $y(t_i) \in \mathcal{M}(t_i)$ for all $t_i \in \mathcal{T}$ (resp. $k \in \mathcal{T}$).* ■

For shorthand of notation, we use in the following $y \in \mathcal{M}$ instead of the formally correct $y(t_i) \in \mathcal{M}(t_i)$ for all $t_i \in \mathcal{T}$ (resp. $k \in \mathcal{T}$) a measurement was taken. Figure 3.1 illustrates the consistency property for some trajectories of a continuous-time model.

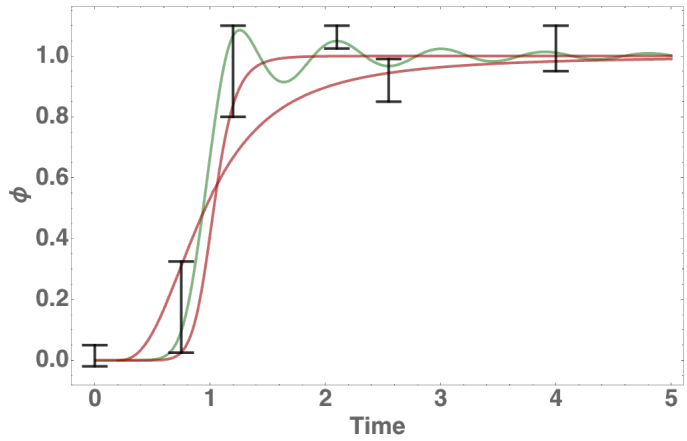

Figure 3.1: Quantitative measurement data: black vertical lines represent measurements given as **ubb** variables. The green line represents an example for a trajectory that is consistent with the data, while the red lines are examples for inconsistent trajectories.

We will denote model consistency with the binary variable $\phi \in \{0, 1\}$, whereas $\phi = 1$ if and only if a model is consistent with the available data, i.e. $(\phi = 1) \iff y \in \mathcal{M}$. This notation will allow in the following section to state not only quantitative measurement data, but also semi-quantitative and qualitative information.

3.3.2 Model Consistency for Semi-Quantitative, Qualitative Data

In practice, one frequently encounters data that are given with respect to some other data. In systems biology this kind of data often appears if e. g. the initial condition can not be defined accurately and all measurement data are only given relative to this initial condition. Furthermore, data might be taken from experiments in which different or even no scaling factors for quantitative/absolute data were available, e. g. when experiments of different cell lines or organisms are considered. However, also in technical systems such data can occur when e. g. in a *bottom-up* approach different experimental setups are compared.

To compare and integrate such data in a model, the data are typically normalized with respect to an initial or maximal value. In terms of **ubb** variables, we can handle this kind of data by applying an order operator, i.e. two **ubb** variables $m_1(t_1), m_2(t_2)$ are said to be relational if there exists a relationship of the form

$$m_1(t_1) \ \diamond \ m_2(t_2), \tag{3.2}$$

where $\diamond \in \{<, \leq, =, \geq, >\}$ denotes a comparative operator. To avoid redefining the order operator for a collection of relational data, we state next model consistency only with respect to two **ubb** variables.

Definition 8 (Model Consistency (Relational Data))**.** *Given relational data* $m_1(t_1), m_2(t_2)$ *of form* (3.2)*. Model* (2.1) *is said to be consistent with* $m_1(t_1), m_2(t_2)$*, if and only if there exists* $x_0 \in \mathcal{X}_0, u \in \Omega$ *and* $p \in \mathcal{P}$ *such that* $y(t_1) = m_1(t_1) \in \mathcal{M}_1(t_1)$ *implies* $y(t_1) \diamond m_2(t_2), \diamond \in \{<, \leq, =, \geq, >\}$*, i. e.* $(\phi = 1) \iff y(t_1) \diamond m_2(t_2)$*.* ∎

An example for relational data can be found in control specifications, such as an overshoot bound of the step response of a system. For instance, if the set-point is equal to A the overshoot is not allowed to be over 1.2 times of the absolute value of A.

A second type of semi-quantitative/qualitative data frequently encountered in practice is temporally uncertain data. For instance, if the measurement sampling rate is uncertain due to measurement restrictions such as human interaction, the precise time-point a measurement was taken might be unknown. Other examples are e. g. different response-times of individual cells in cell populations due to cell-to-cell variability [197] or delay and distortion effects of measurement devices due to an imprecisely known half-life of a fluorescent marker or protein (see e.g. [17]). To capture this kind of

data, we simply allow any value t in the set $\{t_1, \ldots, t_n\}$ for the **ubb** variable $m(t)$. Accordingly, we say $m(t)$ is temporally uncertain.

Definition 9 (Model Consistency (Temporally uncertain data))**.** *Given a temporally uncertain* ***ubb*** *variable* $m(t), \{t_1, \ldots, t_n\}$*. Model* (2.1) *is said to be consistent with* $m(t)$*, if and only if there exists* $x_0 \in \mathcal{X}_0, u \in \Omega$ *and* $p \in \mathcal{P}$ *such that* $y(t_i) \in \mathcal{M}(t_i)$ *for at least one* $t_i \in \{t_1, \ldots, t_n\}$*, i. e.* $(\phi = 1) \Longleftrightarrow y(t_1) \in \mathcal{M}(t_1) \vee \ldots \vee y(t_n) \in \mathcal{M}(t_n)$ ■

Figure 3.2 illustrates temporally uncertain data. Note that we choose a set of distinct time-points $\{t_1, \ldots, t_n\}$ to avoid redefining the $\vee$ operator for infinitesimal time increments, however, in principle Definition 9 holds also for a time-interval $[t_1, t_n]$.

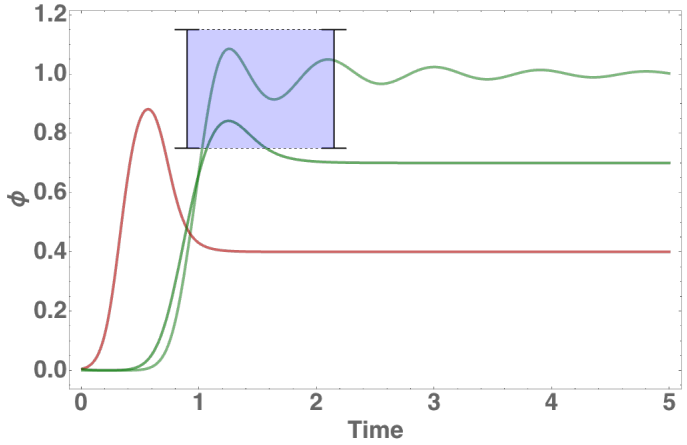

Figure 3.2: Temporally uncertain data: Shaded area represents temporally uncertain data. The green lines represent trajectories that are consistent with the data, while the red line is an example for an inconsistent trajectory.

In many cases, data are available as qualitative, temporal *if-then* information. A classical example for such an *if-then* observation in biology would be the regulation of the *lac*-operon (*diauxie*, e. g. [137] and Figure 3.3): "if fructose is present in the medium then the intake of lactose is inhibited". Other examples can be found in the control of *mixed-logical systems* and system verification, see e. g. [22]. For instance, "if the set-point is changed then the controller has to limit the overshoot".

A variable $m(t_1)$ is said to be conditional, if it is described by logical combinations of **ubb** variables $m_i(t_i), i \in \{1, \ldots, n\}$, i.e. $m(t) = m_1(t_1) \diamond \ldots \diamond m_n(t_n)$, where $\diamond$ is either a *conjunction* $(\wedge)$ or a *disjunction* $(\vee)$.

Definition 10 (Model Consistency (Conditional data))**.** *Given conditional data* ***ubb*** *variable* $m(t_i), i \in \{1, \ldots, n\}$*. Model* (2.1) *is said to be consistent with conditional data* $\mathcal{M}$*, if and only if there exists* $x_0 \in \mathcal{X}_0, u \in \Omega$ *and* $p \in \mathcal{P}$ *such that* $(\phi = 1) \Longleftrightarrow$ $y(t_1) \in \mathcal{M}(t_1) \diamond \ldots \diamond y(t_n) \in \mathcal{M}(t_n),\ \diamond \in \{\wedge, \vee\}$. ■

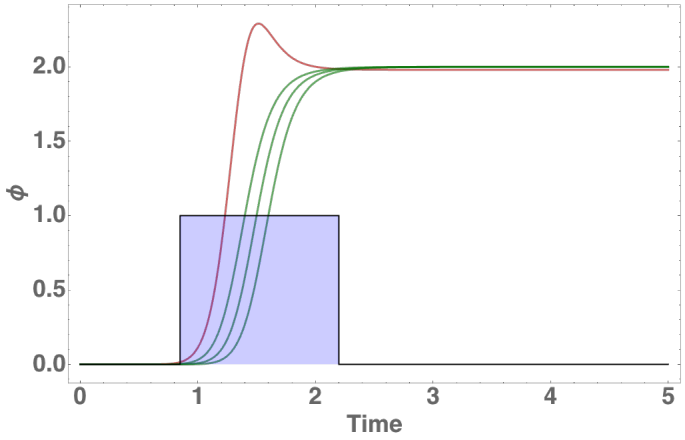

Figure 3.3: Conditional data: if the input u (shaded area) is equal to 1 then the value of the output increases. Note that the red, overshooting trajectory would be considered as inconsistent with the data.

In the next chapters, we show that the concept of model consistency for the considered system class can be verified using semi-definite programming for discrete-time systems of form (2.2). We furthermore show that the same procedure can be used to derive guaranteed outer-bounds on the initial conditions and parameters. Note that the necessary extensions for checking model consistency of continuous-time models are presented in Section 5.5.

3.4 Summary

In this chapter, the considered set-based uncertainty was introduced. Furthermore, we extended the notion of model consistency from quantitative to semi-quantitative data, where data are referred to as measurements on the one hand and as safety requirements on the other hand. Semi-quantitative data is in particular useful to describe requirements of a model that cannot easily be formulated in qualitative terms. Examples include *if-then* conditions in controller verification like: if the set-point is changed then the controller has to stabilize the new set-point in a specific amount of time.

Although, the set-based data description has the advantage to be rather general, it also poses certain challenges. Only few methods are applicable in this case and they are typically computationally rather expensive. For this reason, we propose next computationally efficient methods that can consider set-based (semi-)quantitative data for discrete-time (Chapter 4) and continuous-time systems (Chapter 5). These methods are based on relaxation techniques for polynomial optimization problems as described in the next chapter.

4 Discrete-Time Systems

In this chapter, we provide an overview over set-based estimation for discrete-time systems. The proposed approach is based on reformulating the desired estimation task, e. g. state or parameter estimation, as a nonlinear feasibility problem. As the nonlinear feasibility problem is in general nonconvex and, hence, difficult to solve, it is addressed by relaxation to a semi-definite or linear program employing two approaches known from polynomial optimization. Although, the relaxation strategy might introduce spurious solutions (for a finite relaxation order), it can be certified if the model is inconsistent with the available data or design specifications. For this purpose, two algorithms are derived that further result in a guaranteed outer-approximation of the solution set of the nonlinear feasibility problem, thus solving the estimation task of interest. The presented relaxation strategies and algorithms build the foundation for the results derived in later chapters. One particular focus in the choice of relaxation strategies and in the development of the algorithms is the flexibility to balance theoretical accuracy and computational load.

This chapter is structured as follows. First, we describe the nonlinear feasibility formulation and its relationship to estimation and model validation. Second, we introduce the two employed relaxation approaches that allow addressing the feasibility problem efficiently. Before, concluding this chapter, we present several simulation examples.

4.1 Introduction: Set-based Estimation

In general estimation/identification can be considered as the process of finding an *estimate*, or *approximation*, of a result or solution that reflects the available data. In particular, in the context of mathematical models of a dynamical system, the goal is to extrapolate from the available data certain characteristics and variables of the system. So, for instance, in state estimation the system states are determined based on the made observations of the output. There exist several different approaches to derive such an estimate for every particular instance of the estimation problem, see e. g. [119, 142] for an overview on system identification, [186] for state estimation, or [210, 213] for parameter estimation.

In many cases, such estimation problems are stated as optimization problems, in which some *objective* (or cost) *function* is minimized over appropriate optimization variables, e.g. the model parameters. A commonly employed objective is the minimization of the difference between measurement data and model prediction, evaluated

by least squares or maximum likelihood functions [210]. However, often the construction of the objective is purely based on the experience of the user or deduced from assumptions on the noise, respectively the prior distributions of the interesting variables. Due to the nonlinearities typically arising in models of real systems, the resulting optimization problems are frequently non-convex and hard to solve. As a consequence, common approaches aim at finding locally optimal solutions, instead of globally optimal ones [210]. As the local optimum found strongly depends on some initial guess, such approaches are often combined with stochastic strategies to achieve some desired global property [126, 127].

In a set-based setting the goal is, however, slightly different. Instead of searching one particular optimum one tries to derive the set of all possible solutions. This is particularly useful if guaranteed statements are required. For instance, set-theoretic methods can be used to check whether a process operates within a previously defined safe operating condition or fulfills some design specification. Furthermore, the set-based setting allows a natural consideration of two crucial aspects in the analysis and design of dynamical systems namely uncertainties and constraints. Checking whether the constraints and measurements are represented by the model is typically based on the notion of consistency. Here the interesting variable space is classified into consistent and inconsistent subsets. A classical tool to check consistency is interval arithmetics, see e. g. [83] and references therein.

In this work, we propose a different set-based method for estimation and in particular for model invalidation of discrete-time models based on our previous works (mainly [27, 164, 170]). For this purpose a nonlinear feasibility problem (Section 4.2) is formulated which is relaxed into a semi-definite or linear program (Sections 4.3.1-4.3.3) that can be solved efficiently. On the one hand the presented approaches are able to provide a tight lower bound on a polynomial cost function. On the other hand they allow a guaranteed discrimination of models that are not consistent with the available data and desired specifications. One particular advantage of this approach is the flexibility to incorporate not only quantitative data but also semi-quantitative/qualitative knowledge as introduced in the previous chapter. One further advantage is that the presented approaches can be interchanged to limit the computational requirements, while retaining the obtained guarantees.

4.2 Feasibility Problem Formulation

To verify if a model is consistent with the available qualitative data and semi-quantitative/qualitative information, we have to prove the existence of an initial condition and a parameterization of (2.3) (resp. (2.2)) that leads to a consistent behavior. Note that we formulate this in terms of (2.3), however, an equivalent formulation is possible for (2.2). This can be done by formulating a feasibility (or constraint satisfac-

tion) problem. For this purpose, assume a model (2.3), an initial parameter region $\mathcal{P}$, an admissible input region $\mathcal{U}$, and the available data $\mathcal{M}$ (as defined in Chapter 3) to be given. To ensure that the following problem is well-posed, we further assume that the state space is restricted to the a priori known set $\mathcal{X} \subseteq \mathbb{R}^{n_x}$.

Consider the following set of (semi-)algebraic equations:

$$\begin{aligned} \text{find} \quad & x_0, p \\ \text{subject to} \quad & f_D(x(k+1), x(k), u(k), p) = 0, \\ & h_D(y(k), x(k), u(k), p) = 0, \\ & p \in \mathcal{P}, x_0 \in \mathcal{X}_0, x(k) \in \mathcal{X}, u(k) \in \mathcal{U}, y(k) \in \mathcal{M}, k \in \mathcal{T}. \end{aligned} \tag{4.1}$$

We denote checking whether (4.1) admits a solution or not as a *feasibility problem* (FP). The solution space of FP corresponds to the ability of the model to satisfy the given constraints. If the solution space is empty the model is not consistent with the data as stated in the following theorem.

Theorem 1 (Model Consistency [27])**.** *Model* (2.3) *is consistent with the data* $\mathcal{M}$ *if and only if there exists a feasible point satisfying* (4.1)*.* ■

Proof. The proof follows directly from construction. □

To clarify the notion of solution space (or *feasible region*), it consists of the set of all initial conditions $\mathcal{X}_0^*$, all input signals $\mathcal{U}^*$ and all parameterizations $\mathcal{P}^*$ that lead to trajectories of the state and output such that $x(k) \in \mathcal{X}$ and $y(k) \in \mathcal{M}$ for all $k \in \mathcal{T}$. Furthermore, as the state and output trajectories do not necessarily occupy $\mathcal{X}$ and $\mathcal{M}$ fully, the solution space also consists of the reachable set $\mathcal{R} \subseteq \mathcal{X}$ and the output reachable set $\mathcal{Y}^*$. The projection of the solution space onto a variables in (4.1) estimates its range. In general, we refer to the task of approximating and projecting the solution space of (4.1) onto variables as set-based estimation. However, approximating and projecting the solution space is far from trivial, as it is typically nonconvex. Therefore, we provide in 4.3.1 an approach that solves the set-based estimation problem.

Remark 3 (Variable Selection)**.** *Note that not only projections into the space of* x_0, p*, but the projection into the space of any variable in* (4.1) *can be considered. Such projections can be used e. g. to determine reachable sets [174] or to estimate the states [28] at all time instances* $k \in \mathcal{T}$*.* ■

To state FP, if semi-quantitative information are to be included, special attention has to be paid to the formulation of $\mathcal{M}$. In the following, we give a short overview over the necessary steps and refer the interested reader to [170] (and references therein) for more details.

4.2.1 Formulation of Semi-Quantitative Data

The main difficulty to include semi-quantitative data in (4.1) is to replace the conjunction and disjunction operators needed to describe temporal uncertain and conditional

data as introduced in Chapter 3. Note that the (in-)equalities corresponding to relational data can be directly added to FP.

First, we introduce additional binary variables $\phi \in \{0, 1\}$ indicating whether the model is consistent with the data. To then check if data represented as an **ubb** variable $m(t_k)$ is fulfilled we have to formulate a constraint such that ϕ equals 1 if and only if $y_k \in \mathcal{M}(t_k)$, i. e.

$$(\phi = 1) \iff y_k \in \mathcal{M}(t_k). \tag{4.2}$$

Note that we can state the opposite to (4.2) with help of the *not* operator ($\neg$), i. e. $(\neg\phi = 1) \iff y_k \notin \mathcal{M}(t_k)$, where $\neg\phi = (1 - \phi)$ and ϕ is as in (4.2).

The set-membership, i.e. $y_k \in \mathcal{M}(t_k)$, depends on how the set $\mathcal{M}(t_k)$ is defined. If $\mathcal{M}(t_k)$ is a halfspace, i. e. for $c \in \mathbb{R}^{n_y}$ we have $\mathcal{M}(t_k) = \{\alpha, y_k \in \mathbb{R}^{n_y} : \alpha^T y_k \geq c\}$. Then the implication can be formulated as:

$$\begin{aligned} \phi &\geq \frac{\alpha^T y_k - c}{M}, \\ \phi &\leq \frac{\alpha^T y_k - c}{M} + 1, \end{aligned} \tag{4.3}$$

with $M > \left|\max_{y_k \in \mathcal{M}(t_k)} \alpha^T y_k - c\right|$. Note that on the boundary ϕ can be either 0 or 1. This imprecision can be avoided by introducing further perturbation factors, e. g. [87]. The inequality constraints (4.3) and the equality constraint $\phi = 1$ can then be added to FP to enforce $\alpha^T y_k \geq c$. For more complicated sets it might be necessary to introduce additional binary variables b_i (*atomic propositions*) to formulate the implication. For instance, if for $c \in \mathbb{R}^{n_c}, A \in \mathbb{R}^{n_c \times n_y}$ we have $\mathcal{M}(t_k) = \{y_k \in \mathbb{R}^{n_y} : Ay_k \geq c\}$, which corresponds to a polytope defined by the intersection of halfspaces, every halfspace has to be considered separately. This leads to n_c additional atomic propositions which have to be linked by the conjunction operator, i. e. $(\phi = 1) \iff b_1 \wedge \ldots \wedge b_{n_c}$, e. g.[182]. Sets of the form $\mathcal{M}(t_k) = \{y_k \in \mathbb{R}^{n_y} : a(y_k) \geq 0\}$, where $a : \mathbb{R}^{n_y} \to \mathbb{R}$ is a polynomial or rational function, can simply be included by substituting $\alpha^T y_k - c$ with $a(y_k)$ in (4.3).

To treat the conjunction and disjunction operators, we have to replace them with constraints similar to (4.3). The constraints corresponding to a conjunction of atomic propositions $b_i, i \in I_b = \{1, \ldots, n_b\}$, i. e. $(\phi = 1) \iff b_1 \wedge \ldots \wedge b_n$, are given by

$$\begin{aligned} &\phi \leq b_i, && \forall i \in \mathcal{I}_b \\ &\phi \geq \sum_i b_i - n_b + 1. \end{aligned} \tag{4.4}$$

From (4.4) it follows that $\phi = 1$ can only be true, if all $b_i = 1$. The same consideration is possible for the disjunction of propositions, i. e. $(\phi = 1) \iff b_1 \vee \ldots \vee b_n$,

$$\begin{aligned} &\phi \leq \sum_i b_i, && \forall i \in \mathcal{I}_b, \\ &\phi \geq b_i. \end{aligned} \tag{4.5}$$

Note that Boolean algebra enables us to formulate more complex Boolean functions, include temporal information or to represent truth tables by combinations of not, conjunction and disjunction operators, see e. g. [22, 170]. Furthermore, we can treat other operators frequently used in temporal logic, such as the eventually operator by introducing additional propositions and an appropriate combination of (4.4) and (4.5).

The *integrality constraint* imposed by the binary variables b_i, i. e. $b_i \in \{0, 1\}$, can be formulated in terms of a polynomial equality constraint of form

$$(b_i - 1)b_i = 0, \qquad \forall i \in \mathcal{I}_b. \tag{4.6}$$

The constraints described above can be used to add semi-quantitative/qualitative information (cf. Chapter 3) to FP. Due to the integrality constraints imposed by the binary variables, we denote FP in the remainder with mixed-integer feasibility problem (MIFP). A direct derivation of the solution space of MIFP is not possible as it is non-convex. For this reason, we employ a relaxation technique which results in a convex approximation of MIFP as seen next.

4.3 Reformulation and Relaxation

For the considered system class, we can relax MIFP into a semi-definite program (MISDP) or even further into a linear program MILP. In literature several approaches for reformulating MIFP are known, see e.g. [108, 124, 147, 151, 182]. Note that a detailed comparison of the relaxation and reformulation methods can be found in [113]. In [113], it was also shown that Lasserre's moment relaxation [108] is a refinement of the methods proposed for 0/1 polynomial optimization problems presented in [124, 182]. For more background information we refer to [111, 114] and references therein.

In contrast to previous works [27, 170, 204], we consider mainly Lasserre's moment relaxation [108]. There are mainly two separate reasons. On the one hand this relaxation procedure has advantageous limit properties that allow the derivation of rigorous results regarding the estimation of discrete-time and continuous-time systems. On the other hand, it is the most suited approach if the considered problem is directly formulated in terms of moment conditions to derive closed form approximations of the solution set (see Chapters 5-6). However, depending on the specific problem (resp. example), we also employ Shor's relaxation [185] that is employed in the toolbox ADMIT [204]. This derives from the fact that Shor's relaxation allows the easy derivation of a MILP formulation when semi-quantitative data is considered. Furthermore, having the choice between different relaxations improves the practical applicability in the sense that one can balance between accuracy and computational load (Section 4.3.3).

In the following, we introduce both relaxation techniques and briefly discuss their properties. For convenience, we provide a notation section in Appendix A.1. There, some basic definitions and results on monomial bases, measures, and sequences of

moment matrices are recalled.

4.3.1 Lasserre's Moment Relaxation

The main goal of the relaxation approaches introduced in this and the following subsection, is the reformulation of the nonlinear feasibility problem into an easier manageable problem. Those relaxations further allow an approximation of the solution space of FP. As one of the developed algorithms relies on an objective function, we extend FP to a *polynomial optimization problem* (POP) of form

$$\begin{aligned} \inf \quad & p(x) \\ \text{subject to} \quad & x \in \mathcal{K}, \end{aligned} \tag{4.7}$$

where $p(x)$ is the objective function.

Here $\mathcal{K}$ denotes a not necessarily convex compact set defined by polynomial equalities $h_j(x)$ and inequalities $g_i(x)$, i. e.

$$\mathcal{K} = \{x \in \mathbb{R}^{n_x} : g_i(x) \geq 0, h_j(x) = 0, i \in \mathcal{I}_g, j \in \mathcal{I}_h\}. \tag{4.8}$$

Note that with respect to MIFP the set $\mathcal{K}$ corresponds to the (semi-)algebraic set defined in (4.1), i. e. the constraints regarding the **ubb** variables and the dynamics of (2.2). To get to the standard representation of a semi-definite program, we assume that $g_i(x) \geq 0$ contains non-negativity constraints on all x. Furthermore, for simplicity of notation we consider here only the variable x, but clearly x can be defined such that it covers $x(k), y(k), u(k)$ and p. Concerning the integrality constraints corresponding to the binary variables ϕ_i and b_i introduced for semi-quantitative/qualitative information see Section 4.2. If only feasibility of (4.7) is of interest, then the *cost function* $p(x)$ is set to zero. However, for some applications of POP it is advantageous to have a nonzero cost function, e. g. for the outer-bounding of the feasibility region as described in [27].

In [108], it was shown that for a given real valued polynomial $p(x) \in \mathbb{R}[x]$ a global minimizer x^*_{POP} of POP can be found. The main idea presented in [108] is that POP can be seen as an instance of the following *infinite dimensional linear problem* LP_∞

$$\begin{aligned} \inf \quad & \int_{\mathcal{K}} p(x)\mu(dx) \\ \text{subject to} \quad & x \in \mathcal{K}. \end{aligned} \tag{4.9}$$

Here, the infimum is determined over the positive cone of finite signed Borel measures μ on the constraint set $\mathcal{K}$ denoted by $\mathcal{B}$. It should be noted that by definition of $\mathcal{B}$ the measure of every point outside of $\mathcal{K}$ is equal to zero. It can be shown that POP and LP_∞ are equivalent in the sense that their optimal values are equal, i. e. $x^*_{\text{POP}} = x^*_{\text{LP}_\infty}$ (see Proof of Proposition 2.1 in [111]). A simplified interpretation of (4.9) is that one

calculates the measure of the subset of $\mathcal{K}$ on which $p(x)$ is positive.

The main advantage of LP_∞ is however that it is linear and thus convex. Furthermore, it can be addressed in form of a sequence of finite dimensional semi-definite programs as seen next.

To do so, we first rewrite $\int_\mathcal{K} p(x)\mu(dx) = \sum_\alpha p_\alpha \int x^\alpha \mu(dx)$ in terms of a sequence $(y_\alpha)_{\alpha\in\mathbb{N}^{nx}}$ corresponding to the moments of μ. This leads to the equivalent representation $\sum_\alpha p_\alpha y_\alpha$, with $y_\alpha = \int x^\alpha \mu(dx)$, see also Appendix A.1. Therefore, LP_∞ can be reformulated as

$$\begin{aligned} \inf \quad & \sum_\alpha p_\alpha y_\alpha \\ \text{subject to} \quad & y_0 = 1, \\ & \mathrm{supp}(\mu) \subseteq \mathcal{K}. \end{aligned} \tag{4.10}$$

We can require instead of the support of μ contained in $\mathcal{K}$ that the corresponding moment matrices are positive semi-definite, cf. Lemma 3 in the Appendix. Due to the semi-definite constraints, we get an infinite-dimensional *semi-definite program* (SDP_∞) of form

$$\begin{aligned} \inf \quad & \sum_\alpha p_\alpha y_\alpha \\ \text{subject to} \quad & y_0 = 1 \\ & M(y) \succeq 0, \\ & M(h_j y) = 0, j \in \mathcal{I}_h, \\ & M(g_i y) \succeq 0, i \in \mathcal{I}_g, \end{aligned} \tag{4.11}$$

where the index set $\mathcal{I}_h$ corresponds to the equality constraints in FP and $\mathcal{I}_g$ to the inequalities. For a general introduction to semi-definite programs, see e. g. [23, 214].

From the necessary conditions on moment sequences (see Lemma 3), it follows that SDP_∞ provides a lower bound to the original problem, i. e. $x^*_{\mathrm{SDP}_\infty} \leq x^*_{\mathrm{LP}_\infty} = x^*_{\mathrm{POP}}$. In many cases even $x^*_{\mathrm{SDP}_\infty} = x^*_{\mathrm{POP}}$ holds.

It is yet not possible to solve SDP_∞ as it involves infinite-dimensional matrices. To obtain a finite semi-definite program we consider instead truncated moment matrices. This leads to Lasserre's hierarchy SDP_d of lower bounds for x^*_{POP} given as

$$\begin{aligned} \inf & \sum_\alpha p_\alpha y_\alpha \\ \text{subject to} & y_0 = 1, \\ & M_d(y) \succeq 0, \\ & M_{d-d_h}(h_j y) = 0, j \in \mathcal{I}_h, \\ & M_{d-d_g}(g_i y) \succeq 0, i \in \mathcal{I}_g. \end{aligned} \tag{4.12}$$

The optimal values follow in this hierarchy obviously the inequality $x^*_{\mathrm{SDP}_d} \leq x^*_{\mathrm{SDP}_\infty} \leq$

$x^*_{\mathrm{LP}_\infty} = x^*_{\mathrm{POP}}$.

Theorem 2 (Semi-definite Relaxation [108]). *Let $\mathcal{K}$ be as in (4.8) and for at least one polynomial $f(x)$ defining $\mathcal{K}$ it holds that there exist polynomials $q_j(x)$ such that $f(x) \in \left\{q_0(x) + \sum_{j=1}^m q_j(x) f(x)\right\}$ and the level set $\{x \in \mathbb{R}^{n_x} : f(x) \geq 0\}$ is compact, then*

$$\lim_{d\to\infty} x^*_{SDP_d} \uparrow x^*_{POP},$$

i. e. the optimal value of SDP_∞ is equal to the optimum of POP. ∎

Proof. See e. g. Theorem [108]. □

Under further conditions on the set $\mathcal{K}$ it can be shown that there is even finite convergence of $x^*_{\mathrm{SDP}_d}$ to x^*_{POP}, see e.g. [108] for more details.

We can state the *Lagrangean dual* (dualSDP$_d$) associated with SDP$_d$ for the case that $p(x)$ is set to zero as

$$\begin{aligned} &\sup\ \lambda_0 \\ &\text{subject to } \langle \Lambda_0, B_\alpha\rangle + \sum_{j=1}^{n_g} \langle \Lambda_j, C_{j,\alpha}\rangle + \ldots \\ &\qquad + \lambda_0 e_1 e_1^T + \sum_{j=1}^{n_h} \lambda_j h_{j,\alpha} = 0, \\ &\qquad \lambda_j \geq 0, i \in \mathcal{I}_h \cup \{0\}, \\ &\qquad \Lambda_j \succeq 0, i \in \mathcal{I}_g \cup \{0\}. \end{aligned} \tag{4.13}$$

where B_α and $C_{j,\alpha}$ are appropriate real symmetric matrices such that $M_d(y) = \sum_\alpha B_\alpha y_\alpha$ and $M_{d-d_g}(g_i y) = \sum_\alpha C_{j,\alpha} y_\alpha$. The Lagrangean dual variable λ_0 corresponds to $y_0 = 1$, Λ_0 (resp. Λ_j) to the semi-definite constraints $M_d(y) \succeq 0$ (resp. $M_{d-d_g} \succeq 0$) and λ_j to the equality constraints.

Considering dualSDP$_d$ we can state the following Lemma providing a conclusive proof with respect to model consistency (see also Definition 7-10).

Lemma 1 (Model Invalidation (Sufficient Condition)). *Given a (semi-)algebraic set of form (4.1). If dualSDP$_d$ is unbounded for some $d \geq 1$ then model (2.2) is inconsistent with the semi-quantitative/qualitative measurement data $\mathcal{M}$, i. e. the solution space $\mathcal{S}^*$ of MIFP is empty.* ∎

Proof. The proof follows from the weak duality property of the Lagrangean dual, see e. g. [20, Corollary 4] □

Lemma 1 provides also a possibility to outer-approximate the solution space of MIFP. Namely, if we partition the constraint set $\mathcal{K}$ we can derive a (possibly non-convex) outer-approximation $\widehat{\mathcal{S}}$ of the solution space of MIFP by excluding those partitions $\mathcal{Q}$ for which the dualSDP$_d$ is unbounded, i. e.

$$\widehat{\mathcal{S}} = \mathcal{K} \setminus \bigcup_{\mathcal{Q}\subset\mathcal{K}:\mathrm{dualSDP}_d\to\infty} \mathcal{Q}. \tag{4.14}$$

To derive $\widehat{\mathcal{S}}$ a simple bisectioning procedure can be used as sketched in the following algorithm.

Algorithm 1(Bisectioning Algorithm [164])

`Outer-approximate(`$\mathcal{Q}, \epsilon$`)`

① **if** $dualSDP_d \to \infty$ **return** $\mathcal{Q} = \emptyset$

② **if** $\mu(\mathcal{Q}) \leq \epsilon$ **return** $\mathcal{Q}$

③ Choose $\mathcal{Q}_1, \mathcal{Q}_2$ such that $\mathcal{Q}_1 \cup \mathcal{Q}_2 = \mathcal{Q}$.

- $\mathcal{Q}_1' =$ `Outer-approximate(`$\mathcal{Q}_1, \epsilon$`)`
- $\mathcal{Q}_2' =$ `Outer-approximate(`$\mathcal{Q}_2, \epsilon$`)`

④ **return** $\mathcal{Q}_1' \cup \mathcal{Q}_2'$

A computationally less expensive alternative is to approximate every variable in $\mathcal{K}$ independently as proposed in [25], which results then in an outer-bounding box of the solution space. This is done by defining a nominal direction of every variable of interest (in the simplest case the maximum and minimum of every variable in FP) and minimizing a linear objective function, cf. Algorithm 2 and [204] for a more detailed description. Both algorithms are illustrated in Section 4.4 and in Section 4.5. Note that independently a similar method was proposed in [35, 36].

Algorithm 2(Outer-bounding Algorithm [25])

`Outer-bound(`$\mathcal{Q}, \epsilon$`)`

① **while** $i \in \mathcal{I}_{\mathcal{S}}$

② **set** $p(x) = x_i$

③ **solve** SDP_d

④ update bounds on x_i

⑤ **set** $p(x) = -x_i$

⑥ **solve** SDP_d

⑦ update bounds on x_i

⑧ **set** i = i+1

⑨ **return** $\mathcal{Q}'_1 \cup \mathcal{Q}'_2$

In the following, we introduce the second relaxation procedure employed in this work. In contrast to the moment relaxation approach, the nonlinear feasibility problem is first reformulated into a quadratic problem with a rank constraint. By relaxing this constraint one obtains again a semi-definite program that retains the guarantees of Lemma 1 and can be employed in Algorithm 1 and Algorithm 2. The main reason for presenting this second relaxation approach is to introduce the flexibility of trading some theoretical accuracy with computional efficiency.

4.3.2 Shor's Relaxation

In our previous works, we derived a result similar to Lemma 1 based on a different relaxation strategy, see [25, 27, 104, 151, 184] for details. It should be noted that in the general case this different strategy is complementary to the moment relaxation method, however, without the strict convergence results seen in the previous section, see e.g. [114] for a detailed discussion of the properties of both relaxations.

As at several points of this work the Matlab toolbox ADMIT [204] is employed, the main ideas of this relaxation is sketched in the following. The basic idea here is to express every equation of the vector functions $f_D(x(k+1), x(k), u(k), p) = 0$ and $h_D(y(k), x(k), u(k), p) = 0$ first in a quadratic form $\xi_{\min}^T Q_k^j \xi_{\min} = 0, j \in \{1, \ldots, n_x + n_y\}$ using a minimal monomial basis $\xi_{\min}$ and symmetric matrices $Q_k^j \in \mathbb{R}^{n_\xi \times n_\xi}$. Note that these monomials correspond to the set of monomials $\Xi_{\min}$ appearing in the system equations and constraints in MIFP. For $\Xi_{\min}$ it has to hold that $1 \in \Xi_{\min}$, and for all $q \in \Xi_{\min}$ of degree higher or equal to 2, there exist $q_1, q_2 \in \Xi_{\min}$ with lower degree such that $q = q_1 q_2$. For these n_D higher degree monomials present in ξ_{min}, n_D additional equality constraints in quadratic form have to be introduced.

The sets $\mathcal{P}$, $\mathcal{X}$ and the uncertain measurement data contained in $\mathcal{M}$ are equivalently rewritten as $B\xi_{\min}\xi_{\min}^T e_1 \geq 0$. To get to the standard representation of a semi-definite program, it is assumed that the non-negativity constraint of all elements in $\xi_{\min}$ is included in the matrix B.

Thus, we can relax MIFP into the semi-definite program SDP_{Shor} by introducing $X = \xi_{\min}\xi_{\min}^T$ and replacing the conditions $rank(X) = 1$ and $tr(X) \geq 1$ with the

weaker constraint $X \succeq 0$ (cf. [151]), resulting in the relaxed formulation SDP_{Shor}

$$\begin{aligned} \min\ & tr(CX) \\ \text{subject to}\ & tr(Q_k^j X) = 0, \quad k \in \mathcal{Z}, j \in \mathcal{I}, \\ & tr(e_1 e_1^T X) = 1, \\ & BXe_1 \geq 0, \\ & X \succeq 0, X \in \mathbb{X}. \end{aligned} \tag{4.15}$$

with $\mathbb{X} \subset \mathbb{R}^{n_\xi \times n_\xi}$ denoting the space of the matrix variable X, where entries corresponding to products of discrete variables of $\xi_{\min}$ are treated as discrete variables. For a more detailed discussion of the involved relaxation and reformulation steps, we refer to [151]. As shown in [114], SDP_{Shor} provides a lower bound on SDP_d, i. e. $x^*_{\text{SDP}_{Shor}} \leq x^*_{\text{SDP}_d}$.

It should be noted, although, SDP_{Shor} does not have the same convergence guarantees as SDP_d. However, if the Lagrangean dual of SDP_{Shor} is unbounded, similar to Lemma 1, it can be proven that a model is inconsistent. Therefore, this relaxation approach can be directly employed in Algorithm 1 and Algorithm 2. To demonstrate this, we provide in Section 4.4 an example.

In general, both introduced relaxation approaches can be employed to address the set-based estimation problem, i. e. the approximation of the solution space of FP. However, both approach can be computationally expansive. A natural conclusion is to relax the semi-definite programs even further into linear programs as discussed next.

4.3.3 Remarks on Linear Relaxations

Recent reports suggest that there will be methods for deriving valid cuts for mixed-integer semi-definite programs (e. g. [100]), the maximal number of variables is at the moment restricted to approximately 60. This number has to be reduced further with increasing number of binary variables due to the increasing combinatorial complexity. One possible approach to reduce the computational burden is the use of linear relaxations [109]. An advantage is clearly the availability of efficient solvers for linear programs that are able to handle large problem sizes up to several thousand variables.

As stated in [109] there are some difficulties involved when linear relaxations are used. For instance, the Hausdorff moment conditions are known to be numerically unstable due to the involved binomial coefficients. Furthermore, [109] showed that it is not guaranteed that the optimum of the linear relaxation coincides with x^*_{POP} (even in the limit). Note that combining linear relaxations with a branch-and-bound procedure can overcome some of the aforementioned problems, see e. g. [13] and references therein.

Still we believe that linear relaxations complement the semi-definite method from an algorithmic point of view. When we are interested in proving inconsistency of a model it is done by a sufficient certificate, namely the infeasibility of our semi-definite

program. Therefore, the limitation that the true optimal solution might not be found loses its importance as every feasible solution already suffices. Furthermore, the case when numerical problems occur can be treated as the case that no definite certificate for model invalidity was found. In the following, we discuss a simple linear relaxation for the feasibility problem as used in [204] and refer the interested reader to [109] for a strategy better suited for finding an optimal solution.

A possibility is to substitute the constraint $M_d(y) \succeq 0$ with $M_d(y) \geq 0$ (resp. $X \succeq 0$ with $X \geq 0$). This still leads to a valid constraint, as the constraints $g_i(x) \geq 0$ (resp. $BXe_1 \geq 0$) already guarantee that the variables are positive.

As we are not concerned with finding an optimal solution here, we will discuss the linear program formulation only in terms of SDP_{Shor}. We provide thereafter one simple, but highly nonconvex, example that illustrates the usefulness of this relaxation.

Note that this substitution will lead to an increase of additional solutions, which can be reduced again e. g. by considering further strengthening constraints [5, 182], or by adding constraints for diagonal dominance [215]. Indeed in practice it showed that the constraints mainly responsible for the estimation precision are redundant constraints in the linear program.

Further additional constraints to reduce the relaxation error due to multiplication of integer variables are presented in [173]. A less conservative linear relaxation based on [182] can be found in [109].

Example 1. *The example we are considering is the so-called Folium surface, see e. g. [77]. This surface is the area inside the curve* $r = \sin(2\phi)$ *and can be described as* $\mathcal{K} = \{x \in \mathbb{R}^2 : g(x) = -((2x_1 - 1)^2 + (2x_2 - 1)^2)^3 + 4(2x_1 - 1)^2(2x_2 - 1)^2 \geq 0\}$. *We compare here the results of Lasserre's hierarchy (*SDP_d*), Shor's relaxation and the proposed linear relaxation. For this reason, we compute the area of the Folium surface employing the proposed algorithms for the different relaxations. Furthermore, we compare these results with the results obtained for a close form approximation reported in [77]. The exact area of the Folium surface is* $\pi/6 \approx 0.523$ *(computed employing polar coordinates and Mathematica). The initial bounds on the variables* x_1 *and* x_2 *are chosen to be the interval* $[0, 1]$. *For the bisectioning procedure* ϵ *is chosen to be* $1/6$.

The result of the bisectioning procedure (Algorithm 1), the outer-bounding procedure (Algorithm 2) for $d = 10$, *the closed form solution of [77] as well as the true feasible region are shown in Figure 4.1. Table 4.1 reports the derived areas. The implementation of the example was done with the Matlab toolbox ADMIT [204] and the Matlab toolbox YALMIP [121, 122]. The solution was obtained with the semi-definite programming solver SEDUMI [208] and the mixed integer linear solver CPLEX. Computation times are stated for a desktop computer (with a Intel®Core i5 and 8GB RAM).*

The linear relaxation in combination with Algorithm 1 as well as the closed form solutions approximate the area quite well. Note that by choosing a smaller ϵ *for Algo-*

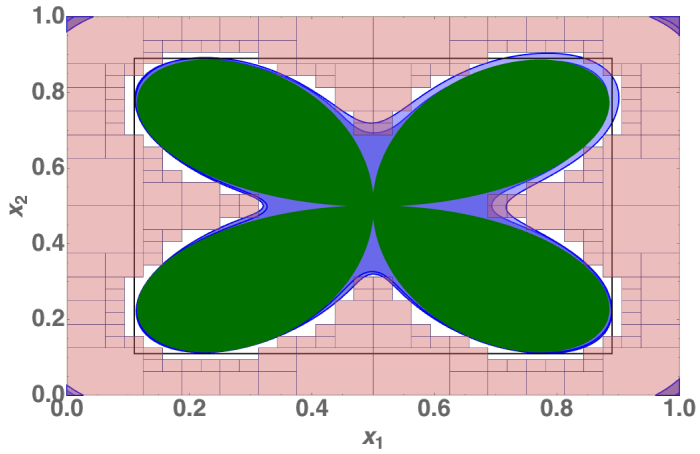

Figure 4.1: Approximation of the two dimensional Folium surface. Blue area depicts the true feasible solution, solid green line depicts the delimiter $r = sin(2\phi)$. Blue regions correspond to closed form solutions approximating the Folium surface (for $d = 4$ und $d = 8$) employing [77]. Red boxes are the excluded regions by means of LP_d (Algorithm 1). Black box corresponds to the result of the outer-bounding procedure (Algoirthm 2) with SDP_{10}.

Table 4.1: Approximate area of the Folium surface

Procedure	Area	Computation Time
Algorithm 1	0.540	$\sim$ 12s
Algorithm 2	0.608	$\sim$ 1s
Closed Form [77] ($d = 4$)	0.538	$\sim$ 5s
Closed Form [77] ($d = 8$)	0.525	$\sim$ 25s

rithm 1 a better approximation can be obtained. Note also that numerical errors for the closed form solution do not allow a better approximation, although in theory for $d \to \infty$ the exact area is derived. To improve the result of the closed form solutions one would have to choose a circular initial area enclosing the Folium surface [77]. ■

The example illustrates the quality of the linear relaxation, when combined with the bisectioning procedure Algorithm 1. However, as explained in the beginning of this section, without adding additional constraints, as they are introduced by the bisectioning, the linear relaxation does in general lead to a weak relaxation. This is in particular true if Algorithm 2 is considered as no guarantee can be given that the bounds on the variables can be derived exactly.

In the following, we present two examples to illustrate the developed set-based estimation approach. Both examples have their origin in systems biology and were

implemented with help of the toolbox ADMIT [204]. The first example only considers quantitative constraints, while the second example also employs constraints based on semi-quantitative data.

4.4 Example: Model Consistency

In this example, we demonstrate how well set-based estimation is situated to handle uncertainties for testing model consistency. This is done with help of Algorithm 1 implemented in the toolbox ADMIT [204]. This section is based on [164].

4.4.1 Model Description

We consider an enzyme-catalyzed reaction, in which an enzyme (E) and a substrate (S) join into an enzyme-substrate complex (C) to form a final product (P). Let the hypotheses proposed for this process be the two models formulated by Henri in 1902, respectively known as the Michaelis-Menten (MM) and the Henri (H) mechanism of enzyme-catalyzed reactions:

$$\text{(MM)}: \qquad E + S \underset{p_2}{\overset{p_1}{\rightleftharpoons}} C \overset{p_3}{\rightarrow} E + P \tag{4.16}$$

$$\text{(H)}: \qquad C \underset{\tilde{p}_2}{\overset{\tilde{p}_1}{\rightleftharpoons}} E + S \overset{\tilde{p}_3}{\rightarrow} E + P, \tag{4.17}$$

where $p_i, i \in \{1,2,3\}$ and $\tilde{p}_i, i \in \{1,2,3\}$ are the rate constants. The relevance of these two models is discussed in [177], in which it is also proved that both models are analytically indistinguishable under steady state conditions, and can only be distinguished in the transient phase.

Both reaction mechanisms are modeled according to the law of mass action. Exploiting two conservation relations fulfilled by both mechanisms, the models can be simplified into second order systems depending only on the concentration of S and C. Considering a simple first order Euler discretization scheme, and fixing the total enzyme concentration $E+C$ to a constant value 1, the difference equations corresponding to the MM mechanism are given by

$$\begin{aligned} s_{k+1} &= s_k + h[(c_k - 1)p_1 s_k + p_2 c_k] \\ c_{k+1} &= c_k + h[(1 - c_k)p_1 s_k - (p_2 + p_3)c_k], \end{aligned}$$

where h is the time-step of the discretization, while for the Henri mechanism we obtain

$$\begin{aligned} s_{k+1} &= s_k + h[(c_k - 1)(\tilde{p}_1 + \tilde{p}_3)s_k + \tilde{p}_2 c_k] \\ c_{k+1} &= c_k + h[(1 - c_k)\tilde{p}_1 s_k - \tilde{p}_2 c_k]. \end{aligned}$$

To show that our approach allows to prove model invalidity, we assume the Henri

Table 4.2: Model invalidation results for the Michaelis-Menten mechanism

Initial Conditions substrate (s_0)	complex (c_0)	**Maximum Error** σ [%]
0.999	0.001	±14.0%
0.990	0.010	±13.0%
0.900	0.100	± 8.5%
0.800	0.100	± 8.0%
0.800	0.200	± 5.0%
0.700	0.300	± 2.5%
0.600	0.400	± 0.5%

mechanism as reference, generate measurements by sampling a simulation during the transient phase, and use the resulting data for model invalidation against the Michaelis-Menten model.

The discrete-time model for the Henri mechanism has been simulated with time-step $h = 0.1$ seconds and parameters $\tilde{p}_1 = \tilde{p}_2 = \tilde{p}_3 = 1$ for several initial conditions $x_0 = (s_0, c_0)$, deriving for each a corresponding sequence of states $x_k = (s_k, c_k)$, for $k = 0, \ldots, 20$. Given a state sequence (x_k) and a measurement error σ, we denote by $\mathcal{Y}^\sigma = \{\mathcal{Y}_k^\sigma \ : \ k = 0, \ldots, 20\}$ the corresponding uncertain measurement sequence, with measurement sets $\mathcal{Y}_k^\sigma = \{x \in \mathbb{R}^2 : |x - x_k| \leq (1 + \sigma)x_k\}$. To test if the sequence $\mathcal{Y}^\sigma$ allows to invalidate the Michaelis-Menten mechanism, we apply Algorithm 1 with precision threshold $\varepsilon = 5\%$, using as bounds for the unknown parameters the interval set $\mathcal{P} = \{p \in \mathbb{R}^3 : p_i \in [0.\bar{3}, 3]\}$. If the resulting parameter set is empty, the Michaelis-Menten mechanism is invalidated.

4.4.2 Results and Discussion

In Table 4.2 we report, for seven different initial conditions, the highest measurement error σ for which our approach allows to invalidate the Michaelis-Menten mechanism. The measurement error decreases as the initial conditions approach the steady state (recall that in the steady state the two systems are indistinguishable [177]). Comparing these results with the practical measurement errors that can be obtained in enzymological assays (see e.g. [115, 131, 163]), invalidation can be achieved when the system is sufficiently excited.

This example illustrates the ability to conclude model consistency under the presence of measurement uncertainty. However, here only quantitative data is considered, to elucidate the proposed set-based approach for semi-quantitative data consider the following example.

4.5 Example: Biological Adaptation Model

In this section, we give a short example how the methods developed in this chapter can be used to invalidate a model based on qualitative information. We demonstrate this with a core adaptation model and a qualitative description of the adaptation process. The results we present in the following are based on [170].

4.5.1 Background: Adaptation

Adaptation is an attribute of a system that allows it to react to changes in its environment. In evolutionary biology this attribute typically corresponds to a natural selection process, while in our work we understand adaptation as a reaction with respect to a change in the input. Perfect adaptation denotes hereby the property of a biological system to return exactly to the pre-stimulus level of the observed variable $y(t)$ for large times after an initial dynamic response, i. e. $\Delta y(t^*) = |y(t_{\text{stimulus}} - y(t^*))| > 0$ for some $t^* > t_{\text{stimulus}}$, and $\Delta y(t) \to 0$ as $t \to \infty$. We say an adaptation is *partial* or *near-perfect* if the output does not converge towards the prestimulus level. The structure and parameterizations of network models that allow adaptation are well studied in literature, see e. g. [16, 21, 50, 129, 193, 222].

For instance in [193] it was shown that a nonlinear system that adapts to a class of inputs, necessarily contains a subsystem that is able to generate inputs of the same class based on the internal model principle. For linear systems adaptation to a constant step function can be achieved with an integral feedback [222]. In models of biochemical networks, an integrator is often obtained by a reaction operating with zero-order kinetics [16, 50]

In [129] two core structures to model adaptation were presented, namely a negative feedback loop with a buffering node and an incoherent feedforward loop with a proportioner node. Using the former, we illustrate in this section that the proposed framework can be used, first, to express adaptation as a mixture of quantitative, semi-quantitative and qualitative data by linear constraints and binary variables following [170]; second, that the data can be used to estimate the parameters of a model. We analyze here only one valid core model, for a more detailed discussion and additional model hypotheses see [170]. The analysis was done using ADMIT [204]. This toolbox can be used to construct the considered feasibility problem, to carry out the necessary mathematical reformulations and relaxations of the constraints, and to perform parameter estimation, state estimation, and model invalidation as described in Section 4.2.

4.5.2 Model Description

The considered reaction scheme comprises of three species that can be either active or inactive and the connection of the species. The species interact with each other

reversibly and enzymatically regulated. By introducing conservation relations that hold for the total amount of three species in both activation states, we obtain the following simplified Euler model.

$$\begin{pmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \end{pmatrix} = \begin{pmatrix} x_1(k) + D(\frac{k_f\, x_2(k)\,(1-x_1(k))}{K_f + (1-x_1(k))} - \frac{k_{iuA}\, u\, x_1(k)}{K_{iuA} + x_1(k)}) \\ x_2(k) + D(\frac{k_{aEB}\, E_B\,(1-x_2(k))}{K_{aEB} + (1-x_2(k))} - \frac{k_{iAB}\, x_1(k)\, x_2(k)}{K_{iAB} + x_2(k)}) \\ x_3(k) + D(\frac{k_{aAC}\, x_1(k)\,(1-x_3(k))}{K_{aAC} + (1-x_3(k))} - \frac{k_{iEC}\, E_C\, x_3(k)}{K_{iEC} + x_3(k)}) \end{pmatrix} \tag{4.18a}$$

$$y(k) = (x_1(k), x_2(k), x_3(k)). \tag{4.18b}$$

Note that here the rather general form of Michaelis-Menten kinetics was chosen for the reactions as for appropriate choices of the parameters, the Michaelis-Menten kinetics approximate either zero- or first-order kinetics. Note further as the kinetics involve rational terms, all equations were multiplied by their respective denominators and the resulting implicit formulation is used for the subsequent analysis.

We assume that the states $x_1(k)$ and $x_2(k)$ are measured and are given as **ubb** variables. Furthermore, we employ a similar qualitative description of adaptation in terms of $x_3(k)$ as in [170] as summarized next.

- The states are in steady-state before the stimulus is removed and $x_3(k)$ is in the initial interval $[0.4, 0.5]$.

- After the stimulus is removed the maximum value of $x_3(k)$ is larger than 115% and is reached within 3 time-steps.

- After the maximum has been reached, $x_3(k)$ returns almost to the initial interval in 5 time-steps, i. e. $x_3(k) \in [0.4, 5.2]$.

- To avoid that $x_3(k)$ shows a strong inverse response behavior additionally $x_3(k)$ is always larger than 0.4.

Employing the formalism introduced in Chapter 3, the qualitative information is first formulated in terms of conjunctions and disjunctions. The resulting formulation is given in Table B.1. The considered intervals of the parameters are given in Table B.2. For generating the quantitative data on $x_1(k)$ and $x_2(k)$, we computed a nominal trajectory for a fixed parameterization and added an absolute error of ± 0.01 to the states, the nominal parameterization is included in Table B.2. Note that the setting described above is based on an example implemented in ADMIT.

4.5.3 Results and Discussion

The obtained simulation results using ADMIT are depicted in Figure 4.2. As can clearly be seen from the calculated bounds on $x_3(k)$ the model is able to reproduce the desired qualitative behavior.

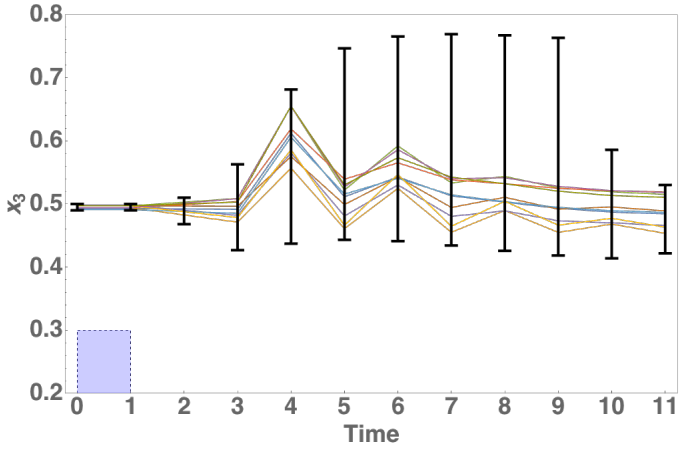

Figure 4.2: Result as produced by ADMIT for the given setup. Intervals depict the estimated ranges of the state $x_3(k)$. Blue shaded area corresponds to the stimulus (not to scale). Colored lines correspond to discrete-time trajectories of (4.18) obtained by sampling the determined intervals of initial conditions and parameters.

The example, therefore, shows that a combination of quantitative and qualitative data can be used to estimate the variables of interest in a set-based setting. What should be kept in mind here is, however, that if measurements instead from a nominal discrete-time model but a continuous-time model are taken the analysis becomes a much harder problem. The main reason is that discrete-time approximations as used here have a tendency to smooth some transient behaviors as seen in the next chapter.

For the considered setting the computation time was only around 10min to compute the results shown in Figure 4.2, but increases rapidly by allowing more uncertainties in the variables. This is, however, to some extend intrinsic to all available set-based methods. Still, employing techniques to reduce the computational load are very important if the proposed methods are applied to real world examples. Suitable extensions were proposed e. g. in [153, 167, 175].

4.6 Summary

In this chapter, the proposed set-based estimation approach for discrete-time systems was introduced. This approach is based on formulating model consistency in terms of a nonlinear feasibility problem incorporating quantitative and semi-quantitative data. In

general, this feasibility problem is nonconvex and, hence, cannot be addressed directly. To overcome this difficulty, we considered two relaxation approaches that transfer the feasibility problem into a semi-definite or linear program. These approaches are based on techniques developed for the solution of polynomial optimization problems. One important property of these relaxation approaches is that they provide a guaranteed model invalidation result. This means if a model is inconsistent with the available data and safety requirements, then a certificate of inconsistency can be derived.

Based on these certificates, we proposed several algorithms to estimate bounds on the variables of interest, thus solving the set-based estimation problem. A major concern hereby was the computational burden. For this reason, both the developed algorithms and the choice of employed relaxation strategies allow a trade-off between computational load and accuracy. Although, this trade-off is possible, the guaranteed invalidation is always retained.

So far only discrete-time models were considered. In practice, determining whether a continuous-time model is able to reproduce the data from a process or not is typically more involved. For this reason, we extend in the following the presented set-based method to continuous-time systems with a focus on the aforementioned flexibility to provide a trade-off between accuracy and applicability.

5 Continuous-Time Systems

In this chapter, we extend the discrete-time estimation procedure of the previous chapter to continuous-time systems. To do so, we identify system properties of discrete-time systems (resp. approximations) that have to be fulfilled such that continuous-time systems can be addressed. In particular, these properties allow to conclude from model invalidity of the discrete-time system to the model invalidity of the continuous-time system. We propose two possibilities to derive suitable discrete-time approximations. The first method is based on the Picard-Lindelöf theorem and the second method is based on augmenting the discrete-time system by an error term. While the first is in general theoretically more rigorous the second one offers a trade-off between accuracy of the model invalidation result with the computational burden.

This chapter is structured as follows. We briefly introduce the difficulties that arise when continuous-time systems are considered in a set-based setting. Then we present several properties that a discrete-time system has to fulfill such that it can be employed to estimate the continuous-time system. Afterwards we present two methods to derive a suitable discrete-time approximation and illustrate with a real world example before concluding this chapter.

5.1 Introduction: Set-based Estimation

As shown in Chapter 4 it is possible to derive an outer-approximation of the feasible set of a discrete-time system. Here, feasible set corresponds to all values of the involved variables, e. g. states and parameters, that lead to a model behavior consistent with the available data and safety requirements. However, for continuous-time systems deriving the feasible set is more involved and only few suitable approaches are available in literature. One possibility is to directly encode the continuous-time dynamics in LP_∞ with the help of *occupation measures*, see e. g. [76, 111] and in particular [203] for parameter estimation. If the main concern is to prove model invalidity one can also employ the concept of *barrier certificates*, see e. g. [3, 156, 157]. These certificates allow to derive a function that separates possible system trajectories with the measurements and, therefore, proving invalidity of the model and can also be applied for parameter estimation [221]. Other approaches rely on additional assumptions. For instance, in [71, 104] the continuous-time system is only considered in steady state. In [52], it is assumed that the derivatives of the states are also available as unknown-but-bounded measurements. The problem then reduces to a (multi-stage) steady-state investigation similar to [104].

The main difficulty in applying the results of Chapter 4 to continuous-time systems derives from the fact that by discretizing typically an approximation error is introduced. *Discretization errors* as they appear in solving initial value problems are a well studied topic in numerical mathematics, see e. g. [46, 198] for an overview. In the context of systems theory several results have been proposed in e. g. [68, 206, 207]. However, typically only initial value problems with fixed parameterizations are considered and a direct application to set-based methods is in general difficult.

One approach dealing with this problem is based on interval arithmetics for deriving validated solutions to ordinary differential equations [38, 97, 116, 123, 139, 140, 161, 216, 217]. One common feature of these works is that they try to minimize the discretization error by employing multi-step discretization schemes or similar ideas. Another approach is related to the classical works of differential inclusions [10, 12, 54], and in particular the discretization of differential inclusions [48, 63, 171]. For instance, in [7] it was shown that the discrete-time approximation of a differential inclusion based on Euler discretization only approximates the differential inclusion for vanishing time-steps.

Before describing the proposed methods, we illustrate the problem of guaranteed model invalidation of a continuous-time system with a simple example.

Example 2. *Consider the linear scalar system* $\dot{x}(t) = -px(t)$ *and the corresponding Euler discretization* $x(k+1) = (1 - p_d D)x(k)$ *with* $D > 0$ *denoting the step-size and* $p, p_d \in \mathbb{R}$ *the system parameter of the continuous-time and discrete-time system, respectively. We compare both systems on the time interval* $\mathcal{T} = [0, 1]$. *We assume further that we have measurements of the continuous-time system at particular time-instances with a constant sampling rate* D^*. *This setup is illustrated in Figure 5.1A.*

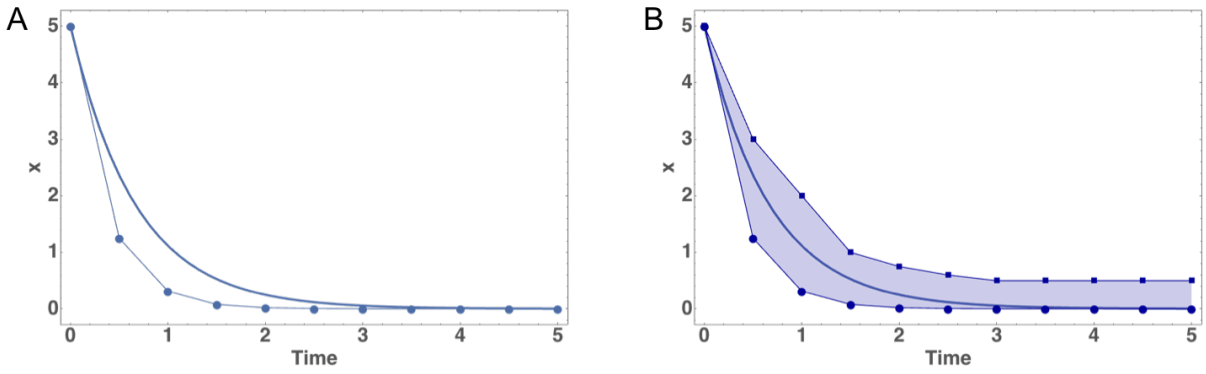

Figure 5.1: A: Euler approximation of an exponential decay function $\dot{x} = -1.5x$ with initial condition $x_0 = 5$ and a fixed step-size $D = 0.5$ on the time interval $[0, 5]$. Measurements of the continuous-time system are depicted as the full blue line, while the Euler approximation is depicted as the dotted line. B: Shaded area depicts the reachable set of an embedding Euler approximation.

For the Euler discretization to be consistent with the measurements, one of two conditions has to be fulfilled. Namely, the step-size D has to be chosen sufficiently small or the parameterization is chosen w.r.t. an exact discretization. In this simple example, we can derive the exact discretization as $x^*(k+1) = e^{-pD^*}x^*(k)$. For an exactly known initial condition $x_0 \in \mathbb{R}$ the parameter p_d of the discrete-time system has to have the value $p_d = (1 - e^{-pD^*})/D$, otherwise the discrete-time system is going to be inconsistent with the data due to the made discretization error.

We, therefore, propose here methods which allow the derivation of a discrete-time system that embeds the trajectories of an uncertain continuous-time system such that model consistency of the continuous-time system can be checked by the embedding system, as illustrated by Figure 5.1B. ■

As seen in the previous example the consistent parameters of continuous-time models and discrete-time approximations are not necessarily equal. This derives from the fact that by discretizing the continuous-time dynamics an error is introduced. To provide a method for which consistency of a continuous-time model can be checked by an approximation, the discretization error has to be considered. For this purpose, we categorize the relationship between continuous-time and discrete-time systems as follows. We denote the continuous-time consistent parameters with $\mathcal{P}^*_{CT}$ and the consistent parameters of the discrete-time system as $\mathcal{P}^*_{DT}$. We have to distinguish three different situations as illustrated by the Venn diagrams in Figure 5.2.

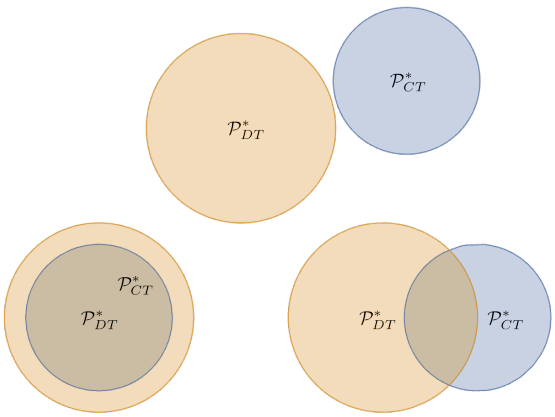

Figure 5.2: Possible relationships of the consistent parameters of continuous-time models and discrete-time approximations of this model. Blue areas represent consistent parameters of the continuous-time model $\mathcal{P}^*_{CT}$, orange areas consistent parameters of the discrete-time model $\mathcal{P}^*_{DT}$. Ideally, $\mathcal{P}^*_{CT} \subseteq \mathcal{P}^*_{DT}$ holds to be able to relate results from the discrete-time model to the continuous-time model. In case, $\mathcal{P}^*_{DT} \cap \mathcal{P}^*_{CT} = \emptyset$ in general results are not related.

If no relationship between the consistent parameters exists, i. e. $\mathcal{P}^*_{CT} \cap \mathcal{P}^*_{DT} = \emptyset$, in general consistency cannot be checked by an approximation, unless a functional dependency between the parameters can be derived. This is the case for exact discretizations

as seen in Section 5.3. The other two cases have to be considered separately. The ideal relationship $\mathcal{P}^*_{CT} \subseteq \mathcal{P}^*_{DT}$, where the parameters of the continuous-time system are bounded by the parameters of the discrete-time system, is considered in Section 5.4. This property is achieved by choosing a time-step size for the approximation such that the reachable set of the continuous-time system are embedded in the reachable set of the discrete-time system. To do so, we employ the Picard-Lindelöf theorem for the existence and uniqueness of differential equations. This leads, however, in most cases to a high computational burden as the step-size will typically be small for uncertain systems. To mitigate this problem, we propose further a method for which $\mathcal{P}^*_{CT} \cap \mathcal{P}^*_{DT} \neq \emptyset$ can be guaranteed (Section 5.5.3). This is done, by constructing a discrete-time system that is augmented with a time dependent error term. This error is then estimated with help of simulation data from the continuous-time model. This leads to a computational advantageous method, that allows us to balance the computational burden with the achievable accuracy depending on the specific process under consideration. That this weaker relationship still provides a useful tool in the analysis of a model is illustrated in Section 5.6 for a real world example. The next section formalizes the considered relationships between consistent parameters and introduces the concept of embedding systems.

5.2 Relationship of Continuous-time and Discrete-time Models

To be able to employ Lemma 1 for continuous-time systems we have to create a relationship between consistent parameters of discrete-time and continuous-time model.

We define such a relationship in terms of consistent parameters as:

Definition 11 (Weak CT/DT-Relationship)**.** *A discrete-time model* (2.2) *is said to be weakly related to a continuous-time model* (2.1)*, if there exists* $D^* > 0$ *and a set of consistent parameters* $\mathcal{P}^*_{DT}$ *of the discrete-time model such that*

$$\mathcal{P}^*_{DT} \cap \mathcal{P}^*_{CT} \neq 0 \tag{5.1}$$

holds for all step-sizes $D \in [0, D^*)$*, where* $\mathcal{P}^*_{CT}$ *denotes the consistent parameters of the continuous-time model.* ∎

As soon as at least one consistent parameterization of the discrete-time model coincides with a consistent parameter of the continuous-time model an application of Lemma 1 provides a necessary condition for model validity. However, this definition is insufficient for guaranteed parameter estimation since an outer-approximation of the consistent parameters of the continuous-time model is not possible. Here a stronger relationship is needed as defined next.

Definition 12 (Strong CT/DT-Relationship)**.** *A discrete-time model* (2.2) *is said to be strongly related to a continuous-time model* (2.1)*, if there exists* $T^* > 0$ *and a set of consistent parameters* $\mathcal{P}^*_{DT}$ *of the discrete-time model such that*

$$\mathcal{P}^*_{CT} \subseteq \mathcal{P}^*_{DT} \tag{5.2}$$

holds for all $T \in [0, T^*)$*, where* $\mathcal{P}^*_{CT}$ *denotes the consistent parameters of the continuous-time model.* ∎

Remark 4 (Variables of Interest)**.** *In this chapter, we are mainly interested in model consistency and we restrict, therefore, the variables of interest, i. e. the variables to be estimated, to the parameters of the discrete-time (resp. continuous-time) model.*

The main question that arises from these definitions is: how can we ensure the desired relationship of discrete-time approximations and model (2.1) exists. We provide in the following some preliminary results that guarantee this relationship and are then used to derive appropriate methods for parameter estimation/model invalidation in Section 5.4 and Section 5.5.

5.2.1 Embedding Systems

We define the term embedding system for discrete-time systems as follows.

Definition 13 (Embedding System)**.** *A discrete-time model* (2.2) *with trajectories* $\Phi^2_D(x(k)|x_0, u(k), p_1)$ *is said to embed a discrete-time model with trajectories* $\Phi^1_D(x(k)|x_0, u(k), p_2)$*, if there exist two sets of parameters* $\mathcal{P}^*_1, \mathcal{P}^*_2$ *such that*

$$\bigcup_{x_0 \in \mathcal{X}_0, u(k) \in \mathcal{U}, p_1 \in \mathcal{P}^*_1,} \Phi^1_D(x(k)|x_0, u(k), p_1) \subseteq \bigcup_{x_0 \in \mathcal{X}_0, u(k) \in \mathcal{U}, p_2 \in \mathcal{P}^*_2, x(k) \in \mathcal{X}} \Phi^2_D(x(k)|x_0, u(k), p_2) \tag{5.3}$$

holds for all time-indices $k = \{1, \ldots, n_k\}$*.* ∎

We further need to extend the same term towards a continuous-time system embedded by a discrete-time system.

Definition 14 (Embedding DT/CT System)**.** *A discrete-time model* (2.2) *is said to embed a continuous-time model* (2.1)*, if*

$$\bigcup_{x_0 \in \mathcal{X}_0, u(t) \in \mathcal{U}, p_1 \in \mathcal{P}^*_1} \Phi^1(x(t_k)|x_0, u(t), p_1) \subseteq \bigcup_{x_0 \in \mathcal{X}_0, u(k) \in \mathcal{U}, p_2 \in \mathcal{P}^*_2,} \Phi^2_D(x(k)|x_0, u(k), p_2) \tag{5.4}$$

holds for all time-points $t_k \in \mathcal{T} := \{t_1, \ldots, t_k\}$*.* ∎

Note that the previous definitions can be extended by an uncertain input sequence, as we assume a zero-order hold for the input of the continuous-time system, cf. Chapter 2.

Furthermore, from these definitions, we have the following implication.

Proposition 1 (Consistency of Embedding Systems)**.** *Given a continuous-time model* (2.1) *and an discrete-time system* (2.2) *for which* (5.4) *holds. If the continuous-time model is consistent with a collection of measurements* $\mathcal{M}$*, i. e.* $y(t_k) \in \mathcal{M}(t_k), \forall t_k \in \mathcal{T}$*, then the discrete-time system is consistent with the measurements.* ∎

It follows that (5.4) is a necessary condition for proving model consistency of a continuous-time system (2.1) with help of a discrete-time system. This necessary condition can also be formulated as a sufficient criterion for model inconsistency of the continuous-time system.

Proposition 2 (CT Model Invalidation)**.** *Under the same assumptions as Proposition 1. If the embedding discrete-time system is inconsistent with the data (verified by Lemma 1), then the continuous-time model is inconsistent.* ∎

In the next section, we provide a class of discrete-time systems that fulfill the embedding property, namely exact discretizations.

5.3 Exact Discretization

Ideally, a functional relationship of the continuous-time system and its discretization is known, to guarantee the embedding property (5.4), in the sense that trajectories of both systems coincide at the sampling points. This allows in principle the exchange of information from one system to the other, e. g. the parameters. However, obtaining such a functional relationship is in general not possible, unless an exact discretization is known. Examples for which such a discretization exist are linear and bilinear systems, see e. g. [47, 168]. For this work, only exact discretizations that preserve the structure of the continuous-time system are of interest. With structure preserving we refer to discretization schemes that starting from a polynomial system lead to polynomial discrete-time systems of the same degree.

To fix terminology consider the following definition.

Definition 15 (Structure-Preserving Exact Discretization)**.** *A (structure-preserving) discretization is called exact, if the discretization is a morphism w.r.t. the continuous-time system and there exists an isomorphism* $\Gamma : \mathbb{R}^{n_p} \to \mathbb{R}^{n_p}$ *such that*

$$\begin{aligned} \Phi(x(t)|x_0, u(t), p) &= \Phi_D(x(t), u(t), \Gamma(p)), \\ \left(\textit{resp. } \Phi(x(t), u(t), \Gamma^{-1}(p)) = \Phi_D(x(t)|x_0, u(t), p)\right), \end{aligned} \tag{5.5}$$

holds for all $t = kD$*.* ∎

As an example, consider the discrete-to-continuous (D2C) reconstructability of separable bilinear systems. A discrete-time bilinear system has the form

$$x(k+1) = Fx(k) + Gu(k) + \sum_{l=1}^{r} u_l(k) N_l x(k), \tag{5.6}$$

where $F \in \mathbb{R}^{n_x \times n_x}$ is the state matrix, $G \in \mathbb{R}^{n_x \times n_u}$ is the input matrix and $N_l \in \mathbb{R}^{n_u \times n_x}$ is the bilinear weighting matrix for the l-th input, $l = 1, \ldots, n_u$, where the matrices depend on the discretization step size D.

Definition 16 (Separable Bilinear Systems)**.** *A bilinear system* (5.6) *is called pairwise separable (or separable) under a constant input $u_{l,k} \neq 0$, $u_{i,k} = 0, i \in \{1, \ldots, r\} \setminus \{l\}$, if there exists a coordinate transformation with non singular transformation matrices Φ_l that transforms $F + u_{l,k} N_l$ into a diagonal form and F is of full rank.* ∎

For such systems, there exists the relationship according to Definition 15 as summarized in the following result. To simplify presentation, we assume that $l = 1$ as well as the state matrix F, the input matrix G and the weighting matrix N_1 in (5.6) are scalars and we refer to [168] for more general cases.

Proposition 3 (D2C Reconstruction (Bilinear) [168])**.** *Consider a discrete-time scalar bilinear system of form* (5.6)*. A continuous-time system $\dot{x} = \alpha x + \beta u + u\rho x$ can be reconstructed from* (5.6) *by choosing the elements of the system matrices as*

$$\begin{aligned} \alpha &= \frac{1}{D} \ln(F), \quad \beta = \frac{G \ln\left[F + uN_1\right]}{D(F + uN_1 - 1)}, \\ \rho &= \frac{\ln\left[\frac{F+uN_1}{F}\right]}{Du}. \end{aligned} \tag{5.7}$$

∎

Proof. The proof is based on [168]. We rewrite the discrete-time bilinear system as follows.

$$x(k+1) = fx(k) + gu(k) + \eta u(k)x(k) = (f + \eta u(k))x(k) + gu(k). \tag{5.8}$$

Given a pair of scalar linear systems

$$\dot{x} = \alpha x + \beta u \text{ and } x(k+1) = fx(k) + gu(k). \tag{5.9}$$

If f, g and D are known, we have the relationships $\alpha = \frac{\ln(f)}{D}$ and $\beta = \frac{g \ln(f)}{D(f-1)}$ for the linear system. Applying these relationships to a separable bilinear system, given f, g,

η and D, we obtain

$$\begin{aligned}\alpha + \rho u &= \frac{1}{D}\ln(f+\eta u)\\ &= \frac{1}{D}\left[\ln(f)+\ln(f+\eta u)-\ln(f)\right]\\ &= \frac{1}{D}\ln(f)+\frac{1}{T}\left[\ln(f+\eta u)-\ln(f)\right]\\ &= \frac{1}{D}\ln(f)+\frac{\left[\ln(f+\eta u)-\ln(f)\right]}{Du}u.\end{aligned} \tag{5.10}$$

Moreover, substituting f with $f+\eta u$ we obtain β. Hence,

$$\alpha = \frac{\ln(f)}{D},\ \rho = \frac{\left[\ln(f+\eta u)-\ln(f)\right]}{Du},\ \beta = \frac{g\ln(f+\eta u)}{D(f+\eta u-1)},$$

gives us the continuous-time system $\dot{x} = \alpha x + \beta u + \rho u x$, which concludes the proof. □

Therefore, separable bilinear systems meet the first requirement of Definition 15 as the continuous-time and the discrete-time system are both bilinear. Furthermore, (5.7) corresponds to the isomorphism, or more precisely the inverse of the isomorphism in Definition 15.

From Definition (5.5) and the previous example it follows that an exact discretization will in general not fulfill the strong CT/DT-relationship as specified before. However, for an exact discretization the following result is immediate.

Theorem 3 (Exact Discretization)**.** *Under the same assumptions as Proposition 1. If the exact discretization is inconsistent with the data, then the continuous-time system is inconsistent. Furthermore, for the consistent parameter regions (as determined by Algorithm 1)*

$$\mathcal{P}_{CT} \subseteq \bigcup_{p\in\mathcal{P}_{DT}} \Gamma^{-1}(p) \tag{5.11}$$

holds. ■

Proof. The proof follows by construction w.r.t. the definition of the isomorphism. □

From the previous result follows that for exact discretizations, we can employ the methods from Chapter 4 to estimate the parameters of the continuous-time system. This is done by first estimating the parameters of the discrete-time system and then employing the back transformation (or reconstruction) to obtain the parameters of the continuous-time system.

As (structure-preserving) exact discretizations are in general not available or do not exist for polynomial systems, we consider in the remainder of this work only discrete-time approximations. We propose next a method to derive an embedding discrete-time system based on the Picard-Lindelöf theorem.

5.4 Guaranteed Parameter Estimation using Discrete-time Approximations

In this section, we show that the estimation methods developed in Chapter 4 can be employed to choose a time-step size for the discretization such that the strong consistency property holds. In other words, the discrete-time system embeds the reachable set of the continuous-time system at some distinct time points. The proposed method is based on the Picard-Lindelöf theorem for the existence and uniqueness of solutions to ordinary differential equation systems [140]. To do so, we develop an algorithm based on the Picard iteration and the constant enclosure method [51] as explained in the following. The main difference to the previous mentioned works is, that we present a method to derive the enclosure based on the continuous-time dynamics. We begin by stating the Picard-Lindelöf theorem and showing that it indeed can be used to derive an embedding discrete-time approximation.

5.4.1 Picard-Lindelöf Theorem

The goal of this section is to compute an enclosure of all trajectories $\phi(x(t)|x_0, u(t), p)$ of (2.1) with an initial condition x_0 in the semi-algebraic set $\mathcal{X}_0$ on a time interval $[0, D]$. For simplicity of presentation we assume that $\mathcal{X}_0$ is convex and $u(t)$ is constant on the time interval $[0, D]$ as well as bounded by the semi-algebraic set $\mathcal{U}$.

Remark 5. *Note that in case $\mathcal{X}_0$ is nonconvex, for instance, Algorithm 2 can be employed to find a convex outer-bounding box or the methods in [73, 110] can be employed to derive a (approximative) semi-definite representation of the convex hull of $\mathcal{X}_0$. Note further that the assumption on $u(t)$ can be relaxed when e. g. $u(t)$ is a polynomial in t, i. e. $u(t) \in \mathbb{R}[t]$, or the input is treated similarly to [68].*

Before stating the main tool for deriving the enclosure, namely the Picard-Lindelöf theorem, recall the so-called Picard iteration for a fixed initial condition

$$x(t) = x(0) + \int_0^t f(x(t), u(t), p). \tag{5.12}$$

The Picard iteration can be used to derive a time-step size D such that the Euler system

$$x(k+1) = x(k) + Df(x(k), u(k), p) \tag{5.13}$$

encloses the trajectories of (2.1) as summarized by the following theorem.

Theorem 4 (Picard-Lindelöf)**.** *Given a convex set $\mathcal{X}_0$ of initial conditions of the continuous-time system* (2.1) *and the semi-algebraic sets $\mathcal{P}$ and $\mathcal{U}$. Furthermore, given a convex enclosure $\hat{\mathcal{X}}$ such that $\cup_{x_0 \in \mathcal{X}_0, u(t) \in \mathcal{U}, p \in \mathcal{P}} \phi(x(t)|x_0, u(t), p) \subseteq \hat{\mathcal{X}}$ and*

$\mathcal{X}_0 \subset \hat{\mathcal{X}}$ *holds. For the Euler system* (5.13) *there exists* $D > 0$ *such that it embeds the continuous-model* (2.1) *on the time interval* $[0, D]$*, i. e.*

$$\bigcup_{x_0 \in \mathcal{X}_0, u(t) \in \mathcal{U}, p \in \mathcal{P}} \phi(x(t)|x_0, u(t), p) \subseteq \bigcup_{x_0 \in \mathcal{X}_0, u(k) \in \mathcal{U}, p \in \mathcal{P}} (x_0 + Df(x_0, u(0), p)) \tag{5.14}$$

holds for all $t \in [0, D]$. ∎

Proof. This proof is by construction and based on [141]. Let $\mathcal{C}$ denote the set of continuous functions with range $\hat{\mathcal{X}}$, i. e. $\mathscr{C} := \{\bar{x} \in \mathscr{C}^0([0, D], \hat{\mathcal{X}})\}$, equipped with the exponential norm $||\bar{x}||_\alpha = \max(e^{-\alpha(t)}||\bar{x}(t)||), \alpha > 0$. Applying the Picard iteration to the continuous function $\bar{x} \in \mathscr{C}$ with fixed initial condition $\bar{x}(0) \in \mathcal{X}_0$, fixed parameter $p \in \mathcal{P}$, and fixed input $u \in \mathcal{U}$, we obtain

$$\begin{aligned} \bar{x}(t) &= \bar{x}(0) + \int_0^D f(\bar{x}(s), u, p) ds \\ &\in \bigcup_{x \in \hat{\mathcal{X}}} (\bar{x}(0) + \int_0^D f(x, u, p) ds) \\ &\subseteq \bigcup_{x \in \hat{\mathcal{X}}} (\bar{x}(0) + Df(x, u, p)) \subseteq \hat{\mathcal{X}} \end{aligned} \tag{5.15}$$

As $\mathcal{C}$ is a bounded subset of $\mathscr{C}^0$ the Picard iteration in (5.15) is a contraction according to the Banach Fixed-Point theorem where the fixed-point is denoted by $\phi(x(t)|x_0, u(t), p)$ that satisfies (2.1) and $\phi(x(t)|x_0, u(t), p) \in \hat{\mathcal{X}}$ for all $t \in [0, D]$. In other words, $\phi(x(t)|x_0, u(t), p)$ exists and is unique. To extend this result to all initial conditions, parameters and piecewise constant inputs, consider the set-valued function

$$f_D(\mathcal{X}_0, \hat{\mathcal{X}}, \mathcal{U}, \mathcal{P}) := \bigcup_{x_0 \in \mathcal{X}_0, x \in \hat{\mathcal{X}}, u \in \mathcal{U}, p \in \mathcal{P}} x_0 + Df(x, u, p). \tag{5.16}$$

If D is chosen such that $f_D(\mathcal{X}_0, \hat{\mathcal{X}}, \mathcal{U}, \mathcal{P}) \subseteq \hat{\mathcal{X}}$, then it follows again from the Banach Fixed-Point theorem that (2.1) has a unique solution $\phi(x(t)|x_0, u(t), p)$ for all $x_0 \in \mathcal{X}_0, u(t) \in \mathcal{U}, p \in \mathcal{P}$. Furthermore, it follows that $\phi(x(t)|x_0, u(t), p) \in \hat{\mathcal{X}}$ for all $x_0 \in \mathcal{X}_0, u(t) \in \mathcal{U}, p \in \mathcal{P}$. Since $f_D(\mathcal{X}_0, \hat{\mathcal{X}}, \mathcal{U}, \mathcal{P}) \subseteq \hat{\mathcal{X}}$, also (5.14) holds. This concludes the proof. Note since $\mathcal{X}_0 \subset \hat{\mathcal{X}}$, D can always be chosen such that this condition holds. Note further that f is a polynomial and, therefore, continuously differentiable such that the conditions of the Banach Fixed-Point theorem can be verified. □

The previous theorem is stated only for the time interval $[0, D]$, however, it can be easily extended to cover a time interval $[0, T], T > D$ by successfully choosing (5.16) as the new initial condition set and again deriving enclosures $\hat{\mathcal{X}}$ and step sizes that fulfill the Picard iteration, cf. also Section 5.4.3. Obviously, this strategy requires that the enclosure $\hat{\mathcal{X}}$ is as tight as possible or otherwise the approximation obtained by (5.16) grows rapidly over time. To counteract this growth, typically, not a Euler system,

but higher order discretization schemes are considered, e. g. [117, 140]. For this work, however, higher order schemes are impracticable as the semi-definite programs in the relaxation approach presented in Section 4.3.1 grow with the order of the polynomials. Therefore, we restrict our attention here to Euler systems. However, as we can employ the guarantees of the relaxation approach, we can derive an algorithm that converges to the maximal allowed time-step that fulfills the Picard-Lindelöf theorem, cf. Section 5.4.3. Furthermore, we can derive an optimal a priori enclosure $\hat{\mathcal{X}}$ based on the continuous-time dynamics as seen next.

5.4.2 Enclosure of the Continuous-Time Reachable Set

In this section, we derive the enclosure $\hat{\mathcal{X}}$ needed for ensuring the embedding property of the Euler system as introduced in the previous section. To do so, we derive directly the reachable set of (2.1) for a distinct time-point D. The method is based on [74, 203], where a method is given to derive the initial conditions that lead to a behavior consistent with a target set, e. g. the measurements. We reverse this problem to derive an approximation of the reachable set of (2.1) starting from a given set of initial conditions.

Remark 6. *In [74] a different terminology is used then in this work. What we refer to as consistent initial conditions and parameters is called in [74] the region of attraction. Furthermore, note that the method presented here was remarked in [74], however, no actual derivation was presented.*

The crucial idea we employ in computing the reachable set of (2.1) is to formulate the dynamics in terms of occupation measures. This has two main advantages. First, it allows us to consider the whole set of initial conditions $\mathcal{X}_0$, inputs $\mathcal{U}$, and parameters $\mathcal{P}$. Second, the formulation allows us despite the nonlinear dynamics to derive an infinite-dimensional linear program that we can address with the relaxation procedure presented in Section 4.3.1. Note that although the employed measures suggest a probabilistic setup the obtained results in this section are deterministic. Note further that since we derive directly a measure problem similar to (4.9) the exposition is different to Section 4.3.1 and all necessary derivation steps are given in the following.

We denote the set of finite Borel measures supported on the set $\mathcal{A}$ by $\mathscr{B}(\mathcal{A})$. These measures can be interpreted as elements of the dual space $\mathscr{C}(\mathcal{A})'$, i. e. as bounded linear functionals acting on the set of continuous functions $\mathscr{C}(\mathcal{A})$. The set $\mathscr{P}(\mathcal{A})$ denotes the probability measures supported on $\mathcal{A}$, i. e. those measures μ of $\mathscr{B}(\mathcal{A})$ which are nonnegative and normalized to $\mu(\mathcal{A}) = 1$.

To derive the connection of the dynamics (2.1) and the occupation measures, we interpret at first the initial condition x_0 as a random variable whose distribution is described by an unknown probability measure $\mu_0 \in \mathscr{P}(\mathcal{X})$. We define this measure as

$$\mu(\mathcal{A} \times \mathcal{B}) := \int_{\mathcal{T}} \int_{\mathcal{X}} I_{\mathcal{A}\times\mathcal{B}}(t, \phi(x(t)|x_0, u, p))\mu_0(dx_0)\, dt \tag{5.17}$$

for all subsets $\mathcal{A}\times\mathcal{B}$ in the Borel σ-algebra of subsets of $\mathcal{T}\times\mathcal{X}$, where $\mathcal{T}\subset\mathbb{R}$ is the time interval $[0, D]$. $I_{\mathcal{A}}(x)$ denotes the indicator function of the set $\mathcal{A}$, i. e. the indicator function is equal to one if $x \in \mathcal{A}$, and zero otherwise. For simplicity of notation, we assume that the time invariant parameters p are introduced in the dynamics as additional states, i. e. $\dot{x}_i = 0$ if x_i corresponds to a parameter.

We refer to $\mu \in \mathscr{P}(\mathcal{T}\times\mathcal{X})$ as an occupation measure, whereas this term is motivated by the observation that the value $\int_{\mathcal{T}} \mu(dt, \mathcal{B}) = \mu(\mathcal{T}\times\mathcal{B})$ is equal to the total time the trajectory spends in the set $\mathcal{B}\subset\mathcal{X}$. In addition, note that μ encodes the system trajectories, in the sense that if $v \in \mathscr{C}^\infty(\mathcal{T}\times\mathcal{X};\mathbb{R})$ is a smooth test function, and $\mu_0 = \delta_{x_0}$ is the Dirac measure at x_0, integration of v w.r.t. μ amounts to time integration along the system trajectory starting at x_0:

$$\int_{\mathcal{T}}\int_{\mathcal{X}} v(t,x)\mu(dt,dx) = \int_{\mathcal{T}} v\big(t, x(t|x_0)\big)\,dt. \tag{5.18}$$

With these notations, for all sufficiently regular test functions $v \in \mathscr{C}^1(\mathcal{T}\times\mathcal{X};\mathbb{R})$, it holds that

$$\begin{aligned}\int_{\mathcal{X}} v(T,x)\mu_T(dx) - \int_{\mathcal{X}} v(0,x)\mu_0(dx) = \\ \int_{\mathcal{T}}\int_{\mathcal{X}} \frac{d}{dt}\, v\big(t, x(t|x_0)\big)\mu_0(dx_0),\end{aligned} \tag{5.19}$$

which corresponds to the evolution of all trajectories along v starting from an initial condition x_0 as specified by the distribution μ_0. The right-hand-side of the above equation can be rewritten as

$$\begin{aligned}&\int_{\mathcal{T}}\int_{\mathcal{X}} \Big(\frac{\partial}{\partial t} v\big(t, x(t|x_0)\big) + \\ &\qquad \operatorname{grad} v\big(t, x(t|x_0)\big)\cdot f\big(t, x(t|x_0)\big)\Big)\,\mu_0(dx_0)\,dt \\ &= \int_{\mathcal{T}}\int_{\mathcal{X}} \Big(\frac{\partial}{\partial t} v(t,x) + \operatorname{grad} v(t,x)\cdot f(t,x)\Big)\,\mu(dt,dx).\end{aligned} \tag{5.20}$$

To simplify notation, we introduce the Liouville operator $\mathcal{L} : \mathscr{C}^1(\mathcal{T}\times\mathcal{X}) \to \mathscr{C}(\mathcal{T}\times\mathcal{X})$ as $\mathcal{L}v := \frac{\partial v}{\partial t} + \operatorname{grad} v \cdot f$ and its adjoint $\mathcal{L}' : \mathscr{C}(\mathcal{T}\times\mathcal{X})' \to \mathscr{C}^1(\mathcal{T}\times\mathcal{X})'$ such that for the bilinear form $\langle \mathcal{L}v, \mu\rangle = \langle v, \mathcal{L}'\mu\rangle$ holds for all $v \in \mathscr{C}^1(\mathcal{T}\times\mathcal{X})$, i. e. $\mathcal{L}'\mu := -\frac{\partial \mu}{\partial t} - \operatorname{div}(\mu f)$.

With these notations, we write (5.19) concisely as

$$\langle \mathcal{L}v, \mu\rangle = \langle v, \delta_D\mu_D\rangle - \langle v, \delta_0\mu_0\rangle \tag{5.21}$$

for all $v \in \mathscr{C}^1(\mathcal{D}\times\mathcal{X})$, where δ_0 and δ_T refers to $t = 0$ and $t = D$, respectively. Equivalently, we can write

$$\mathcal{L}'\mu = \delta_D\mu_D - \delta_0\mu_0. \tag{5.22}$$

Note that as the initial conditions are not known, the measures are unknown as well.

In the next sections we derive an optimization problem that allows us to determine the unknown measures.

In [74] it was proved that the the set of initial conditions x_0 consistent with the dynamics (2.1) and the constraints $x(t) \in \mathcal{X}$, $t \in \mathcal{T}$ and a constraint $\hat{\mathcal{X}}$ at $t = D$, is the support of the measure μ_0 solving the infinite-dimensional linear programming (LP) problem

$$\begin{aligned} \sup \quad & \langle 1, \mu_0 \rangle \\ \text{subject to} \quad & \hat{\mu}_0 + \mu_0 = \lambda, \\ & \mathcal{L}'\mu + \delta_0\mu_0 - \delta_D\mu_D = 0, \\ & \hat{\mu}_0 \geq 0,\ \mu_0 \geq 0,\ \mu_D \geq 0,\ \mu \geq 0, \end{aligned} \tag{5.23}$$

where λ is the Lebesgue measure restricted to $\mathcal{X}$, i. e. the standard n_x-dimensional volume. The supremum is here w.r.t. the measures $\hat{\mu}_0 \in \mathscr{P}(\mathcal{X})$, $\mu_0 \in \mathscr{P}(\mathcal{X})$, $\mu_D \in \mathscr{P}(\mathcal{X})$ and $\mu \in \mathscr{P}(\mathcal{T} \times \mathcal{X})$. Note that the slack measure $\hat{\mu}_0$ results from the inequality $\mu_0 \leq \lambda$ as further explained in [74].

However, we are not interested in the set of consistent initial conditions, but in an approximation of the reachable set $\hat{\mathcal{X}}$. Therefore, we reverse the role of μ_0 and μ_D and derive the following infinite-dimensional linear program describing the measure μ_D at time D.

$$\begin{aligned} \sup \quad & \langle 1, \mu_D \rangle \\ \text{subject to} \quad & \hat{\mu}_D + \mu_D = \lambda, \\ & \mathcal{L}'\mu + \delta_0\mu_0 - \delta_D\mu_D = 0, \\ & \hat{\mu}_0 \geq 0,\ \mu_0 \geq 0,\ \mu_D \geq 0,\ \mu \geq 0. \end{aligned} \tag{5.24}$$

The following result describes the relationship between the reachable set $\mathcal{R}$ at time D of (2.1) (cf. Chapter 2) and the linear program (5.24).

Theorem 5 (Hypervolume)**.** *The optimal value l^* of* (5.24) *is equal to the hypervolume of $\mathcal{R}$ at time D, denoted by $\mathcal{R}_D$, i. e.*

$$l^* = \int_{\mathcal{R}_D} \lambda(dx). \tag{5.25}$$

Furthermore, the supremum is attained by the restriction of the Lebesgue measure to the reachable set $\hat{\mathcal{X}}$. ∎

Proof. The proof is based on the proof of [74, Thm.1]. Note that by definition of the reachable set there is an initial condition $x_0 \in \mathcal{X}_0$ such that $\phi(x(D)|x_0, u(t), p) \in \hat{\mathcal{X}}$. Accordingly, for any measure μ_D with support in $\mathcal{X}$ there exists a measure $\hat{\mu}_D, \mu_0$ and μ that fulfills the constraints in (5.24). Thus, $l^* \geq \int_{\mathcal{R}_D} \lambda(dx)$. From the constraint $\mu_D + \hat{\mu}_D = \lambda$ it follows that $l^* \leq \int_{\mathcal{R}_D} \lambda(dx)$ and consequently $l^* = \int_{\mathcal{R}_D} \lambda(dx)$. □

By applying the same ideas as presented in Section 4.3.1 we can derive a converging hierarchy of semi-definite programs to solve (5.24). However, as we are interested in deriving a closed form approximation of the reachable set $\hat{\mathcal{X}}$, we first derive the dual

of (5.24) over the space of continuous functions. To do so, we state the algebraic dual pair

$$\langle A(x), y\rangle = \langle x, A^*(y)\rangle, \tag{5.26}$$

where $x = (\mu_D, \hat{\mu}_D, \mu, \mu_0), A(x) = (\mathcal{L}'\mu - \mu_T + \mu_0, \mu_D + \hat{\mu}_D), y = (v, w), A^*(y) = (w - v, w, \mathcal{L}v, v), w \in \mathscr{C}(\mathcal{X}), v \in \mathscr{C}^1(\mathcal{T} \times \mathcal{X})$.

The dual pair allows us to state the dual LP as

$$\begin{aligned} \inf \quad & \langle w, \lambda\rangle & & \\ \text{subject to} \quad & w(x) \geq 0, & \forall x \in \mathcal{X}, \\ & w(x) \geq 1 + v(D, x), & \forall x \in \mathcal{X}, \\ & v(0, x) \geq 0, & \forall x \in \mathcal{X}_0, \\ & \mathcal{L}v(t, x) \geq 0, & \forall (t, x) \in \mathcal{T} \times \mathcal{X}, \end{aligned} \tag{5.27}$$

where the infimum is w.r.t. continuous functions $w \in \mathscr{C}(\mathcal{X})$ and $v \in \mathscr{C}^1(\mathcal{T} \times \mathcal{X})$. Further information on dual formulations of infinite dimensional linear programs can be found in [2, 18].

The above LPs (5.24) and (5.27) are infinite-dimensional, because the equations are required to hold for all test functions v. One can solve these LPs by a converging hierarchy of finite dimensional linear matrix inequality (LMI) problems using semidefinite programming. At a given relaxation order d, the primal LMI is a moment relaxation of primal LP (5.24), whereas the dual LMI is a polynomial sum-of-squares (SOS) restriction of dual LP (5.27). As said before we are interested in deriving a closed form approximation $\hat{\mathcal{X}}$, therefore, we provide only the SOS restriction formulated according to Putinar's positivstellensatz.

$$\begin{aligned} \inf \quad & w_c{}'l & & \\ \text{subject to} \quad & \mathcal{L}v(t,x) = & & q_0(t,x) + q_{m_x+1}(t,x)t(D-t) + \sum_{i=1}^{m_x} q_i(t,x)g_{x,i}(x), \\ & w(x) = & & v(D,x) + 1 + r_0(x) \\ & & & \quad + \sum_{i=1}^{m_x} r_i(x)g_{x,i}(x), \\ & v(0,x) = & & s_0(x) + \sum_{i=1}^{m_{x_0}} s_i(x)g_{x_0,i}(x), \\ & w(x) = & & z_0(x) + \sum_{i=1}^{m_x} z_i(x)g_{x,i}(x), \end{aligned} \tag{5.28}$$

where l is the vector of Lebesgue moments over $\mathcal{X}$ indexed in the same basis in which the polynomial $w(x)$ with coefficients w_c is expressed. The minimum is over polynomials $v(t, x)$ and $w(x)$. The polynomials $g_{x,i}$ and $g_{x_0,i}$ correspond to the m_x (resp. m_{x_0}) inequalities defining $\mathcal{X}$ and $\mathcal{X}_0$, respectively The polynomials $q_i(t, x)$, $r_i(x)$, $s_i(x)$, and z_i are sum-of-square polynomials and of appropriate degrees. The constraints that polynomials are sum-of-squares can be written explicitly as LMI constraints, and the

objective is linear in the coefficients of the polynomial $w(x)$. Therefore, problem (5.28) can be formulated as a SDP [151].

By defining the set $\hat{\mathcal{X}} := \{w_c' l \geq 1\}$, we can state the following result. Note that $\hat{\mathcal{X}}$ does implicitly depend on the degree r of the sum-of-square polynomials.

Theorem 6. *The function $w := w_c' l$ that minimizes (5.28) has the property that for the degree of the sum-of-squares polynomials $r \to \infty$ the set $\hat{\mathcal{X}}$ converges (almost uniformly) to the reachable set $\mathcal{R}$ of (2.1) at time D and $\mathcal{R} \subseteq \hat{\mathcal{X}}$ holds for all r.* ∎

Proof. The proof is equivalent to the proof of [74, Thm.6]. □

The outer-approximation $\hat{\mathcal{X}}$ encloses, therefore, the reachable set of the continuous-time system. In the following, we give a brief example.

Example 3. *Consider the continuous-time model*

$$\begin{aligned} \dot{s}(t) &= (c(t) - 1)s(t) + c(t) \\ \dot{c}(t) &= (1 - c(t))s(t) - 2c(t). \end{aligned} \tag{5.29}$$

The considered setup is equivalent to the discrete-time model considered in Section 4.4 and as summarized by the following constraint set

$$\begin{aligned} \mathcal{K} := \{s(t), c(t) : 0 \leq s(t) \leq 1, 0 \leq c(t) \leq 1, \forall t \in [0,1], \\ 0.85 \leq s(0) \leq 0.95, 0.5 \leq c(0) \leq 0.55\}. \end{aligned} \tag{5.30}$$

Given these constraints we implemented (5.28) with the help of YALMIP [121, 122]. The involved semi-definite programs were solved using MOSEK 7.0. Figure 5.3 depicts the enclosure of the reachable set of (5.29) at time $t = 1$ as an example. ∎

In the next section, we use the outer-approximation to provide a procedure to guarantee that the Euler system (5.13) fulfills the strong CT/DT relationship property and, therefore, can be used to approximate the consistent parameters of the continuous-time system. Note further that the approximation $\hat{\mathcal{X}}$ is used in Section 5.5.3 for directly analyzing model consistency of a continuous-time system, in case that the consistent parameters do not have to be derived by a discrete-time approximation.

5.4.3 Reachability and Strong Consistency

As seen in the previous section, we can derive an enclosure $\hat{\mathcal{X}}$ as employed in the Picard-Lindelöf theorem. However, we still need to determine an approximation of the set-valued function (5.16) to determine a suitable step-size D. This can be achieved by stating the set-valued function as the following feasibility problem

$$\begin{aligned} \text{find} \quad & x^+ \\ \text{subject to} \quad & x^+ = x_0 + Df(x, u, p), \\ & p \in \mathcal{P}, x_0 \in \mathcal{X}_0, x \in \hat{\mathcal{X}}, u \in \mathcal{U}, x^+ \in \mathcal{X} \end{aligned} \tag{5.31}$$

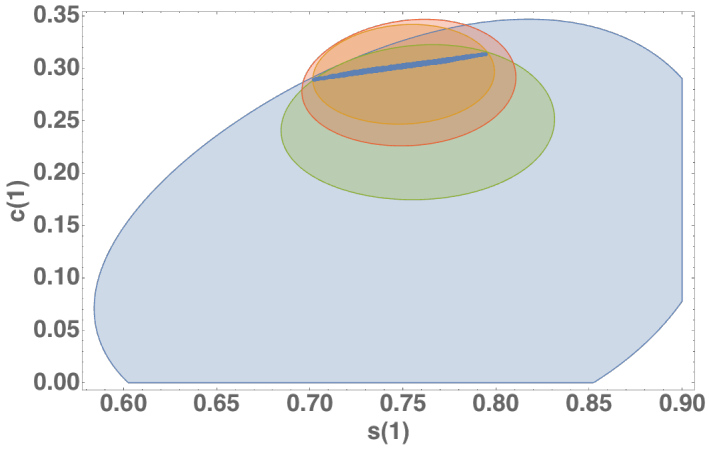

Figure 5.3: Enclosure of the reachable set for the Michaelis-Menten model (5.29) obtained by the sum-of-squares restriction (5.28) for relaxation degrees $r = 4, 6, 8, 10$ at $t = 1$ (blue, green, red, orange area). Blue dots represent 1000 Monte Carlo samples by numerically simulating (5.29) with initial conditions taken as uniformly distributed samples from $\mathcal{K}$.

where $\hat{\mathcal{X}}$ is an initially assumed approximation and D is fixed. Note that in case $\hat{\mathcal{X}}$ is nonconvex, first a convex approximation has to be derived following Remark 5.

In Section 4.3.1 it was shown that the feasibility problem (5.31) can be addressed with Algorithm 1 and Algorithm 2. Therefore, an appropriate D can be computed by checking whether the solution set of (5.31) projected on the space of variable x^+ is contained in the initially assumed enclosure $\hat{\mathcal{X}}$. For simplicity of presentation, assume that Algorithm 2 is used to derive an approximation of (5.31). We denote the obtained result, i. e. the outer-bounding box of x^+, by $\bar{\mathcal{X}}^+$ corresponding to some relaxation degree r. If $\bar{\mathcal{X}}^+ \subseteq \hat{\mathcal{X}}$ does not hold the step-size D has to be reduced as summarized by the following algorithm.

Algorithm 4 (Maximal Time-step Size)

① **Given constraints on initial conditions** $\mathcal{X}_0$

② **Set** ϵ, T_l, T_u**, where** $T_l < T_u$

determineStepSize(T_l, T_u, ϵ)

① $T = (T_u + T_l)/2$

② **Compute enclosure** $\hat{\mathcal{X}}$ **with** (5.28) **for** $t = T$

③ **Determine** $\hat{\mathcal{X}}^+$ **with** (5.31)
- **if** $\bar{\mathcal{X}}^+ \subseteq \hat{\mathcal{X}}$
  - **if** $(T_u - T_l)/2 \leq \epsilon$ **return** $T^* = T$
  - **else** $T_l = T$
- **else** $T_u = T$

④ **determineStepSize**(T_l, T_u, ϵ)

We can state the following property of Algorithm 4, when Algorithm 2 is employed for the approximation of the set-valued function.

Proposition 4. *Given the enclosure $\hat{\mathcal{X}}$ obtained by* (5.28)*. The time-step size derived by Algorithm 4 employing Algorithm 1, converges to the maximal (up to precision ϵ) time-step size for which the existence and uniqueness of solutions to* (2.1) *can be guaranteed according to the Picard-Lindelöf theorem w.r.t. the outer-bounding box $\bar{\mathcal{X}}^+$ approximating the set-valued function.* ∎

Proof. Follows by construction, as convergence to the minimal outer-bounding box of $f_D(\mathcal{X}_0, \hat{\mathcal{X}}, \mathcal{U}, \mathcal{P})$ is guaranteed by Lasserre's relaxation hierarchy. The convergence of D is guaranteed by the convergence of the bisectioning. □

To summarize, Algorithm 4 computes a step size D for which the Picard-Lindelöf theorem guarantees that all trajectories of the continuous-time system are contained in the set $\bar{\mathcal{X}}^+$ at time D. In other words, the Euler system embeds the reachable set of (2.1) at time D (5.13).

By repeating Algorithm 4 to compute all needed time-steps to cover the time interval of interest we can determine an outer-approximation of the reachable set of the continuous-time system using the Euler system (5.13). We denote the sequence of time steps as $\mathcal{T}_D := \{t_{D,1}, \ldots, t_{D,n_{\mathcal{T}_D}}\}$ needed to cover a time-interval $[0, T]$, such that on every interval $[t_{D,i}, t_{D,i+1}]$ the Picard-Lindelöf theorem holds. We denote the enclosure provided by the Euler system (5.13) at a time point $t_{D_i} \in \mathcal{T}_D$ by $\bar{\mathcal{X}}(t_{D_i})$.

By construction we have the following property.

Theorem 7. *Given an initial set $\mathcal{X}_0$, a piecewise constant input sequence bounded by $\mathcal{U}$, and a time interval time $[0, T]$. If the sequence of step sizes $\mathcal{T}_D$ is derived by successively employing Algorithm 4, then the Euler system embeds the trajectories of the continuous-time system* (2.1)*, i. e.*

$$\mathcal{R} \subset \hat{\mathcal{X}}(t_D) \tag{5.32}$$

holds for all distinct time points $t_D \in \mathcal{T}_D$. ∎

Proof. Follows by construction and the properties of the employed algorithms. □

From [117] and the previous result it follows that the consistent parameters and initial conditions of the continuous-time system can be studied by a discrete-time approximation. To derive an approximation of the consistent initial conditions and parameters of the continuous-time model (2.1) given some measurement data, we can employ the procedure presented in [117]. This procedure is as follows. At first the time step sizes for the discrete-time system are determined by Algorithm 4. Afterwards, a feasibility problem of form (4.1) is constructed including the data and the discrete-time system employing the derived time-step sizes. Then, Algorithm 1 or Algorithm 2 are employed for estimating the parameters based on the constructed feasibility problem as presented in Chapter 4. Additionally, the estimated parameters for the Euler system outer-approximate the parameters of the continuous-time system and the strong CT/DT relationship holds, see [117] for details.

To demonstrate the problem inherent to this procedure of deriving a discrete-time enclosure fulfilling the strong CT/DT relationship, we show next how Algorithm 4 performs for the Michaelis-Menten example as described in Example 3, cf. also Section 4.4.

Example 4. *Consider the continuous-time Michaelis-Menten model from Example 3 By successively reducing the time step size and recomputing the enclosure according to Algorithm 4, initialized with* $\epsilon = 10^{-8}$, $T_l = 0$ *and* $T_u = 5 \cdot 10^{-6}$, *we determined a time step size of* $7.5 \cdot 10^{-7}$ *for which the Picard-Lindelöf theorem is fulfilled and the continuous-time system is embedded by the Euler system, as depicted in Figure 5.4. From an exhaustive Monte Carlo analysis a practical bound on the step size was determined to be* 10^{-6}, *which shows that Algorithm 4 can be employed to derive the desired step-size. Furthermore, this example also illustrates that small step sizes are needed to ensure that the Euler system embeds the continuous-time model.* ■

Example 3 demonstrates the problem of only using an Euler model to derive an outer-approximation of the reachable set for a continuous-time system (2.1). Namely one is limited to so-called Euler steps to derive the approximation of the reachable set. These Euler steps become rather small due to the allowed uncertainties in the initial conditions, inputs, and parameters. Therefore, to cover a time interval $[0, T]$ many Euler steps might be needed. Although, this can be difficult for practical implementation of the proposed procedure, it is necessary to be able to estimate the consistent parameters of a continuous-time system with help of a discrete-time approximation.

In case only model invalidation is of interest a simplified method can be employed that does not require the consideration of Euler steps. The proposed method, is based on the approximation $\hat{\mathcal{X}}$ of the reachable set of a continuous-time system as derived in Section 5.4.2. However, as this method is only based on the continuous-time system, the connection between discrete-time and continuous-time system is lost. To regain

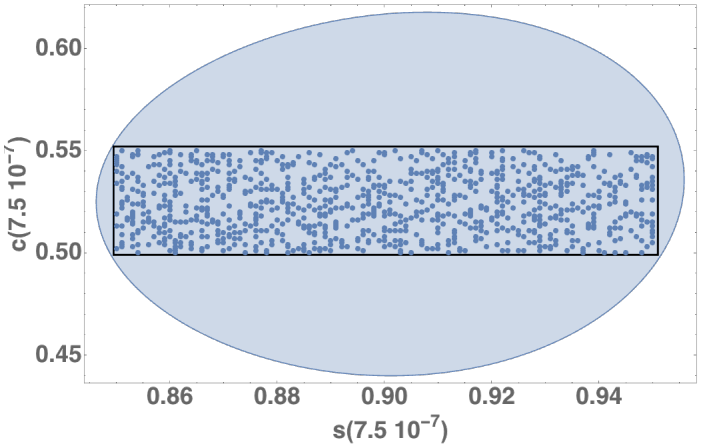

Figure 5.4: Enclosure of the reachable set for the Michaelis-Menten model (5.29) obtained by the sum-of-squares restriction (5.28) for a relaxation degree $r = 8$ at $t = 7.5 \cdot 10^{-7}$ (blue area). Blue dots represent 1000 Monte Carlo samples by numerically simulating the discrete-time model with initial conditions taken as uniformly distributed samples from $\mathcal{K}$. Black box corresponds to the embedding $\bar{\mathcal{X}}$ obtained with Algorithm 2.

the relationship, we furthermore sketch how this method can in principle be employed to study the error made by discretizing a continuous-time system.

5.5 Model Inconsistency

We have shown in the previous section that we can employ a discrete-time system to approximate the reachable set of a continuous-time system. In this case, the discrete-time system fulfills the strong CT/DT relationship property, i. e. from model consistency of the discrete-time system follows consistency of the continuous-time system. However, the derived discrete-time system is restricted to time steps that guarantee existence and uniqueness of trajectories of the continuous-time system according to the Picard-Lindelöf theorem. These time steps are in general relatively small for uncertain systems in comparison to the time interval at which a process is operated. This means checking model consistency becomes a computationally difficult problem since the size of the corresponding semi-definite programs as introduced in Section 4.3.1 depend on the amount of considered time steps. Therefore, it is necessary to consider a computationally more efficient solution. This solution is based on the approximation of the reachable sets presented in Section 5.4.2. Afterwards, we discuss a less rigorous method in case the computation of the reachable sets are not directly possible. Hereby, we focus on the principle possibility to study the error made by discretizing a continuous-time system.

5.5.1 Strong Inconsistency

Recall that the sum-of-squares restriction (5.28) provides an outer-approximation of the reachable set of the continuous-time system at a time $t \in \mathbb{R}$. Assume that we have measurements of the output $\mathcal{Y}(t_i)$ at time points $t_i \in \mathcal{T} := \{t_1, \ldots, t_{n_\mathcal{T}}\}$. If we compute with help of (5.28) outer-approximations $\hat{\mathcal{X}}(t_i)$ for all $t_i \in \mathcal{T}$, we can immediately test for model consistency employing the following feasibility problem.

$$\begin{aligned} \text{find} \quad & x(k) \\ \text{subject to} \quad & y(x(k), u(k), p) - y(k) = 0 \\ & x(k) \in \hat{\mathcal{X}}(t_i), u(k) \in \mathcal{U}, p \in \mathcal{P}, y(k) \in \mathcal{Y}(t_i), kD \in \mathcal{T}, t_i \in \mathcal{T}. \end{aligned} \tag{5.33}$$

The solution space of (5.33) corresponds to the mapping of $\hat{\mathcal{X}}(t_i)$ to the output space, therefore, we can state model inconsistency of the continuous-time system as follows.

Theorem 8 (Strong Inconsistency)**.** *Given the measurement data $\mathcal{Y}(t_i)$ from the real process at $t_i \in \mathcal{T} := \{t_1, \ldots, t_{n_\mathcal{T}}\}$ and the approximations $\hat{\mathcal{X}}(t_i)$ of the reachable set of* (2.1) *determined by* (5.28)*. If* (5.33) *does not admit a solution verified by Lemma 1, then the continuous-time system* (2.1) *is inconsistent with the data.* ■

Proof. Follows directly from Lemma 1 and the construction of (5.33). □

The previous result shows that the model invalidation task for continuous-time systems can be solved by employing the algorithms presented in Chapter 4. Furthermore, as the computation of the reachable sets of (2.1) is independent of a time step size one is no longer restricted to time steps fulfilling the Picard-Lindelöf theorem.

However, as this method is independent of a discrete-time system, the connection between discrete-time and continuous-time system cannot be studied any longer. To regain the ability to study this connection consider the augmented Euler system as introduced in the next section.

5.5.2 Augmented Euler System

For studying the connection of a continuous-time system and its discrete-time approximation, we derive an augmented discrete-time system. This augmented discrete-time system is based on the following consistency property following [166].

Definition 17 (One-step consistency[106])**.** *Given a discrete-time approximation f_D^a of continuous-time system* (2.1)*, i. e. a discrete-time system obtained by numerical integration of form* (2.2)*. The system f_D^a is said to be one-step (strongly) consistent if given any pair of strictly positive real numbers (Δ_x, Δ_u), there exists a class function $\rho \in \mathcal{K}_\infty$ and $D^* > 0$ such that*

$$|\phi(x(t_i)|x_0, u(t_i), p) - f_D^a(x(k), u(k), p)| \leq D\rho(D) \tag{5.34}$$

holds for all times $t_i = kD$*, for all* $D \in (0, D^*)$ *and* $|x_k| \leq \Delta_x, |u_k| \leq \Delta_u, p \in \mathcal{P}$. ∎

To derive the augmented discrete-time system, we choose for f_D^a the Euler discretization

$$x(k+1) = x(k) + Df(x(k), u(k), p) := f_D^{\text{Euler}}(x(k), u(k), p). \tag{5.35}$$

On the one hand the Euler discretization is one-step consistent, see e. g. [207] and on the other hand the Euler system neither changes the system class nor the degree of the polynomials of the continuous-time system. The latter is beneficial for limiting the size of the semi-definite programs employed for proving model consistency.

By rewriting (5.34) as an equality of form

$$\phi(x(t_i)|x_0, u(t_i), p) - x(k+1) + f_D^{\text{Euler}}(x(k), u(k), p) + \epsilon(k) = 0, \tag{5.36}$$

with $||\epsilon(k)|| \leq \delta := \max_{p\in\mathcal{P}, D\in(0,D^*)} ||D\rho(D)|| \leq \Delta_x + \Delta_u$, we see that the Euler system and the continuous-time trajectory differ by the discretization error $\epsilon(k)$.

By introducing the discretization error $\epsilon(k)$ as a time-varying parameter, we can state the augmented system as

$$x(k+1) = f_D^{\text{Euler}}(x(k), u(k), p) + \epsilon(k). \tag{5.37}$$

We denote this system in the following as an augmented Euler system. Note that a similar augmented system was introduced in [183] to study discretization errors with the help of interval arithmetics.

In principle, if $\epsilon(k)$ is known the augmented Euler system embeds a continuous-time trajectory (follows from (5.36)). However, computing this error for a fixed parameter p and a known input sequence $u(k)$ can already be challenging for nonlinear systems as considered in this work [46, 207]. Instead of computing $\epsilon(k)$ explicitly we can treat $\epsilon(k)$ as another unknown-but-bounded variable in the considered set-based framework, as $\epsilon(k)$ is bounded on the state constraint set $\mathcal{X}$ for a finite time interval. This, however, leads to another problem. Namely, that the reachable set of (5.37) becomes quite large, as the additive structure leads to an increase in the uncertainties of states, parameters and inputs. In this work, to limit the uncertainties introduced by the discretization error, we employ the following simplified system description.

$$\begin{aligned} \tilde{x}(k+1) &= f_D^{Euler}(\tilde{x}(k), \tilde{u}(k), \tilde{p}), \\ \hat{\mathcal{X}}(k+1) &= \tilde{x} \oplus \mathcal{E}(k), \end{aligned} \tag{5.38}$$

where $\oplus$ denotes the Minkowski sum, and the tilde sign $\sim$ corresponds to a nominal value of the initial condition, the input, and the parameters. The set $\mathcal{E}(k)$ has to be determined such that $\hat{\mathcal{X}}(k) \supseteq \cup_{x_0\in\mathcal{X}_0, u(t)\in\mathcal{U}, p\in\mathcal{P}} \phi(x(kD)|x_0, u(t), p)$ holds as seen in the following section. The main advantage of the simplified augmented Euler formulation

comes into play when a model has to be (in-)validated w.r.t. available data from multiple experiments (resp. semi-quantitative observations). This derives from the fact that after determining the sets $\hat{\mathcal{X}}(k)$, these sets can be tested against the data without recomputing the error sets.

5.5.3 Weak Consistency

In this section, we study how the error set $\mathcal{E}(k)$ in the simplified augmented Euler formulation introduced in the previous section can be approximated. This approach is based on simulation data taken from the continuous-time model. Here, only a weaker relationship can be guaranteed, however, it allows to consider not only polynomial but also rational systems as demonstrated with a real world example in Section 5.6.

We present first how the error set can be derived for given training sets $\bar{\mathcal{X}}(t_i)$ representing the trajectories of the continuous-time system at distinct time points $t_i \in \mathcal{T} := \{t_1, \ldots, t_{n_\mathcal{T}}\}$. Afterwards, we discuss possible methods to determine these training sets. In accordance to Section 4.2, we can formulate a sequence of feasibility problems to estimate the error set w.r.t. the training data as

$$\begin{aligned} \text{find} \quad & \epsilon(k) \\ \text{subject to} \quad & x(k+1) = f_D^{\text{Euler}}(\tilde{x}(k), \tilde{u}(k), \tilde{p}) + \epsilon(k), \\ & x(k+1) \in \bar{\mathcal{X}}(t_i), \epsilon(k) \in \mathcal{E}(k), (k+1)D = t_i, \end{aligned} \tag{5.39}$$

where $\mathcal{E}(k) := \{\epsilon(k) \in \mathbb{R}^{n_x} \mid ||\epsilon(k)|| \leq \Delta_x + \Delta_u\}$ denotes the initial error set, and $\tilde{x}(k), \tilde{u}(k), \tilde{p}$ are nominal values taken from the training set of the previous step $\bar{\mathcal{X}}(kD)$, the piecewise constant input set $\mathcal{U}(k)$, and the parameter set $\mathcal{P}$, respectively. Note that the training set at time t_1 is only used to initialize the sequence. The projection of the solution space of (5.39) on the variable $\epsilon(k)$ determines the error set. We summarize this by the following statement for Algorithm 1.

Proposition 5 (Error Set [166])**.** *Given the feasibility problem* (5.39)*. The error introduced by the discretization can be outer-approximated by*

$$\hat{\mathcal{E}}(k) = \mathcal{E}(k) \setminus \bigcup_{\mathcal{Q} \subseteq \mathcal{E}(k): dualSDP_d \to \infty} \mathcal{Q}, \tag{5.40}$$

with Algorithm 1. ∎

Proposition 5 shows that the error set can be outer-approximated, if the training sets $\bar{\mathcal{X}}(t_i)$ are available. By replacing $\mathcal{E}(k)$ in (5.38) with the estimation $\hat{\mathcal{E}}(k)$, we can guarantee that (5.38) embeds the training sets and, therefore, the continuous-time system. Note that the embedding depends on the choice of the nominal values and the time step size, cf. also to Remark 7 for a more detailed discussion.

To test if a continuous-time model is consistent with measurement data $\mathcal{Y}(t_i)$ at the time points $t_i \in \mathcal{T} := \{t_1, \ldots, t_{n_\mathcal{T}}\}$, we can investigate the following feasibility problem

similar to (5.33).

$$\begin{aligned} \text{find} \quad & x(k) \\ \text{subject to} \quad & y(x(k), \tilde{u}, \tilde{p}) - y(k) = 0 \\ & x(k) \in \hat{\mathcal{X}}(k), y(k) \in \mathcal{Y}(t_i), kD \in \mathcal{T}, t_i \in \mathcal{T}, \end{aligned} \tag{5.41}$$

where $\hat{\mathcal{X}}(k)$ represents the embedding derived from (5.38). Note that in principle the training sets can be derived by the sum-of-squares restriction (5.28), however, as this approach is not applicable for the real world example in Section 5.6 due to the rational dynamics of the continuous-time model, we discuss in the following other suitable approaches.

To derive appropriate training sets $\hat{\mathcal{X}}(k)$ several methods are known in literature. For instance, in [180, 187] representative trajectories are employed to derive convex lower and upper bounds on the trajectories of a continuous-time system. Representative trajectories were also used, for instance, in [129] to investigate network topologies that support adaptation in cells. This was done by a simulation study by partitioning the parameter space into boxes and then simulating one representative parameterization for each box. Other suitable approaches can be found e. g. in [53, 118]. The aforementioned approaches can easily adapted to generate a training set for the augmented Euler system. However, it should be noted that the aforementioned approaches in general assume certain characteristics of the model under investigation, e. g. that the continuous-time trajectories have to stay close in terms of a chosen norm or small disturbances in the parameters do not change the qualitative behavior of the model. Another commonly employed assumption is that the trajectories are derived analytically or the precision of the numerical integration is high enough. These assumptions can be rather restrictive and choosing the correct method depends strongly on the system under investigation. Still from a practical perspective these methods provide another degree of freedom to balance accuracy and computational effort. However, the made assumptions in the derivation of the training set lead in the set-based setting only to a weak relationship of continuous-time and discrete-time system. To illustrate this behavior, assume the nominal trajectory $\phi_{nom}(x(t)|x_0, u(t), p)$ is a solution of (2.1), Theorem 8 can then be formulated as:

Proposition 6 (Weak Inconsistency (Nominal Model))**.** *Given the measurement data $\mathcal{Y}(t_i), t_i \in \mathcal{T}$ from the real process. Assume that the error set $\hat{\mathcal{E}}(k)$ has been determined according to Lemma 5 for training sets based on $\phi_{nom}(x(t)|x_0, u(t), p)$. If* (5.41) *does not admit a solution verified by Lemma 1, then the nominal trajectory $\phi_{nom}(x(t)|x_0, u(t), p)$ of* (2.1) *is inconsistent with the data.* ∎

Clearly, the previous proposition can be extended to more then one trajectory or more precisely to training sets. However, in general the guarantees of Theorem 8 can only be recovered for an infinite amount of considered trajectories. However, one

advantage of this simplified treatment is the possibility to investigate the influence of the discretization error made. This can in principle be done by comparing the Lebesgue measure of the error sets $\hat{\mathcal{E}}(k)$ for different step sizes, cf. also to Remark 7.

Before demonstrating that this derivation of training sets by simulation data can still be of practical use, we provide in the following remark a discussion of the limitations of the presented approach and possible improvements are sketched.

Remark 7. *The main limitation of the simplified Euler system derives from the fact that the error set depends not only on the time step size, but also from the chosen nominal parameterization and the nominal input sequence. This increase in uncertainty might lead to the wrong conclusion that a continuous-time model is consistent with the data. To limit this increase either one has to determine a in some sense optimal nominal parameterization/input sequence or one has to estimate the dynamics of the discretization error. The former can be achieved by means of optimal parameter estimation [152, 210] with an objective appropriately chosen for the process under consideration. The latter can be done by employing system identification methods. Applicable approaches are the method presented in [26] for the identification of growth phases and the set-membership system identification approach presented in [135]. The advantage of employing system identification methods is that in principle one can derive a connection of the parameters of the continuous-time system and parameters of the discrete-time system similar to the exact discretization considered in Section 5.3.*

5.6 Example: Bacterial Growth Model

We illustrate in the following how the methods developed in the previous sections can be employed in the identification of an appropriate model for the control of a biotechnological process. Namely, the steady-state cultivation of *Rhodospirillum rubrum* as introduced in [33]. We first provide a short overview over the model organism and describe the applied process control scheme. Thereafter, we introduce the considered model hypotheses and then present how a suitable model was identified.

5.6.1 Background: *Rhodospirillum rubrum*

The anoxygenic, photosynthetic bacteria *Rhodospirillum rubrum* have a high potential for biotechnological applications. Especially, the products related to the formation of intracytoplasmatic, photosynthetic membranes (PM) are of industrial interest. Examples are the heterologous expression of membrane proteins [31], coenzyme Q10 [211], biohydrogen [223], biopolymers [188], bacteriochlorophyll and its precursors [44]. The industrial application, however, is still impeded by very low productivities, high energy efforts and complex setups. This is mainly due to the difficulties involved in defining and maintaining microaerobic conditions in continuously operated fed-batches.

Investigating microaerobic conditions is, however, not only interesting for industrial applications but also for several other research fields including experimental medicine and pharmacology. Microaerobic environments are often breeding grounds for pathogenic organisms and in medical research such environments have to be mimicked ex situ [150]. Also the expression of virulence factors would benefit from controlled low oxygen conditions, see e. g. [132, 155]. Phototrophic bacteria, as *R. rubrum*, present interesting model organisms for studying such conditions [65] as the growth mode can be easily monitored by e. g. fluorescence spectroscopy [33].

To be able to use *R. rubrum* as a case study several challenges have to be overcome. For instance, *R. rubrum* allows a maximal production of PM only when grown on a culture medium with the two carbon sources succinate and fructose [61]. This renders the application of established process control strategies like [1] inapplicable. Note that the need for a multi-carbon source medium is unique to *R. rubrum* and has not been observed in related photosynthetic purple bacteria. A possible explanation might be connected to the cellular redox signaling at the level of the cytosolic glutathione and the membrane localized ubiquinone pool [32, 64].

Furthermore, due to the lack of an exact stoichiometrical knowledge of PM production in relation to oxygen consumption does not allow for a respiratory control scheme, see [225]. Even though the regulation of PM production in response to oxygen for other anoxygenic photosynthetic bacteria [19, 62, 148] are rather well investigated, the regulatory mechanisms in *R. rubrum* harbor a lot of questions.

In [33], we proposed an experimental setup based on the *culture redox potential* that overcomes most of these challenges. In contrast to previously used setups like [61, 65] this approach provides stationary microaerobic conditions. Attaining the desired oxygen-limited steady-state condition in continuous cultivations requires, however, a feedback-control strategy. We proposed, therefore, in [33] a feedforward variant of the classical model-based two-degree-of-freedom (2DOF) controller [6]. The main advantage of this strategy is clearly that the tradeoff between set-point tracking and disturbance rejection can be avoided. In other words, this approach allows to reach the steady-state without undesired oscillations which reduces the time needed to ensure the culture is in a steady-state with respect to the retention time. The 2DOF controller relies on spectroscopic real-time measurements of the biomass concentration. This is done by measuring the optical density of the cultivation medium by a transflexion probe and then calculating the biomass concentration. Nevertheless, it should be noted that the good performance achieved with 2DOF control schemes typically depends on the availability of an adequate model describing the process dynamics [45]. However, if a good model is available this strategy can be easily adapted to other processes and other organisms [43].

In the following, we describe the model hypotheses identified in [33] and demonstrate how we were able to determine which model is the most suitable for the control strategy. Note that we consider here only two out of three model hypotheses as we were able

to show that one of the hypotheses is not identifiable [33]. Note further that the algorithms developed in this work were used to decide which of the remaining two hypotheses was best suited for the control strategy and that the identified model was successfully employed in steady-state cultivations. We focus here mainly on the technical aspects in the development and validation process of the models and refer to [33] for a more detailed discussion.

5.6.2 Model Description

The following equations represent the time evolution of the concentrations of biomass (x), succinate (x_s) and fructose (x_f). We developed in [33] three alternative model hypotheses by choosing specific growth, biomass-substrate-yield coefficients and with this uptake rate correlations. The resulting model has the form

$$\begin{aligned} \dot{x} &= \mu(x_s, x_f)x - Dx, \\ \dot{x}_s &= -q_s(x_s)x + D(F_s - x_s), \\ \dot{x}_f &= -q_f(x_f)x + D(F_f - x_f). \end{aligned} \tag{5.42}$$

Here, μ denotes the specific growth rate of biomass, D denotes the dilution rate, q_s, q_f denote the consumption rate of succinate and fructose, F_s, F_f denote the substrate concentrations of the feed solution respectively.

We consider, in this work, only model hypotheses II and III, since hypothesis I is not identifiable, cf. [33] for details. In hypothesis II the specific growth rate μ and the consumption rates q_s, q_f are defined as follows.

$$\begin{aligned} \mu(x_s, x_f) &= v_{b,CRPmax}\left(\frac{x_s}{x_s + k_s} + \frac{x_f}{x_f + k_f}\right), \\ q_s(x_s) &= \frac{1}{Y_{b/s} + Y_{b/f}}\frac{x_s}{x_s + k_s}, \\ q_f(x_f) &= \frac{1}{Y_{b/s} + Y_{b/f}}\frac{x_f}{x_f + k_f}. \end{aligned} \tag{5.43}$$

In hypotheses III the growth and consumption rates become:

$$\begin{aligned} \mu(x_s, x_f) &= v_{b,CRPmax}\left(\frac{x_s}{x_s + k_s + 0.2x_2^2} + \frac{x_f}{x_f + k_f}\right), \\ q_s(x_s) &= \frac{1}{Y_{b/s} + Y_{b/f}}\frac{x_s}{x_s + k_s + 0.2x_2^2}, \\ q_f(x_f) &= \frac{1}{Y_{b/s} + Y_{b/f}}\frac{x_f}{x_f + k_f}. \end{aligned} \tag{5.44}$$

The goal is to discriminate between hypotheses II and III employing the measurement data obtained from batch cultivation at a CRP set point of -50mV for all three

states [224]. To represent batch cultivation, the dilution rate D is set to zero in (5.42). We show in the following the derivation of the simplified augmented Euler system only for hypothesis III, as the derivation is analogue for hypothesis II.

From the available measurement data, we derived the following constraint set on which the error term in the augmented Euler system has to be determined.

$$\begin{aligned}\mathcal{K} := \{x(t), x_f(t), x_s(t) : 0 \leq x(t) \leq 2, 0 \leq x_f(t) \leq 3, 0 \leq x_s(t) \leq 4, \forall t \in [0,3]\text{h}, \\ 0.8 \leq x(0) \leq 0.9, 3.8 \leq x_s(0) \leq 4, 2.5 \leq x_f(0) \leq 2.8, \\ 0.9 \leq k_s \leq 1.1, 1.9 \leq k_f \leq 2.1\}.\end{aligned} \tag{5.45}$$

Note that the parameters in $\mathcal{K}$ are set to ranges corresponding to a reference parameterization of the system as obtained in [224]. This derives from the fact that for larger uncertainties in the parameters hypotheses II and III were not distinguishable w.r.t. the available measurements from the real process. Note that the initial conditions for the states are chosen in accordance to the relative errors of the measurement data reported in [224].

For generating the training data, we employed numerical simulations as an application of Theorem 8 is not possible due to the rational dynamics. Therefore, we consider only the weak inconsistency property of Proposition 6. For the numerical simulations, we took 10000 uniformly distributed Monte Carlo samples for the initial conditions and parameters as defined by $\mathcal{K}$. The samples were then simulated, i. e. the corresponding trajectories $\phi(x(t)|x_0, p)$ were determined, with Mathematica and the absolute accuracy goal was set to machine precision. As training data $\bar{\mathcal{X}}(k)$ in (5.39), we then used the minimal and maximal value of all trajectories at the time points $t_i = \{0, 0.25, 0.5, \ldots, 3\}$. The time points, were chosen such that the measurement data from the real process available for x_f and x_s are covered. From the sensitivity analysis conducted in [33], we further know that the behavior of hypothesis III does not change qualitatively for small perturbations of the nominal parameterization and initial conditions. This justifies the choice of the employed training sets in the model consistency analysis. The obtained intervals corresponding to the training data are depicted in Figure 5.5.

5.6.3 Results and Discussion

For estimating the error set according to Lemma 5, we employed the toolbox ADMIT [204] for the training set defined above. The estimated error sets for the states x, x_s and x_f are given in Table 5.1.

From Table 5.1 it is apparent that the standard Euler discretization of (5.42) for the chosen step size $D = 0.25$ approximates the continuous-time system already well as the corresponding error sets are relatively small. Additionally, one notices that the

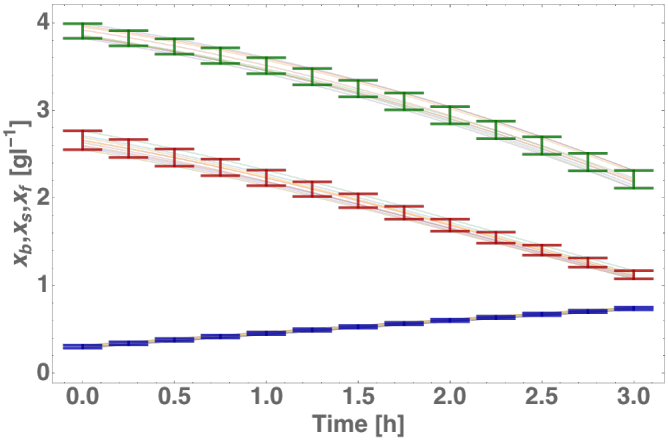

Figure 5.5: An example trainingset for the determination of the discretization errors as employed in the augmented Euler model. Intervals depict the artificial measurements, while continuous lines represent trajectories resulting from sampling initial conditions and parameters uniformly in $\mathcal{K}$. Blue corresponds to x, green to x_s, and red to x_f.

needed error sets for the substrates x_s and x_f are larger as for the biomass x, which is intuitively clear if one compares the slopes of the trajectories of the three states in Figure 5.5. This suggests that the step size should be reduced if the error sets for x_s and x_f are too large for the model invalidation task. Note that for hypothesis II similar ranges for the error sets can be reported.

In Table 5.2, we report the interval enclosure as well as the measurement data for the biomass x. The measurement data corresponds to measurements taken by fluorescence spectroscopy, where a relative error was added according to the tolerances specified for the employed measurement device. The measurement device was a two-component-system composed of a fiber optic transflexion probe and a UV-Vis spectrometer as reported in [33]. For testing whether hypothesis II and hypothesis III are consistent with the measurements, we solved (5.41) with help of the toolbox ADMIT and were able to provide a conclusive proof of model inconsistency of hypothesis II w.r.t. the training data. This example illustrates nicely that the simplified augmented Euler method can be used for model invalidation. For reference, we included in Figure 5.6 an example steady-state cultivation, whereas hypothesis III is employed in the controller strategy. Figure 5.6 shows that hypothesis III does represent the measurement data well, and that the derived optimal dilution rate is only slightly adapted by the PID controller. Both factors suggest that hypothesis III is well suited for the control of steady-state cultivations of *R. rubrum*.

Table 5.1: Estimated error sets $\hat{\mathcal{E}}(k)$ for hypothesis III.

Time [h]:	Error in x:	Error in x_s:	Error in x_f:
0.25	[-0.009,0.014]	[-0.083,0.0714]	[-0.175,0.083]
0.5	[-0.010,0.014]	[-0.091,0.0664]	[-0.172,0.079]
0.75	[-0.114,0.014]	[-0.099,0.0652]	[-0.170,0.076]
1	[-0.013,0.013]	[-0.108,0.0645]	[-0.175,0.073]
1.25	[-0.014,0.013]	[-0.117,0.0641]	[-0.180,0.070]
1.5	[-0.015,0.012]	[-0.126,0.0642]	[-0.183,0.067]
1.75	[-0.017,0.012]	[-0.135,0.0648]	[-0.184,0.065]
2	[-0.018,0.012]	[-0.144,0.0660]	[-0.184,0.064]
2.25	[-0.019,0.011]	[-0.154,0.0677]	[-0.182,0.062]
2.5	[-0.021,0.011]	[-0.163,0.0702]	[-0.177,0.061]
2.75	[-0.022,0.010]	[-0.173,0.0733]	[-0.170,0.061]
3	[-0.023,0.009]	[-0.181,0.0772]	[-0.161,0.061]

Table 5.2: Estimated enclosure of the biomass concentration x for hypothesis III, and the employed measurement data from the real process.

Time [h]:	Enclosure of x:	Measurements of x (fluorescence spectroscopy):
0.25	[0.323,0.357]	[0.351,0.360]
0.5	[0.362,0.395]	[0.360,0.373]
0.75	[0.401,0.432]	[0.400,0.415]
1	[0.439,0.470]	[0.450,0.465]
1.25	[0.476,0.507]	[0.460,0.476]
1.5	[0.514,0.543]	[0.501,0.516]
1.75	[0.551,0.579]	[0.542,0.556]
2	[0.587,0.615]	[0.594,0.611]
2.25	[0.622,0.651]	[0.648,0.661]
2.5	[0.656,0.685]	[0.660,0.670]
2.75	[0.690,0.719]	[0.709,0.727]
3	[0.722,0.752]	[0.709,0.727]

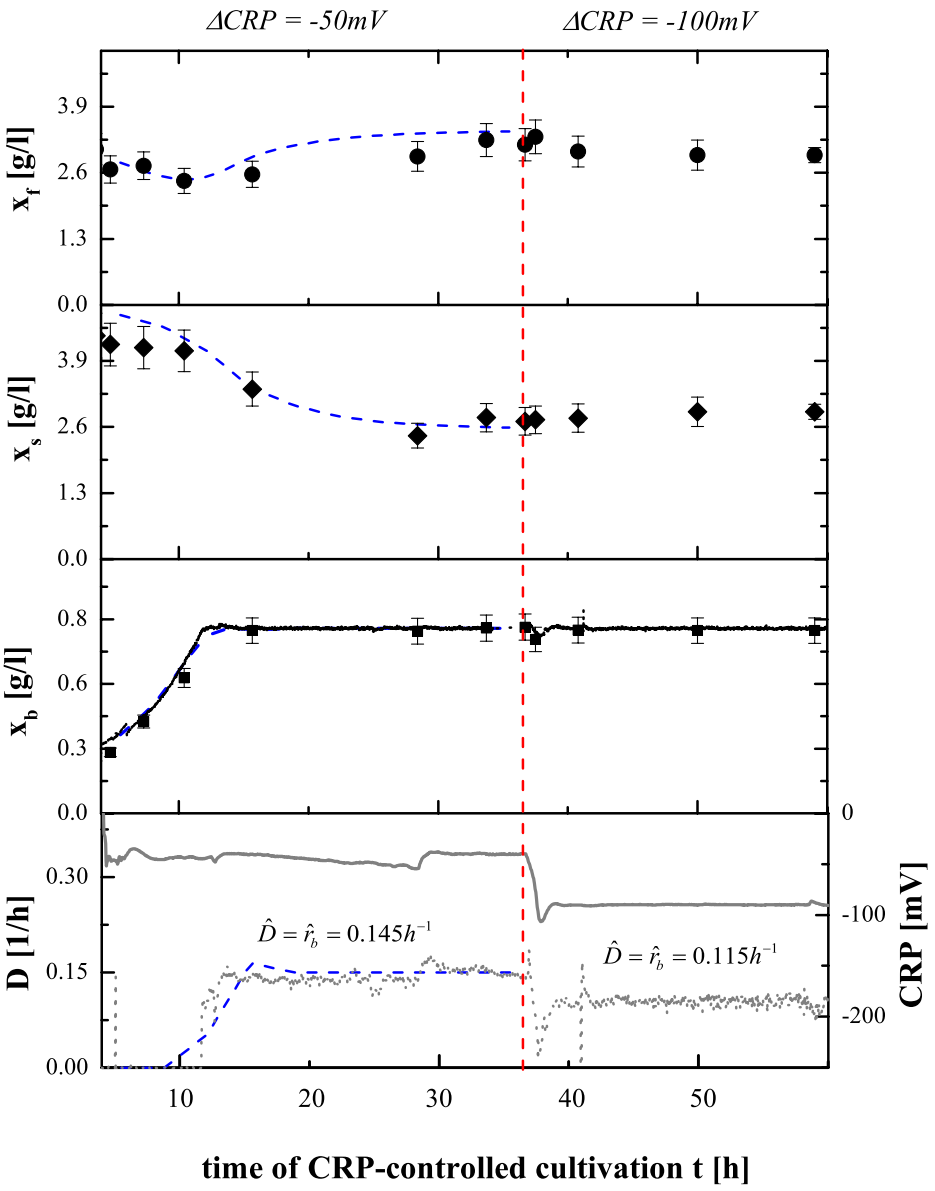

Figure 5.6: Example of a microaerobic steady-state cultivation with a CRP stepsize switch from -50mV to -100mV (red dashed line). Time course of the offline measured concentrations x_s(■),x_s(◇), x_f(●). $\hat{D}$ represents the dilution rate. The grey line corresponds to the CRP. The blue dashed line illustrates the quality of the derived model. This figure was taken from an earlier version of [33].

5.7 Summary

In this chapter, several extensions to the set-based estimation approach were proposed that allow the consideration of continuous-time systems. We focused hereby on the derivation of conditions that are needed such that estimation problems for continuous-time systems can be addressed by discrete-time systems.

The most important properties to do so are the CT/DT relationship, i. e. the condition that consistent parameters of the continuous-time system are also consistent parameters of the discrete-time system, and the embedding property, i. e. the trajectories of the continuous-time system are enclosed by the trajectories of the discrete-time system. We showed that for exact discretizations both relationships are fulfilled, however, such discretizations are in general not available for polynomial systems. Therefore, we presented two approaches to achieve these conditions for discrete-time approximations. The first approach is based on the Picard-Lindelöf theorem for the existence and uniqueness of solutions to initial value problems. For this reason, we developed a method that results in an guaranteed enclosure of the reachable set of the continuous-time system. Based on this enclosure we presented an algorithm to derive a time step size such that the Picard-Lindelöf theorem is fulfilled. The resulting discrete-time system then fulfills the strong CT/DT relationship and can be directly employed for the estimation of the continuous-time system. However, the time step size is typically relatively small such that the estimation procedures from Chapter 4 become computationally demanding.

To tackle this problem, we proposed first a procedure that is independent of a time-step size by approximating the reachable set of the continuous-time directly. However, as this approach does not allow to study the connection of discrete-time and continuous-time systems, we proposed a simplified Euler formulation that encloses the trajectories of a continuous-time system with respect to a training set. We discussed several approaches to generate appropriate training sets, and we provided a strategy to test consistency of a continuous-time system. Furthermore, we exemplified with a real world example, namely the steady-state cultivation of *Rhodospirillum rubrum*, that the simplified Euler system can be applied to derive a quantitative and predictive model.

6 Observability of Polynomial Systems

In this chapter, we investigate observability of polynomial systems. In general, observability refers to the possibility to reconstruct from given measurement data the initial conditions of a system. To investigate observability we consider at first the algebraic conditions for local-at-a-point observability presented by Kawano and Ohtsuka in [94]. We show that these conditions can be addressed by polynomial optimization for nominal systems and that the optimization problem is finite-dimensional in case the system is local-at-a-point observable. Moreover, we show that the local conditions are not sufficient to address observability in the uncertain case. This derives from the fact that observability for uncertain systems corresponds to the ability to reduce the initial bounds on the involved variables, in particular, the states and parameters. To consider uncertain systems, we have to rely on a stronger condition, namely global observability as presented in [94]. However, as we are working with states and parameters that are unknown-but-bounded the algebraic conditions of [94] become semi-algebraic. To handle these semi-algebraic conditions, we provide a sufficient condition that can again be checked with polynomial optimization. Another difficulty in considering uncertain systems derives from the fact that the algebraic conditions are not necessary. For instance, we cannot conclude from a system not being globally observable to a loss in the ability to reduce the initial bounds. To study this problem, we provide a systematic procedure to investigate the state and parameter space for partitions that lead to the loss of observability.

To complement the extended algebraic conditions, we present a second method that links the energy visible at the output to the initial conditions. This idea is based on the idea of observability Gramians, i. e. a high output energy corresponds to an easier observable system. The link between energy and initial conditions is then used to derive a set-based observability notion and an algorithmic procedure to study this notion is presented.

This chapter is structured as follows. First, the considered algebraic framework of Kawano and Ohtsuka is introduced and it is shown that this framework can be addressed by polynomial optimization. Second, the algebraic conditions are extended to uncertain systems and the involved difficulties are illustrated. Third, based on output norms an alternative approach to investigate observability of uncertain systems is introduced. Both approaches for uncertain systems are exemplified by a two-tank system, before concluding this chapter.

6.1 Introduction: Observability

To ensure the requirements of modern industry, such as high product quality or safety, modelling, monitoring and control becomes increasingly important. Typically, as the considered process becomes more complex, more information on the process characteristics is needed to complete these three tasks. However, not all of these characteristics are equally easy accessible. For instance, while for pressure and pH cheap and accurate sensors are available, for other key variables like the biomass concentration more sophisticated measurement devices are needed that might have limitations as long sampling times, maintenance costs, long processing times and similar difficulties.

A possibility to circumvent some of these problems is to employ state observers to reconstruct the desired information from the available measurement setup with help of a dynamical model instead of directly measuring all process characteristics. Several observers or soft sensors have been proposed in literature and design approaches can be found in many articles and standard text books, see e. g. [24, 56, 60, 85, 102, 125, 134, 159, 186]. The crucial point is, however, that the employed model has to be observable, i. e. the states of the model have to be reconstructable from ideal measurements.

When speaking of observability, one typically speaks about a nominal system, i. e. a system for which the parameters are known. In this case observability is quite well understood for linear systems, see e. g. [29, 84, 138]. For nonlinear systems, it is typically more involved to define suitable observability notions [78, 192] and only few computational techniques are known, e. g. [192]. For polynomial systems the situation is more favorable, see e. g. [15, 55, 72, 189, 190, 194] for general definitions and requirements, and [89–92, 94, 143, 212] for constructive algebraic conditions.

For uncertain systems, however, only few results exist and are mainly restricted to the field of differential inclusions [11, 88]. Here, the main focus is to investigate what properties an uncertain system has to exhibit such that observability can be deduced if the uncertainty is removed. This is mainly due to the difficulty of defining observability and distinguishability notions for uncertain systems due to the lack of a sharp uniqueness notion in this case. One possibility to avoid this lack is to relate the norm of the output to the norm of the states. This builds the foundation for controller design methods based on the output-to-state stability concept, see e. g. [8, 79, 196] and references therein. Other applications of output norms are observability Gramians in observability analysis. Such Gramians can be used to quantify observability of linear systems e. g. as needed in model reduction. Several extensions of Gramians have been proposed in the literature, i. a. [70, 81, 107, 199] for nonlinear systems and [154, 172, 189, 218] for uncertain linear systems.

In this work, two approaches are considered for observability analysis. The first one is based on the algebraic conditions for the analysis of nominal systems derived in [89–94]. There are mainly two reasons why to base a consideration on these works. First of all their methods can be verified by polynomial optimization as employed in

this work. Second, the general description of observability in terms of ideals allows a uniform treatment of continuous-time and discrete-time systems. However, we show that by extending these results to uncertain systems only a sufficient condition for observability can be derived. The second presented approach is based on the energy of the output [66]. For this reason, we derive a relationship between the output energy and the initial conditions similar to the norm-observability concept in [79]. This allows us to partition the state space into subsets that lead to high output energy and into subsets that lead to low output energy. Here, low output energy corresponds to a system that is difficult to observe in general.

6.2 Observability

In this section several conditions to determine observability for uncertain systems are derived. In the nominal case observability of a polynomial system corresponds to properties of the variety corresponding to an ideal generated by the output map. We present in the following a short summary of these properties following [92]. One of the main advantages to derive such observability conditions is that after the derivation of the ideal continuous-time and discrete-time systems can be handled in a unified way. The methods presented here are, therefore, kept general to avoid unnecessary repetitions. For the obtained results, we have to distinguish between the nominal case and the uncertain case. To demonstrate that observability of a uncertain system can be addressed with the methods presented in this work, we first show that the nominal case can be tackled by polynomial optimization as seen next.

6.2.1 Nominal Model

To clarify, a nominal system refers to a system of form (2.1) or (2.2) for which the initial condition and the parameters are fixed, i. e. $x_0 \in \mathbb{R}^{n_x}, p \in \mathbb{R}^{n_p}$. Furthermore, for simplicity of notation, the input is assumed to be zero, however, considering a nonzero input is straightforward, see e. g. [94]. Another simplification is that the parameters p in the above equations are added to the state vector . In other words, we consider continuous-time systems of form

$$\begin{aligned} \dot{x}(t) &= f(x(t)), \quad x(t_0) = x_0, \\ y(t) &= h(x(t)), \end{aligned} \tag{6.1}$$

and discrete-time systems of form

$$\begin{aligned} x(k+1) &= f_D(x(k)), \quad x(0) = x_0, \\ y(k) &= h_D(x(k)), \end{aligned} \tag{6.2}$$

where $\dot{x}_i(t) = 0$ (resp. $x_i(k+1) = x_i(k)$) holds for every state x_i corresponding to a time invariant parameter.

We begin with local-at-a-point observability. Recall that local-at-a-point observability corresponds to the following (cf. also to Definition 3 in Section 2.2.3): if for an initial condition α there exists an open neighborhood $\mathcal{N}$ such that for every initial condition $\beta \in \mathcal{N} \setminus \{a\}$ the output trajectories of the considered system fulfill $\phi_h(x(t)|\alpha, u(t), p) \neq \phi_h(x(t)|\beta, p)$ for some $t \in [0, T]$ (resp. $\phi_{D,y}(x(k)|\alpha, p) \neq \phi_{D,y}(x(k)|\beta, u(k), p)$ for some $k \in \{1, \ldots, n\}, n \in \mathbb{N}$). In other words, the initial conditions α and β are distinguishable according to Definition 2 and for the set of indistinguishable initial conditions $\mathscr{O}(\alpha, \mathcal{A}) \cap \mathcal{N} = \{\alpha\}$ holds. We follow here the most common approach to test this property. For continuous-time systems this test employs Lie derivatives, see e. g. [78], and the so-called *observability mapping*

$$O(x) = \begin{pmatrix} h(x) \\ L_f h(x) \\ L_f^2 h(x) \\ \vdots \\ L_f^{n_O} h(x) \end{pmatrix}. \tag{6.3}$$

Here, the Lie derivative $L_f h(x)$ is defined as $L_f h(x) := \left(\frac{\partial}{\partial x} h(x)\right) f(x)$ and the notation $L_f^n h(x)$ corresponds to $L_f(L_f^{n-1} h(x))$.

For discrete-time systems the equivalent mapping is defined by

$$O_{DT}(x) = \begin{pmatrix} h_T(x) \\ h_T(f_D(x)) \\ h_T(f_D^2(x)) \\ \vdots \\ h_T(f_D^{n_O}(x)) \end{pmatrix}, \tag{6.4}$$

where the notation $f_D^n(x)$ corresponds to $f_D(f_D^{n-1}(x))$.

Note that for simplicity and to avoid redundancies we sometimes drop the index DT, whenever it is clear from the context to which observability mapping we refer. Furthermore, we drop the time argument and time index, respectively.

If the observability mapping is locally invertible, then a system has the local distinguishability property. Consider the following sufficient condition based on the inverse function theorem, which holds for continuous-time and discrete-time systems.

Theorem 9 (Observability Rank Condition [78],[146])**.** *The system* (2.1) *(resp.* (2.2)*) has the local distinguishability property at an initial condition $a \in \mathcal{K}$ if $rank(O(a)) = n_x$ (resp. $rank(O_{DT}(a)) = n_x$).* ∎

Based on the introduced observability mappings, it was shown in [94] that necessary algebraic conditions can be derived that extend Theorem 9 and can be easily verified

with symbolic software, e. g. Macaulay2, Singular. The main idea to do so is to form an ideal from the observability mappings based on the notion of indistinguishability. This leads to the following two ideals for two initial conditions α and β

$$\mathscr{I} := \langle O_1(\alpha) - O_1(\beta)\rangle + \ldots + \langle O_{n_O}(\alpha) - O_{n_O}(\beta)\rangle \subseteq \mathbb{R}[\alpha, \beta] \tag{6.5}$$

and

$$\mathscr{I}_{DT} := \langle O_{DT,1}(\alpha) - O_{DT,1}(\beta)\rangle + \ldots + \langle O_{DT,n_O}(\alpha) - O_{DT,n_O}(\beta)\rangle \subseteq \mathbb{R}[\alpha, \beta], \tag{6.6}$$

where the index i refers to the i-th row of O or O_{DT}, respectively. The associated varieties are denoted by $\mathcal{V}(\mathscr{I}) \subseteq \mathbb{R}^{n_y}$ and $\mathcal{V}(\mathscr{I}_{DT}) \subseteq \mathbb{R}^{n_y}$, i. e. the algebraic sets where every polynomial $p \in \mathscr{I}$ (resp. $p \in \mathscr{I}_{DT}$) vanishes. For a detailed discussion of ideals and varieties the interested reader is referred to [9, 40].

For both ideals, we have the following fact from Hilbert's basis theorem.

Lemma 2. *The ideals $\mathscr{I}$ and $\mathscr{I}_{DT}$ are finitely generated.* ∎

This means for some large enough n_O, the observability of system (2.1) (resp. (2.2)) can be investigated via a finite set of polynomials. This derives from the fact that the ideals are formed by an ascending chain of ideals, e. g. for $\mathscr{I}_{DT}$ we have:

$$\begin{aligned} \mathscr{I}_{DT,1} &:= \langle h(\alpha) - h(\beta)\rangle, \\ \mathscr{I}_{DT,2} &:= \mathscr{I}_{DT,1} + \langle h(f_D(\alpha)) - h(f_D(\beta))\rangle, \\ &\vdots \\ \mathscr{I}_{DT,n_O} &:= \mathscr{I}_{DT,n_O-1} + \langle h(f_D^n(\alpha)) - h(f_D^n(\beta))\rangle, \end{aligned} \tag{6.7}$$

where $\mathscr{I}_{DT,1} \subset I_{DT,2} \subset \cdots \subset \mathscr{I}_{DT,n_O} = \mathscr{I}_{DT,n_O+1} = \mathscr{I}_{DT}$ holds.

Kawano and Othsuka showed in [89, 90, 92] that these ideals and the corresponding varieties allow to consider all introduced observability notions (cf. also to Section 2.2.3). To do so, consider the prime ideal

$$\mathfrak{p} = \langle x_1 - \alpha_1, \ldots, x_n - \alpha_n\rangle, \tag{6.8}$$

which allows the study of local properties w.r.t. the point $\alpha = (\alpha_1, \ldots, \alpha_{n_x})^T \in \mathbb{R}^{n_x}$.

The first result of interest is local-at-a-point observability. For this reason, consider a ring homomorphism $\gamma_\beta : \mathbb{R}[a, b] \to \mathbb{R}[a]$ which substitutes the variable b in a polynomial of $\mathbb{R}[a, b]$ with a fixed initial condition $\beta \in \mathbb{R}^{n_x}$. Assume that the ideal $\mathscr{I}$ constructed from the observability mapping is as in (6.5) or (6.6), respectively. The localization of $\mathscr{I}$ at $\mathfrak{p}$ is defined as

$$\mathscr{I}_{\mathfrak{p}} := \{a/b : a \in \mathscr{I}, b \in \mathbb{R}[x], b \notin \mathfrak{p}\}. \tag{6.9}$$

For an introduction to local rings, see e. g. [9]. In [94] it is shown that using a ring

homomorphism and the localization of an ideal local-at-a-point observability can be analyzed as follows.

Theorem 10 (Local-at-a-point Observability [94])**.** *A system* (2.1) *(resp.* (2.2)*) is locally observable at initial condition* $\beta \in \mathcal{K} \subseteq \mathbb{R}^{n_x}$ *if and only if*

$$\mathcal{V}(\gamma_\beta(\mathscr{I})_{\gamma_\beta(\mathfrak{p})}) = \mathcal{V}(\gamma_\beta(\mathfrak{p})) = \{\beta\}, \tag{6.10}$$

where $\mathscr{I}$ *is defined as in* (6.5) *(resp.* (6.6)*).* ∎

Theorem 10 basically describes that if only a single point β is investigated then a system is observable if $a \equiv \beta$ is the only zero of the localization $\gamma_\beta(\mathscr{I})_{\gamma_\beta(\mathfrak{p})}$. Therefore, local-at-a-point observability is equivalent to the determination of the solution set of polynomial equalities describing the variety of the localization.

This theorem can be illustrated by considering the actual computation of a localization at a prime ideal as seen in the following example. Note that the computation is done in two-steps: first, a primary decomposition is computed and second, the ideal-membership of the decomposition w.r.t. to $\mathfrak{p}$ is tested.

Example 5 (Local-at-a-point Observability [92])**.** *Consider the discrete-time system*

$$\begin{aligned} x(t+1) &= \begin{pmatrix} 2x_3^2(t) - x_3^3(t) \\ x_1(t) + x_2^2(t) \\ x_1^2(t) \end{pmatrix}, \\ y &= x_2(t). \end{aligned} \tag{6.11}$$

Then $\mathscr{I} = \langle a_1 - b_1, a_1 + a_2^2 - (b_1 + b_2^2), 2a_3^2 - a_3^3 + (a_1 + a_2^2)^2 - (2b_3^2 - b_3^3 + (b_1 + b_2^2)^2)\rangle$ *and* $\mathfrak{p}$ *is as in* (6.8) *for* $n = 3$*. As a first step a primary decomposition of* $\mathscr{I}$ *is derived, e. g. with Macaulay2. For this example, we get*

$$\mathscr{I} = \langle a_1 - b_1, a_2 - b_2, a_3 - b_3\rangle \cap \langle -2a_3 + a_3^2 - 2b_3 + a_3 b_3 + b_3^2\rangle. \tag{6.12}$$

Note that the first ideal in the primary decomposition (6.12) *is* $\mathfrak{p}$*. Now the second ideal of the primary decomposition is tested for ideal-membership in* $\mathfrak{p}$ *to derive the localization [9]. We denote the second ideal in the following by* $p(a_3, b_3)$*. It is easily verified that the polynomial* $p(a_3, b_3)$ *is not in* $\mathfrak{p}$ *with help of a Gröbner basis, therefore, it follows that* $\mathscr{I}_\mathfrak{p} = \mathfrak{p}$ *and the condition of Theorem 10 trivially holds for every every replacement of* b *with* β*. This example illustrates nicely that the localization at a prime ideal allows to study the behavior of an ideal close to a point as the dependency* $p(a_3, b_3)$ *is filtered out.* ∎

The previous theorem shows that observability for polynomial systems can be derived from the zeros of finitely many polynomials. Determining this set of polynomials, deriving the localization and quotients can in this setting be done by Gröbner bases

and related algorithms, see e. g. [40]. To determine the varieties (the zeros (or roots) of the polynomial equalities) of a polynomial equation system several possibilities are known. For an overview and relevant notions we refer to the text books [39, 40]. For an interesting interpretation in terms of eigenvalue computations and realization theory we refer to [49]. However, as soon as the algebraic setting is left and polynomial inequalities are introduced in the formulation the problems become semi-algebraic and, therefore, computationally more involved as seen in the next section. We follow, therefore, the general idea of this work and formulate the conditions on the varieties in terms of a feasibility problem of form

$$\begin{aligned} \text{find} \quad & a \\ \text{subject to} \quad & f(a,\beta) = 0, \forall f \in \gamma_\beta(\mathscr{I})_{\gamma_\beta(\mathfrak{p})}, \\ & a \in \mathcal{K}, \end{aligned} \tag{6.13}$$

where the constraint set $\mathcal{K} \subseteq \mathbb{R}^{n_x}$ is again a semi-algebraic set as defined in previous chapters.

From construction it follows that the solution set of (6.13) describes the zeros of all polynomials in $\gamma_\beta(\mathscr{I})_{\gamma_\beta(\mathfrak{p})}$, i. e. it is equal to the variety $\mathcal{V}(\gamma_\beta(\mathscr{I})_{\gamma_\beta(\mathfrak{p})}) := \{a \in \mathcal{K} : f(a,\beta) = 0, \forall f \in \gamma_\beta(\mathscr{I})_{\gamma_\beta(\mathfrak{p})}\}$. From Lemma 2 we can rewrite (6.13) to obtain a system with finitely many constraints

$$\begin{aligned} \text{find} \quad & a \\ \text{subject to} \quad & g_i(a,\beta) = 0, i \in \{1,\dots,n_{\mathscr{I}}\}, \\ & a \in \mathcal{K}, \end{aligned} \tag{6.14}$$

where g_i are the generators of the localization of $\mathscr{I}$ at $\mathfrak{p}$, e. g. a Gröbner basis.

Following the relaxation procedure in Section 4.3.1, we derive next an approximation of the solution set. To be able to state the results, we shortly recall the idea of outer-bounding, see also Algorithm 2. In (6.14) the variable of interest is a. Determining an outer-bounding box of the admissible values of a constitutes in minimizing (resp. maximizing) every entry in the vector a as follows.

$$\begin{aligned} \min \quad & c^T a \\ \text{subject to} \quad & g_i(a) = 0, i \in \{1,\dots,n_{\mathscr{I}}\}, \\ & a \in \mathcal{K}, \end{aligned} \tag{6.15}$$

where c^T is a vector with one non-zero entry in the i-th column, whereas the non-zero entry is either 1 or -1. This then corresponds to finding a lower respective upper bound for a_i. Doing so for all a_i yields the outer-bounding box. As (6.15) is in the same form as the polynomial optimization problem (4.7) in Section 4.3.1 a detailed derivation is omitted. We refer to the optimal solution of (6.15) as a^*_{POP} and of the d-th semi-definite relaxation as $a^*_{SDP_d}$. The lower and upper bounds obtained by solving the semi-definite relaxation are denoted by $\underline{a}^*_{SDP_d}$ and $\overline{a}^*_{SDP_d}$. We can state the following

strong convergence result.

Theorem 11 (Finite Convergence of Relaxation)**.** *If $\mathcal{V}(\mathscr{I})$ is a finite-dimensional variety, i. e. the generators g_i, $i \in \{1, \ldots, n_{\mathscr{I}}\}$ have finitely many common zeros, there exists $d \in \mathbb{R}$ such that $a^*_{POP} = a^*_{SDP_d}$.* ∎

Proof. The proof is based on Theorem 6.15 of [114]. Let y be a feasible solution (4.9). We show that $p^T y \geq p^*$ for a sufficiently large relaxation degree. Note that from the localizing matrix we know that the relaxation degree has to fulfill $t \geq 2d_{F_j}$. Furthermore, for a sufficiently large t, the kernel of $M_t(y)$ contains any given finite set of polynomials in the ideal of all functions vanishing on the observability mapping (see also Claim 6.16 in [114]). Let $\{f_1, \ldots, f_L\}$ be a Gröbner basis of the ideal of the variety (denoted with $\mathscr{I}(\mathcal{V}(\mathscr{I}))$) with a degree monomial ordering and $\mathscr{B}$ a basis of the radical of the quotient algebra $\mathbb{R}[x] \setminus \mathscr{I}(\mathcal{V}(\mathscr{I}))$ with maximal degree $d_{\mathcal{B}} := \max_{b \in \mathscr{B}} deg(b)$. Note that the variety is finite, therefore, $d_{\mathscr{B}}$ is well defined. Given $\epsilon > 0$ and considering Putinar's Positivstellensatz, we can write

$$p - p^* + \epsilon = s_0 + \sum_{j=1}^{m} s_j F_j + q, \tag{6.16}$$

where s_0, s_j are sum of squares and $q \in \mathscr{I}$. Let $s_j = \sum_i s_{i,j}^2$, where $s_{i,j} = r_{i,j} + q_{i,j}$, $r_{i,j} = \sum_k a_k b_k$ with $deg(r_{i,j}) \leq d_{\mathscr{B}}$ for $a_k \in \mathbb{R}$ and b_k some member of $\mathscr{B}$, and $q_{i,j} \in \mathscr{I}$. We can determine in this way a second decomposition of form (6.16) with s'_0, s'_j being sum of squares of degree at most $2d_{\mathscr{B}}$ and $q' \in \mathscr{I}(\mathscr{V}(\mathscr{I}))$. If we now define $T_0 := \max(2d_p, 2d_{\mathscr{B}}, 2d_{\mathscr{B}} + F_1, \ldots, 2d_{\mathscr{B}} + F_j)$, then we have $deg(s'_0), deg(s'_j), deg(q') \leq T_0$. If we write $q' = \sum_{l=1} L u_l f_l$ with $deg(u_l f_l) \leq T_0$, then we can find a t such that $u_l f_l$ is in the kernel of $M_t(y)$ (see also Claim 6.16 in [114]) and thus q' is also in the kernel of $M_t(y)$. From Putinar's Positivstellensatz we have $1^T M_t(y) vec(s_j F_j) = 1^T M_t(y) vec(p - p^* + \epsilon) \geq 0$. Therefore, $1^T M_t(y) p \geq p^* - \epsilon$, thus $p_t \geq p^* - \epsilon$ for vanishing ϵ and noting that p_t is lower bound of p^*, i. e. $p_t \leq p^*$, we have the desired result $p_t = p^*$. □

There are some limitations to the previous theorem. Namely, the assumption that the variety corresponding to the observability mapping, i. e. $\mathcal{V}(\gamma_\beta(\mathscr{I}))_{\gamma_\beta(\mathfrak{p})}$ is finite, can be restrictive as seen next for the uncertain case. However, if β is indeed an locally observable initial condition for (2.1) (resp. (2.2)) we have the following verifiable necessary and sufficient condition.

Corollary 1 (Local-at-a-point Observability (Nominal System))**.** *A system (2.1) (resp. (2.2)) is locally observable at initial condition a, if and only if the lower and upper bounds of each entry a_i are equal, i. e. $\underline{a}^*_{SDP_d} = \overline{a}^*_{SDP_d}$. Furthermore, this can be determined by $2n_x$ finite-dimensional semi-definite programs of form (4.12).* ∎

Corollary 1 simply states that for an observable initial condition β the condition of Theorem 11 holds and observability can, therefore, be verified. It should be noted

that Corollary 1 is formulated for Algorithm 2, however, in practice the condition $a^*_{POP} = \beta$ can also be checked by extracting the global minimizer of (6.15) employing the methods described in [75]. Unfortunately, to prove that a system is not observable is in general harder as finite convergence can in general not be assumed. However, based on the Curto-Fialkow flat extension criterion [41], we can state a sufficient criterion for a system not being observable as follows, see also [75] and Appendix A.1 for more details.

Proposition 7 (Sufficient Condition for local-at-a-point Unobservability (Nominal System))**.** *A system* (2.1) *(resp.* (2.2)*) is not observable if for some s the conditions rank* $M_{s-v}(y^*) = rank\ M_s(y^*)$ *and* $\underline{a}^*_{SDP_d} \neq \overline{a}^*_{SDP_d}$ *hold.* ∎

Proof. The rank condition ensures that the global optimum has been obtained [75], therefore, if $\underline{a}^*_{SDP_d} \neq \overline{a}^*_{SDP_d}$ the condition $\mathcal{V}(\gamma_\beta(\mathscr{I})_{\gamma_\beta(\mathfrak{p})}) = \mathcal{V}(\gamma_\beta(\mathfrak{p}))$ of Theorem 10 does not hold. □

Note that the rank condition in Proposition 7 is a condition that is verified for the moment formulation (4.12) in Section 4.3.1.

The previous results show that the observability of a nominal system can be checked by solving polynomial optimization problems. Furthermore, the local-at-a-point observability property guarantees finite convergence of the relaxation approach presented in Section 4.3.1. The next section deals with the problem, when the variety corresponding to the prime ideal $\mathfrak{p}$ is no longer a point, i. e. if an uncertain system is considered. To do so, we incorporate in the observability analysis methods used for approximating the volume of sets following [77]. Furthermore, we show that based on the global observability condition presented in [94], observability can be studied on semi-algebraic sets.

6.2.2 Uncertain Systems

If we proceed to the uncertain case, we basically leave the algebraic conditions considered before and several adaptations have to be made. In the uncertain case, observability corresponds to the principle possibility to reduce the initial bounds on the considered unknown-but-bounded variables, i. e. states and parameters, when measurements are available. We denote this possibility in the following as set-observability.

The first intuition to consider uncertainties, would be to change the variety in Theorem 10 to

$$\mathcal{V}(\gamma_{\mathcal{B}}(\mathscr{I})) := \bigcup_{\beta \in \mathcal{B}} \mathcal{V}(\gamma_\beta(\mathscr{I})_{\gamma_\beta \mathfrak{p}})), \tag{6.17}$$

where $\mathcal{B}$ is a semi-algebraic set. However, in this case we would come to a wrong conclusion regarding set-observability. To exemplify this circumstance consider again Example 5. There it is shown that the nonlinear connection of the initial conditions in the primary decomposition is filtered out by the localization at the prime ideal $\mathfrak{p}$.

This nonlinear connection is exactly what limits the ability to reduce the bounds on a variable in the end. Before providing an illustration, we summarize the stronger condition needed for uncertain systems.

Theorem 12 (Global Observability [92])**.** *A continuous-time system* (2.1) *(resp. discrete-time system* (2.2)*) is globally observable, then the following holds*

$$\mathcal{V}(\sqrt{\mathscr{I}} : \mathfrak{p}) = \mathcal{V}(\mathfrak{p}), \tag{6.18}$$

where $\sqrt{\mathscr{I}} : \mathfrak{p}$ *is the colon ideal of the radical of* $\mathscr{I}$ *and the prime ideal* $\mathfrak{p}$*, i. e.* $\sqrt{\mathscr{I}} := \{f \in \mathbb{R}[x] : f^m \in \mathscr{I}, m \geq 1\}$ *and* $\sqrt{\mathscr{I}} : \mathfrak{p} := \{f \in \mathbb{R}[x] : fg \in \sqrt{\mathscr{I}}, \forall g \in \mathfrak{p}\}$. ∎

The following example illustrates why global observability is needed for uncertain systems and local observability notions are not sufficient.

Example 6 (Global Observability [92])**.** *Consider again the system*

$$\begin{aligned} x(t+1) &= \begin{pmatrix} 2x_3^2(t) - x_3^3(t) \\ x_1(t) + x_2^2(t) \\ x_1^2(t) \end{pmatrix}, \\ y &= x_2(t). \end{aligned} \tag{6.19}$$

As seen in Example 5 the system is local-at-a-point observable as the nonlinear dependency $p(a_3, b_3) = -2a_3 + a_3^2 - 2b_3 + a_3b_3 + b_3^2$ *between the initial conditions for state* x_3 *vanishes from the observability ideal when localized at the prime ideal. Note that the system is local-at-a-point observable as soon as an initial condition is fixed, in fact for any chosen point the system is local-at-a-point observable, i. e. the system is locally observable according to Definition 4.*

We study now global observability according to Theorem 12. The colon ideal for this example is given by

$$\sqrt{\mathscr{I}} : \mathfrak{p} = \langle a_1 - b_1, a_2 - b_2, p(a_3, b_3) \rangle. \tag{6.20}$$

Clearly, the condition for the varieties does not hold as the third generator $p(a_3, b_3)$ *allows more zeros than the third generator of* $\mathfrak{p}$*. Indeed,* $p(a_3, b_3) = 0$ *describes an ellipse in the space spanned by* a_3, b_3 *and every initial condition chosen on this ellipse leads to the same output trajectory. This dependency is depicted in Figure 6.1.*

We illustrate next that local observability is not sufficient to be able to reduce the initial bounds on the variables by measurements. For this reason, consider the sets

$$\begin{aligned} \mathcal{A}_0 &:= \{a_3 : -2 \leq a_3 \leq 2\}, \\ \mathcal{B}_0 &:= \{b_3 : -0.5 \leq b_3 \leq 0\}, \end{aligned} \tag{6.21}$$

depicted in Figure 6.1 as the blue dashed box. We assume that the set $\mathcal{B}_0$ *describes the influence of a measurement on* x_3*. As the system is locally observable by estimating the admissible values of* x_3 *one expects the same bounds as* $\mathcal{B}_0$*. However, from* $p(a_3, b_3)$

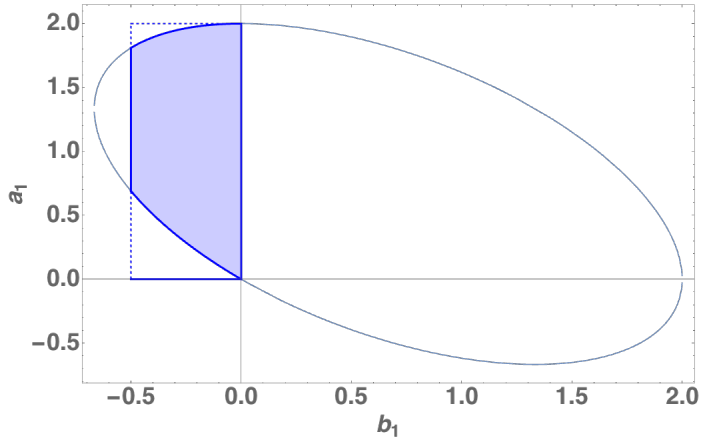

Figure 6.1: Illustration of the nonlinear dependency $p(a_3, b_3)$. Dashed lines indicate the bounding box given by $\mathcal{A}_0$ and $\mathcal{B}_0$. Shaded area depicts a practical loss of global observability, due to uncertainties.

follows that x_3 cannot be reduced further then the interval $[0, 2]$ as seen in Figure 6.1. Therefore, local observability is not sufficient to deduce the ability to reduce the initial bounds on x_3. ∎

As seen in the previous example, we need to consider the varieties of Theorem 12. Furthermore, we can state the following trivial consequence.

Corollary 2 (Set-Observability)**.** *If a continuous-time system* (6.1) *(resp. discrete-time system* (6.2)*) is globally observable, it is set-observable.*

The previous result follows from the fact that by reducing the uncertainty of the output to zero, basically a nominal system is considered. However, if a system is not globally observable it does not follow that the system is not set-observable. To study this circumstance, we provide in the following two suitable procedures based on the determination of $\mathcal{V}(\sqrt{\mathscr{I}} : \mathfrak{p})$.

Determining the variety $\mathcal{V}(\sqrt{\mathscr{I}} : \mathfrak{p})$ in the uncertain case is considerably harder as determining the varieties in the nominal case. In particular, if convergence of the approximation to the variety is required. The first procedure is based on the integration of polynomial functions over semi-algebraic sets as presented in [77]. This method deviates considerably from the methods presented in Section 4.3.1. We, therefore, give a short overview of the involved steps. The employed methods to do so are similar to the determination of the one-step reachable sets in Section 5.4.2, whereas the main difference derives from the lack of a dynamic component.

We consider here the constraint set $\mathcal{K}$ which contains the inequality constraints as before and in addition the constraints corresponding to the variety $\mathcal{V}(\sqrt{\mathscr{I}} : \mathfrak{p})$, i. e.

$$\mathcal{K} := \{x \in \mathbb{R}^{n_x} : g_i(x) = 0, g_j(x) \geq 0, i \in \{1, \ldots, n_{\mathscr{I}}\}, j = \{1, \ldots, n_{\mathcal{K}}\}\}. \tag{6.22}$$

As the procedure requires the Lebesgue moments of a second set, we assume without loss of generality that $\mathcal{K} \subset \bar{\mathcal{K}} \subseteq [-1,1]^{n_x}$, where $\bar{\mathcal{K}}$ is a set for which the Lebesgue measure can be easily determined, e. g. the unit box or the unit ball. Note that the polynomials defining $\mathcal{K}$ can always be scaled such that the assumed condition holds.

As before let $\mathscr{B}$ denote the Borel σ-algebra of Borel subsets of $\bar{\mathcal{K}}$. We denote the set of finite Borel measures on $\mathcal{K}$ with $\mathscr{M}(\mathcal{K})$. The normalized Lebesgue measure of $\bar{\mathcal{K}}$ is denoted by μ_2, where normalized means that $2^n \mu_2(\bar{\mathcal{K}}) = \lambda(\bar{\mathcal{K}})$ and $\lambda(\bar{\mathcal{K}})$ is the Lebesgue measure. The second measure $\mu_1 \in \mathscr{M}(\mathcal{K})$ corresponds to the variety we want to determine. Note that by the Radon-Nykodym theorem, there exists a nonnegative measurable function $f(x)$ that fulfills $\mu_1(\mathcal{B}) = \int_{\mathcal{B}} f(x)\mu_2(dx)$ for some subset $\mathcal{B} \in \mathscr{B}$. Consider the infinite-dimensional linear program:

$$\begin{aligned} \sup \quad & \int_{\mathcal{K}} 1 \mu_1(dx) \\ \text{subject to} \quad & \mu_1 \leq \mu_2, \\ & \mu_1 \in \mathscr{M}(\mathcal{K}), \end{aligned} \tag{6.23}$$

where $\mu_1 \leq \mu_2$ ensures that μ_1 is absolutely continuous.

Similar to Theorem 5 we can state the correspondence between (6.23) and the hypervolume of $\mathcal{K}$.

Theorem 13 (Hypervolume of Observability Variety)**.** *The optimal value l^* of* (6.23) *is equal to the hypervolume of $\mathcal{K}$. Moreover, the supremum is attained by the restriction $\mu_1^*(\mathcal{B}) = \mu_2(\mathcal{B} \cap \mathcal{K})$, $\forall \mathcal{B} \in \mathscr{B}$ to $\mathcal{K}$. In particular, the optimal value is equal to the Lebesgue measure divided by the normalization, i. e. $l^* = \lambda(\mathcal{K})/2^n$.* ∎

Proof. For a proof see [77, Thm 3.1] and setting $p = 1$. □

As the hypervolume is not a criterion that can be uniquely related to observability we next derive the dual to find an approximation of the indicator function of $\mathcal{K}$.

$$\begin{aligned} \inf \quad & \int v(x)\ \mu_2(dx) \\ \text{subject to} \quad & v(x) \geq 1 \text{ on } \operatorname{supp}\mu_2, \end{aligned} \tag{6.24}$$

where $v(x) \in \mathscr{C}$ is a continuous test function.

For completeness, we also state the dual of the moment relaxation as its sum-of-squares restriction, which is used for the computation of the example in the next section. For details on the primal moment problem we refer to Section 4.3.1 and the references therein.

$$\begin{aligned} \min \quad & l v_c \\ \text{subject to} \quad & v(x) - 1 = s_0(x) + \sum_{i=1}^{n_{\mathcal{X}}} s_i(x) g_i(x), \\ & v(x), \sigma_i \in \Sigma_d^2[x]. \end{aligned} \tag{6.25}$$

Here l are the Lebesgue moments and v_c the coefficients of a polynomial of degree corresponding to the relaxation degree r.

We define $\hat{\mathcal{V}} := \{v(x) \geq 1\}$ and we have the following result.

Theorem 14 (Convergence). *The function $v = lv_c$ that minimizes* (6.25) *has the property that for $r \to \infty$ the set $\hat{\mathcal{V}}$ it defines converges (almost uniformly) to the variety $\mathcal{V}(\sqrt{\mathscr{I}} : \mathfrak{p})$ and $\mathcal{V}(\sqrt{\mathscr{I}} : \mathfrak{p}) \subseteq \hat{\mathcal{V}}$ holds for all r.* ∎

Proof. The proof is equivalent to the proof of Theorem 6. □

We can relate this result to set-observability with help of the primal moment relaxation as follows.

Proposition 8 (Sufficient Conditions for Set-Observability). *A system* (6.1) *(resp.* (6.2)*) is set-observable on $\mathcal{B} \subset \mathcal{K}$ if for some s the conditions rank $M_{s-v}(y^*) =$ rank $M_s(y^*)$ and $\hat{\mathcal{V}} = \mathcal{B}$ hold.*
Equivalently, a system (6.1) *(resp.* (6.2)*) is not set-observable observable if for some s the conditions rank $M_{s-v}(y^*) =$ rank $M_s(y^*)$ and $\hat{\mathcal{V}} \supset \mathcal{B}$ hold.* ∎

Proposition 8, therefore, allows to systematically check a partition of the constraint set $\mathcal{K}$ if it leads to a loss of observability. To check several partitions consecutively a different formulation of the varieties is computationally more efficient. However, we can then only determine that a system is not set-observable and set-observability cannot be concluded.

Consider, for this reason, the following feasibility problem similar to the previous section.

$$\begin{aligned} \text{find} \quad & a \\ \text{subject to} \quad & g_i(a, \beta) = 0, i \in \{1, \ldots, n_{\mathscr{I}}\}, \\ & a \in \mathcal{K}, \beta \in \mathcal{B} \subset \mathcal{K}, \end{aligned} \tag{6.26}$$

where g_i are the generators of the colon ideal in Theorem 12. Formulating local observability in this way has the advantage that we can employ Algorithm 1 and Algorithm 2 from Section 4.3.1. For Algorithm 1 we have the following result equivalent to Proposition 8 based on Lemma 2.

Proposition 9 (Sufficient Unobservability Condition). *Given a semi-definite program of order d obtained by relaxing* (6.26) *according to Section 4.3.1 and denoted by SDP_d. If a system* (2.1) *(resp.* (2.2)*) is not set-observable on $\mathcal{K}$, then there exists a subset $\mathcal{Q} \subset \mathcal{K} \setminus \mathcal{B}$ for which the Langrangean dual of SDP_d is not unbounded.*

Proposition 9 is equivalent to Theorem 12 in the sense that if there exists a partition $\mathcal{Q}$ as defined in Proposition 9 then (6.18) cannot hold. We illustrate Proposition 6.18 with the following example.

Example 7 (Set-Observability [92]). *Consider again the discrete-time system from Example 5 and Example 6. We consider the following constraints on the initial conditions for analyzing set-observability.*

$$\begin{aligned} \mathcal{A}_0 &:= \{a : -1 \leq a_i \leq 2.2, i \in \{1, 2, 3\}\}, \\ \mathcal{B}_0 &:= \{b : -1 \leq b_i \leq 2.2, i \in \{1, 2, 3\}\}. \end{aligned} \tag{6.27}$$

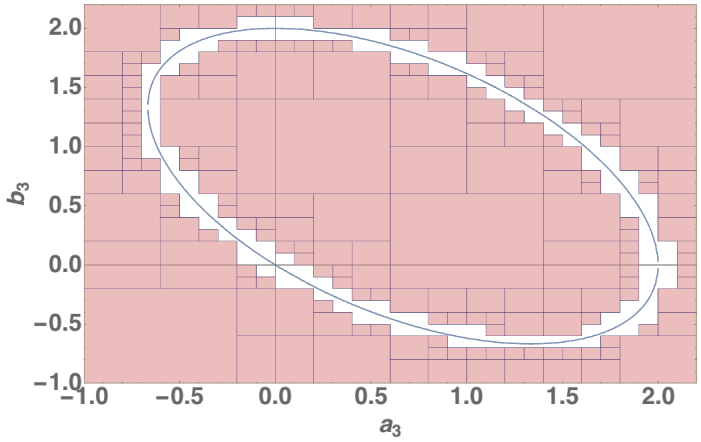

Figure 6.2: Red partitions correspond to set-observable partitions, i. e. sets for which the restriction of $b \in \mathcal{Q} \subset \mathcal{B}_0$ leads to $\hat{\mathcal{A}} = \mathcal{Q}$. The white area could not be classified according to Proposition 9 is, therefore, considered as not set-observable. The blue ellipse, corresponds to the nonlinear connection of initial conditions a_3 and b_3.

Recall that the colon ideal for the considered example for global observability analysis is as follows.

$$\sqrt{\mathscr{I}} : \mathfrak{p} = \langle a_1 - b_1, a_2 - b_2, -2a_3 + a_3^2 - 2b_3 + a_3 b_3 + b_3^2 \rangle. \tag{6.28}$$

The goal is to test whether the system is set-observable. To do so, we have to estimate the variety $\mathcal{V}(\sqrt{\mathscr{I}} : \mathfrak{p})$ when the initial conditions b_i are restricted to a subset of $\mathcal{B}_0$. We implemented the feasibility problem 6.26 and a bisectioning algorithm similar to Algorithm 1 with help of the ADMIT toolbox [204]. Then, the for each partition $\mathcal{Q} \in \mathcal{B}_0$, we tested whether the estimate $\hat{\mathcal{A}}$ as the projection of 6.26 on the variable a is equal to $\mathcal{Q}$. The result is depicted in Figure 6.2 as a projection on the variables a_3 and b_3. By determining the partitions that are set-observable it is possible to decide a priori if measurement data is helpful to estimate the states and parameters or not. ∎

To complement the notion of set-observability we derive next a similar notion based on the energy visible at the output quantified in terms of the L_2-norm for continuous-time systems.

6.3 Output Energy Measure for Practical Observability Analysis

In this section, we present a set-based output energy measure for constrained polynomial systems with parameter uncertainties following [165]. The output energy is measured in terms of the L_2-norm on a finite-time interval while the initial conditions

and parameters are allowed to take values from a set. By specifying a bound on the output norm, the measure allows further to determine the set of initial conditions and parameters which lead to satisfaction of this bound.

Furthermore, this set characterizes whether an uncertain system can be estimated by a norm-observer and, therefore, can be applied for observability analysis complementing the approach presented in the previous section. The derivation of the set is based on recasting a nonlinear program with embedded differential equations into an infinite-dimensional linear program similar to the procedure for estimating the reachable set of a continuous-time system as presented in Section 5.4.2.

6.3.1 Output Energy

We quantify the energy visible at the output $\phi_h(x(t)|x_0)$ starting from a specific initial condition using the L_2-norm. The goal is to link an initial condition to the output energy, in the sense that the norm of the output is used to derive bounds on the initial condition similar to the concept of norm-observability, see e. g. [79]. This work considers the output energy over a finite-time interval as defined next.

Definition 18 (Finite-Time Output Energy Measure)**.** *The output energy measure* $\boldsymbol{M}_T(x_0)$ *of an initial condition* $x_0 \in \mathbb{R}^{n_x}$ *depends on the resulting output trajectory* $\phi_h(t|x_0) \in \mathbb{R}^{n_y}$ *of* (6.1) *on the finite-time interval* $[0, T]$*,* $T > 0$ *by*

$$\boldsymbol{M}_T(x_0) \coloneqq \int_0^T \|y(t|x_0)\|^2 \, dt. \qquad \blacksquare$$

In [66] an equivalent infinite-time measure is used for observability analysis for linear systems and systems for which the so-called zero-state assumption holds. Note that Definition 18 only allows to quantify the output energy of a single initial condition.

If there are no uncertainties, the quantification of the energy could be determined by deriving the initial condition x_0 that maximizes the L_2-norm with the following polynomial optimization problem for a continuous-time system.

$$\begin{aligned} \boldsymbol{M}_T^{min} \coloneqq \min_{x_0} \quad & \textstyle\int_0^T \|y(t|x_0)\|^2 \, dt \\ \text{s.t.} \quad & \dot{x}(t) = f(x(t), u_s(t)), \\ & y(t) = h(x(t), u_s(t)), \\ & x_0 \in \mathcal{X}_0, \\ & x(t) \in \mathcal{X}, \quad \forall t \in (0, T], \end{aligned} \tag{6.29}$$

where $u_s(t) \in \mathbb{R}[t]$ is a known input signal. Note that this problem can be solved by the methods presented in [76].

In case of uncertainties, however, $\boldsymbol{M}_T^{min}$ and $\boldsymbol{M}_T^{max}$ do not provide much insight, as typically uniqueness of the optima is lost resulting in entire sets of initial conditions that lead to the same output energy. In addition, a system that is unobservable in the

classical sense might still have high output energy.

For this reason, the subsequent definition is more suitable for uncertain systems.

Definition 19. *(Initial Conditions with Bounded Output Energy Measure). For a given $\boldsymbol{m} \in \mathbb{R}$. Determine the set of initial conditions $\mathcal{X}_0^*$ for which every initial condition $x_0 \in \mathcal{X}_0^*$ satisfies $\boldsymbol{M}_T(x_0) \leq \boldsymbol{m}$.* ∎

In the following, a computational approach is proposed to determine the set $\mathcal{X}_0^*$ for a given bound $\boldsymbol{m}$. To do so, we modify (6.29) as follows.

$$\begin{aligned} \mathcal{X}_0^* \coloneqq \quad & \text{find} \quad x_0 \\ & \text{s.t.} \quad \dot{x}(t) = f(t, x(t), u_s(t)), \\ & \qquad\; y(t) = h(t, x(t), u_s(t)), \\ & \qquad\; \textstyle\int_0^T \|y(t|x_0)\|^2 \, dt \leq \boldsymbol{m}, \\ & \qquad\; x_0 \in \mathcal{X}_0, \\ & \qquad\; x(t) \in \mathcal{X}, \quad \forall t \in (0, T]\,. \end{aligned} \tag{6.30}$$

With the feasibility problem (6.30), we aim to find initial conditions/parameter values x_0 which lead to a L_2-norm of the output that is bounded by $\mathbf{m}$. Clearly, addressing (6.30) is difficult due to the present nonconvexity and the embedded differential equations. The purpose of the subsequent sections is to derive a convex optimization problem that directly takes the dynamics into account and considers entire sets of parameters and initial conditions.

Following Section 5.4.2, we reformulate the nonlinear optimization problem with embedded differential equation in terms of occupation measures. The occupation measures contain the needed information about the initial condition and parameter values of the system, as well as the nonlinear dynamics of the system, i. e. they encode the system trajectories. The main advantage of this reformulation is the resulting linear relationship between the occupation measures as seen in Section 5.4.2. The only major difference in the derivation of the infinite-dimensional problem derived in Section 5.4.2 is the constraint corresponding to the output energy. To be able to state (6.30) in terms of occupation measures, the constraint on the output energy has to be treated as shown next.

6.3.2 Reformulation of the Output Norm

To link the L_2-norm of the output to the occupation measure, consider the mapping H that maps the output space to $[0, \infty]$ and

$$H(\mu(\mathcal{A})) = \mu(h^{-1}(\mathcal{A})) \tag{6.31}$$

holds, where $h^{-1}(\mathcal{A}) := \{(t, x) \in \mathcal{T} \times \mathcal{X} : h(t, x) \in \mathcal{A}\}$ and μ denotes the occupation measure as defined in (5.17) of Section 5.4.2. Note that H is commonly called a

pushforward operator in measure theory as it transports a measure from a measurable space to another according to function h.

To be able to represent the previous statement also in terms of continuous functions as needed in the following, consider the canonical basis of monomials up to degree r:

$$m_r(x) := (1, x_1, \ldots, x_{n_x}, x_1^2, x_1x_2, \ldots, x_1^{r-1}x_2, \ldots, x_{n_x}^r)^\mathsf{T} \tag{6.32}$$

and the Riesz functionals

$$\begin{aligned} z_0 &:= \textstyle\int_{\mathcal{X}} m_r(x)\mu_0(dx), & z_T &:= \textstyle\int_{\mathcal{X}} m_r(x)\mu_T(dx), \\ z &:= \textstyle\int_{\mathcal{T}\times\mathcal{X}} m_r(x)\mu(dt,dx), & \bar{z} &:= \textstyle\int_{\mathcal{T}\times\mathcal{X}} m_r(x)\circ m_r(x)\mu(dt,dx), \end{aligned} \tag{6.33}$$

where $\circ$ is the Hadamard product. Then we can simply define a vector $c \in \mathbb{R}^{n_{m_r}}$ such that $h(t,x,u_s) = c^\mathsf{T} m_r(x)$ and $\int_{\mathcal{T}} h(t,x,u_s) = \int_{\mathcal{T}\times\mathcal{X}} c^\mathsf{T}\bar{z}\mu(dt,dx)$.

6.3.3 Infinite-Dimensional Linear Program and Relaxation

By applying the same steps as presented in Section 5.4.2 to the optimization problem (6.30), we obtain the infinite-dimensional linear program in the positive cone of the space of finite signed Borel measures:

$$\begin{aligned} \sup_{\mu_0} \quad & \langle 1, \mu_0\rangle \\ \text{subject to} \quad & \delta_T\mu_T - \mathcal{L}'\mu = \delta_0\mu_0, \\ & \boldsymbol{m} - H(\mu) \geq 0, \\ & \mu_0 + \hat{\mu}_0 = \lambda, \\ & \mu_0, \mu_T, \mu, \hat{\mu}_0 \geq 0, \end{aligned} \tag{6.34}$$

where $\mathcal{L}'$ is the Liouville operator as introduced in (5.22) and the constraint $\mu_0+\hat{\mu}_0 = \lambda$, where $\hat{\mu}_0$ is a slack variable (or complementary measure) ensures that the derived measure μ_0 is dominated by the Lebesgue measure λ.

The optimization (6.34) describes basically the hyper-volume of $\mathcal{X}_0^*$. To derive an exact, explicit description of $\mathcal{X}_0^*$ we represent (6.34) as its infinite dimensional dual problem over the space of continuous functions in terms of nonnegative polynomials for an appropriate algebraic dual pair as

$$\begin{aligned} \inf_{w} \quad & \langle w(x), \lambda\rangle \\ \text{subject to} \quad & w(x) - v(0,x,u_s) \geq 1, \forall x \in \mathcal{X}, \\ & -\mathcal{L}v(t,x,u_s) - h^2(t,x,u_s) + \boldsymbol{m} \geq 0, \\ & \qquad \forall x \in [0,T]\times\mathcal{X}, \\ & v(T,x,u_s), w(x) \geq 0, \forall x \in \mathcal{X}, \end{aligned} \tag{6.35}$$

where $w(x) \in \mathscr{C}(\mathcal{X}), v \in \mathscr{C}^1([0,T]\times\mathcal{X})$ are continuous (resp. continuously differentiable) functions. As both LPs are infinite-dimensional they cannot be solved directly.

We employ here Lasserre's hierarchy, e. g. [111], to derive a solution similar to Section 5.4.2 and we obtain the following sum-of-squares strengthening:

$$\begin{aligned} \inf_{w_{c,r}} \quad & w_{c,r}^{\mathsf{T}} l \\ \text{subject to} \quad & w_r(x) = v(0,x) + 1 + r_0(x) + \\ & \quad + \textstyle\sum_{i=1}^{m_x} r_{0,i}(x) g_{x,i}(x), \\ & -\mathcal{L}v(t,x) - h^2(t,x,u_s) + \boldsymbol{m} = p(t,x) \\ & \quad + q_0(t,x)t(T-t) + \textstyle\sum_{i=1}^{m_x} q_i(t,x) g_{x,i}(x), \\ & w_r(x) = p_0(x) + \textstyle\sum_{i=1}^{m_x} q_{0,i}(x) g_{x,i}(x), \\ & v(1,x) = p_1(x) + \textstyle\sum_{i=1}^{m_{x_k}} q_{1,i}(x) g_{x,i}(x), \end{aligned} \tag{6.36}$$

where l is the vector of Lebesgue moments over $\mathcal{X}$ indexed in the same basis in which the polynomial $w_r(x)$ with coefficients $w_{c,r}$ is expressed. The minimum is over polynomials $v(t,x)$ and $w_r(x)$, and polynomial sum-of-squares $p(t,x)$, $q_0(t,x)$, $q_i(t,x)$, $p_0(x) \in \Sigma_r[t,x]$, $q_{0,i}(x)$, $p_1(x)$, $r_0(x), r_{0,i}(x) \in \Sigma_r[x], \forall i = 1,\ldots,m_x$ and $q_{1,i}(x), \forall i = 1,\ldots,m_{x_k}$ of appropriate degrees. The constraints that polynomials are sum-of-squares can be written explicitly as LMI constraints, and the objective is linear in the coefficients of the polynomial $w_r(x)$. Therefore, (6.36) can be formulated as a semi-definite program. Furthermore, the set $\mathcal{W}_r := \{x : w_r(x) \geq 1\}$ is an outer-approximation of $\mathcal{X}_0^*$, i. e. $\mathcal{X}_0^* \subseteq \mathcal{W}_r$, and the Lebesgue measure of $\mathcal{W}_r$ converges to the Lebesgue measure of $\mathcal{X}_0^*$ for $r \to \infty$, see also Theorem 6.

In the next section, it is illustrated how $\mathcal{X}_0^*$ can be applied in observability analysis.

6.3.4 Application to Observability Analysis

This section illustrates that the set $\mathcal{X}_0^*$ can contain more information on the initial conditions than the original assumed set of initial conditions $\mathcal{X}_0$. In particular, if the set $\mathcal{X}_0^*$ is not contained in the interior of $\mathcal{X}_0$ then $\mathcal{X}_0^*$ does not provide any additional information on the location of possible initial conditions. Therefore, to strengthen the expressiveness of $\mathcal{X}_0^*$ the following additional requirement is defined.

Definition 20 (Set-Observability (Output Energy))**.** *A state x_i of system* (6.1) *with initial conditions in the set $\mathcal{X}_0$ is said to be set-observable w.r.t. a bound* $\mathbf{m}$ *on the output energy $\boldsymbol{M}^T$ if the projection of $\mathcal{X}_0^*$ onto x_i has a smaller Lebesgue measure λ than the projection of $\mathcal{X}_0$ onto x_i, i. e. $\lambda(\perp_{x_i} \mathcal{X}_0^*) < \lambda(\perp_{x_i} \mathcal{X}_0)$. System* (6.1) *is said to be set-observable if this condition holds for all $i = 1,\ldots,n_x$.* ∎

The idea behind this set-observability notion is to compare the length of the projections of $\mathcal{X}_0$ and $\mathcal{X}_0^*$ onto different state directions. If the former length is larger than the latter then this means that the bound on the output energy can be used to reduce the initially present uncertainties. According to Definition 20 the observability of a system corresponds, therefore, not only to the output energy but also to the shape and

size of $\mathcal{X}_0^*$. With the help of (6.36) we can analyze the set-observability of system (6.1) according to Definition 20. We can state the following result:

Theorem 15 (Sufficient Condition for Set-Observability [165])**.**
Given an outer-approximation $\mathcal{W}_r$. Assuming $\mathcal{X}_0^* \neq \emptyset$, then the following statements are equivalent.

1. *The system is set-observable according to Definition 20.*
2. *$\exists r \in \mathbb{R} \cup \{\infty\}$ and $\exists \boldsymbol{m} \in \mathbb{R}$ such that $\lambda(\perp_{x_j} \mathcal{X}_0^*) \leq \lambda(\perp_{x_j} \mathcal{W}_r) < \lambda(\perp_{x_j} \mathcal{X}_0)$, $\forall x_j$.*

■

Proof: Follows the same lines as the proof of [74, Thm. 6]. In the case of finite convergence the same argumentation as provided in the following can be employed by replacing ∞ with a sufficiently large constant $R \in \mathbb{R}$. From the convergence of the relaxation it follows that w_r converges to the indicator function $I_{\mathcal{X}_0^*}$ of the set $\mathcal{X}_0^*$. Furthermore, at every relaxation order r we have $\mathcal{X}_0^* \subset \mathcal{W}_r$, i. e. $w_r \geq I_{\mathcal{W}_r} \geq I_{\mathcal{X}_0^*}$. Therefore, we have

$$\begin{aligned}\lambda(\perp_{x_j} \mathcal{X}_0^*) &= \int_{\mathcal{X}} \perp_{x_j} I_{\mathcal{X}_0^*} d\lambda \\ = \lim_{r\to\infty} \int_{\mathcal{X}} \perp_{x_j} w_r d\lambda &\geq \lim_{r\to\infty} \int_{\mathcal{X}} \perp_{x_j} W_r d\lambda = \lambda(\perp_{x_j} \mathcal{W}_r).\end{aligned}$$

As $\mathcal{X}_0^* \subset \mathcal{W}_r$, it follows that $\lambda(\perp_{x_j} \mathcal{X}_0^*) \leq \lambda(\perp_{x_j} \mathcal{W}_r)$. Therefore, $\lambda(\perp_{x_j} \mathcal{X}_0^*) = \lambda(\perp_{x_j} \mathcal{W}_r)$ must hold, which concludes the proof. □

Note that there are certain similarities of the proposed concept of set-observability and norm-observability introduced in [79, 80]. A system is small-time initial-state norm-observable (SINO), if $\forall \tau$ there exists $\gamma \in \mathcal{K}_\infty$ such that the Euclidean norm of the initial state is upper bounded by the infinity norm of the output, i. e. $|x(0)| \leq \gamma(||y||_{\infty,[0,\tau]})$, see [79]. It is obvious that set-observability and SINO are related in the case that parameters are not unknown-but-bounded. The main difference derives from the fact that this work considers only bounded state space regions, therefore, a strict relationship exists only for $\boldsymbol{m} = 0$. In this case, if a system is not set-observable ($\forall r$) it follows that no γ exists on $\mathcal{X}$.

The application of Theorem 15 to observability analysis is exemplified next.

Example 8 (Two-Tank System [165])**.** *We consider the polynomial model of a two-tank as derived in [105]. The process consists of two water tanks and one pump modeled by:*

$$\begin{aligned}\dot{x}_1 &= 0.073x_1^2 - 1.6x_1 - 0.047x_2^2 + 0.2x_2, \\ \dot{x}_2 &= 0.33x_2^2 - 1.4x_2, \\ y &= h(x_1, x_2),\end{aligned} \tag{6.37}$$

where the states x_1, x_2 correspond to the water levels in the two tanks. The output $h(x_1, x_2)$ is assumed to be either x_1 or x_2. As state constraints we consider $\mathcal{X} = \mathcal{X}_0 = [0,1] \times [0,1]$ and the end-time T is set to one for the L_2-norm of the output.

For given bounds $\mathbf{m} \in \{0, 0.01, \ldots, 0.1, 0.2, 0.3\}$ on the output energy measure, the set of initial conditions is derived with (6.36), cf. Figure 6.3. Furthermore, the results show that the system is set-observable for $h(x_1, x_2) = x_1$. However, this is not the case for $h(x_1, x_2) = x_2$ according to Theorem 15 as the corresponding outer-approximation $\mathcal{W}_r$ is up to numerical optimality equal to $\mathcal{X}_0$.

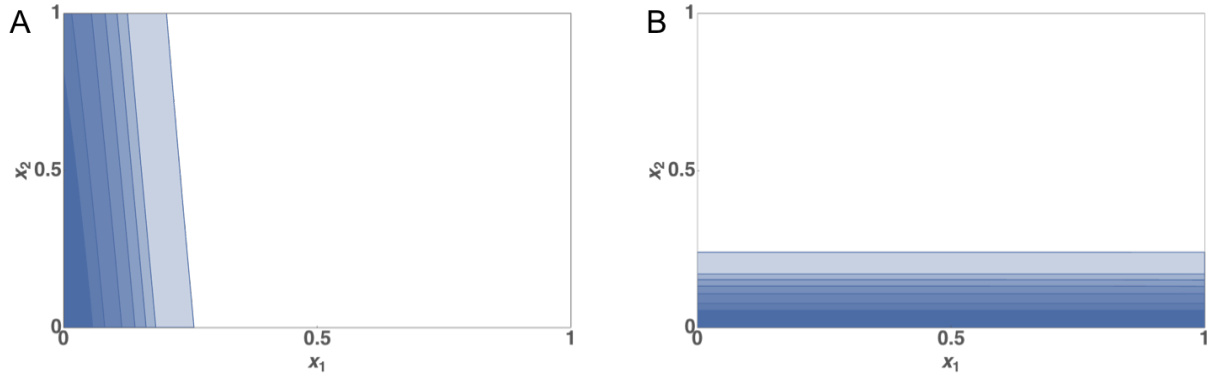

Figure 6.3: Illustration of the observability analysis for the two-tank example. A: The system is set-observable for $y = x_1$ w.r.t. the output energy. B: The system is not set-observable for $y = x_2$ w.r.t. the output energy. Computation time ∼1min with MOSEK 7.0/YALMIP [122] (for each case). Relaxation order $r = 5$ to derive the outer-approximation $\mathcal{W}_r$, which is considered by MOSEK 7.0 as numerically optimal.

Example 9 (Mass-Spring System [165])**.** *We consider a mass-spring system with a softening spring (for details see [96]) scaled to the unit box:*

$$\begin{aligned} \dot{x}_1 &= (x_2 - 0.5), \\ \dot{x}_2 &= -(p+1)(x_1 - 0.5) - 0.5(x_2 - 0.5) \\ &\quad + 4(p+1)(x_1 - 0.5)^3, \\ y &= x_2. \end{aligned} \tag{6.38}$$

The state constraints are $\mathcal{X} = \mathcal{X}_0 = [0,1] \times [0,1]$, $p \in [0,1]$ and the end-time T is set to one.

Instead of varying $\mathbf{m}$ as in the previous example we modified (6.36) such that we can consider a lower and an upper bound on the L_2-norm of the output, namely $\mathbf{m}_1 = 0.6 \leq \int_{\mathcal{T}} ||y(t|x_0)||^2 dt \leq \mathbf{m}_2 = 1.2$. In this case, the output energy measure shows that the states are set-observable and the parameter is not set-observable w.r.t. $\boldsymbol{m}$ as only the projections of $\mathcal{X}_0^*$ onto the axes of x_1, x_2 have a smaller Lebesgue measure than the projections of $\mathcal{X}_0$ (see also Thm. 15) as illustrated by Fig. 6.4.

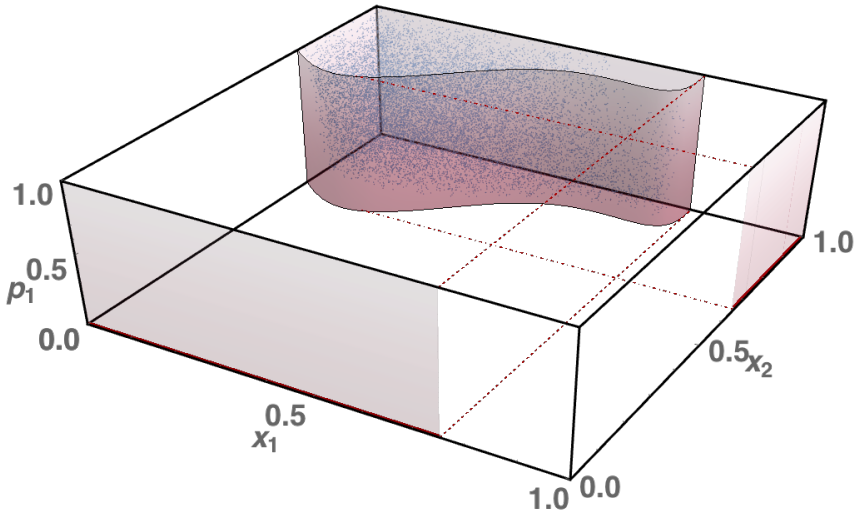

Figure 6.4: Mass-spring example: Projections of the computed initial conditions (indicated by dotted lines) show the system is partially set-unobservable. Full red lines correspond to the projections onto the x_1 and x_2 axis, respectively. Computation time ∼30s with MOSEK 7.0/YALMIP. Relaxation order $r = 5$. Dots represent consistent initial conditions obtained by Monte-Carlo sampling.

6.4 Summary

In this chapter, we proposed several methods to study observability of uncertain polynomial systems. For this reason, we extended the algebraic observability analysis of Kawano and Othsuka. At first we demonstrated that the algebraic conditions for local-at-a-point observability of nominal discrete-time and continuous-time systems can be tested efficiently with semi-definite programs. We, furthermore, showed that observability of a system ensures finite convergence of Lasserre's hierarchy and, therefore, only finite dimensional semi-definite programs have to be considered.

For uncertain systems, we demonstrated that local observability properties of a system in general do not allow a reduction of the initially assumed bounds on the uncertainties when measurements are considered. The possibility of reducing the uncertainties we denote by set-observability. To study set-observability a stronger global observability property is needed. However, for uncertain systems, if global observability does not hold, it cannot be directly concluded that a system is not set-observable. To study this circumstance, we provided two methods. The first method, is based on finding a normed Lebesgue measure of a semi-algebraic set, while the second method is based on relaxed feasibility problems.

To complement the notion of set-observability, we, furthermore, studied the connection of output energy and possible initial conditions of an uncertain continuous-time system. The output energy in this case is measured in terms of the L_2-norm on a

finite-time interval. By specifying a bound on the output norm, we were able to derive a converging hierarchy of semi-definite programs that outer-approximate the corresponding initial conditions.

7 Summary and Conclusions

To satisfy the steadily increasing requirements for safe operation and quality in industry [30, 57], model-based control and system theoretical methods become increasingly important. The basis of successfully employing such methods is the derivation of a quantitative and predictive mathematical model of the process under consideration. For this reason, system identification is an important research area, i. e. the automatic construction of a mathematical model from measurement data [120, 135, 142]. However, most approaches applied in practice consider only the identification and analysis of linear systems, while in reality every system behaves nonlinearly [58, 96]. Therefore, a strong interest in the development of versatile and rigorous analysis methods that can handle nonlinear systems exists. One particular problem in the derivation of such methods is the fact that nowadays digital sensors are used to gather data on a process at distinct time points. Hence, it is typically easier to identify a discrete-time model [120], although, real processes evolve in general continuously in time.

Another important aspect in the development of system identification methods is the consideration of uncertainties [133, 149, 170, 202]. For instance, if the available measurement data is uncertain, e. g. due to limited accuracy of the employed measurement device or noise, it becomes difficult to decide whether a model reproduces the data or not. This task, also called model (in-)validation, is a crucial step in verifying that a model is indeed suitable for analyzing the process under consideration. One promising direction in system identification of uncertain systems is, therefore, the development of set-based methods, as uncertainties on different levels can be considered and guaranteed results on model validity can be derived [83, 164, 170, 203].

Motivated by the aforementioned difficulties in the derivation of system identification methods, this work presented set-based methods for the analysis and verification of uncertain continuous-time and discrete-time polynomial systems. Such systems arise, for instance, in the fields of systems biology, chemical engineering and mechatronics [98, 164, 173]. The fundamental basis of the derived results is the reformulation of model analysis and verification tasks as polynomial feasibility and polynomial optimization problems. This allows on the one hand the consideration of uncertainties in a quantitative or semi-quantitative description and on the other hand the unified treatment of the considered problems. In particular, the problems of proving model (in-)validity, estimating parameters and states of dynamical systems, deriving reachable sets of continuous-time and discrete-time systems, and analyzing the observability of uncertain systems were considered.

With respect to model (in-)validity, we show that by relaxing the polynomial feasi-

bility problem into a converging hierarchy of semi-definite programs a conclusive proof for model inconsistency of discrete-time systems can be derived. This is possible since the solution space of the semi-definite program provides an outer-approximation of the solution space of the original feasibility problem. Therefore, it is sufficient to show that the solution space of the relaxation is empty to conclude model inconsistency of the model. Based on this principle we, furthermore, provide an approach to approximate the solution set of the feasibility problem. Hence, we are able to estimate the possible values of initial conditions and parameters leading to a consistent behavior of a model. Moreover, as the feasibility problem in its general form not only contains information on initial conditions and parameters the approximation can as well be used to estimate the states and their reachable sets. However, the performed derivations also point to a series of open questions: How can continuous-time systems be considered in this framework? Furthermore, under which conditions are discrete-time approximations sufficient to relate model (in-)consistency from the discrete-time to the continuous-time system?

To provide answers to these question, we study the properties a discrete-time approximation has to possess such that this relationship is possible. Based on these properties, two approaches are studied that allow the desired relationship. The first approach is based on the existence and uniqueness of initial value problems according to the Picard-Lindelöf theorem. To be able to employ the theorem, we provide a procedure to compute the reachable set of a continuous-time system at distinct time points. This allows the derivation of a time step size for the discrete-time approximation such that the approximation embeds the trajectories of the continuous-time system. Thus, solving the problem of proving consistency of the continuous-time system. The second approach is motivated by the one-step consistency property of Euler discretizations and can be employed to reduce the computational demands of the first approach as demonstrated for a real world example.

With respect to observability analysis, we outline how the algebraic observability conditions presented in [94] can be extended to uncertain polynomial systems. To do so, we establish at first that the algebraic conditions can be addressed by polynomial feasibility problems in the nominal case. In a second step, based on the previous results, we provide an extension to uncertain systems. The, hereby, necessary notion of set-observability corresponds to the ability to reduce the bounds on the initial conditions and parameters by measurements. Furthermore, to complement this notion, we link the energy visible at the output of a system to the initial conditions and parameters similar to a Gramian. This again allows a statement regarding the ability to estimate the initial conditions and parameters of a continuous-time system.

In summary, the results of this thesis provide solutions to a series of practically and theoretically important problems in system identification and analysis for a rather general class of systems. In particular, the results on reachability and observability of uncertain systems provide a well-based foundation for further research.

7.1 Outlook

Several topics and results addressed in this work offer the opportunity for further research. The focus so far was on the development of a rigorous theoretical foundation which lead inevitably to a limited consideration of practical examples. The main problem to overcome here is the computational demand of the employed semi-definite and linear programs. First results in this direction are, for instance, the decomposition of semi-definite programs into smaller subproblems by symmetry [59, 167], the reduction of the monomial basis [153], or the abstraction of dynamical subsystem via outer-approximation in [175]. These methods depend, however, on the specific process under consideration. To study, whether a generalization of these approaches is possible appears still to be promising.

Another promising research direction is the study of the convergence rate of the employed relaxation. First theoretical results indicate a slower convergence rate [145, 178] than actually is observed in practical applications, i. a. [169, 170, 204, 205]. The main question hereby from a system theoretical perspective is the question whether and how the system structure can be employed to improve the convergence rate. A second interesting question is, how the convergence rate is influenced by additional constraints as e. g. imposed by the bisectioning procedure employed for the estimation tasks in this work.

From a control theoretical point of view there are also other interesting possibilities for extensions. Quite recently, several groups started to work on the combination of set-based and probabilistic methods. Those works allow, for instance, the consideration of probabilistic data [200] or to make a tradeoff between guaranteed and probabilistic results [42]. Another important area that is missing so far is controller and input design. In particular, [101] would fit directly in the presented problem formulation and could, for instance, be applied for optimal parameter estimation or active fault diagnosis based on the observability results obtained in this work.

A Moment Relaxation

A.1 Mathematical Notation and Definitions

To be able to relate a general polynomial optimization problem to a generalized moment problem and the involved relaxations steps, we define the notion of monomials and polynomials. Then we give some classical results on measures and the relationship between polynomials and measures.

We denote a *monomial* in the variables $x = (x_1, \dots, x_{n_x})$ with x^α, i. e. $x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_{n_x}^{\alpha_{n_x}}$ for $n \in \mathbb{N}^{n_x}$. The degree of x^α is $|\alpha| = \sum_{i=1}^{n_x} \alpha_i$. The set of all monomials is $\Xi = \{x^\alpha \in \mathbb{R}^{n_x} : \alpha \in \mathbb{N}^{n_x}\}$. The set of all monomials up to a degree d is $\Xi_d = \{x^\alpha \in \mathbb{R}^{n_x}, \alpha \in \mathbb{N}^{n_x} : |\alpha| \leq d\}$. We define a *complete monomial basis* $\xi_d \in \Xi$ to be the vector containing all monomials up to degree d, i. e.

$$\xi_d = (1, x_1, x_2, \dots, x_{n_x}, x_1^2, x_1 x_2, \dots, x_1 x_{n_x}, x_2^2, x_2 x_3, \dots, x_{n_x}^2, \dots, x_1^d, \dots x_{n_x}^d)^T.$$

A polynomial $p(x) : \mathbb{R}^{n_x} \to \mathbb{R}$ is a sum of monomials, i.e.

$$p(x) = \sum_{\alpha \in \mathbb{N}^{n_x}} p_\alpha \, x^\alpha, \tag{A.1}$$

where there are only finitely many nonzero coefficients $p_\alpha \in \mathbb{R}$. We denote with p the coefficient vector consisting of the nonzero elements p_α, i. e. $p = \{p_\alpha\}$. Note that we choose this notation to make the relationship between polynomials and the needed notions from measure theory clear.

It is out of the scope of this work to provide the reader with a complete introduction to measure theory, see e. g. [209], however, we will give some definitions and results that play a central role in the moment relaxation method in the next section. Note that in the following we consider only nonnegative Borel measures μ on $\mathbb{R}^{n_x}$, whereas $\mu(\mathbb{R}^{n_x}) = 1$, i. e. μ is a *propability measure*.

Given such a measure μ on $\mathbb{R}^{n_x}$, its *support* $\text{supp}(\mu)$ is the smallest closed subset $\mathcal{Z} \subset \mathbb{R}^{n_x}$ for which $\mu(\mathbb{R}^{n_x} \setminus \mathcal{Z}) = 0$. We say μ is a measure supported by the constraint set $\mathcal{K}$ if $\text{supp}(\mu) \subseteq \mathcal{K}$. The quantity $y_\alpha = \int x^\alpha \mu(dx)$ is called the *moment* of μ with order α. We denote the sequence of moments of the measure μ with $(y_\alpha)_{\alpha \in \mathbb{N}^{n_x}}$ and the truncated sequence $(y_\alpha)_d$ where d is the maximal degree of x^α.

The *moment matrix* $M(y)$ to a corresponding sequence $(y_\alpha)_{\alpha \in \mathbb{N}^{n_x}}$ is indexed by $\mathbb{N}^{n_x}$ and the (α, β)th entry is $y_{\alpha,\beta}, \alpha, \beta \in \mathbb{N}^{n_x}$. Analogously, $M_d(y)$ denotes the moment matrix for a truncated sequence and $|\alpha|, |\beta| \leq d$. To clarify this notation consider the

case that $n_x = d = 2$, in this case we have

$$M_2(\xi) = \xi_2 \xi_2^T = \begin{bmatrix} 1 & x_1 & x_2 & x_1^2 & x_1x_2 & x_2^2 \\ x_1 & x_1^2 & x_1x_2 & x_1^3 & x_1^2x_2 & x_1x_2^2 \\ x_2 & x_1x_2 & x_2^2 & x_1^2x_2 & x_1x_2^2 & x_2^3 \\ x_1^2 & x_1^3 & x_1^2x_2 & x_1^4 & x_1^3x_2 & x_1^2x_2^2 \\ x_1x_2 & x_1^2x_2 & x_1x_2^2 & x_1^3x_2 & x_1^2x_2^2 & x_1x_2^3 \\ x_2^2 & x_1x_2^2 & x_2^3 & x_1^2x_2^2 & x_1x_2^3 & x_2^4 \end{bmatrix}. \tag{A.2}$$

Given a sequence $(y_\alpha)_d$, a polynomial $p(x)$ and a moment matrix M_d we can define the so-called *localizing matrix* $M_d(p(x)y)$ as

$$M_d(p(x)y) = M_d(y)p, \tag{A.3}$$

where $p = \{p_\alpha\}$ denotes the coefficient vector of $p(x)$. For instance, if $n_x = d = 1$ and $p(x) = x^2$, then

$$M_1(u(x)y) = \begin{bmatrix} p(x) & p(x)x \\ p(x)x & p(x)x^2 \end{bmatrix} = \begin{bmatrix} x^2 & x^3 \\ x^3 & x^4 \end{bmatrix}. \tag{A.4}$$

Let us further define the linear functional L_y as

$$L_y(p(x)) = \sum_{\alpha \in \mathbb{N}^{n_x}} p_\alpha y_\alpha, \tag{A.5}$$

where $p(x)$ as in (A.1). Then we can express the multiplication of two polynomials $p(x), q(x)$ with coefficient vectors $p = \{p_\alpha\}, q = \{q_\alpha\}$ with respect to the moment matrix as

$$\begin{aligned} L_y(p(x)q(x)) &= p^T M_d(y) q, \\ L_y(p(x)^2) &= p^T M_d(y) p, \end{aligned} \tag{A.6}$$

with the clear implication that

$$L_y(p(x)^2) \geq 0 \iff M_d(y) \succeq 0 \tag{A.7}$$

i. e. $L_y(p(x)^2)$ is positive definite if and only if $M_d(y) \succeq 0$, cf. also Lemma 4.1 in [114].

Lemma 3 (Necessary Conditions for Moment Sequences [114])**.** *Let $g(x)$ be a polynomial as in* (A.1) *with degree $2d_g$.*

(i) If $(y_\alpha)_{2d}$ is the sequence of moments of a measure μ, then $M_d(y) \succeq 0$ and $rank M_d(y) \leq |supp(\mu)|$.

(ii) If $(y_\alpha)_{2d}$ is the sequence of moments of a measure μ supported by the set $\{x \in \mathbb{R}^n : g(x) \geq 0\}$, then $M_{d-d_g}(g(x)y) \succeq 0$. ∎

B Tables of the Examples

B.1 Adaptation

Table B.1: Qualitative data used in the adaptation example. Notation see Section 5.5.

Verbal description of the qualitative data	**Logical constraints representing the data**
E_1 : The system is in steady-state. Initially, the output is at least 10% of the maximum value possible.	$b_{1,k} \Longleftrightarrow (x_{1,k} \geq x_{1,l} - \epsilon x_{1,l})$ $b_{2,k} \Longleftrightarrow (x_{1,k} \leq x_{1,l} + \epsilon x_{1,l})$ $b_{3,k} \Longleftrightarrow (x_{2,k} \geq x_{2,l} - \epsilon x_{2,l})$ $b_{4,k} \Longleftrightarrow (x_{3,k} \geq 0.1 C_{tot})$ $\phi_1 \Longleftrightarrow \bigwedge_{\mathcal{I}_k^{\phi_1}} (b_{1,k} \wedge b_{2,k} \wedge$ $\wedge b_{3,k} \wedge b_{4,k})$ $k \in \{1,2,3\}, l = 1$
E_2 : After the stimulus is removed the maximum value of $x_{3,k}$ is larger than 115% and is reached within 3 time-steps.	$b_{8,k} \Longleftrightarrow (u \leq 1)$ $b_{9,k} \Longleftrightarrow (x_{3,k} \geq 1.15 x_{3,0})$ $\phi_2 \Longleftrightarrow \bigwedge_{\mathcal{I}_k^{\phi_3}} b_{8,k}$ $\phi_3 \Longleftrightarrow \bigvee_{\mathcal{I}_k^{\phi_2}} b_{9,k}$ $\phi_4 \Longleftrightarrow (\phi_1 \wedge \phi_2 \wedge \phi_3)$ $k \in \{4,5,6,7\}$
E_3 : After the maximum has been reached, $x_{3,k}$ returns almost to the initial interval in 5 time-steps, i. e. $x_{3,k} \in [0.4, 5.2]$.	$\phi_5 \Longleftrightarrow \bigwedge_{\mathcal{I}_k^{\phi_5}} (b_{3,k}) \quad k \in \{8,9,10\}$
R : To avoid that $x_{3,k}$ shows a strong inverse response behavior additionally $x_{3,k}$ is always larger than 0.4.	$b_{10,k} \Longleftrightarrow (x_{3,k} \geq 0.4)$ $\phi_6 \Longleftrightarrow \bigwedge_{\mathcal{I}_k^{\phi_6}} b_{10,k}$ $k \in \{4, \ldots, 10\}$

Table B.2: Nominal parameters

Parameter:	Nominal Value:	Interval:
Kf	100	[90,110]
$KiuA$	100	[90,110]
$KaAC$	100	[90,110]
EB	0.5	-
EC	0.5	-
$kiuA$	100	-
$kaAC$	100	-
$kiAB$	1	-
$kaEB$	1	-
$kiEC$	.5	-
kf	200	-
$KiAB$	0.0001	-
$KiEC$	0.0001	-
$KaEB$	0.0001	-

Bibliography

[1] S. Alexeeva, K. J. Hellingwerf, and M. J. T. de Mattos. Quantitative assessment of oxygen availability: Perceived aerobiosis and its effect on flux distribution in the respiratory chain of *Escherichia coli*. *Journal of Bacteriology*, 184(5):1402–1406, 2002.

[2] E. J. Anderson and P. Nash. *Linear Programming in Infinite-dimensional Spaces*. theory and applications. John Wiley & Sons, 1987.

[3] J. Anderson and A. Papachristodoulou. On validation and invalidation of biological models. *BMC Bioinformatics*, 10(1):132, 2009.

[4] D. Angeli and E. D. Sontag. Forward completeness, unboundedness observability, and their Lyapunov characterizations. *Systems & Control Letters*, 38(4):209–217, 1999.

[5] K. M. Anstreicher. Semidefinite programming versus the reformulation-linearization technique for nonconvex quadratically constrained quadratic programming. *Journal of Global Optimization*, 43(2):471–484, 2009.

[6] M. Araki and H. Taguchi. Two-degree-of-freedom PID controllers. *International Journal of Control Automation and Systems*, 1:401–411, 2003.

[7] Z. Artstein. First-order approximations for differential inclusions. *Set-Valued Analysis*, 2(1):7–17, 1994.

[8] A. Astolfi and L. Praly. Global complete observability and output-to-state stability imply the existence of a globally convergent observer. *Mathematics of Control, Signals, and Systems*, 18(1):32–65, 2006.

[9] M. F. Atiyah and I. G. Macdonald. *Introduction to Communicative Algebra*. Addison-Wesley Publishing Company, 1969.

[10] J. P. Aubin and A. Cellina. *Differential Inclusions*, volume 264 of *Grundlehren der mathematischen Wissenschaften*. Springer, 1984.

[11] J. P. Aubin and H. Frankowska. Observability of systems under uncertainty. *SIAM Journal on Control and Optimization*, 27(5):949–975, 1989.

[12] J. P. Aubin and H. Frankowska. *Set-Valued Analysis*. Birkhäuser, 2008.

[13] C. Audet, P. Hansen, B. Jaumard, and G. Savard. A branch and cut algorithm for nonconvex quadratically constrained quadratic programming. *Mathematical Programming*, 87(1):131–152, 2000.

[14] 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.

[15] J. Baillieul. Controllability and observability of polynomial dynamical systems. *Nonlinear Analysis, Theory, Methods & Applications*, 5(5):543–552, 1981.

[16] N. Barkai and S. Leibler. Robustness in simple biochemical networks. *Nature*, 387(6636):913–917, 1997.

[17] M. Barrio, K. Burrage, A. Leier, and T. Tianhai. Oscillatory regulation of Hes1: Discrete stochastic delay modelling and simulation. *PLoS Computational Biology*, preprint(2006):e117, 2005.

[18] A. Barvinok. *A Course in Convexity*. American Mathematical Society, 2002.

[19] C. Bauer, S. Elsen, L. R. Swem, D. L. Swem, and S. Masuda. Redox and light regulation of gene expression in photosynthetic prokaryotes. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 358(1429):147–154, 2003.

[20] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty. *Nonlinear Programming*. Theory and Algorithms. John Wiley & Sons, 2nd edition, 1993.

[21] M. Behar, N. Hao, H. G. Dohlman, and T. C. Elston. Mathematical and computational analysis of adaptation via feedback inhibition in signal transduction pathways. *Biophysical Journal*, 93(3):806–821, 2007.

[22] A. Bemporad and M. Morari. Control of systems integrating logic, dynamics, and constraints. *Automatica*, 35:407–428, 1999.

[23] A. Ben-Tal and A. Nemirovski. *Lectures on Modern Convex Optimization*. Analysis, Algorithms, and Engineering Applications. SIAM, 2001.

[24] D. Bestle and M. Zeitz. Canonical form observer design for non-linear time-variable systems. *International Journal of Control*, 38(2):419–431, 1983.

[25] S. Borchers, S. Bosio, P. Rumschinski, A. Savchenko, R. Weismantel, and R. Findeisen. Set-Membership Estimation and Analysis of Polynomial Systems: A Relaxation-based Framework. Technical Report 2, Magdeburg, Germany, 2012.

[26] S. Borchers, S. Freund, A. Rath, 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*, 8(8):11, August 2013.

[27] 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 *Proceedings of the IEEE Conference on Decision and Control*, pages 6786–6792, Shanghai, China, 2009.

[28] S. Borchers, P. Rumschinski, S. Bosio, R. Weismantel, and R. Findeisen. Model invalidation and system identification of biochemical reaction networks. In *Proceedings of the Vienna Conference on Mathematical Modelling*, pages 787–795, Vienna, Austria, 2009.

[29] T. Boukhobza, F. Hamelin, and D. Sauter. Observability of structured linear systems in descriptor form: A graph-theoretic approach. *Automatica*, 42(4):629–635, 2006.

[30] T. Brogårdh. Present and future robot control development—An industrial perspective. *Annual Reviews in Control*, 31(1):69–79, 2007.

[31] N. C. Butzin, H. A. Owen, and M. L. P. Collins. A new system for heterologous expression of membrane proteins: *Rhodospirillum rubrum*. *Protein Expression and Purification*, 70(1):88–94, 2010.

[32] A. B. Carius, M. Henkel, and H. Grammel. A glutathione redox effect on photosynthetic membrane expression in *Rhodospirillum rubrum*. *Journal of Bacteriology*, 193(8):1893–1900, 2011.

[33] L. Carius, P. Rumschinski, T. Faulwasser, D. Flockerzi, H. Grammel, and R. Findeisen. Model-based derivation, analysis and control of unstable microaerobic steady-states - considering *Rhodospirillum rubrum* as an example. *Biotechnology and Bioengineering*, 111(4):734–747, 2014.

[34] R. J. Carroll, D. Ruppert, L. A. Stefanski, and C. M. Crainiceanu. *Measurement Error in Nonlinear Models*, volume 105 of *Monographs on Statistics and Applied Probability*. Chapman & Hall, 2nd edition, 2006.

[35] V. Cerone, D. Piga, and D. Regruto. Polytopic outer approximations of semialgebraic sets. In *Proccedings of the IEEE Conference on Decision and Control*, pages 7793–7798, Maui, USA, 2012.

[36] V. Cerone, D. Piga, and D. Regruto. Set-Membership error-in-variables identification through convex relaxation techniques. *IEEE Transactions on Automatic Control*, 57(99):517–522, 2012.

[37] E. W. Cheney. *Multivariate Approximation Theory*. Selected Topics. SIAM, 1986.

[38] G. F. Corliss. Guaranteed Error Bounds for Ordinary Differential Equations. In M. Ainsworth, J. Levesley, W. A. Light, and M. Marletta, editors, *Theory and Numerics of Ordinary and Partial Differential Equations*. Oxford University Press, 1995.

[39] D. A. Cox, J. Little, and D. O'Shea. *Using Algebraic Geometry*. Springer, 2005.

[40] D. A. Cox, J. B. Little, and D. O'Shea. *Ideals, Varieties, and Algorithms*. An Introduction to computational algebraic Geometry and commutative Algebra. Springer, 3rd edition, 2007.

[41] R. Curto and L. Fialkow. The truncated complex K-moment problem. *Transactions of the American Mathematical Society*, 352(6):2825–2855, 2000.

[42] F. Dabbene, M. Sznaier, and R. Tempo. A probabilistic approach to optimal estimation part I: Problem formulation and methodology. In *Proceedings of the IEEE Conference on Decision and Control*, pages 190–195, Maui, USA, 2012.

[43] M. Dabros, M. M. Schuler, and I. W. Marison. Simple control of specific growth rate in biotechnological fed-batch processes based on enhanced online measurements of biomass. *Bioprocess and Biosystems Engineering*, 33(9):1109–1118, 2010.

[44] C. M. Davis, P. L. Bustamante, and P. A. Loach. Reconstitution of the bacterial core light-harvesting complexes of *Rhodobacter sphaeroides* and *Rhodospirillum rubrum* with isolated-and-polypeptides, bacteriochlorophyll a, and carotenoid. *Journal of Biological Chemistry*, 270(11):5793, 1995.

[45] H. De Battista, J. Picó, and E. Picó-Marco. Nonlinear PI control of fed-batch processes for growth rate regulation. *Journal of Process Control*, 22(4):789–797, 2012.

[46] P. Deuflhard and F. Bornemann. *Scientific Computing with Ordinary Differential Equations*, volume 42 of *Texts in Applied Mathematics*. Springer, 2002.

[47] F. Ding, L. Qiu, and G. Chesi. Reconstruction of continuous-time systems from their non-uniformly sampled discrete-time systems. *Automatica*, 45(2):324–332, 2009.

[48] A. Dontchev and F. Lempio. Difference methods for differential inclusions: A survey. *SIAM Review*, pages 263–294, 1992.

[49] P. Dreesen and B. De Moor. Polynomial optimization problems are eigenvalue problems. In *Model-Based Control*, pages 49–68. Springer, 2009.

[50] T. Drengstig, H. R. Ueda, and P. Ruoff. Predicting perfect adaptation motifs in reaction kinetic networks. *The Journal of Physical Chemistry B*, 112(51):16752–16758, 2008.

[51] P. Eijgenraam. *The Solution of Initial Value Problems using Interval Arithmetics*. 1981.

[52] D. Fey and E. Bullinger. Limiting the parameter search space for dynamic models with rational kinetics using semi-definite programming. In *Proceedings of the Symposium on Computer Applications in Biotechnology*, pages 150–155, 2010.

[53] K. A. Fichthorn and W. H. Weinberg. Theoretical foundations of dynamical Monte Carlo simulations. *The Journal of chemical physics*, 95(2):1090–1096, 1991.

[54] A. F. Filippov. *Differential Equations with Discontinuous Righthand Sides*. Mathematics and Its Applications (Soviet Series). Kluwer Academic Publishers, 1988.

[55] M. Fliess and S. T. Glad. An Algebraic Approach to Linear and Nonlinear Control. In *Essays on Control: Perspectives in Theory and Its Applications*, page 45. Birkhäuser, 1993.

[56] P. M. Frank and X. Ding. Survey of robust residual generation and evaluation

methods in observer-based fault detection systems. *Journal of Process Control*, 7(6):403–424, 1997.

[57] M. Fujiwara, Z. K. Nagy, J. W. Chew, and R. D. Braatz. First-principles and direct design approaches for the control of pharmaceutical crystallization. *Journal of Process Control*, 15(5):493–504, 2005.

[58] H. Garnier and L. Wang. *Identification of Continuous-time Models from Sampled Data*. Advances in Industrial Control. Springer, 2008.

[59] K. Gatermann and P. A. Parrilo. Symmetry groups, semidefinite programs, and sums of squares. *Journal of Pure and Applied Algebra*, 192(1):95–128, 2004.

[60] J.-P. Gauthier and I. Kupka. *Deterministic Observation Theory and Applications*. Cambridge University Press, 2001.

[61] R. Ghosh, A. Hardmeyer, I. Thoenen, and R. Bachofen. Optimization of the Sistrom culture medium for large-scale batch cultivation of *Rhodospirillum rubrum* under semiaerobic conditions with maximal yield of photosynthetic membranes. *Applied and Environmental Microbiology*, 60(5):1698–1700, 1994.

[62] L. Gomelsky, O. V. Moskvin, R. A. Stenzel, D. F. Jones, T. J. Donohue, and M. Gomelsky. Hierarchical regulation of photosynthesis gene expression by the oxygen-responsive PrrBA and AppA-PpsR systems of *Rhodobacter sphaeroides*. *Journal of Bacteriology*, 190(24):8106–8114, 2008.

[63] G. Grammel. Towards fully discretized differential inclusions. *Set-Valued Analysis*, 11(1):1–8, 2003.

[64] H. Grammel and R. Ghosh. Redox-state dynamics of ubiquinone-10 imply cooperative regulation of photosynthetic membrane expression in *Rhodospirillum rubrum*. *Journal of Bacteriology*, 190(14):4912–4921, 2008.

[65] H. Grammel, E. D. Gilles, and R. Ghosh. Microaerophilic cooperation of reductive and oxidative pathways allows maximal photosynthetic membrane biosynthesis in *Rhodospirillum rubrum*. *Applied and Environmental Microbiology*, 69(11):6577–6586, 2003.

[66] W. S. Gray and J. P. Mesko. Controllability and observability functions for model reduction of nonlinear systems. In *Proceedings of the Conference on Information Sciences and Systems*, pages 1244–1249, 1996.

[67] W. S. Gray and J. P. Mesko. Observability functions for linear and nonlinear systems. *Systems & Control Letters*, 38(2):99–113, 1999.

[68] L. Grüne. *Asymptotic Behavior of Dynamical and Control Systems under Perturbation and Discretization*, volume 1783 of *Lecture Notes in Mathematics*. Springer, 2002.

[69] G. Häckl. *Reachable Sets, Control Sets and Their Computation*. PhD thesis, Universität Augsburg, Augsburg, Germany, 1996.

[70] J. Hahn, T. F. Edgar, and W. Marquardt. Controllability and observability

covariance matrices for the analysis and order reduction of stable nonlinear systems. *Journal of Process Control*, 13(2):115–127, 2003.

[71] 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):1076–1083, 2010.

[72] U. Helmke, C. F. Martin, and Y. Zhou. Genericity of Algebraically Observable Polynomial Systems. Technical Report 36, 2003.

[73] D. Henrion. Semidefinite representation of convex hulls of rational varieties. *Acta applicandae mathematicae*, 115(3):319–327, 2011.

[74] D. Henrion and M. Korda. Convex computation of the region of attraction of polynomial control systems. In *Proceedings of the European Control Conference*, pages 676–681, Zürich, Switzerland, 2013.

[75] D. Henrion and J. B. Lasserre. Detecting global optimality and extracting solutions in GloptiPoly. *Positive Polynomials in Control*, pages 581–581, 2005.

[76] D. Henrion, J. B. Lasserre, and C. Savorgnan. Nonlinear optimal control synthesis via occupation measures. In *Proceedings of the IEEE Conference on Decision and Control*, pages 4749–4754, Cancun, Mexico, 2008.

[77] D. Henrion, J. B. Lasserre, and C. Savorgnan. Approximate volume and integration for basic semi-algebraic sets. *SIAM Review*, 51(4):722–743, 2009.

[78] R. Hermann and A. J. Krener. Nonlinear controllability and observability. *IEEE Transactions on Automatic Control*, 22(5):728–740, 1977.

[79] J. P. Hespanha, D. Liberzon, D. Angeli, and E. D. Sontag. Nonlinear norm-observability notions and stability of switched systems. *IEEE Transactions on Automatic Control*, 50(2):154–168, 2005.

[80] J. P. Hespanha, D. Liberzon, and E. D. Sontag. Nonlinear observability and an invariance principle for switched systems. In *Proceedings of the IEEE Conference on Decision and Control*, pages 4300–4305, Las Vegas, USA, 2002.

[81] T. C. Ionescu and J. M. Scherpen. Nonlinear Cross Gramians. In *System Modeling and Optimization*, pages 293–306. Springer, 2009.

[82] A. Isidori. *Nonlinear Control Systems*. Springer, 1995.

[83] L. Jaulin, M. Kieffer, O. Didrit, and E. Walter. *Applied Interval Analysis*. With Examples in Parameter and State Estimation, Robust Control and Robotics. Springer, 2001.

[84] R. E. Kalman. On the general theory of control systems. *IRE Transactions on Automatic Control*, 4(3):110–110, 1959.

[85] R. E. Kalman. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(1):35–45, 1960.

[86] I. Karafyllis and C. Kravaris. Global stability results for systems under sampled-data control. *International Journal of Robust and Nonlinear Control*, 19(10):1105–1128, 2009.

[87] S. Karaman and E. Frazzoli. Linear temporal logic vehicle routing with applications to multi UAV mission planning. *International Journal of Robust and Nonlinear Control*, 21(12):1372–1395, 2011.

[88] K. Kassara. Observability by using viability kernels. *Journal of Control Theory and Applications*, 10(3):303–308, 2012.

[89] Y. Kawano and T. Ohtsuka. Global observability of discrete-time polynomial systems. In *Proceedings of the IFAC Symposium on Nonlinear Control Systems*, pages 203–207, Bologna, Italy, 2010.

[90] Y. Kawano and T. Ohtsuka. Global observability of polynomial systems. pages 2038–2041, 2010.

[91] Y. Kawano and T. Ohtsuka. An algebraic approach to local observability at an initial state for discrete-time polynomial systems. In *Proceedings of the IFAC World Congress*, pages 6449–6453, Milano, Italy, 2011.

[92] Y. Kawano and T. Ohtsuka. Necessary condition for local observability of discrete-time polynomial systems. In *Proceedings of the American Control Conference*, pages 6757–6762, Montréal, Canada, 2013.

[93] Y. Kawano and T. Ohtsuka. Observability analysis of nonlinear systems using pseudo-linear transformation. In *Proceedings of the IFAC Symposium on Nonlinear Control Systems*, pages 606–611, Toulouse, France, 2013.

[94] Y. Kawano and T. Ohtsuka. Observability at an initial state for polynomial systems. *Automatica*, 49(5):1126–1136, 2013.

[95] Y. Kawano and T. Ohtsuka. Simple sufficient conditions for reachability of discrete-time polynomial systems. *IEEE Transactions on Automatic Control*, 58(12):3203–3206, 2013.

[96] H. K. Khalil. Nonlinear Systems. Prentice-Hall, 2002.

[97] M. Kieffer and E. Walter. Guaranteed estimation of the parameters of nonlinear continuous time models: Contributions of interval analysis. *International Journal of Adaptive Control and Signal Processing*, 25(3):191–207, 2011.

[98] M. Kishida, P. Rumschinski, R. Findeisen, and R. D. Braatz. Efficient polynomial-time outer bounds on state trajectories for uncertain polynomial systems using skewed structured singular values. In *Proceedings of the IEEE Symposium on Computer-Aided Control System Design*, pages 216–221, Denver, USA, 2011. IEEE.

[99] H. Kitano. Systems biology: A brief overview. *Science*, 295(5560):1662–1664, 2002.

[100] F. W. Kong, D. Kuhn, and B. Rustem. A cutting-plane method for mixed-

logical semidefinite programs with an application to multi-vehicle robust path planning. In *Proceedings of the IEEE Conference on Decision and Control*, pages 1360–1365, Atlanta, USA, 2010. IEEE.

[101] M. Korda, D. Henrion, and C. N. Jones. Controller design and region of attraction estimation for nonlinear dynamical systems. *arXiv.org*, 2013.

[102] A. J. Krener and W. Respondek. Nonlinear observers with linearizable error dynamics. *SIAM Journal on Control and Optimization*, 23(2):197–216, 1985.

[103] W. Kühn. Rigorously computed orbits of dynamical systems without the wrapping effect. *Computing*, 61(1):47–67, 1998.

[104] L. Küpfer, U. Sauer, and P. A. Parrilo. Efficient classification of complete parameter regions based on semidefinite programming. *BMC Bioinformatics*, 8(1):12, 2007.

[105] B. Labibi, H. J. Marquez, and G. Chesi. Decentralized robust output feedback control for control affine nonlinear interconnected systems. *Journal of Process Control*, 19(5):865–878, 2009.

[106] D. S. Laila. *Design and Analysis of Nonlinear Sampled-data Control Systems*. PhD thesis, University of Melbourne, Melbourne, Australia, 2003.

[107] S. Lall, J. E. Marsden, and S. Glavaški. A subspace approach to balanced truncation for model reduction of nonlinear control systems. *International Journal of Robust and Nonlinear Control*, 12(6):519–535, 2002.

[108] J. B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization*, 11(3):796–817, 2001.

[109] J. B. Lasserre. Semidefinite programming vs. LP relaxations for polynomial programming. *Mathematics of Operations Research*, pages 347–360, 2002.

[110] J. B. Lasserre. Convex sets with semidefinite representation. *Mathematical Programming*, 120(2):457–477, 2009.

[111] J. B. Lasserre. *Moments, Positive Polynomials and Their Applications*, volume 1 of *Imperial College Press Optimization Series*. Imperial College Press, 2010.

[112] 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(4):1643–1666, 2008.

[113] M. Laurent. A comparison of the Sherali-Adams, Lovász-Schrijver, and Lasserre relaxations for 0-1 programming. *Mathematics of Operations Research*, 28(3):470–496, 2003.

[114] M. Laurent. Sums of squares, moment matrices and optimization over polynomials. In M. Putinar and S. Sullivant, editors, *IMA Volumes in Mathematics and its Applications*, pages 157–270. Springer, 2009.

[115] P. S. Lietman, T. J. White, and W. V. Shaw. Chloramphenicol: An enzymolog-

ical microassay. *Antimicrobial Agents and Chemotherapy*, 10(2):347–353, 1976.

[116] Y. Lin and M. A. Stadtherr. Validated solution of initial value problems for ODEs with interval parameters. In *Proceedings of the NSF Workshop on Reliable Engineering Computing*, 2006.

[117] Y. Lin and M. A. Stadtherr. Guaranteed state and parameter estimation for nonlinear continuous-time systems with bounded-error measurements. *Industrial & Engineering Chemistry Research*, 46(22):7198–7207, 2007.

[118] J. S. Liu and R. Chen. Sequential Monte Carlo Methods for Dynamic Systems. *Journal of the American Statistical Association*, 93(443):1032–1044, 1998.

[119] L. Ljung. *System Identification: Theory for the User*. Prentice-Hall, 1987.

[120] L. Ljung. System identification. 1999.

[121] J. Löfberg. YALMIP: A toolbox for modeling and optimization in MATLAB. In *Proceedings of the IEEE Symposium on Computer Aided Control Systems Design*, pages 284–289. IEEE, 2004.

[122] J. Löfberg. Pre-and post-processing sum-of-squares programs in practice. *IEEE Transactions on Automatic Control*, 54(5):1007–1011, 2009.

[123] R. Lohner. Einschließungen bei Anfangs- und Randwertaufgaben gewöhnlicher Differentialgleichungen. In U. Kulisch, editor, *Wissenschaftliches Rechnen mit Ergebnisverifikation*. Vieweg Verlag, Braunschweig, 1989.

[124] L. Lovász and A. Schrijver. Cones of matrices and set-functions and 0–1 optimization. *SIAM Journal on Optimization*, 1(2):166–190, 1991.

[125] D. G. Luenberger. Observing the state of a linear system. *IEEE Transactions on Military Electronics*, 8(2):74–80, 1964.

[126] D. G. Luenberger. *Optimization by Vector Space Methods*. Wiley-Interscience, 1969.

[127] D. G. Luenberger and Y. Ye. *Linear and Nonlinear Programming*. Springer, 2008.

[128] J. Lygeros, C. Tomlin, and S. Sastry. Controllers for reachability specifications for hybrid systems. *Automatica*, 35(3):349–370, 1999.

[129] W. Ma, A. Trusina, H. El-Samad, W. A. Lim, and C. Tang. Defining network topologies that can achieve biochemical adaptation. *Cell*, 138(4):760–773, 2009.

[130] T. Maiwald and J. Timmer. Dynamical modeling and multi-experiment fitting with PottersWheel. *Bioinformatics*, 24(18):2037–2043, 2008.

[131] N. R. Marquis and I. B. Fritz. Enzymological determination of free carnitine concentrations in rat tissues. *Journal of Lipid Research*, 5(2):184–187, 1964.

[132] B. Marteyn, F. B. Scorza, P. J. Sansonetti, and C. Tang. Breathing life into pathogens: the influence of oxygen on bacterial virulence and host responses in the gastrointestinal tract. *Cellular Microbiology*, 13(2):171–176, 2011.

[133] I. Mezić and T. Runolfsson. Uncertainty propagation in dynamical systems. *Automatica*, 44(12):3003–3013, 2008.

[134] H. Michalska and D. Q. Mayne. Moving horizon observers and observer-based control. *IEEE Transactions on Automatic Control*, 40(6):995–1006, 1995.

[135] M. Milanese and C. Novara. Set membership identification of nonlinear systems. *Automatica*, 40(6):957–975, 2004.

[136] M. Milanese and A. Vicino. Optimal estimation theory for dynamic systems with set membership uncertainty: an overview. In M. Milanese, J. Norton, H. Piet-Lahanier, and E. Walter, editors, *Bounding Approaches to System Identification*. Springer, 1996.

[137] J. Monod. The growth of bacterial cultures. *Annual Reviews in Microbiology*, 3(1):371–394, 1949.

[138] B. Moore. Principal component analysis in linear systems: Controllability, observability, and model reduction. *IEEE Transactions on Automatic Control*, 26(1):17–32, 1981.

[139] R. E. Moore, R. B. Kearfott, and M. J. Cloud. *Introduction to Interval Analysis*. SIAM, 2009.

[140] N. S. Nedialkov. *Computing Rigorous Bounds on the Solution of an Initial Value Problem for an Ordinary Differential Equation*. PhD thesis, University of Toronto, Toronto, Canada, 1999.

[141] N. S. Nedialkov, K. R. Jackson, and G. F. Corliss. Validated solutions of initial value problems for ordinary differential equations. *Applied Mathematics and Computation*, 105(1):21–68, 1999.

[142] O. Nelles. *Nonlinear System Identification*. From Classical Approaches to Neural Networks and Fuzzy Models. Springer, 2000.

[143] D. Nešić. A note on observability tests for general polynomial and simple Wiener–Hammerstein systems. *Systems & Control Letters*, 35(4):219–227, 1998.

[144] D. Nešić and L. Grüne. Lyapunov-based continuous-time nonlinear controller redesign for sampled-data implementation. *Automatica*, 41(7):1143–1156, 2005.

[145] J. Nie and M. Schweighofer. On the complexity of Putinar's positivstellensatz. *Journal of Complexity*, 23(1):135–150, 2007.

[146] H. Nijmeijer. Observability of autonomous discrete time non-linear systems: a geometric approach. *International Journal of Control*, 36(5):867–874, 1982.

[147] I. Nowak. *Relaxation and Decomposition Methods for Mixed Integer Nonlinear Programming*, volume 152 of *International Series of Numerical Mathematics*. Birkhäuser, 2005.

[148] J.-I. Oh and S. Kaplan. Generalized approach to the regulation and integration of gene expression. *Molecular Microbiology*, 39(5):1116–1123, 2001.

[149] H. Owhadi, C. Scovel, T. J. Sullivan, M. McKerns, and M. Ortiz. Optimal uncertainty quantification. *SIAM Review*, 55(2):271–345, 2013.

[150] S. A. Park, A. Ko, and N. G. Lee. Stimulation of growth of the human gastric pathogen Helicobacter pylori by atmospheric level of oxygen under high carbon dioxide tension. *BMC Microbiology*, 11(1):96, 2011.

[151] P. A. Parrilo. Semidefinite programming relaxations for semialgebraic problems. *Mathematical Programming*, 96(2):293–320, 2003.

[152] 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.

[153] F. Permenter and P. A. Parrilo. Selecting a monomial basis for sums of squares programming over a quotient ring. In *Proceedings of the IEEE Conference on Decision and Control*, pages 1871–1876, Maui, USA, 2012.

[154] I. R. Petersen. Notions of observability for uncertain linear systems with structured uncertainty. *SIAM Journal on Control and Optimization*, 41(2):345–361, 2002.

[155] T. G. Popova, B. Millis, M. C. Chung, C. Bailey, and S. G. Popov. Anthrolysin O and fermentation products mediate the toxicity of *Bacillus anthracis* to lung epithelial cells under microaerobic conditions. *FEMS Immunology & Medical Microbiology*, 61(1):15–27, 2011.

[156] S. Prajna. *Optimization-based Methods for Nonlinear and Hybrid Systems Verification*. PhD thesis, California Institute of Technology, Pasadena, USA, 2005.

[157] S. Prajna. Barrier certificates for nonlinear model validation. *Automatica*, 42(1):117–126, 2006.

[158] J. O. Ramsay, G. Hooker, D. Campbell, and J. Cao. Parameter estimation for differential equations: A generalized smoothing approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 69(5):741–796, 2007.

[159] J. B. Rawlings and B. R. Bakshi. Particle filtering and moving horizon estimation. *Computers and Chemical Engineering*, 30(10-12):1529–1541, 2006.

[160] J. Richalet. Industrial applications of model based predictive control. *Automatica*, 29(5):1251–1274, 1993.

[161] R. Rihm. Interval methods for initial value problems in ODEs. *Topics in Validated Computations*, pages 173–207, 1994.

[162] M. Rodriguez-Fernandez and J. R. Banga. SensSB: A software toolbox for the development and sensitivity analysis of systems biology models. *Bioinformatics*, 26(13):1675–1676, 2010.

[163] C. Rojas, S. T. Frazier, J. Flanary, and B. S. Slusher. Kinetics and inhibition of glutamate carboxypeptidase II using a microplate assay. *Analytical Biochemistry*, 310(1):50–54, 2002.

[164] P. Rumschinski, S. Borchers, S. Bosio, R. Weismantel, and R. Findeisen. Set-base dynamical parameter estimation and model invalidation for biochemical reaction networks. *BMC Systems Biology*, 4(1):69, 2010.

[165] P. Rumschinski, R. Findeisen, and S. Streif. Finite-time observability measure for polynomial systems. In *Proceedings of the IFAC World Congress*, pages 2800–2805, Cape Town, South Africa, 2014.

[166] P. Rumschinski, D. S. Laila, S. Borchers, and R. Findeisen. Influence of discretization errors on set-based parameter estimation. In *Proceedings of the IEEE Conference on Decision and Control*, pages 296–301, Atlanta, USA, 2010. IEEE.

[167] P. Rumschinski, D. S. Laila, and R. Findeisen. Set-based parameter estimation for symmetric network motifs. In *Proceedings of the IFAC World Congress*, pages 10454–10459, Milan, Italy, 2011.

[168] P. Rumschinski, D. S. Laila, and R. Findeisen. Discrete-to-continuous dynamics reconstruction for bilinear systems. In *Proccedings of the IEEE Conference on Decision and Control*, pages 172–177, Maui, USA, 2012.

[169] P. Rumschinski, J. Richter, A. Savchenko, S. Borchers, J. Lunze, and R. Findeisen. Complete fault diagnosis of uncertain polynomial systems. In *Proceedings of the IFAC International Symposium on Dynamics and Control of Process Systems*, pages 127–132, Leuven, Belgium, 2010.

[170] P. Rumschinski, S. Streif, and R. Findeisen. Combining qualitative information and semiquantitative data for guaranteed invalidation of biochemical network models. *International Journal of Robust and Nonlinear Control*, 22(10):1157–1173, 2012.

[171] M. Sandberg. Convergence of the forward Euler method for nonconvex differential inclusions. *SIAM Journal on Numerical Analysis*, 47:308, 2008.

[172] S. Sastry and C. Desoer. The robustness of controllability and observability of linear time-varying systems. *IEEE Transactions on Automatic Control*, 27(4):933–939, 1982.

[173] A. Savchenko, P. Rumschinski, and R. Findeisen. Fault diagnosis for polynomial hybrid systems. In *Proceedings of the IFAC World Congress*, pages 2755–2760, Milan, Italy, 2011.

[174] A. Savchenko, P. Rumschinski, S. Streif, and R. Findeisen. Complete diagnosability of abrupt faults using set-based sensitivities. In *Proceedings of the IFAC International Symposium on Fault Detection, Supervision and Safety of Technical Processes*, pages 595–600, Mexico City, Mexico, 2012.

[175] A. Savchenko, P. Rumschinski, S. Streif, and R. Findeisen. Structural problem reduction for set-based fault diagnosis. In *Proceedings of the IFAC International Symposium on Dynamics and Control of Process Systems*, pages 595–600, Mumbai, India, 2013.

[176] C. Savorgnan, J. B. Lasserre, and M. Diehl. Discrete-time stochastic optimal control via occupation measures and moment relaxations. In *Proceedings of the IEEE Conference on Decision and Control*, pages 519–524, Shanghai, China, 2009.

[177] 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(1):51–61, 2006.

[178] M. Schweighofer. On the complexity of Schmüdgen's positivstellensatz. *Journal of Complexity*, 20(4):529–543, 2004.

[179] F. C. Schweppe. Recursive state estimation: Unknown but bounded errors and system inputs. *IEEE Transactions on Automatic Control*, 13(1):22–28, 1968.

[180] J. K. Scott. *Reachability Analysis and Deterministic Global Optimization of Differential-Algebraic Systems*. PhD thesis, Massachusetts Institute of Technology, Boston, USA, 2012.

[181] J. K. Scott and P. I. Barton. Bounds on the reachable sets of nonlinear control systems. *Automatica*, 49(1):93–100, 2013.

[182] H. D. Sherali and W. P. Adams. *A Reformulation-Linearization Technique for Solving Discrete and Continuous Nonconvex Problems*, volume 31 of *Nonconvex Optimization and Its Applications*. Springer, 1999.

[183] Y. I. Shokin. *Interval analysis*. Science Press, Siberian Division, Novosibirsk, 1981.

[184] N. Z. Shor. An approach to obtaining global extremums in polynomial mathematical programming problems. *Cybernetics*, 23(5):695–700, 1988.

[185] N. Z. Shor. *Nondifferentiable Optimization and Polynomial Problems*. Springer, 1998.

[186] D. Simon. *Optimal State Estimation*. Kalman, H Infinity, and Nonlinear Approaches. Wiley-Interscience, 2006.

[187] A. B. Singer and P. I. Barton. Bounding the solutions of parameter dependent nonlinear ordinary differential equations. *SIAM Journal on Scientific Computing*, 27(6):2167–2182, 2006.

[188] R. L. Smith, T. West, and W. R. Gibbons. *Rhodospirillum rubrum*: Utilization of condensed corn solubles for poly-(3-hydroxybutyrate-co-3-hydroxyvalerate) production. *Journal of Applied Microbiology*, 104(5):1488–1494, 2008.

[189] S. Sojoudi, J. Lavaei, and A. G. Aghdam. Robust controllability and observability degrees of polynomially uncertain systems. *Automatica*, 45(11):2640–2645, 2009.

[190] E. D. Sontag. On the observability of polynomial systems, I: Finite-time problems. *SIAM Journal on Control and Optimization*, 17(1):139–151, 1979.

[191] E. D. Sontag. A concept of local observability. *Systems & Control Letters*, 5(1):41–47, 1984.

[192] E. D. Sontag. *Mathematical Control Theory.* Deterministic Finite Dimensional Systems. Springer, 1998.

[193] E. D. Sontag. Adaptation and regulation with signal detection implies internal model. *Systems & Control Letters*, 50(2):119–126, 2003.

[194] E. D. Sontag and Y. Rouchaleau. On discrete-time polynomial systems. *Nonlinear Analysis, Theory, Methods & Applications*, 1(1):55–64, 1976.

[195] E. D. Sontag and Y. Wang. Detectability of nonlinear systems. *Proceedings of the Conference on Information Sciences and Systems*, pages 1031–1036, 1996.

[196] E. D. Sontag and Y. Wang. Output-to-state stability and detectability of nonlinear systems. *Systems & Control Letters*, 29(5):279–290, 1997.

[197] S. L. Spencer, S. Gaudet, J. G. Albeck, J. M. Burke, and P. K. Sorger. Non-genetic origins of cell-to-cell variability in TRAIL-induced apoptosis. *Nature*, 459(7245):428–432, 2009.

[198] J. Stoer and R. Bulirsch. *Introduction to Numerical Analysis*, volume 12 of *Texts in Applied Mathematics.* Springer, 2002.

[199] S. Streif, R. Findeisen, S. Waldherr, and F. Allgöwer. Steady state sensitivity analysis of biochemical reaction networks: a brief review and new methods. In *Systems Analysis of Biological Networks*, pages 129–148. Artech House MIT Press, 2009.

[200] S. Streif, D. Henrion, and R. Findeisen. Probabilistic and set-based model invalidation and estimation using LMIs. *Proceedings of the IFAC World Congress*, pages 4110–4115, 2014.

[201] S. Streif, M. Karl, and R. Findeisen. Outlier analysis in set-based estimation for nonlinear systems using convex relaxations. In *Proceedings of the European Control Conference*, pages 2921–2926, 2013.

[202] S. Streif, K.-K. K. Kim, P. Rumschinski, M. Kishida, D. E. Shen, R. Findeisen, and R. D. Braatz. Robustness analysis, prediction and estimation for uncertain biochemical networks. In *Proceedings of the IFAC International Symposium on Dynamics and Control of Process Systems*, pages 1–20, Mumbai, India, 2013.

[203] S. Streif, P. Rumschinski, D. Henrion, and R. Findeisen. Estimation of consistent parameter sets for continuous-time nonlinear systems using occupation measures and LMI relaxations. In *Proceedings of the IEEE Conference on Decision and Control*, pages 6379–6384, Florence, Italy, 2013.

[204] 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):1290–1291, 2012.

[205] S. Streif, N. Strobel, and R. Findeisen. Inner approximations of consistent pa-

rameter sets via constraint inversion and mixed-integer linear programming. In *Proceedings of the IFAC International Symposium on Computer Applications in Biotechnology*, pages 326–331, Mumbai, India, 2013.

[206] A. M. Stuart and A. R. Humphries. Numerical analysis of dynamical systems. *Acta Numerica*, 3(1):467–572, 1994.

[207] A. M. Stuart and A. R. Humphries. *Dynamical Systems and Numerical Analysis*, volume 2 of *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge University Press, 1998.

[208] J. F. Sturm. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization methods and software*, 11(1-4):625–653, 1999.

[209] T. Tao. *An Introduction to Measure Theory*, volume 126 of *Graduate Studies in Mathematics*. American Mathematical Society, 2011.

[210] A. Tarantola. *Inverse Problem Theory and Methods for Model Parameter Estimation*. SIAM, 2005.

[211] Y. Tian, T. Yue, Y. Yuan, P. K. Soma, P. D. Williams, P. A. Machado, H. Fu, R. J. Kratochvil, C.-i. Wei, and Y. M. Lo. Tobacco biomass hydrolysate enhances coenzyme Q10 production using photosynthetic *Rhodospirillum rubrum*. *Bioresource Technology*, 101(20):7877–7881, 2010.

[212] B. Tibken. Observability of nonlinear systems-an algebraic approach. In *Proceedings of the IEEE Conference on Decision and Control*, pages 4824–4825. IEEE, 2004.

[213] A. van den Bos. *Parameter Estimation for Scientists and Engineers*. Wiley-Interscience, 2007.

[214] L. Vandenberghe and S. P. Boyd. Semidefinite programming. *SIAM Review*, 38(1):49–95, 1996.

[215] R. S. Varga. On recurring theorems on diagonal dominance. *Linear Algebra and its Applications*, 13(1-2):1–9, 1976.

[216] 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(8):911–915, 1989.

[217] E. Walter and H. Piet-Lahanier. Estimation of parameter bounds from bounded-error data: A survey. *Mathematics and Computers in Simulation*, 32(5-6):449–468, 1990.

[218] K. Wang and A. N. Michel. Necessary and sufficient conditions for the controllability and observability of a class of linear, time-invariant systems with interval plants. *IEEE Transactions on Automatic Control*, 39(7):1443–1447, 1994.

[219] H. Witsenhausen. Sets of possible states of linear systems given perturbed observations. *IEEE Transactions on Automatic Control*, 13(5):556–558, 1968.

[220] O. Wolkenhauer. Systems biology: The reincarnation of systems theory applied in biology? *Briefings in Bioinformatics*, 2(3):258–270, 2001.

[221] T.-M. Yi, M. Fazel, X. Liu, T. Otitoju, J. Goncalves, A. Papachristodoulou, S. Prajna, and J. Doyle. Application of robust model validation using SOSTOOLS to the study of G-protein signaling in yeast. In *Proceedings of the Conference on Foundations of Systems Biology in Engineering*, 2005.

[222] T.-M. Yi, Y. Huang, M. I. Simon, and J. Doyle. Robust perfect adaptation in bacterial chemotaxis through integral feedback control. *Proceedings of the National Academy of Sciences*, 97(9):4649, 2000.

[223] H. Younesi, G. Najafpour, K. S. Ku Ismail, A. R. Mohamed, and A. H. Kamaruddin. Biohydrogen production in a continuous stirred tank bioreactor from synthesis gas by anaerobic photosynthetic bacterium: *Rhodopirillum rubrum*. *Bioresource Technology*, 99(7):2612–2619, 2008.

[224] L. Zeiger and H. Grammel. Model based high cell density cultivation of *Rhodospirillum rubrum* under respiratory dark conditions. *Biotechnology and Bioengineering*, 105(4):729–739, 2010.

[225] A. Zeng, T. G. Byun, C. Posten, and W. D. Deckwer. Use of respiratory quotient as a control parameter for optimum oxygen supply and scale-up of 2, 3-butanediol production under microaerobic conditions. *Biotechnology and Bioengineering*, 44(9):1107–1114, 1994.