

**Band
426**

Verlagsschriftenreihe des Heinz Nixdorf Instituts
Prof. Dr.-Ing. Ansgar Trächtler (Hrsg.)
Regelungstechnik und Mechatronik

Michael Hesse

Interaktive Inbetriebnahme von Steuerungen und Regelungen für partiell bekannte dynamische Systeme mittels Gauß-Prozess-Regression

Michael Hesse

***Interaktive Inbetriebnahme von
Steuerungen und Regelungen für par-
tiell bekannte dynamische Systeme
mittels Gauß-Prozess-Regression***

***Interactive Commissioning of
Control Applications for partially
known dynamic Systems using
Gaussian Process Regression***

Bibliografische Information Der Deutschen Bibliothek

Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie detaillierte bibliografische Daten sind im Internet über <http://dnb.ddb.de> abrufbar

Band 426 der Verlagsschriftenreihe des Heinz Nixdorf Instituts

© Heinz Nixdorf Institut, Universität Paderborn – Paderborn – 2024

ISSN (Online): 2365-4422
ISBN: 978-3-947647-45-3

Das Werk einschließlich seiner Teile ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung der Herausgeber und des Verfassers unzulässig und strafbar. Das gilt insbesondere für Vervielfältigung, Übersetzungen, Mikroverfilmungen, sowie die Einspeicherung und Verarbeitung in elektronischen Systemen.

Als elektronische Version frei verfügbar über die Digitalen Sammlungen der Universitätsbibliothek Paderborn.

Satz und Gestaltung: Michael Hesse

**Interaktive Inbetriebnahme von Steuerungen und Regelungen
für partiell bekannte dynamische Systeme mittels
Gauß-Prozess-Regression**

zur Erlangung des akademischen Grades
DOKTOR DER INGENIEURWISSENSCHAFTEN (Dr.-Ing.)
der Fakultät für Maschinenbau
der Universität Paderborn

genehmigte
DISSERTATION

von
Michael Hesse, M. Sc.
aus Bünde

Tag des Kolloquiums: 7. Oktober 2024
Referentin: Dr.-Ing. Julia Timmermann
Korreferent: Prof. Dr. Eyke Hüllermeier

Danksagung

Der Inhalt der vorliegenden Dissertation entstand während meiner Tätigkeit als wissenschaftlicher Mitarbeiter in der Nachwuchsgruppe „Datengetriebene Methoden in der Regelungstechnik“ am Heinz Nixdorf Institut, sowie am Lehrstuhl für Regelungstechnik und Mechatronik der Universität Paderborn und am Fraunhofer-Institut für Entwurfstechnik Mechatronik.

Ein herzliches Dankeschön richte ich an Dr.-Ing. Julia Timmermann für ihre langjährige Unterstützung in Form von konstruktivem Feedback und wertvollen Ratschlägen. Das entgegengebrachte Vertrauen und die herausragende Betreuung weiß ich sehr zu schätzen. Herrn Prof. Dr. Eyke Hüllermeier von der Ludwig-Maximilians-Universität München danke ich für die Übernahme des Korreferats. Sein Interesse an meiner Arbeit macht mich stolz. Gleichermaßen bedanke ich mich bei Frau Prof. Dr. Iryna Mozgova für ihr Mitwirken in der Promotionskommission. Mein ganz besonderer Dank gilt Herrn Prof. Dr.-Ing. habil. Ansgar Trächtler, welcher mich durch seine ausgezeichneten Vorlesungen für die Regelungstechnik begeisterte und mir später die Möglichkeit gab meine Forschungsinteressen an seinem Lehrstuhl weiter zu vertiefen. Die zahlreichen Gespräche haben mich nicht nur fachlich, sondern auch persönlich wachsen lassen – vielen Dank!

Außerdem möchte ich einen herzlichen Dank an meine Kollegen und Kolleginnen für die positive Arbeitsatmosphäre aussprechen, die in dieser Form einzigartig ist. Auf die tatkräftige Unterstützung von Martin Leibenger und Jörg Schaffrath bei Problemen an Prüfständen konnte ich mich jederzeit verlassen – danke dafür. Des Weiteren gilt mein Dank allen studentischen Hilfskräften und Abschlussarbeitern, die mich mit ihren Beiträgen und Tätigkeiten bei meiner Forschung weitergebracht haben. An dieser Stelle möchte ich insbesondere Changcheng Li, Fabian Reiling, Jan Herbermann und meinen langjährigen Freund Ahmed Ismail namentlich nennen. Für den engeren fachlichen Austausch und die erfolgreiche Zusammenarbeit danke ich Dr.-Ing. Annika Junker, Dr.-Ing. Ricarda Götte und meinem ehemaligen Bürokollegen Cederic Lenz. In besonderer Erinnerung werde ich die Tennisduelle gegen Niklas Fittkau behalten, welche für den nötigen Ausgleich während der Erstellung dieser Arbeit sorgten – danke dir.

Abschließend widme ich meinen größten Dank meiner liebevollen Partnerin Madleen, welche mich in jeder Lebenslage unterstützt und einen wesentlichen Beitrag zum Gelingen meiner Promotion beigetragen hat. Danke für deine Zuversicht und deinen Rückhalt!

Zusammenfassung

Die Inbetriebnahme von Steuerungen und Regelungen stellt sicher, dass ein mechatronisches System ordnungsgemäß funktioniert und den Anforderungen gerecht wird. Der modellbasierte Entwurf basiert auf einem genauen Simulationsmodell. Allerdings ist dieser klassische Weg bei komplexen Systemen oft nicht praktikabel, da die analytische Modellierung zu kompliziert und zeitaufwendig ist. Diese Forschungslücke wird durch Verfahren adressiert, die eine effiziente und sichere Inbetriebnahme ermöglichen. Diese Verfahren kombinieren Regelungstechnik und Reinforcement Learning und nutzen vorhandenes Wissen über die Regelungsaufgabe, um Korrekturen basierend auf Messdaten und der probabilistischen Gauß-Prozess-Regression vorzunehmen. Das Vorwissen kann als teilweise bekanntes physikalisches Modell oder als Steuerungsfunktion vorliegen.

Anwendungsbeispiele sind der Ultraschalldrahtbondprozess, verschiedene Pendelsysteme und ein Hexapod. Eine angepasste Bayessche Optimierung wird zur Identifikation einer Steuerparametrisierung für das Ultraschallboden eingesetzt. Außerdem wird eine hybride Optimalsteuerung für das Doppelpendel auf einem Wagen entwickelt und erfolgreich validiert. Für einen Hexapod zur Fahrzeugachsprüfung wird eine hybride Zustandslinearisierung formuliert und ein Funktionsnachweis im Rahmen einer Simulation erbracht. Die Einhaltung technischer Rahmenbedingungen und stabiles Systemverhalten werden durch probabilistische Prädiktionen gewährleistet. In allen Anwendungsfällen wird eine Steigerung der Effizienz und Güte erzielt.

Abstract

The commissioning of control systems ensures that a mechatronic system functions properly and meets the requirements. Model-based design is based on a precise simulation model. However, this classic approach is often impractical for complex systems, as analytical modeling is too complicated and time-consuming. This research gap is addressed by methods that enable efficient and safe commissioning. These methods combine control engineering and reinforcement learning and use existing knowledge about the control task to make corrections based on measurement data and probabilistic Gaussian process regression. The prior knowledge can be available as a partially known physical model or as a control function.

Application examples include the ultrasonic wire bonding process, various pendulum systems and a hexapod. An adapted Bayesian optimization is used to identify a control parameterization for ultrasonic bonding. In addition, a hybrid optimal control for the double pendulum on a cart is developed and successfully validated. A hybrid state linearization is formulated for a hexapod for vehicle axle testing and a proof of concept is provided in a simulation. Compliance with technical framework conditions and stable system behavior are ensured by probabilistic predictions. An increase in efficiency and quality is achieved in all use cases.

Vorveröffentlichungen

- [HTHT18] HESSE, M.; TIMMERMANN, J.; HÜLLERMEIER, E.; TRÄCHTLER, A.: A Reinforcement Learning Strategy for the Swing-up of the Double Pendulum on a Cart. *Procedia Manufacturing* 24 (2018). *4th International Conference on System-Integrated Intelligence: Intelligent, Flexible and Connected Systems in Products and Production (SysInt)*, S. 15–20.
- [HHTT22] HESSE, M.; HUNSTIG, M.; TIMMERMANN, J.; TRÄCHTLER, A.: Batch Constrained Bayesian Optimization for Ultrasonic Wire Bonding Feed-Forward Control Design. *Proceedings of the 11th International Conference on Pattern Recognition, Applications and Methods (ICPRAM)*. 2022, S. 383–394.
- [HTT23] HESSE, M.; TIMMERMANN, J.; TRÄCHTLER, A.: Hybrid Optimal Control for Dynamical Systems using Gaussian Process Regression and Unscented Transform. *European Control Conference (ECC)*. 2023, S. 1–8.
- [HSTT24] HESSE, M.; SCHWARZER, L.; TIMMERMANN, J.; TRÄCHTLER, A.: Robust and Efficient Hybrid Optimal Control via Gaussian Process Regression and Multiple Shooting with Experimental Validation on a Double Pendulum on a Cart. *Proceedings of the 94th Annual Meeting of the International Association of Applied Mathematics and Mechanics (GAMM)*. 2024.

Interaktive Inbetriebnahme von Steuerungen und Regelungen für partiell bekannte dynamische Systeme mittels Gauß-Prozess-Regression

Inhaltsverzeichnis

1 Einleitung und Motivation	1
1.1 Einführung in den regelungstechnischen Entwurfsprozess	1
1.2 Identifikation der Forschungslücke und Zielsetzung	9
1.3 Aufbau der Arbeit	17
2 Mathematische Grundlagen	19
2.1 Regelungstechnik	19
2.2 Wahrscheinlichkeitstheorie	23
2.2.1 Diskrete und kontinuierliche Zufallsvariablen	23
2.2.2 Erwartungswert und Varianz	27
2.2.3 Transformationssatz und Bayessche Regel	30
2.2.4 Univariate und multivariate Normalverteilung	31
2.2.5 Numerische Approximation	33
2.3 Anwendung der Wahrscheinlichkeitstheorie auf Pendeldynamik . .	35
2.3.1 Parameteridentifikation mittels Bayesscher linearer Regression	35
2.3.2 Zustandspropagation mittels Moment Matching, Linearisierung und Unscented Transform	40
2.3.3 Annäherung an exakte Folgezustandsverteilung	46
2.4 Maschinelles Lernen mittels Gauß-Prozess-Regression	47
2.4.1 Partitionierung einer multivariaten Gaußverteilung	49
2.4.2 Lösung der Regressionsaufgabe	51
2.4.3 Kovarianzfunktion und Hyperparameter	54
2.5 Ergänzende theoretische und praktische Aspekte	55
2.5.1 Matérn und Piecewise Squared Exponential Kernel	55
2.5.2 Interpretation als radiales Basisfunktionen-Netz	57
2.5.3 Details zur Anwendung und Implementierung	60

3 Steuerungsentwurf mittels Bayesscher Optimierung	65
3.1 Aufgabenstellung des parametrisierten Steuerungsentwurfs	66
3.2 Bayessche Optimierung als Lösungsansatz	67
3.3 Anwendungsbeispiel: Ultraschalldrahtbondprozess	78
3.4 Physikalische Modellbildung für das Ultraschalldrahtbonden	82
3.5 Angepasste Bayessche Optimierung für den Steuerungsentwurf beim Ultraschalldrahtbonden	89
3.6 Anwendung und Ergebnisanalyse	96
3.7 Weiterführender Entwurf mit Bayesscher Optimierung	102
4 Hybride Optimalsteuerung dynamischer Systeme	107
4.1 Klassischer modellbasierter Optimalsteuerungsentwurf	107
4.2 Mehrfachpendel auf einem Wagen als Anwendungsbeispiel	110
4.3 Probabilistic Inference for Learning Control	116
4.4 Entwicklung der hybriden Optimalsteuerung und Erprobung an einem simulierten voll-aktuierten Doppelpendel	124
4.5 Erweiterung um Zwei-Freiheitsgrade-Regelungsstruktur und Validierung am Doppelpendel auf einem Wagen	136
5 Hybride Zustandslinearisierung für mechanische Mehrkörpersysteme	153
5.1 Klassischer Regelungsentwurf durch Zustandslinearisierung	154
5.2 Herleitung der hybriden Zustandslinearisierung	155
5.3 Konzeptionierung einer Regelung für einen Hexapod mittels hybrider Zustandslinearisierung	161
6 Zusammenfassung und Ausblick	169
Literaturverzeichnis	177

Anhang

A1 Mathematischer Anhang	188
A2 Sequentielle Quadratische Programmierung	194
A3 Details zu Experimenten	196

Abkürzungsverzeichnis

ARD	Automatic Relevance Determination
BFGS	Broyden–Fletcher–Goldfarb–Shanno
BO	Bayessche Optimierung
DGL	Differentialgleichung
EHVI	Expected Hypervolumen Improvement
EI	Expected Improvement
EKF	Extended Kalman Filter
ES	Entropy Search
FEM	Finite Elemente Methode
GP	Gauß-Prozess
HiL	Hardware in the Loop
HMCMC	Hamilton-Markov-Chain-Monte-Carlo
HV	Hypervolumen
ISE	Informed Squared Exponential
KI	Künstliche Intelligenz
KKT	Karush-Kuhn-Tucker
LCB	Lower Confidence Bound
LSL, USG	Lower Specification Limit, Untere Spezifikationsgrenze
MCMC	Markov-Chain-Monte-Carlo
MES	Max-Value Entropy Search
ML	Machine Learning, Maximum Likelihood
MM	Moment Matching
MPC	Model Predictive Control

PCA	Principle Component Analysis
PDF	Probability Density Function
PF	Partikelfilter
PI	Probability of Improvement
PILCO	Probabilistic Inference for Learning COntrol
PSE	Piecewise Squared Exponential
RBF	Radiales Basisfunktionen-Netz
RL	Reinforcement Learning
SE	Squared Exponential
SEI	Scaled Expected Improvement
SQP	Sequentielle Quadratische Programmierung
SVD	Singular Value Decomposition
UB, LB	Upper Bound, Lower Bound
UKF	Unscented Kalman Filter
UT	Unscented Transform

Symbolverzeichnis

Name	Beschreibung	Einheit
A	Dynamikmatrix, Anbindungsfläche	m^2
B	Eingangsmatrix	—
C_{pK}	Prozessfähigkeitsindex	—
C	Menge der differenzierbaren Funktionen	—
D	Daten, Menge	—
F	Kraft	$\text{N}, \text{cN}, \text{dN}$
G	Gradient	—
H	Maximale Anzahl von zeitlichen Schritten, Entropie	—
J	Güßtemaß, Trägheitsmoment, Jacobimatrix	kg m^2
K	Reglerverstärkung, Kovarianz-/Daten-Matrix	—
L	Lagrangefunktion	—
P	Wahrscheinlichkeit, Leistung, Riccatische Matrix	$\%, \text{W}$
S	(Ko-)Varianzmatrix, Schwerpunkt	—
T	Gesamtdauer, Tool Center Point	s
U	Elektrische Spannung, potentielle Energie, Umgebung	V, J
V	Kinetische Energie	J
W	Gewichtungsmatrix, Energie	—
Ω	Ereignisraum	—
Φ	Integrationsschema, Grammsche Matrix	—
	Standardverteilungsfunktion einer Normalverteilung	—
Σ	(Ko-)Varianzmatrix	—
α	Akquisitionsfunktion, Gewichtung eines RBF, Winkel	rad
δ	Binäre Variable, Diracfunktion	—
η	Hyperparameter	—
γ	Reinigungsgrad	—
κ	Entwurfsparameter, Konditionszahl	—
λ	Eigenwert	—
\mathbb{C}	Kreuzkovarianzoperator	—
\mathbb{E}	Erwartungswertoperator	—
\mathbb{L}	Matrix einer Choleskyzerlegung	—
\mathbb{N}	Menge der natürlichen Zahlen	—
\mathbb{R}	Reelle Zahlenmenge	—
\mathbb{V}	Varianzoperator	—
\mathbb{Z}	Ganze Zahlenmenge	—

Name	Beschreibung	Einheit
\mathcal{A}	Realisierungen	—
\mathcal{GP}	Gauß-Prozess	—
\mathcal{O}	O-Notation	—
\mathcal{P}	Paretofront	—
μ	Mittel-/Erwartungswert, Reibwert	—
ϕ	Standardnormalverteilung, Transformationsfunktion, Basisfunktion	—
σ	Standardabweichung	—
τ	Zeitliches Diskretisierungsgitter, Scherkraft(-festigkeit)	N
θ	Parametrisierung einer Steuerung oder Regelung, Optimierungsvektor	—
φ	Straffunktion, Winkel	rad
ξ	Grenzwert	—
c	Federsteifigkeit, mechanische Modellparameter	$\frac{\text{N}}{\text{m}}$
d	Dämpfungskonstante, Distanzmaß	$\frac{\text{kg}}{\text{s}}, \frac{\text{kg m}^2}{\text{s}}$
f	(Dynamik-)Funktion, Frequenz	kHz
g	Ausgangsfunktion, (Un-)Gleichheitsnebenbedingung, Gravitationskonstante	$\frac{\text{m}}{\text{s}^2}$
h	Gleichheitsnebenbedingung	—
i, j	Index	—
k	Diskretisierungsindex, Kovarianz-/Kernelfunktion	—
l	Lengthscale	—
m	Mittel-/Erwartungswert, Masse	kg
n	Dimension	—
p	Parametrierung eines physikalischen Modells, Wahrscheinlichkeitsverteilung	—
q	Generalisierte Koordinaten	—
r	Regelgesetz	—
s	Position	m
$t, \Delta t$	Zeit, zeitliche Schrittweite	ms, s
u	Stellgröße	—
w	Sollwert, Gewichtung	—
x	Zustand, Eingang eines ML-Modells	—
y	Ausgangsgröße	—
z	Höhenkoordinate	m

1 Einleitung und Motivation

Im Allgemeinen ist die Inbetriebnahme ein Prozess, welcher sicherstellt, dass ein technisches System ordnungsgemäß funktioniert, sowie den vorhergehenden Anforderungen und Erwartungen gerecht wird. Der Fokus der vorliegenden Arbeit liegt im Speziellen auf mechatronischen Systemen, welche durch ein enges synergetisches Zusammenwirken der Fachdisziplinen Maschinenbau, Elektrotechnik und Informatik gekennzeichnet sind. Mechatronische Systeme besitzen zudem Sensorik und Aktorik, womit über eine informationsverarbeitende Einheit eine gewünschte Beeinflussung des Systems vorgenommen wird. Hierfür sind Steuerungs- und Regelungsalgorithmen notwendig, welche bei der Inbetriebnahme erstmalig erprobt und ordnungsgemäß eingestellt werden müssen. Im Rahmen dieser Arbeit werden die dabei auftretenden Probleme benannt, die sich daraus ableitende Konsequenz und die Forschungslücke identifiziert und weiterführend konkrete Lösungsansätze vorgelegt. Die Lösungsansätze werden in Form von zielgerichteten Verfahren anhand von passenden Anwendungsbeispielen erläutert und basieren im Kern auf der Kombination von bewährter klassischer Regelungstechnik mit den modernen Methoden des *Reinforcement Learning* (RL). Die Grundidee dieses Ansatzes und die zugeordneten beiden Fachdisziplinen werden in Unterabschnitt 1.2 weiter besprochen. Zunächst wird die Problematik bei der Inbetriebnahme von Steuerungen und Regelungen anhand des regelungstechnischen Entwurfsprozesses näher beschrieben und erklärt.

1.1 Einführung in den regelungstechnischen Entwurfsprozess

Die Regelungstechnik ist eine ingenieurtechnische und wissenschaftliche Disziplin, deren übergeordneten Aufgabe es ist, einer Ausgangsgröße eines technischen Systems durch die Stellgröße ein bestimmtes Sollverhalten, d. h. ein gewünschtes Verhalten, aufzuprägen, und zwar gegen den Einfluss einer Störgröße, die nur unvollständig bekannt ist [FKL⁺22]. Die Dynamik des Systems, also dessen zeitliches Verhalten, und ihre gezielte Beeinflussung sind daher von zentraler Bedeutung in der Regelungstechnik. Eine Charakterisierung erfolgt über den Zustand des Systems, welcher dieses zeitlich eindeutig beschreibt. Im Fall eines mechanischen Systems setzt sich der Zustand beispielsweise aus den Positionen und Geschwindigkeiten der verschiedenen Massen zusammen. Der Zustand spiegelt die Wirkung der Stell- und Störgröße auf die Dynamik wider. Soll das System von einem bestimmten Zustand in einen anderen überführt werden, dann ist dazu die passende Vorgabe der Stellgröße notwendig. Ist diese in Form einer vorab bekannten Steuerung gefunden, ist dies nicht zwangsläufig ausreichend, um den Zielzustand zu erreichen, denn die unbekannte Störgröße wirkt sich negativ aus und bringt das System von seinem vorgesehenen Kurs - der Soll-Trajektorie - ab. Um diese Problematik zu lösen, wird der Zustand durch

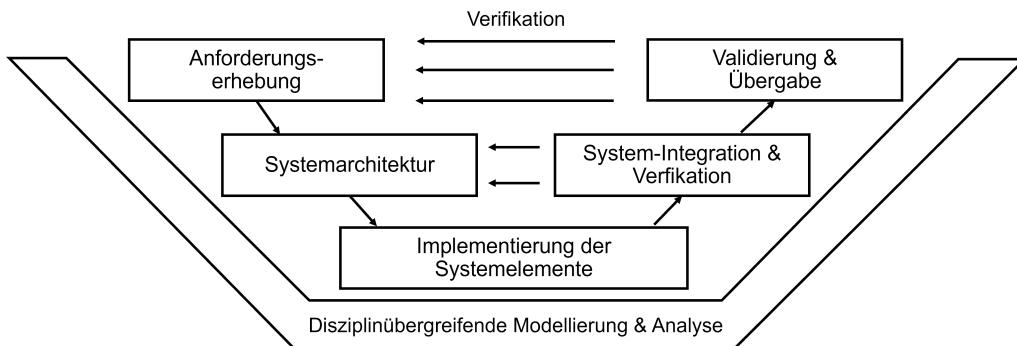


Bild 1-1: Vorgehensmodell für den Entwurf eines mechatronischen Systems nach VDI/VDE 2206 [GM02].

die Sensorik des Systems laufend erfasst und auf Basis dieser Information eine Korrektur der Stellgröße durch eine Regelung vorgenommen. Dieser Lösungsweg ist erfolgreich, da sich die Störung indirekt im gemessenen Zustand bemerkbar macht und die Regelung somit eine Kompensation vornehmen kann.

Der Regelungstechniker¹ steht neben der ganzheitlichen Konzipierung und Konstruktion des mechatronischen Gesamtsystems auch vor etwaigen Herausforderungen, welche die Entwicklung der Steuerung und Regelung betreffen. Zur Erläuterung dieser Herausforderungen wird als Ausgangspunkt des regelungstechnischen Entwurfsprozesses das Vorgehensmodell (kurz V-Modell) nach Richtlinie VDI 2206 [GM02] (s. Bild 1-1) herangezogen. Im frühen Entwicklungsstadium erfolgt eine Anforderungserhebung, in der beispielsweise die technischen Rahmenbedingungen geklärt werden, und die Festlegung der übergeordneten Systemarchitektur. Anschließend erfolgt die Implementierung der Systemelemente, wobei die physikalischen Wirkungszusammenhänge definiert und miteinander verknüpft werden. Innerhalb der späten Entwicklungsphase wird die Systemintegration, also die reale Umsetzung des Entwurfs, umgesetzt und eine ausgiebige Validierung vorgenommen. Eine besondere Charakteristik des V-Modells ist die fortlaufende Verifikation mit den früheren Entwicklungsschritten, wodurch sich insgesamt ein iteratives Vorgehen ergibt. Das V-Modell lässt somit die Rückkehr zu einem früheren Schritt zu, um Entwurfsfehler oder Verbesserungen vorzunehmen. Eine weitere Charakteristik ist die disziplinübergreifende Modellierung, Simulation und Analyse, welche parallel zu den bereits genannten Schritten durchgeführt wird und das V-Modell absichert. Die Implementierung einer Simulationsumgebung erscheint zunächst kosten- und zeitaufwändig, führt in der Gesamtbetrachtung jedoch zu einem positiven Ergebnis, auch bezüglich dieser beiden Zielgrößen. Mit Hilfe des Simulationsmodells lassen sich bereits vor der rea-

¹Aus Gründen der besseren Lesbarkeit wird auf die gleichzeitige Verwendung der Sprachformen männlich, weiblich und divers verzichtet. Sämtliche Personenbezeichnungen gelten gleichermaßen für alle Geschlechter.

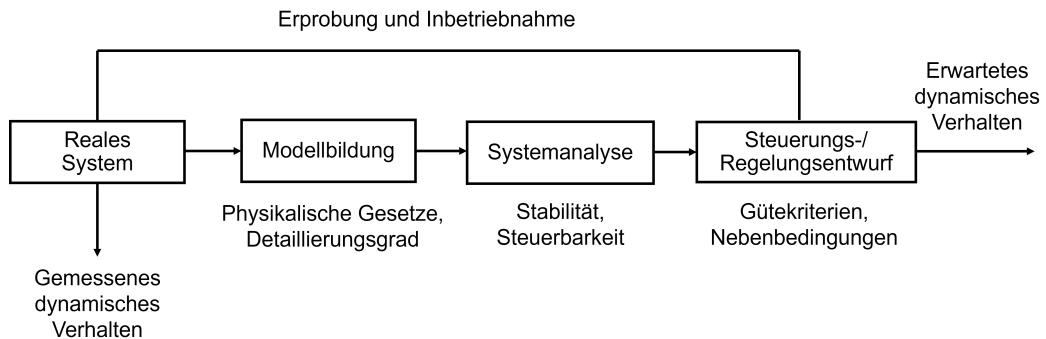


Bild 1-2: Modellbasierter Steuerungs- und Regelungsentwurf

len Umsetzung ausgiebige Auswertungen, welche beispielsweise die Dynamik betreffen, vornehmen. Auf diese Weise kann der Ingenieur weitere konstruktive Maßnahmen vornehmen, die für das Verhalten des Prototypen als sinnvoll erachtet werden. Des Weiteren wird anhand des Modells bereits der Steuerungs- und Regelungsentwurf vorgenommen. Für bestimmte Regelungskonzepte stellt das Simulationsmodell nicht nur eine geeignete Testumgebung dar, sondern ist ein essentieller Bestandteil des Konzeptes selbst. Aus den genannten Gründen ist die disziplinübergreifende Modellierung und Analyse ein wesentliches und hilfreiches Element des V-Modells.

Das Hauptaugenmerk der vorliegenden Arbeit liegt auf dem letzten Entwicklungsschritt, bzw. der Validierung. In diesem Schritt findet die Inbetriebnahme der Steuerung oder Regelung statt. Bild 1-2 stellt den modellbasierten Entwurf genauer dar. Im Rahmen des V-Modells wird davon ausgegangen, dass das reale System erst am Ende des Entwicklungsprozesses existiert. Neben diesem Szenario kann der Regelungstechniker auch vor der Aufgabe stehen, dass es bereits ein existierendes System gibt, wofür der Entwurf durchgeführt werden soll. Unabhängig vom betrachteten Szenario ist der erste Schritt die Modellbildung auf der Grundlage von physikalischen Gesetzmäßigkeiten. Ein wesentlicher Punkt ist hierbei die Frage nach dem Detaillierungsgrad der Modellierung, welcher in Hinblick auf die Komplexität der Anwendung bzw. Aufgabe optimal gewählt werden sollte. Ein zu geringer Grad kann das reale System womöglich nicht genau genug beschreiben, womit die durchgeführten Analysen und Auswertungen unbrauchbar werden. Ein zu hoher Grad ist in dieser Metrik mutmaßlich geeigneter, hat allerdings den Nachteil, dass er zu zeit- und kostenintensiv ist und damit den gesamten Entwicklungsprozess verzögern kann. Die angesprochenen Vorteile des modellbasierten Entwurfs würden auf diese Weise revidiert werden. Nach der Aufstellung des Modells erfolgt die Systemanalyse. Der Regelungstechniker erhält hierbei tiefere Einsichten in die Eigenschaften des Systems bspw. in Bezug auf dessen Stabilität oder Steuer- und Beobachtbarkeit. Anschließend wird das passende Regelungskonzept ausgewählt und, anhand von festgelegten Gütekriterien und technischen Nebenbedingungen, ausgelegt. Über eine weitere Si-

mulation, in welcher das Modell des Systems mit der entwickelten Regelung angesteuert wird, resultiert eine bestimmte Erwartung des Regelungstechnikers an das dynamische Verhalten des realen Systems. Mit dieser Erwartungshaltung wird die Erprobung und Inbetriebnahme durchgeführt, wobei das bisher verwendete Simulationsmodell durch das reale System ersetzt wird. Bei der Erprobung wird das wahre dynamische Verhalten gemessen, so dass rückblickend ein Vergleich zur Erwartung gezogen und eine Bewertung vorgenommen werden kann.

An diesen Punkt angelangt, zeigt sich eine mögliche Problematik des modellbasierten Entwurfs, welche den Hauptgegenstand dieser Arbeit bildet. Stimmen das gemessene und erwartete dynamische Verhalten nicht überein, spiegelt sich dies in der Qualität der Regelung wider. Die Regelung kann in diesem Fall von einem suboptimalen Ergebnis bis hin zum kompletten Fehlschlag durch instabiles Verhalten führen. Der ausschlaggebende Punkt ist die Voraussetzung mit welcher der modellbasierte Entwurf durchgeführt wird: Das aufgestellte Modell bildet das reale System genau genug nach. Wenn diese Voraussetzung nicht erfüllt ist, kommt es zu der angesprochenen Problematik. Der Regelungstechniker steht entsprechend vor der Herausforderung, den rein theoretischen Entwurf mit der Wirklichkeit des realen Versuchs, also der Praxis, in Einklang zu bringen. Die *IEEE Control Society* greift unter anderem diese Herausforderung in [AAA⁺23] auf und verweist auf die immer komplexer werdenden mechatronischen Systeme, sowie deren zunehmende Funktionalität und dafür notwendigen Regelalgorithmen. Die Probleme bei der Inbetriebnahme müssen entsprechend durch geeignete regelungstechnische Werkzeuge adressiert werden, welche die aufklaffende Lücke zwischen Theorie und Praxis überbrücken. Als gedanklicher Anstoß wird in [AAA⁺23] dazu das folgende Zitat angeführt:

*„The gap between theory and practice is a lot bigger
in practice than in theory.“*

Objektiv betrachtet, sollte die Lücke aus beiden Perspektiven im Grunde gleich weit entfernt sein. Während der Inbetriebnahme ist die subjektive Wahrnehmung des Regelungstechnikers jedoch eine andere, so dass die Umsetzung deutlich schwieriger erscheint als sie es nach den theoretischen Vorüberlegungen normalerweise sein dürfte. Als Fazit bleibt an dieser Stelle festzuhalten, dass die Problematiken bei der Inbetriebnahme bereits beim theoretischen Entwurfsprozess berücksichtigt werden müssen, um ein zufriedenstellendes Endergebnis erzielen zu können.

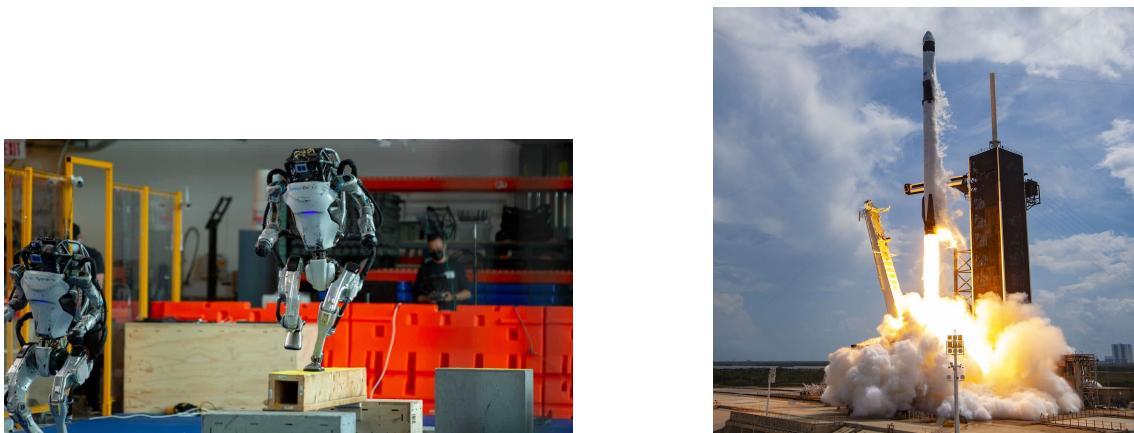


Bild 1-3: Humanoider Roboter (© Boston Dynamics) (links), © SpaceX Rakete (rechts).

Bild 1-3 zeigt beispielhaft zwei aktuell medienpräsente Systeme, die die Schwierigkeiten bei der Inbetriebnahme verdeutlichen sollen. Das erste Beispiel bezieht sich auf die humanoiden Roboter des Unternehmens *Boston Dynamics*, die sowohl für Forschungszwecke als auch industrielle Anwendungen gedacht sind. Aufmerksamkeit erhält das Unternehmen durch seine mittlerweile zahlreichen Videos, in denen die Roboter Parkour-ähnliche Bewegungsabfolgen (u. a. Rückwärtssaltos) ausführen. Aufgrund der zahlreichen (und z. T. schweren) Massen, sowie der hohen Dynamik der Manöver, sind hierfür präzise Steuerungs- und Regelungsvorgaben notwendig. In begleitenden Dokumentationen wird gezeigt, wie viel Aufwand für die Inbetriebnahme der Bewegungsabfolgen erforderlich ist. Der erfolgreichen Umsetzungen gehen eine Vielzahl von Fehlschlägen voraus, bei denen einzelne Systemkomponenten teilweise stark beschädigt werden und es zu erheblichen Leckagen des Hydrauliköls kommt. Jede fehlerhafte Erprobung geht daher mit aufwendigen Reparaturarbeiten einher, für die ein großes Team von spezialisierten Ingenieuren unerlässlich ist. Das zweite Beispiel sind die wiederverwendbaren Raketen des Unternehmens *SpaceX*. Die Zielsetzung ist es, die gleiche Rakete sowohl für den Start als auch die Landung zu nutzen, um die Kosten gegenüber von herkömmlichen Raumfahrtmissionen deutlich zu senken. Im Jahr 2006 wurden die ersten Erprobungen der anspruchsvollen Start- und Landungs-Trajektorien durchgeführt, welche sich bis ins Jahr 2015 erstreckten. Im Dezember 2015 konnte die erfolgreiche Landung zum ersten Mal realisiert werden. Die vorhergehenden Misserfolge sind jedoch mit z. T. enormen Explosionen verbunden gewesen, wobei die Kosten für eine Erprobung bis zu mehrere Milliarden US-Dollar betrugen. Die beiden angeführten Beispiele verdeutlichen, dass ein iteratives Vorgehen, wie es auch durch das V-Modell vorgegeben wird, unabdingbar für die Inbetriebnahme ist. Die gesammelten Erfahrungen und Messdaten, welche bei jedem Versuch akkumuliert werden, tragen zum letztendlichen Erfolg der Inbetriebnahme bei. Nichtsdestotrotz zeigen die angeführten Beispiele, dass die Inbetriebnahme ein komplizierter Prozess ist und die

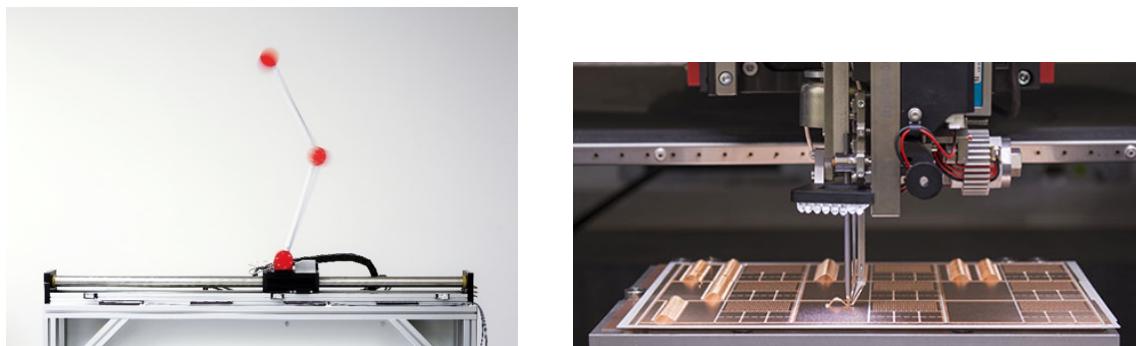


Bild 1-4: Doppelpendel auf einem Wagen (© Heinz Nixdorf Institut) (links), Ultraschall-drahtbondprozess (© Hesse Mechatronics) (rechts).

Anzahl der Versuche so gering wie möglich gehalten werden sollte, um Zeit und Kosten bzgl. Verschleiß, Reparaturen und Arbeitseinsatz einzusparen.

Zwei weitere Beispiele sollen die Problematiken bei der Inbetriebnahme näher erläutern. In Bild 1-4 sind hierzu das Doppelpendel auf einem Wagen und der Ultraschalldrahtbondprozess abgebildet, welche in dieser Arbeit als Anwendungsbeispiele herangezogen werden. Das Doppelpendel auf einem Wagen ist ein beliebtes System, an welchem sich die Regelungstechnischen Zusammenhänge in der Lehre sehr gut erklären lassen und welches zudem besonders geeignet für Forschungszwecke sowie die Erprobung neuartiger Regelungskonzepte ist. An einem Linearmotor (Wagen), welcher sich lediglich horizontal verfahren lässt, sind zwei rotatorische Pendelarme befestigt. Das Regelungsziel ist es, die beiden anfangs nach unten hängenden Pendelarme durch eine passende Ansteuerung des Linearmotors aufzuschwingen und zu stabilisieren. Bei diesem Manöver bewegt sich das System aus seiner unteren stabilen in seine obere instabile Ruhelage. Eine Besonderheit des Doppelpendels ist sein chaotisches Bewegungsverhalten, welches sich in einer hohen Sensitivität bzgl. verschiedener Einflussfaktoren widerspiegelt, und damit nur schwer vorhersagbar ist.

Wird der modellbasierte Regelungsentwurf für das Doppelpendel auf einem Wagen durchgeführt, so erhält man mit verhältnismäßig wenig Aufwand über die Anwendung des *Lagrange-Formalismus* ein einfaches mechanisches Simulationsmodell des Systems. Mit Hilfe des Modells lässt sich anschließend ein Optimalsteuerungsentwurf für das Aufschwungmanöver durchführen. Alleine mit der berechneten Steuerung lässt sich am realen System bzw. dem Prüfstand der Aufschwung jedoch nicht realisieren. Ausschlaggebend hierfür sind die äußeren Störeinflüsse und insbesondere die vorhandenen Modelfehler, deren Auswirkungen vor dem Hintergrund der chaotischen Dynamik besonders ausgeprägt sind. Ein Lösungsweg zur erfolgreichen Realisierung ist die Hinzunahme einer Regelung, deren Ziel es ist, den Aufschwung entlang der Trajektorie zu stabilisieren, bzw. auf Kurs zu halten. Im Allgemeinen funktioniert dieses Vorgehen robust und reproduzier-

bar. Durch die Wahl der Optimierungsziele ist es möglich, verschiedene Aufschwungmanöver für das System zu berechnen [Tim13]. Für einen Teil dieser Trajektorien, insbesondere solche mit einer langen (Prädiktions-)Dauer, ist der oben beschriebene Ansatz jedoch nicht umsetzbar. Die Modellfehler, welche beispielsweise die Reibungseffekte in den Pendelgelenken oder zwischen dem Linearmotor und dem Verfahrweg betreffen, sind zu signifikant, wodurch eine zeit- und kostenintensive Nachmodellierung notwendig wird. Bei dieser Modellierung wird das bestehende einfache mechanische Modell ausdetailliert und mit weiteren physikalischen Effekten verbessert. Anschließend wird eine Neuberechnung der Steuerung und Regelung vorgenommen und führt idealerweise zum erwünschten Ergebnis. Je nach Komplexität des System ist dieses Vorgehen für den Regelungstechniker allerdings nicht praktikabel und unter Umständen gar nicht durchführbar. Dies liegt in der Tatsache begründet, dass der Aufwand für immer detailliertere Modelle sehr rasch ansteigt und im Besonderen mechatronische Systeme sich durch eine Vielzahl von unterschiedlichsten Komponenten zusammensetzen, für deren jeweilige Modellierung ein eigener Fachexperte notwendig ist. Stellt die Verbesserung der Modellbildung daher keine Option dar, befindet sich der Regelungstechniker mangels Alternativen in einer Situation, in der keine weiteren Entwicklungspfade verfügbar sind. Wünschenswert wäre in diesem Szenario ein regelungstechnisches Werkzeug in Form eines Verfahrens, das trotz des unausgereiften Modells eine Strategie vorgibt, um die entwickelte Steuerung oder Regelung dennoch in Betrieb nehmen zu können.

Ein weiteres Anwendungsbeispiel dieser Arbeit stellt der Ultraschalldrahtbondprozess [Har10], auch Drahtboden genannt, dar (Bild 1-4, rechts). Dies ist eine Verbindungstechnik bei der elektrische Bauteile über dünne Drähte miteinander verbunden werden. Das Verfahren ist u. a. im Bereich der Mikroelektronik angesiedelt und nutzt hochfrequente Schwingungen eines Piezoaktors, um eine Kaltschweißung durchzuführen, bei welcher die Schmelztemperatur der Verbindungspartner nicht überschritten wird. Ein aktuelles Anwendungsgebiet ist die Verbindung von mehreren Batteriepackungen für Elektrofahrzeuge. Die Stellgrößen des Systems sind die elektrische Spannung, mit welcher der Piezoaktor angesteuert und in Schwingung versetzt wird, und eine vertikal wirkende Kraft, die das sogenannte Bondwerkzeug auf die Verbindungsfläche drückt. Das Steuerungsziel besteht darin, die passenden Verläufe der Stellgrößen zu finden, um eine möglichst zuverlässige Verbindung, die robust gegenüber äußeren Einwirkungen ist, zu erhalten. Eine wesentliche Herausforderung dabei ist, dass das richtige Maß an eingebrachter Energie gefunden werden muss. Ein zu niedriger Eintrag führt dazu, dass die Festigkeit der Verbindung zu gering ausfällt, wohingegen ein zu hoher Eintrag zwar die Festigkeit stark erhöht, aber gleichzeitig auch die Gefahr von sogenannten Werkzeugaufsetzern erhöht. Werkzeugaufsetzer gilt es unbedingt zu verhindern, denn sie beschädigen nicht nur die

Werkzeugspitze, sondern auch den empfindlichen Untergrund, womit das Bauteil insgesamt unbrauchbar wird.

Aus Regelungstechnischer Sicht wäre es ideal, einen beschränkten Optimalsteuerungsentwurf, bei dem die technischen Rahmenbedingungen Berücksichtigung finden, für das System umzusetzen. Ebenso wie im Falle des Doppelpendels ist hierfür jedoch ein genaues Modell notwendig. Zwar lassen sich unterschiedlich komplexe Modelle für das System aufstellen, allerdings sind diese für den Entwurf ungeeignet. Einfache physikalische Modelle erklären die Zusammenhänge des Systems zwar nachvollziehbar, lassen sich allerdings schlecht mit realen Messungen validieren und weisen daher eine sehr geringe Prädiktionsfähigkeit für verschiedene Ansteuerungen auf. Detailliertere Modelle können über die *Finite Elemente Methode* (FEM) aufgestellt werden, besitzen allerdings den Nachteil, dass sie eine lange Berechnungsdauer aufgrund ihrer hohen Komplexität benötigen. Erschwerend kommt hinzu, dass eine gewisse Unklarheit darüber herrscht, wie die Zustände des Systems genau zu definieren sind, da der Bondprozess auf mikroskopischer Ebene abläuft und sich daher mit der klassischen Mechanik nur schwer beschreiben lässt. Unter der theoretischen Voraussetzung, dass die Zustände klar definiert sind, ist ein weiteres Problem, dass es keine geeignete Sensorik gibt, die einen Zugang zu diesen Zuständen ermöglichen würde. Die Identifikation der Parameter des zugehörigen Modells und der Einsatz einer stabilisierenden (vom Zustand abhängigen) Regelung (wie sie beim Doppelpendel auf einem Wagen genutzt wird) sind damit nicht umsetzbar. In Bezug auf die Inbetriebnahme einer Steuerung für den Ultraschalldrahtbondprozess steht der Regelungstechniker damit vor der Problematik diese ohne ein Modell, und damit ohne einem großen Teil des Regelungstechnischen Methodenspektrums, durchzuführen. Als Folge dessen muss die Steuerung direkt am realen System entworfen werden, was ohne das Vorhandensein von geeigneten Verfahren zu langen Entwicklungszeiten und suboptimalen Ergebnissen führt.

Als Fazit lässt sich feststellen, dass die Inbetriebnahme von Steuerungen und Regelungen einen entscheidenden Entwicklungsschritt darstellt und sowohl die theoretischen Vorüberlegungen als auch die Realität der praktischen Umsetzung miteinander verbinden muss. Vor diesem Hintergrund verdeutlichen die vorgestellten Beispiele, dass die Inbetriebnahme je nach betrachtetem System mitunter aufwändig sein kann und daher so effizient wie möglich gestaltet werden sollte. Kommt es beim modellbasierten Entwurf zu einer Diskrepanz zwischen erwartetem und gemessenem dynamischen Verhalten (vgl. Bild 1-2), so fehlen in vielen Fällen die Regelungstechnischen Werkzeuge, bzw. Verfahren und Methodiken, um eine erfolgreiche Realisierung zu erreichen. In der vorliegenden Arbeit wird dieser Sachverhalt daher als vorhandene Problematik bei der Inbetriebnahme

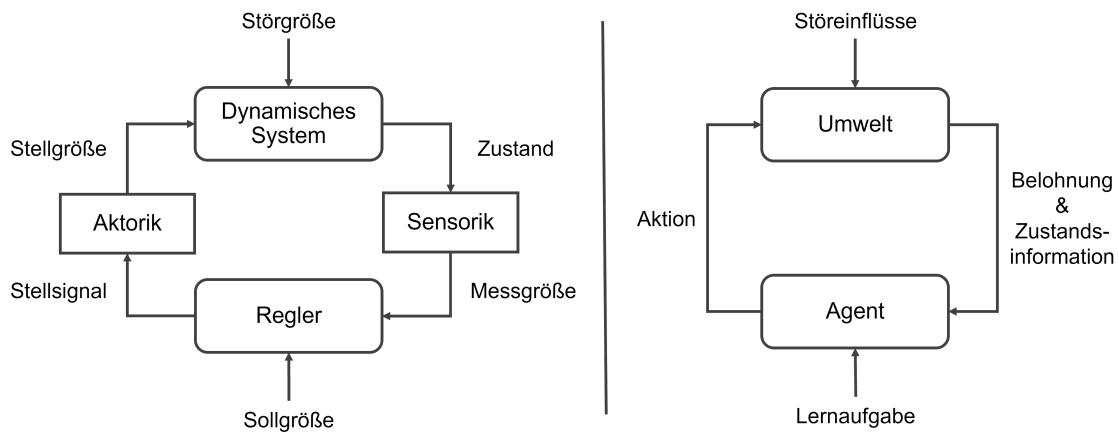


Bild 1-5: Schematischer Aufbau in der Regelungstechnik (links) und beim Reinforcement Learning (rechts).

diagnostiziert und im nachfolgenden Abschnitt hieraus die Forschungslücke abgeleitet, sowie Lösungsansätze und Zielsetzungen formuliert.

1.2 Identifikation der Forschungslücke und Zielsetzung

Eine Möglichkeit, die Probleme bei der Inbetriebnahme zu adressieren, bietet das RL, welches dem Bereich des maschinellen Lernens zugeordnet wird und sich mit intelligenten Algorithmen befasst, die eigenständig Entscheidungen treffen und zielorientiert ausführen. Die grundlegende Vorgehensweise, welche jeder RL-Algorithmus nutzt, ist die des *Trial-and-Error* Prinzips. Bei dieser heuristischen Methode werden verschiedene angestrebte Lösungsmöglichkeiten nach und nach ausprobiert, bis das übergeordnete Ziel erreicht wird. Ein wesentlicher Baustein ist dabei die Rückführung von Informationen in Form eines *Feedbacks*, wodurch eine Bewertung der ausgeführten Entscheidungen in Bezug auf die Zielsetzung erfolgen kann.

An dieser Stelle wird erneut Bezug auf den Entwurf am Doppelpendel als anschauliches Beispiel genommen. Wie bereits erwähnt, steht mit dem RL eine Alternative zum herkömmlichen modellbasierten Entwurf zur Verfügung. In [HTHT18] wurde beispielsweise der Aufschwung durch einen lernfähigen Algorithmus namens *Probabilistic Inference for Learning COnrol* (kurz PILCO) realisiert, wobei kein besonderes domänen-spezifisches Expertenwissen genutzt wurde. Durch die Interaktion mit dem realen System sammelt der Algorithmus Messdaten, mit denen ein rein datengetriebenes Dynamikmodell gelernt wird. Das Dynamikmodell verwendet das Verfahren wiederum, um eine parameterabhängige Regelung innerhalb eines Optimierungsproblems zu berechnen. Nach jedem neuen Versuch am Prüfstand liegen neue Messdaten vor, die den bestehenden Messdaten hinzugefügt und womit das Dynamikmodell, sowie die zugehörige Regelung, aktualisiert werden. Nach einer bestimmten Anzahl von Erprobungen kann die Regelungs-

aufgabe nach diesem Vorgehen erfolgreich gelöst und die Inbetriebnahme als erfolgreich abgeschlossen werden. Als Fazit wird festgehalten, dass das RL ein Potential bietet, um die Inbetriebnahme zu verbessern, indem aufgenommene Messdaten stärker in den Fokus gerückt und aus ihnen zusätzliche Informationen zum Erfüllen des Regelungsziels gezogen werden.

Nachfolgend wird ein Vergleich zwischen den beiden Herangehensweisen des regelungstechnischen Entwurfs und des RLs aufgestellt und die inhärente Forschungslücke definiert. Bild 1-5 zeigt hierfür den schematischen Aufbau beider Disziplinen. Es fällt auf, dass sich der Aufbau stark ähnelt und im Zentrum die Beeinflussung des dynamischen Systems bzw. der Umwelt durch den Regler bzw. den Agenten steht. Vor dem Hintergrund der gleichen Aufgabe und unter der Voraussetzung eines abgeschlossenen Lernvorgangs stellen die jeweiligen Begriffe damit lediglich Synonyme dar. Das Grundprinzip der Rückkopplung von Informationen ist zudem beiden Disziplinen immanent. Die Regelungstechnik ist grundsätzlich technisch und praxisorientiert geprägt, sodass die Begriffe Aktorik und Sensorik explizit als Bausteine auftauchen; wohingegen sie beim RL entfallen. Das RL ist insgesamt allgemeiner gefasst und kann sich unter anderem auch auf Spiele, wie bspw. Schach oder Go, beziehen. Das liegt vor allem darin begründet, dass das RL in der Informatik und Psychologie angesiedelt ist und der Lernvorgang des Agenten im Vordergrund der Entwicklung und Untersuchung steht. Diese beiden Prägungen werden auch durch die beiden Eingangsbeispiele aus der maßgeblichen Standardliteratur [FKL⁺22] und [SB18] deutlich. Während in [FKL⁺22] die regelungstechnischen Zusammenhänge an einem Servomotor erläutert werden, wird in [SB18] zunächst das Spiel *Tic-Tac-Toe* herangezogen. Weiterführend ist ein wesentlicher Untersuchungsgegenstand der Regelungstechnik die Gewährleistung der Sicherheit eines technischen Systems durch Stabilitätsuntersuchungen und interpretierbare Schlussfolgerungen. Das Themengebiet der Wahrscheinlichkeitstheorie spielt (mit Ausnahme des Beobachterentwurfs) nur selten eine Rolle. Im Gegensatz dazu, baut das RL sehr wesentlich auf der Wahrscheinlichkeitstheorie auf, da der Lernprozess zusammenfassend eine kontinuierliche Entscheidungsfindung unter Unsicherheit der Umwelt darstellt. Hinzu tritt das *Exploitation-Exploration-Dilemma* [SB18], welches die Findung eines Kompromisses zwischen vermuteten bestmöglichen und ausprobierenden Entscheidungen beschreibt. In vielen Fällen wird das Dilemma über wahrscheinlichkeitstheoretische Ansätze adressiert mit denen u. a. Beweise zum Konvergenzverhalten des lernenden Agenten geführt werden.

Im konkreten Bezug auf die Inbetriebnahme von Steuerungen und Regelungen gibt es aus der Regelungstechnik heraus kaum Handlungsanweisungen für die Behebung von auftretenden Problemen. Ein mutmaßlicher Grund hierfür könnte sein, dass die auftretenden Probleme sehr vielseitig sind und Lösungen immer auf das individuelle System abge-

stimmt sein müssen. Daher erscheint es schwierig allgemeingültige Handlungsanweisungen auszusprechen. Ein weiterer Grund ist die bereits praxisorientierte Theorie der Regelungstechnik. Unvorhergesehene Störungen sind die Hauptmotivation für den Einsatz einer Regelungen, welche immer in der Lage sein muss, vorhandene Abweichungen und Ungenauigkeiten zu kompensieren. Das behandelte Szenario dieser Arbeit geht jedoch davon aus, dass es bei der Inbetriebnahme zu tieferliegenden Problemen kommen kann, die im Rahmen einer Regelung nicht durch diese ausgeglichen werden können. Ein Indikator für mangelnde Verfahren zur effizienten Inbetriebnahme findet sich zudem in dem regelungstechnischen Standardwerk [Ada14], in welchem eine Vielzahl von nichtlinearen Regelungsverfahren ausführlich besprochen werden. Zwar findet eine theoretische Aufarbeitung statt, allerdings wird im Bezug auf die reale Umsetzung bzw. die Inbetriebnahme in den meisten Fällen auf das rudimentäre "Ausprobieren" verwiesen. Aus Perspektive der Regelungstechnik zeigt sich somit eine vorhandene Forschungslücke, welche die Probleme des modellbasierten Entwurfs bei der Inbetriebnahme umfasst und mit geeigneten regelungstechnischen Werkzeugen in Form von intelligenten Verfahren adäquat adressiert. Die Grundideen und Methoden des RL bieten hierbei großes Potential, um solche Verfahren zu entwickeln, die schlussendlich eine Kombination beider Disziplinen darstellen und in dieser Arbeit, aufgrund ihrer teils physikalisch geprägten und teils daten-getriebenen Ansätze, als *hybrid* bezeichnet werden.

Als Gegenstück kann die bereits angeführte Forschungslücke auch aus der Perspektive des RL definiert, bzw. betrachtet werden. Das RL gilt im Allgemeinen als daten-ineffizient, wenn es um die Anwendung auf reale Systeme geht. Die Ineffizienz ist besonders problematisch, da sie sich in einem enormen Zeitaufwand für die Generierung der Daten bemerkbar macht und die meisten RL Verfahren damit unpraktisch, bzw. unmöglich umsetzbar sind. Der angesprochene PILCO Algorithmus stellt eine Ausnahme dar und wird dem sogenannten modellbasierten RL zugeordnet. Das heißt, dass die Strategie des Agenten anhand eines gelernten Modells der Umwelt optimiert wird. Die Strategie wird also nur indirekt für die Umwelt entworfen, indem in einer ähnlichen Simulationsumgebung ein Großteil der Interaktion stattfindet. Wie beim RL üblich, wird das Modell alleine auf der Basis von Messdaten vom realen System gelernt, wobei die Herausforderung darin besteht, aus einer ggf. geringen Datenmenge das dynamische Verhalten zu verallgemeinern. Ein spezielles Vorwissen (a-priori der Datenaufnahme) über die Dynamik wird in der Regel nicht berücksichtigt. Hierfür gibt es zwei Gründe. Zum einen wird von einem RL Algorithmus eingefordert, dass er ohne menschliche Hilfestellung funktionieren kann, um ggf. als *künstliche Intelligenz* (KI) bezeichnet werden zu können. Der zentrale Forschungsgegenstand des RL ist der Lernvorgang des Agenten, welcher aus Sichtweise der Informatik und Psychologie möglichst unvoreingenommen sein sollte. Zum anderen ist das nützliche Vorwissen unter Umständen nicht vorhanden. Die Beschreibung

dynamischer, bzw. physikalischer Effekte spielt in den genannten Fachdisziplinen eine untergeordnete Rolle, sodass auf dieses Fachwissen nicht zurückgegriffen werden kann. Vor dem Hintergrund der Inbetriebnahme deuten diese Punkte ebenfalls die bearbeitete Forschungslücke an, wobei das RL zwar ein mächtiges Werkzeug für eine flexible Inbetriebnahme darstellt, jedoch die zielgerichtete Einbringung von physikalischem, bzw. Regelungstechnischem Vorwissen bislang wenig bis keine Berücksichtigung findet.

Zusammenfassend fehlen in der Regelungstechnik geeignete Ansätze und Verfahren, um den Prozess der Inbetriebnahme zu unterstützen, bzw. bei auftretenden Problemen allgemeingültige Lösungsmöglichkeiten aufzuzeigen. Ein Potential, diese Lücke von fehlenden Regelungstechnischen Werkzeugen zu schließen, bieten die Ideen und Verfahren des primär datenbasierten RLs. Im Zusammenhang mit mechatronischen Systemen und deren Sensorik stehen die Daten unkompliziert zur Verfügung und enthalten bereits inhärent die notwendigen Ein-Ausgangsdatenpaare, wenn es um das Erlernen des dynamischen Systemverhaltens geht. Eine Erweiterung der bestehenden Regelungstechnischen Verfahren mit Hilfe von RL ist daher besonders sinnvoll und bietet die erforderliche Flexibilität, um eine breite Klasse von (nichtlinearen) Systemen zu adressieren. Diese übergeordnete Zielsetzung kann alternativ auch aus der Perspektive des RL betrachtet werden, indem die Verfahren durch Regelungstechnisches Fachwissen über die zu lösende Steuerungs- oder Regelungsaufgabe erweitert werden, um in Folge dessen die Effizienz von RL Methoden bei der Anwendung auf reale Systeme deutlich zu steigern. Bereits in [SB98] wird hierzu die enge Verwandtschaft beider Fachgebiete thematisiert und ein Ausblick auf hybride Entwicklungsmethoden geschaffen:

„One of the larger trends of which reinforcement learning is a part is that toward greater contact between artificial intelligence and other engineering disciplines. Not all that long ago, artificial intelligence was viewed as almost entirely separate from control theory and statistics. ... Over the last decades this view has gradually eroded. Modern artificial intelligence researchers accept statistical and control algorithms, for example, as relevant competing methods or simply as tools of their trade. The previously ignored areas lying between artificial intelligence and conventional engineering are now among the most active, including new fields such as neural networks, intelligent control, and our topic, reinforcement learning.“ [SB98, S. 5]

Die Zielsetzung der vorliegenden Arbeit ist damit, hybride Verfahren zur Inbetriebnahme von Steuerungen und Regelungen zu entwickeln, die sowohl die Regelungstechnik, als auch das RL miteinander in Verbindung setzen. Auf diese Weise lassen sich die Vorteile beider Disziplinen ausnutzen und sich ihre Nachteile, in gewissen Grenzen, gegenseitig kompensieren. Konkret bedeutet dies, die Mängel des modellbasierten Entwurfs für

bspw. schwer zu modellierende Systeme durch den flexiblen datengetriebenen Grundgedanken des RL zu beseitigen. Auf der anderen Seite sind RL Ergebnisse für den Menschen nur sehr schwer interpretierbar und damit vor dem Hintergrund von sicherheitskritischen technischen Systemen äußerst risikobehaftet. Die Einbringung von Regelungstechnischem Wissen, welches durch nachvollziehbare Stabilitätsuntersuchungen geprägt ist und damit für den Menschen deutlich einfacher verstanden werden kann, bietet die Möglichkeit RL Verfahren nicht nur transparenter zu machen, sondern auch die Effizienz in Bezug auf die notwendige Datenlage wesentlich zu steigern. Die Begründung hierfür ist, dass bspw. das dynamische Verhalten des Systems nicht von Grund auf erlernt, sondern nur partiell korrigiert werden muss. Des Weiteren kann davon ausgegangen werden, dass das a-priori physikalische Modell deutlich bessere Extrapolationseigenschaften aufweist, als eine untrainierte rein datengetriebene Variante. Um der Zielsetzung im geeigneten Maße zu begegnen, werden im Folgenden verschiedene Aspekte aufgelistet und bestimmte Anforderungen für die entwickelten Verfahren erläutert.

Einheitliche mathematische und begriffliche Formulierung

Wie bereits angemerkt, weisen die Regelungstechnik und das RL einige Unterschiede und Gemeinsamkeiten auf. Jede Disziplin wird in Verbindung mit ihrer ganz eigenen mathematischen Formulierung bezüglich der Deklaration von Variablen und Funktionen gebracht. So wird in der Regelungstechnik bspw. der Buchstabe x für den Zustandsvektor eines dynamischen Systems benutzt, wohingegen im RL der gleiche Buchstabe für den Eingang in eine Regressionsfunktion, bzw. die zugehörigen Trainingsdaten, fungiert. Ein Ziel dieser Arbeit ist es daher, einen einheitlichen Rahmen für hybride Methoden zu schaffen, der eine Konsistenz bezüglich der Benennung von mathematischen Ausdrücken aufweist. Außerdem gilt es, bestimmte Zusammenhänge im jeweils anderen Kontext zu betrachten. Ein Beispiel ist hierfür die Parameteridentifikation für ein dynamisches Modell in der Regelungstechnik, welche im engen Zusammenhang mit dem Training von Gewichtungen eines künstlichen neuronalen Netzes steht. In beiden Fällen wird zur Bestimmung ein Mehrzieloptimierungsproblem gelöst, wobei ein Kompromiss zwischen der genauen Nachbildung der Daten und der Verallgemeinerungsfähigkeit bzgl. noch unbeobachteter Zustände/Daten gefunden wird. Dieses Problem ist im *Machine Learning* (ML) unter den Begriffen *Over-* und *Underfitting* [Bis06] bekannt und spielt somit auch in der Regelungstechnik eine wesentliche Rolle, obwohl hier für das zugrundeliegende Problem keine speziellen Begriffe definiert sind.

Berücksichtigung von Vorwissen über die Steuerungs-/Regelungsaufgabe

Im Allgemeinen werden beim maschinellen Lernen keine Angaben über möglicherweise vorhandenes Vorwissen getätigt; auch wenn dieses unter Umständen zur Verfügung stehen würde. Der Grund hierfür liegt in der Tatsache begründet, dass dies gegen den Grundge-

danken einer KI gerichtet wäre. Denn autonomes Lernen setzt ein Entscheiden auf der Grundlage von gesammelten Daten und eigenen Erfahrungen voraus. Auch wenn dies im Rahmen des ML eine valide Herangehensweise darstellt, um den Lernprozess als solchen zu untersuchen und möglichst allgemeingültige Algorithmen zu entwickeln, so ist die Nichtberücksichtigung von Vorwissen im Zusammenhang mit einer Regelungstechnischen Aufgabe kontraproduktiv. Jegliches Vorwissen bietet nicht nur das Potenzial zumindest einen Teil der Lösung für den Menschen interpretierbar, im Sinne eines *White- und Black-Box*-Modells, zu machen, sondern kann auch den Lernvorgang als solchen beschleunigen und effizienter gestalten. Dieser Punkt steht in enger Verbindung mit dem Ziel einer effizienten Ausnutzung der Daten und wird im nächsten Abschnitt näher erläutert. Darüber hinaus besitzt ein hybrides Modell, welches sich aus dem ursprünglich aufgestellten physikalischen Modell und einem erweiternden flexiblen datengetriebenen Teil zusammensetzt, vorteilhafte Eigenschaften in Hinblick auf die Prädiktionsgenauigkeit. Datengetriebene Modelle alleine betrachtet, sind in der Regel in der Nähe der Trainingsdaten zuverlässig und können diese gut nachbilden, weisen allerdings Defizite auf, wenn sie in unbekannten Gebieten Vorhersagen treffen sollen. Ihre Extrapolations-, sind damit im Gegensatz zu ihren Interpolationseigenschaften, als schwach einzustufen. Ein hybrides Modell kann hierbei bessere Ergebnisse erzielen, da dem physikalischen Modellteil vergleichsweise eine gute Extrapolation zuzuschreiben ist und gleichzeitig die guten Interpolationseigenschaften des datengetriebenen Teils erhalten bleiben. Neben der Modellbetrachtung, bietet die Regelungstechnik eine Vielzahl von möglichen Regelungsstrukturen, die auf bestimmte Systemklassen zugeschnitten sind und damit besondere Eigenschaften ausnutzen. Die klare Struktur grenzt dabei verschiedene Aufgaben einer Regelung voneinander ab und macht es dem Menschen damit einfacher sie nachzuvollziehen und auftretende Probleme leichter beheben zu können. Dieses Wissen kann in der richtigen Kombination mit einer RL Methode die Regelgüte und dessen Interpretierbarkeit signifikant verbessern, womit die Berücksichtigung von Vorwissen ein wesentliches Ziel dieser Arbeit darstellt.

Dateneffizienz

Im Kontext eines mechatronischen Systems werden Daten durch die Messeinrichtungen bereitgestellt. Diese Daten können ohne größeren Aufwand direkt während des regulären Betriebs aufgenommen und für den Lernalgorithmus verwendet werden. In diesem Szenario spielt die Dateneffizienz eine untergeordnete Rolle. Soll jedoch eine neu entwickelte Regelung in Betrieb genommen werden und sind dafür Erprobungen und Experimente notwendig, dann ist aus Zeit- und Kostengründen eine hohe Dateneffizienz erforderlich. Je weniger Experimente und damit weniger Messdaten für das Erfüllen der Regelungsaufgabe benötigt werden, desto schneller kann die Inbetriebnahme erfolgreich abgeschlossen werden. Jedes zusätzliche weitere Experiment, welches als Absicht hat, die Datenmenge zu erhöhen, ist nicht nur mit einem Zeitaufwand für den Regelungstechniker verbunden,

sondern verhindert auch, dass das System in einem wirtschaftlichen Sinne genutzt werden kann. Hiermit sind insbesondere Produktionsanlagen gemeint. Zudem nimmt die Beanspruchung der Systemkomponenten und der Verschleiß der Bauteile über die Zeit zu, was einen weiteren Kostenfaktor darstellt. Die in dieser Arbeit entwickelten Verfahren haben daher eine hohe Dateneffizienz während der Inbetriebnahme als Ziel, so dass die Anzahl der benötigten Experimente und Daten minimiert und der Informationsgewinn aus ihnen maximiert wird.

Flexibilität und Verständlichkeit bei der Anwendung

Mechatronische Systeme weisen ein ähnliches Grundprinzip auf, unterscheiden sich z. B. jedoch klar in ihren Eigenschaften. Das passende Regelungskonzept muss daher für jedes System sorgfältig abgewägt und realisierbar umgesetzt werden. Da ML Ansätze im Allgemeinen als flexibel eingestuft und sie oftmals nur in Abhängigkeit von Daten formuliert werden, soll dieser Ansatz gleichermaßen auf die hybriden Methoden dieser Arbeit übertragen werden. Die Verfahren sollen dabei auf eine breite Klasse von Systemen anwendbar sein und möglichst wenig Voraussetzungen haben. Auf diese Weise bieten die Verfahren ein hohes Maß an Praxisrelevanz, und stellen Werkzeuge dar, welche für die Inbetriebnahme vieler Regeleinrichtungen in Erwägung gezogen werden können. Vor diesem Hintergrund spielt auch die Verständlichkeit der Verfahren eine Rolle. Der Anwender muss nachvollziehen können, wie ein Lernalgorithmus arbeitet, welchen Prinzipien er folgt und wie seine Wirkungszusammenhänge sind. Dies trägt dazu bei, dass der Anwender die Entscheidungen des Algorithmus bzgl. der verschiedenen Experimente versteht und ihren Zweck richtig einordnen kann. Im Idealfall kann der Anwender interaktiv Einfluss auf die ansonsten gewissermaßen automatisierte Inbetriebnahme nehmen, und so weiteres Expertenwissen über das System und die Regelungsaufgabe einbringen. Die Voraussetzung hierfür ist, dass der Lernalgorithmus, sein genereller Aufbau und seine Vorgehensweise bei der Inbetriebnahme leicht nachvollzogen werden kann.

Robustheit und Konvergenzverhalten

Zusätzlich zu den bereits angeführten Anforderungen sollen sich die Algorithmen durch eine hohe Robustheit und ein schnelles stabiles Konvergenzverhalten auszeichnen. Hierbei sollen Störeinflüsse eine möglichst geringe (und im Idealfall keine) Auswirkung auf den Lernprozess an sich und das Lernergebnis haben. Als Beispiel sei dabei das Messrauschen genannt, wodurch die aufgenommenen Daten mit unterschiedlicher Ausprägung verfälscht werden. Der Lernalgorithmus muss sich dem Messrauschen gegenüber robust verhalten, kann das Rauschniveau identifizieren und nach Möglichkeit aus den Daten entfernen, bzw. herausrechnen. Neben dem Messrauschen ist auch eine Robustheit gegenüber verschiedenen Rahmenbedingungen gefordert. Die Aufnahme der Daten und die Durchführung der Inbetriebnahme sind als stochastische Prozesse zu betrachten, die nur

schwerlich genau zu reproduzieren sind. Diese Schwankungen bzgl. leicht abweichender Anfangsbedingungen oder einer etwas anderen Position der Daten im Zustandsraum sollen den Lernvorgang nur geringfügig beeinflussen. Infolgedessen soll das Regelungsziel zuverlässig erfüllt werden und sich in einem stabilen zügigen Konvergenzverhalten widerspiegeln. Dies kann wiederum durch die bereits angesprochene Dateneffizienz und das bereitgestellte Vorwissen erreicht werden. Dennoch impliziert das Vorwissen bereits die Tatsache, dass es sich nicht um vollständiges Wissen handelt, womit die Regelungsaufgabe direkt gelöst werden könnte. Es beinhaltet teilweise fehlerhafte oder gar nicht vorhandene Terme im Bezug auf das mathematische Dynamikmodell. Die hybriden Lernverfahren müssen dementsprechend ebenso eine Robustheit gegenüber eben diesen potentiell möglichen Modellen und Formen von Vorwissen haben. Die letztendliche Lösung, welche während der Inbetriebnahme entwickelt wird, soll eindeutig und stabil gefunden werden und soll nicht vom verwendeten Vorwissen abhängen.

Gewährleistung der Sicherheit

Der Ingenieur muss sich bereits vor der Inbetriebnahme intensiv Gedanken über die Sicherheitsaspekte machen, um potenzielle menschliche und maschinelle Schäden zu vermeiden. Bei der Inbetriebnahme spielt die Unsicherheit über das Systemverhalten eine große Rolle, da sie andernfalls nicht notwendig wäre aufgrund eines idealen Modells. Um die Sicherheit zu gewährleisten, müssen bestimmte Zustands- und Stellgrößenbeschränkungen eingehalten werden. Diese können durch manuell implementierte Abfragen von Sensorsignalen überwacht werden. Diese Sicherheitsvorkehrungen stehen über dem automatisierten Lernvorgang und können von ihm nicht beeinflusst werden. Im Falle einer Verletzung der Bedingungen, wird der Lernvorgang oder das Experiment abgebrochen und das System in einen sicheren Zustand (bspw. Stillstand) überführt. Die währenddessen aufgezeichneten Messdaten können jedoch im Rahmen des Lernalgorithmus weiter verarbeitet und für zukünftige Experimente berücksichtigt werden. In der Regel können Beschränkungen auch bereits als Vorwissen in den Lernprozess integriert werden. Es gibt jedoch Anwendungen, in denen die Grenzen erst durch Testläufe ermittelt werden müssen, sodass eine Berücksichtigung im Vorfeld nicht möglich ist. Probabilistische Modelle erweisen sich hier als besonders hilfreich, da sie mit wahrscheinlichkeitsbasierten Angaben arbeiten und damit die Unsicherheit zumindest quantifizieren können. Mit dieser Information kann der Einrichter besser abschätzen, wie erfolgversprechend ein Experiment ist, ob eine Verletzung der Sicherheitsbeschränkungen zu erwarten ist und wann diese während des Experiments eintritt. Außerdem sind im Zusammenhang mit dem RL probabilistische Modelle nicht nur erforderlich, um das erwähnte Exploitation-Exploration-Dilemma zu bewältigen, sondern auch, um die Dateneffizienz und Robustheit zu erhöhen. Vor diesem Hintergrund ist allerdings wichtig zu beachten, dass die Angaben zur Sicherheit immer nur so zuverlässig sind, wie die Vorhersagequalität des (hybriden)

Modells. Obwohl probabilistische Modelle darauf abzielen, Modellfehler angemessen zu berücksichtigen, unterliegen auch sie bestimmten Annahmen. Diese Annahmen können die Vorhersagequalität negativ beeinflussen, dienen jedoch gleichzeitig der praktischen Anwendbarkeit, beispielsweise in Bezug auf die Berechnungskomplexität. Daher ist es entscheidend, dass der Entwurfsingenieur die Sicherheitsangaben stets im Zusammenhang mit dem Modell und den Sicherheitsvorkehrungen kritisch bewertet.

Die formulierten Anforderungen werden bei der Erarbeitung der Lernverfahren berücksichtigt. Die Vision sind daten-effiziente flexible, sowie robuste und sicherere Algorithmen zur Unterstützung bei der Inbetriebnahme eines Steuerungs- und/oder Regelungskonzepts. Im nachfolgenden Abschnitt werden die entwickelten Verfahren vorgestellt und den Kapiteln dieser Arbeit zugeordnet.

1.3 Aufbau der Arbeit

Zum besseren Verständnis der Hauptkapitel werden in Kapitel 2 zunächst die mathematischen Grundlagen vorgestellt. Hierbei wird eine ausführliche Einführung in die regelungstechnischen Zusammenhänge (2.1), die Begrifflichkeiten und Regeln der Wahrscheinlichkeitstheorie (2.2) und das maschinelle Lernen durch die Gauß-Prozess-Regression (2.4) gegeben. Neben den formalen Definitionen, wird das Grundlagenkapitel durch einfache Anwendungsbeispiele und Überlegungen zur praktischen Implementierung abgerundet. Dadurch soll eine Verbindung der unterschiedlichen Fachdisziplinen untereinander entstehen, welche in den nachfolgenden Hauptkapiteln vertiefend ausdetailliert wird.

In Kapitel 3 wird das Verfahren der Bayesschen Optimierung vorgestellt, welches im Kontext der Regelungstechnik zum Steuerungsentwurf eingesetzt wird. Hierbei wird eine parametrisierte Steuerung vorausgesetzt, welche durch die Bayessche Optimierung effizient und direkt am realen System entworfen wird. Dieser Ansatz ist besonders geeignet, wenn der herkömmliche modellbasierte Steuerungsentwurf aufgrund einer zu schwierigen oder praktisch unmöglichen physikalischen Modellbildung an seine Grenzen stößt. Die Bayessche Optimierung wird vor diesem Hintergrund ausführlich eingeführt und im weiteren Verlauf des Kapitels ein breiter Überblick zum Stand der Forschung gegeben (3.2). Hierbei wird insbesondere ein Fokus auf bestehende Veröffentlichungen gelegt, welche die Bayessche Optimierung innerhalb eines regelungstechnischen Kontextes verwenden. Hierauf aufbauend wird eine angepasste Variante der Bayesschen Optimierung entworfen, welche für den komplizierten Steuerungsentwurf eines Ultraschalldrahtbondprozesses ausgelegt ist. Im Vorhinein findet in Unterkapitel 3.3 eine Einführung zum Ultraschalldrahtbonden statt, welche anschließend durch eine Vorstellung des physikalischen Modellbildungsprozesses (3.4) abgerundet wird. Basierend auf der Modellbildung kann dann der Einsatz der Bayesschen Optimierung weiter motiviert werden. Nach der Ent-

wicklung des Verfahrens in Abschnitt 3.5, findet eine detaillierte Validierung am realen System statt (3.6). Das erste Hauptkapitel wird durch einen Ausblick zu weiterführenden Themengebieten abgeschlossen.

Die Bayesschen Optimierung ist ein geeignetes Werkzeug für komplexe mechatronische Systeme, deren Zustand nicht eindeutig definiert werden kann oder es kein Sensorkonzept gibt, um Zustandsdaten aufnehmen zu können. Die zwei weiteren Hauptkapitel konzentrieren sich auf Szenarien, in denen diese Rahmenbedingungen nicht vorliegen. In Kapitel 4 wird vor diesem Hintergrund die hybride Optimalsteuerung für die Inbetriebnahme eines nur teilweise unbekannten Systems vorgestellt. Dabei wird die Grundidee verfolgt, die vorhandenen Modellfehler durch aufgenommene Messdaten und die Gauß-Prozess-Regression zu kompensieren, um so einen erweiterten modellbasierten Optimalsteuerungsentwurf durchführen zu können. Als Anwendungsbeispiel werden in diesem Kapitel Pendelsysteme betrachtet, welche in Form einer Modellbildung für das Doppelpendel auf einem Wagen in Abschnitt 4.2 auszugsweise behandelt werden. Im Anschluss wird die Auswirkung eines unvollständigen modellbasierten Steuerungsentwurfs gezeigt und ein rein datengetriebenes Verfahren zur Lösung vorgestellt (4.3). Das datengetriebene Verfahren dient als Referenz und ermöglicht in den nachfolgenden Betrachtungen einen Vergleich zum entwickelten hybriden Optimalsteuerungsverfahren. Das Ziel ist es, den positiven Einfluss von vorhandenen Vorwissen in Form des teilweise bekannten physikalischen Modells genauer zu verifizieren. Hierzu wird der Stand der Forschung erläutert und sowohl simulationsgestützte als auch reale Untersuchungen an einem Doppelpendelprüfstand vorgenommen (4.5).

Das letzte Hauptkapitel (5) beschäftigt sich mit der Herleitung eines hybriden Konzepts zur klassischen Zustandslinearisierung, welche ein nichtlineares System in ein lineares überführt. Hierfür muss die Voraussetzung gelten, dass es sich um ein eingangsaffines System mit separierten Nichtlinearitäten handelt. Die Voraussetzung und die weiteren Zusammenhänge, sowie die Problematik eines fehlerbehafteten Modells werden in Unterabschnitt 5.1 erläutert. Hierauf aufbauend wird die hybride Zustandslinearisierung für allgemeine mechanische Systeme entwickelt (5.2). Des Weiteren wird das Konzept anhand eines hydraulischen Hexapoden erläutert und ein simulationsbasierter Funktionsnachweis für ein Ersatzmodell erbracht (5.3). Zum Abschluss der Arbeit wird in Kapitel 6 eine umfängliche Zusammenfassung gegeben. Zudem wird ein Ausblick für weiterführende Forschungsfragen beschrieben.

2 Mathematische Grundlagen

In diesem Kapitel werden die Grundlagen, die für das Verständnis der weiteren Kapitel und die zugehörigen Verfahren notwendig sind, erläutert. Da davon ausgegangen wird, dass der Leser einen regelungstechnischen Hintergrund besitzt, werden zunächst in Abschnitt 2.1 in verkürzter Weise die wesentlichen Begriffe der Regelungstechnik definiert. Damit wird ein einheitlicher Rahmen für die nachfolgenden Abschnitte geschaffen. Der anvisierte Fokus liegt auf den Grundlagen der Wahrscheinlichkeitstheorie und des maschinellen Lernens. In Abschnitt 2.2 werden daher schwerpunktmäßig die Grundbegriffe Zufallsvariable, Wahrscheinlichkeitsverteilung und Erwartungswertoperator, sowie diverse Rechenregeln eingeführt. Zur Verdeutlichung der Zusammenhänge in einem regelungstechnischen Kontext dienen einfache Anwendungsbeispiele in Form von Pendelsystemen. Das Ziel ist den herkömmlichen deterministischen Modellbegriff um probabilistische Überlegungen zu erweitern. Hierbei werden Fragen geklärt, wie eine Unsicherheit bzgl. der Parametrierung durch Wahrscheinlichkeitsverteilungen abgebildet werden kann oder sich eine probabilistische Zustandsprädiktion (auch Zustandspropagation genannt) berechnen lässt. Den Abschluss bildet Unterkapitel 2.4. Hierin wird die für diese Arbeit wesentliche Gauß-Prozesse-Regression behandelt. Diese stellt ein außerordentliches maschinelles Lernverfahren dar, womit die Unsicherheit über eine gesuchte Funktion alleine auf der Basis von Messdaten quantifiziert werden kann. Hierfür wird die Schätzung der gesuchten Funktion als parameterabhängige Zufallsvariable aufgefasst und ihr Verlauf mittels der Bayesschen Regel bestimmt. Im Kontext der Inbetriebnahme lassen sich so Unsicherheitsfaktoren genau abbilden und drauf aufbauend geeignete Maßnahmen für die erfolgreiche Umsetzung vornehmen.

2.1 Regelungstechnik

Ein allgemeines nichtlineares dynamisches System wird durch mehrere Differentialgleichungen erster Ordnung beschrieben mit

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t); \mathbf{p}), \quad (2-1)$$

wobei $t \in \mathbb{R}$ die Zeit, $\mathbf{x}(t) \in \mathbb{R}^{n_x}$ den Systemzustand, $\mathbf{u}(t) \in \mathbb{R}^{n_u}$ den Stelleingang und $\mathbf{p} \in \mathbb{R}^{n_p}$ die Parametrierung der Dynamikfunktion $\mathbf{f} : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \mapsto \mathbb{R}^{n_x}$ darstellen. Eine explizite Abhängigkeit zwischen \mathbf{f} und t wird in dieser Arbeit nicht betrachtet. Der Vollständigkeit halber wird (2-1) in der Regelungstechnik üblicherweise zusammen mit der Ausgangsgleichung $\mathbf{y}(t) = \mathbf{g}(\mathbf{x}(t))$ genannt, welche bspw. die Abbildung des Zustands auf die Messungen oder Regelgrößen beschreibt. Wenn nicht weiter angegeben,

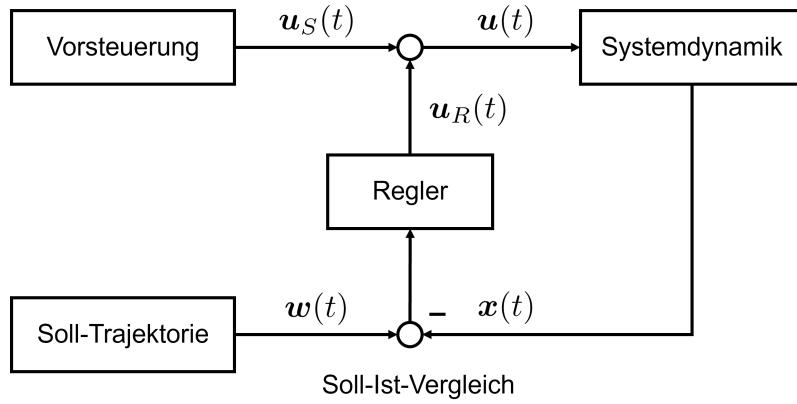


Bild 2-1: Zwei-Freiheitsgrade-Regelungsstruktur

soll in dieser Arbeit davon ausgegangen werden, dass der Zustand vollständig gemessen, bzw. durch einfache Ansätze berechnet werden kann. Aus diesem Grund wird die Ausgangsgleichung vernachlässigt.

Soll das dynamische System (2-1) kontrolliert werden, ist dazu eine geeignete Steuerung bzw. Regelung, welche gleich der Stellgröße $u(t)$ gesetzt wird, notwendig. Ein allgemeiner Ansatz hierfür lautet

$$u(t) = r(t, x(t), w(t); \theta), \quad (2-2)$$

wobei das Regelgesetz $r : \mathbb{R} \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_w} \mapsto \mathbb{R}^{n_u}$ neben der Zeit und dem Zustand auch von der Soll- oder Führungsgröße $w(t) \in \mathbb{R}^{n_w}$ und einer Parametrisierung $\theta \in \mathbb{R}^{n_\theta}$ abhängt. Werden für alle Zustände gewisse Sollgrößen ($n_w = n_x$) vorgegeben, handelt es sich um eine vollständige Zustandsregelung. Entfällt die Abhängigkeit vom zurückgeführten Ist-Zustand $x(t)$ und den Sollwerten $w(t)$, dann handelt es sich bei r um eine reine Steuerung, die keine Anpassungen bzgl. äußerer Störungen vornehmen kann.

Bild 2-1 zeigt beispielhaft den klassischen Aufbau einer Zwei-Freiheitsgrade-Regelungsstruktur [FKL⁺22], welche sich aus einem Steuerungs- ($u_S(t)$) und Regelungs-Zweig ($u_R(t)$) zusammensetzt. Gegeben dieser Struktur, lässt sich (2-2) genauer spezifizieren zu

$$r(t, x(t), w(t); \theta) = u_S(t) + u_R(x(t), w(t); \theta),$$

wodurch die additive Trennung ersichtlich wird. Des Weiteren lässt sich der Anteil der Regelung bspw. mit einen zeitinvarianten linearen Regler der Form

$$u_R(x(t), w(t); \theta) = K(\theta)(w(t) - x(t)),$$

ausdrücken, wobei die Verstärkungsmatrix $\mathbf{K}(\boldsymbol{\theta}) \in \mathbb{R}^{n_u \times n_x}$ in diesem Fall von einer Parametrisierung $\boldsymbol{\theta}$ abhängt, welche die Elemente der Matrix betrifft.

Für die Simulation von (2-1) mit gegebenenfalls (2-2) ist eine zeitliche Diskretisierung erforderlich. Hierzu wird beispielhaft das äquidistante Diskretisierungsgitter

$$\tau := \{t_k = k\Delta t \mid k = 0, \dots, H \in \mathbb{N}, \Delta t \in \mathbb{R}_+\},$$

mit H Zeitschritten und Δt als konstanter Schrittweite definiert, woraus sich die gesamte Simulationsdauer $T \in \mathbb{R}_+$ durch $H\Delta t$ ausdrücken lässt. Anschließend wird entlang τ ein Integrationsverfahren Φ bzgl. (2-1) angewandt, sodass sich für $\mathbf{x}_k = \mathbf{x}(k\Delta t)$, $\mathbf{u}_k = \mathbf{u}(k\Delta t)$ und $\mathbf{w}_k = \mathbf{w}(k\Delta t)$ die diskretisierten Gleichungen

$$\begin{aligned} \mathbf{x}_{k+1} &= \mathbf{x}_k + \Delta t \Phi(\mathbf{f}(\mathbf{x}_k, \mathbf{u}_k; \mathbf{p}), \Delta t), \\ \mathbf{u}_k &= \mathbf{r}(t_k, \mathbf{x}_k, \mathbf{w}_k; \boldsymbol{\theta}), \end{aligned} \quad (2-3)$$

mit z. B. $\Phi(\mathbf{f}(\mathbf{x}_k, \mathbf{u}_k; \mathbf{p}), \Delta t) = \mathbf{f}(\mathbf{x}_k, \mathbf{u}_k; \mathbf{p})$ für das explizite Euler-Verfahren, ergeben.

Aktuiertes Pendel als Einführungsbeispiel

Um die Zusammenhänge von (2-3) tiefergehend zu erläutern und weitere relevante Begriffe aus der Regelungstechnik einzuführen, wird im Folgenden ein im Gelenk aktuiertes Pendel betrachtet. Der Zustand des Systems wird durch den zweidimensionalen Vektor $\mathbf{x} = [x_1, x_2]^T = [\varphi, \dot{\varphi}]^T$, welcher den Winkel φ und die zugehörige Winkelgeschwindigkeit $\dot{\varphi}$ beinhaltet, beschrieben. Die Dynamik folgt der nichtlinearen Zustandsraumgleichung

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, u; \mathbf{p}) = \begin{bmatrix} x_2 \\ -\frac{g}{l} \sin(x_1) - \frac{d}{ml^2} x_2 + \frac{1}{ml^2} u \end{bmatrix}, \quad (2-4)$$

wobei u das einstellbare Drehmoment, welches im Gelenk wirkt, darstellt. Des Weiteren wird das Pendel durch die mechanische Parametrierung $\mathbf{p} = [l, m, d, g]^T$, mit l und m als Länge und Masse des Pendels, dem Reibungskoeffizienten d und der Gravitationskonstanten g , charakterisiert.

Bild 2-2 zeigt die Phasenebene inkl. durch (2-4) induziertem Vektorfeld (blaue Pfeile) des Pendels. Auf der linken Seite gilt für die Stellgröße $u \equiv 0$, wohingegen auf der rechten Seite ein linearer Zustandsregler nach $u = \mathbf{K}(\mathbf{w} - \mathbf{x})$ mit dem konstanten Sollwert $\mathbf{w} = [\pi, 0]^T$ verwendet wird. Die schwarzen Linien zeigen beispielhaft zwei Trajektorien, wobei die schwarzen Punkte den Anfangszustand andeuten. Die farbigen Kreuze

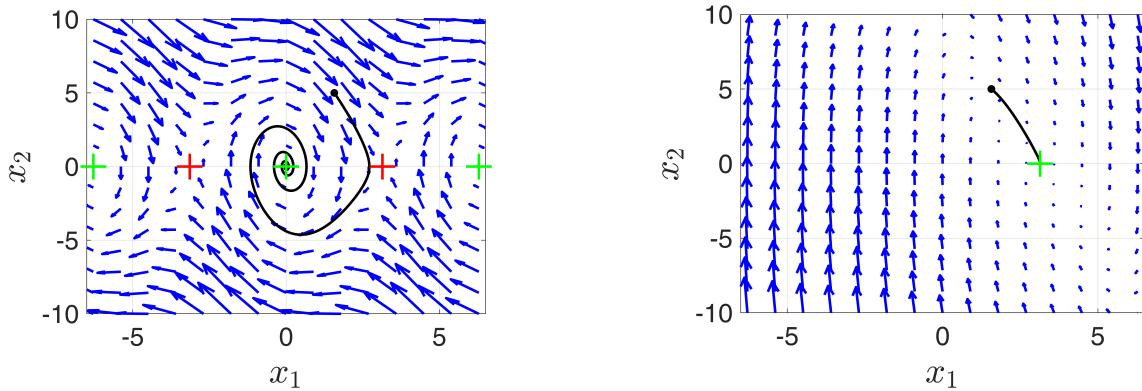


Bild 2-2: Phasenebene der Pendeldynamik (2-4). Freies unaktuiertes (links) und linear geregeltes System (rechts).

kennzeichnen die verschiedenen Ruhelagen bzw. Fixpunkte des Systems. Alle Ruhelagen liegen unter der Voraussetzung von

$$\dot{\mathbf{x}}_R = \mathbf{f}(\mathbf{x}_R, u_R) = \mathbf{0}$$

vor. Gegeben $u_R = 0$, folgen die Ruhelagen des Pendels dem Zusammenhang $\mathbf{x}_R = [n\pi, 0]^T$ mit $n \in \mathbb{Z}$. Eine Charakterisierung findet anhand ihrer Stabilität statt. Nach der Definition von Ljapunov [Ada14], heißt eine Ruhelage stabil, wenn es zu jeder ε -Umgebung

$$U_\varepsilon = \{\mathbf{x} \in \mathbb{R}^{n_x} \mid \|\mathbf{x}\| < \varepsilon\}, \quad \text{eine } \delta\text{-Umgebung} \quad U_\delta = \{\mathbf{x} \in \mathbb{R}^{n_x} \mid \|\mathbf{x}\| < \delta\} \quad (2-5)$$

gibt, sodass alle Trajektorien des Systems, die in der δ -Umgebung beginnen, d. h. $\mathbf{x}(t=0) \in U_\delta$, in ihrem weiteren Verlauf in der ε -Umgebung bleiben, d. h. $\mathbf{x}(t) \in U_\varepsilon$ für $t > 0$. Asymptotische Stabilität liegt vor, wenn außerdem die Trajektorien in die Ruhelage streben. Im Spezialfall eines linearen dynamischen Systems

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}u,$$

kann die Stabilität anhand der Eigenwerte $\lambda_i, i = 1, \dots, n_x$ der Dynamikmatrix \mathbf{A} bestimmt werden. Ist mindestens ein λ_i größer als null, so ist das System instabil.

Für das betrachtete Beispiel des Pendels kann neben dem Stabilitätsnachweis nach Ljapunov alternativ eine Linearisierung in den Ruhelagen durchgeführt werden und die lokale Stabilität anhand der zugehörigen Eigenwerte abgelesen werden. Die roten Kreuze in Bild 2-2 deuten hierzu instabile Ruhelagen an, wohingegen grüne Kreuze für stabile Fixpunkte stehen. Anhand der gezeigten Trajektorie wird ersichtlich, dass sich das System zeitlich zunächst entlang einer Mannigfaltigkeit auf die instabile Ruhelage zubewegt, dann aber

von dieser abgestoßen wird und letztendlich in die benachbarte stabile untere Ruhelage verläuft. Im rechten Bild kommt ein linearer Regler zum Einsatz. Dadurch wird dem System eine neue Dynamik aufgeprägt, die an der Veränderung des Vektorfelds ersichtlich wird. Die ursprünglich instabile Ruhelage wird durch den Regler stabilisiert, sodass sich das System nun ausgehend von dem gleichen Startzustand der linken Seite auf sie zubewegt und dort gehalten wird.

2.2 Wahrscheinlichkeitstheorie

Die Wahrscheinlichkeitstheorie befasst sich formal mit der Untersuchung des Zufalls und der Unsicherheit von bestimmten Ereignissen. Sie bietet einen Rahmen für die quantitative Analyse von probabilistischen Phänomenen und spielt eine entscheidende Rolle in verschiedenen Bereichen, bspw. dem Finanzwesen und der Naturwissenschaften. In Hinblick auf das Thema dieser Arbeit, der Inbetriebnahme einer Steuerung oder Regelung, eröffnet die Wahrscheinlichkeitstheorie eine systematische Sichtweise auf die mit Unsicherheit behafteten Gegebenheiten. Dies kann sowohl die Ungenauigkeiten des Dynamikmodells betreffen, als auch die Gefahr einer möglichen Instabilität des Systems während des Betriebs oder des Inbetriebnahmeprozesses. In diesem Zusammenhang liefert die Wahrscheinlichkeitstheorie, die notwendigen formalen Begriffe und zugehörigen Rechenregeln zur Beschreibung und adäquaten Behandlung der auftretenden Vorgänge. Im Folgenden werden daher hilfreiche stochastische Werkzeuge eingeführt und in Bezug zur Regelungstechnik erläutert. Das Ziel ist es, den herkömmlichen deterministischen Modellbegriff durch probabilistische Überlegungen zu erweitern, sodass er im Rahmen der Inbetriebnahme gewinnbringend eingesetzt werden kann.

2.2.1 Diskrete und kontinuierliche Zufallsvariablen

Zunächst wird der Begriff der Zufallsvariable eingeführt. Hierbei wird zwischen dem diskreten und kontinuierlichen Fall unterschieden. Anhand des diskreten Falls werden anschaulich die grundlegenden Zusammenhänge bei der Behandlung von Zufallsvariablen verdeutlicht und anschließend auf den kontinuierlichen Fall übertragen. Die nachfolgenden Erläuterungen dieses Unterkapitels lehnen an die Darstellungen in [Bis06] und [DFO20] an. Formal ist eine Zufallsvariable eine Abbildung der Form $X : \Omega \mapsto \mathcal{A}$, wobei Ω den sogenannten Ereignisraum und \mathcal{A} die Realisierungen der Zufallsvariable darstellen. Die Eintrittswahrscheinlichkeit der Realisierungen wird über die Funktion $P : \mathcal{A} \mapsto [0, 1]$ beschrieben. Zusammen bilden (Ω, \mathcal{A}, P) den Wahrscheinlichkeitsraum.

Bild 2-3 (oben, links) zeigt den zufälligen Beschuss einer Platte mit $N = 25$ Kugeln. Die Treffer werden bestimmten Gebieten auf der Platte zugeordnet. Das Beispiel soll dazu

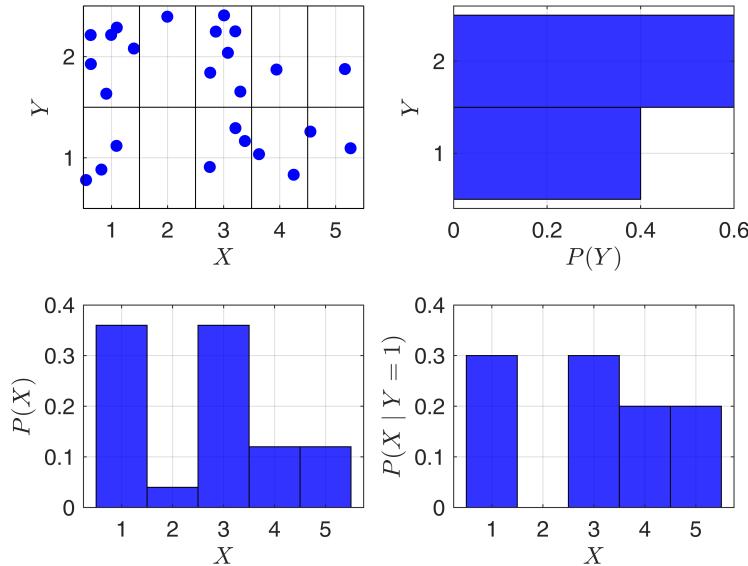


Bild 2-3: Beispiel für eine zweidimensionale diskrete Wahrscheinlichkeitsverteilung

dienen die Begriffe der Wahrscheinlichkeitstheorie einzuführen, wobei die zugrundeliegenden Axiome und Rechenregeln im Anschluss besprochen werden. Die Koordinaten (X, Y) der Felder stellen in diesem Beispiel zwei Zufallsvariablen dar, welche mit

$$\Omega_x = \{x_i = i \mid i = 1, \dots, 5\}, \quad \Omega_y = \{y_j = j \mid j = 1, 2\},$$

nur bestimmte diskrete Werte aus ihrem jeweiligen Ereignisräumen (Ω_x, Ω_y) annehmen können. Die Ereignisräume und die Realisierungen sind für dieses Beispiel identisch. Die Zufallsvariablen (X, Y) beschreiben damit Realisierungen (x, y) , welche einer bestimmten Eintrittswahrscheinlichkeit $P(X = x, Y = y) : \Omega_x \times \Omega_y \mapsto [0, 1]$ folgen. Der Ereignisraum stellt in diesem Zusammenhang einen Bezug zu den Werten her, welche von der Zufallsvariable eingenommen werden können. Im Beispiel stellt eine Kugel eine Realisierung der beiden Zufallsvariablen dar, womit jede Kugel nur ein bestimmtes Gebiet auf der Platte zugeordnet werden kann und die Realisierung somit jeweils einen Wert aus Ω_x und Ω_y zugeordnet ist. In der Wahrscheinlichkeitstheorie besteht das größte Interesse in der Angabe von Wahrscheinlichkeiten, mit denen bestimmte Ereignisse eintreten. Genauer ausgedrückt, wird die sogenannte Wahrscheinlichkeitsverteilung in Abhängigkeit der Zufallsvariablen gesucht. Im vorliegenden Fall kann die (Verbund-)Wahrscheinlichkeitsverteilung

$$P(X = x_i, Y = y_j) = \frac{n_{ij}}{N},$$

durch das Zählen der Kugeln n_{ij} in einem bestimmten Feld und Division durch die Gesamtanzahl N angegeben werden. Diese Größe wird auch gemeinsame Wahrschein-

lichkeitsverteilung genannt, da sie jedem möglichen Ereignispaar eine bestimmte Wahrscheinlichkeit zuordnet. Im Bezug auf Bild 2-3 lauten also bspw. $P(X = 1, Y = 1) = \frac{3}{25}$ und $P(X = 2, Y = 1) = \frac{0}{25} = 0$ für die ersten zwei Felder. In diesem Zusammenhang sind zwei Axiome der Wahrscheinlichkeitstheorie wesentlich und stellen strikte Anforderungen an die Verteilungsfunktion $P(\cdot, \cdot)$ dar. Das erste Axiom lautet

$$0 \leq P(X = x_i, Y = y_j) \leq 1,$$

und schließt negative und unplausible Wahrscheinlichkeiten aus. Das zweite Axiom definiert die Summe über alle Eintrittswahrscheinlichkeiten auf den Wert 1,

$$\sum_{i=1}^{|\Omega_x|} \sum_{j=1}^{|\Omega_y|} P(X = x_i, Y = y_j) = 1,$$

sodass als Konsequenz immer mindestens ein Ereignis aus dem Raum aller Möglichkeiten eintreten muss. Diese Axiome sind für den Fall von zwei Zufallsvariablen allgemeingültig, können auf mehrere Zufallsvariablen verallgemeinert werden und sind nicht nur vor dem Hintergrund des Beispiels zu betrachten.

Neben der Verbundverteilung stellt die Marginal-/Randverteilung für einen Teil der Zufallsvariablen, bzw. eine bestimmte Zufallsvariable

$$P(X = x_i) = \sum_{j=1}^{|\Omega_y|} p(X = x_i, Y = y_j) = \frac{\sum_j n_{ij}}{N},$$

eine wichtige Größe dar. Dieser Zusammenhang wird als Summenregel (Marginalisierung) bezeichnet und findet insbesondere dann Anwendung, wenn nicht alle Zufallsvariablen benötigt oder vernachlässigbar sind. In Bild 2-3 unten links und oben rechts sind diese Verteilungen $P(X)$ und $P(Y)$ durch die Summation über die entsprechend andere Zufallsvariable abgebildet. Weiterführend kann aus den bereits eingeführten Verteilungen die bedingte Wahrscheinlichkeitsverteilung über

$$P(X = x_i | Y = y_j) = \frac{P(X = x_i, Y = y_j)}{P(Y = y_j)} = \frac{n_{ij}}{\sum_i n_{ij}},$$

angegeben werden, wobei auf die Anzahl der Kugeln in den Bereichen des fixierten Subindex j normiert wird, um dem zweiten Axiom der Wahrscheinlichkeitstheorie (Summe über alle Wahrscheinlichkeiten ergibt 1) gerecht zu werden. Diese Regel wird Produktregel genannt. In Bild 2-3 unten rechts ist die bedingte Verteilung $P(X = x_i | Y = 1)$ zu sehen, wobei die spezielle Wahrscheinlichkeit $P(X = 2 | Y = 1)$ mit 0 angegeben ist, da sich im zweiten Feld keine Kugel befindet.

Ein Spezialfall liegt vor, wenn die Verbundwahrscheinlichkeit über die Marginalverteilungen der Zufallsvariablen

$$P(X = x_i, Y = y_j) = P(X = x_i) P(Y = y_j) \text{ oder } P(X = x_i | Y = y_j) = P(X = x_i),$$

angegeben werden kann bzw. die bedingte Wahrscheinlichkeitsverteilung mit der Marginalverteilung übereinstimmt. Dann sind die Zufallsvariablen unabhängig voneinander und das Wissen über den Wert einer Zufallsvariable birgt keine Informationen über die andere Zufallsvariable. Für das anschauliche Beispiel aus Bild 2-3 kann diese Aussage jedoch verneint werden. Besonders deutlich wird dies durch das zweite Feld, welches keine Kugel besitzt, und das genau darüber liegende Feld, welches zumindest eine Kugel aufweist. Die Kenntnis über Y hat damit also direkten Einfluss auf die bedingte Wahrscheinlichkeitsverteilung für X . Ein weiterer Spezialfall ergibt sich bei der Betrachtung von mindestens drei Zufallsvariablen (X, Y, Z) . Dabei kann es zu einer bedingten Unabhängigkeit kommen, die sich formal mit

$$P(X = x_i, Y = y_j | Z = z_k) = P(X = x_i | Z = z_k) P(Y = y_j | Z = z_k)$$

beschreiben lässt. Ein beliebtes Beispiel hierfür ist, dass das Ereignis X das Wetter in Berlin, das Ereignis Y das Wetter in München darstellen und das Ereignis Z das Eintreten eines Hochdruckgebiets ist. X und Y könnten unter der Bedingung des Hochdruckgebiets bedingt unabhängig sein, da das Wetter in Berlin nicht direkt das Wetter in München beeinflusst, wenn bereits bekannt ist, dass ein Hochdruckgebiet existiert. Zur Vereinfachung wird in der Wahrscheinlichkeitstheorie häufig keine explizite Unterscheidung zwischen der Zufallsvariable und ihrer annehmbaren Werte gemacht. In der Regel erschließt sich diese Formalität aus dem Kontext, sodass beim Übergang $(X, Y) \rightarrow (x, y)$ verkürzt für die Wahrscheinlichkeitsverteilungen

$$P(x, y), \quad P(x) \quad \text{und} \quad P(x | y),$$

geschrieben wird. Die hier vorgestellten Grundregeln der Wahrscheinlichkeitstheorie sind äquivalent für $P(y)$ und $P(y | x)$ anwendbar. Des Weiteren sind sie gleichermaßen auf den Fall mit mehr als zwei Zufallsvariablen übertragbar.

Beim Übergang zu kontinuierlichen Zufallsvariablen $x, y \in \mathbb{R}$ bleiben die elementaren Zusammenhänge des diskreten Falls erhalten. Allerdings treten an die Stellen der Sum-

men nun Integrale und die ursprünglichen Wahrscheinlichkeitsverteilungen werden durch Wahrscheinlichkeitsdichteveerteilungen (*Probability Density Function (PDF)*)

$$p(x, y) : \mathbb{R} \times \mathbb{R} \mapsto \mathbb{R}_+,$$

ersetzt. Die Wahrscheinlichkeitsdichteveerteilungen gibt damit nicht mehr direkt die Wahrscheinlichkeit für ein Ereignispaar an, sondern ist als Wahrscheinlichkeit pro Einheit der Zufallsvariablen zu interpretieren. Für den kontinuierlichen Fall lauten die beiden Axiome

$$p(x, y) \geq 0 \quad \text{und} \quad \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x, y) dx dy = 1. \quad (2-6)$$

Da es sich um eine Dichte handelt, kann die Verteilungsfunktion im Vergleich zum diskreten Fall nun auch Werte oberhalb von 1 annehmen. Aus diesem Grund wird zur Unterscheidbarkeit der Kleinbuchstabe p anstelle von P für die Verteilung benutzt. Konkrete Wahrscheinlichkeiten P werden mit einer Integration über ein bestimmtes Intervall berechnet

$$P(x \in [x_l, x_r], y \in [y_l, y_r]) = \int_{y_l}^{y_r} \int_{x_l}^{x_r} p(x, y) dx dy,$$

wobei l für die linke und r für die rechte Grenze stehen. Abschließend ergeben sich die Summen- und Produktregel zu

$$p(x) = \int_{-\infty}^{\infty} p(x, y) dy, \quad \text{und} \quad p(x | y) = \frac{p(x, y)}{p(y)}. \quad (2-7)$$

Die Überlegungen zur (bedingten) Unabhängigkeit sind weiterhin gültig. Soweit nicht weiter angegeben, wird in dieser Arbeit weitestgehend von kontinuierlichen Zufallsvariablen ausgegangen. Die nachfolgenden Abschnitte vertiefen die Handhabung von Zufallsvariablen und zugehörigen PDFs.

2.2.2 Erwartungswert und Varianz

In diesem Unterkapitel wird der Erwartungswertoperator definiert und weitere Größen, die sich hieraus ableiten. Mit Hilfe dieses Operators lassen sich bestimmte Kennwerte, die in der Wahrscheinlichkeitstheorie einen wichtigen Stellenwert haben, angeben. Zunächst wird auf den eindimensionalen Fall mit $x \in \mathbb{R}$ eingegangen und danach der mehrdimensionale Fall $\mathbf{x} \in \mathbb{R}^n$ besprochen.

Der Erwartungswertoperator ist für beliebige Funktionen $f(x)$ definiert und lautet

$$\mathbb{E}_x[f(x)] = \int_{-\infty}^{\infty} f(x)p(x)dx,$$

und bildet damit formal die Kurzschreibweise für ein Integral über das Produkt der Funktion $f(x)$ und der Verteilung der Zufallsvariable $p(x)$. Es gelten daher auch die gleichen Rechenregeln, die auf Integrale zutreffen, also beispielsweise solche, die sich aus der Linearität ergeben. Der Subindex des Operators gibt an, bezüglich welcher Zufallsvariable das Integral gebildet wird. Dies ist hilfreich und trägt dem Verständnis bei, wenn es mehrere Zufallsvariablen gibt. Er wird in vielen Fällen jedoch vereinfachend weggelassen, wenn sich die Operation aus dem Kontext ergibt. Einige Funktionen $f(x)$ sind von besonderer Bedeutung. Für die Identität $f(x) = x$ folgt der sogenannte Mittelwert, oder auch Erwartungswert μ über

$$\mathbb{E}[x] = \int_{-\infty}^{\infty} xp(x)dx =: \mu_x,$$

welcher auch als Moment erster Ordnung bezeichnet wird. Entsprechend folgen für $f(x) = x^n, n \in \mathbb{Z}_{n>1}$ die höheren Momente n -ter Ordnung. Weiterführend kann eine Verkettung des Erwartungswertoperators vorgenommen und eine weitere wichtige Kennzahl über die Funktion $f(x) = (x - \mu)^2$ abgeleitet werden: die Varianz σ^2 , bzw. zentrales Moment zweiter Ordnung aufgrund der Zentrierung durch den Mittelwert. Für die Varianz wird im Rahmen dieser Arbeit ein eigener Operator $\mathbb{V}[\cdot]$ definiert über

$$\mathbb{V}[x] = \mathbb{E}[(x - \mu)^2] = \int_{-\infty}^{\infty} (x - \mu)^2 p(x)dx = \mathbb{E}[x^2] - \mu^2 =: \sigma_x^2,$$

wobei dieser auf das Moment zweiter Ordnung und den quadratischen Mittelwert zurückzuführen ist und hierbei die Eigenschaft der Linearität ausgenutzt wurde. Anschaulich gibt die Varianz damit an, wie weit die Realisierungen der Zufallsvariablen um ihren Mittelwert herum streuen bzw. wie weit sie von diesem entfernt liegen. Sie kann daher einen quantitativen Aufschluss darüber liefern, wie es um die Genauigkeit, Präzision und Robustheit eines Prozesses, bspw. einer *Pick-and-Place* Aufgabe in der Robotik, bestellt ist. Neben der Varianz wird häufig auch die sogenannte Standardabweichung über σ angegeben, da diese die gleiche Einheit wie die Zufallsvariable besitzt.

Der Übergang zu einer vektoriellen Zufallsvariable $\mathbf{x} \in \mathbb{R}^n$ kann als Verbund mehrerer skalarer Zufallsvariablen $x_i, i = 1, \dots, n$ interpretiert werden, womit für die Wahrschein-

lichkeitsdichtefunktion die Kurzschreibweise $p(\mathbf{x}) = p(x_1, \dots, x_n)$ gilt. Der vorherige Erwartungswert wird zum Erwartungsvektor und wird komponentenweise über

$$\mathbb{E}[\mathbf{x}] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \mathbf{x} p(\mathbf{x}) dx_1 \dots dx_n = \begin{bmatrix} \mathbb{E}[x_1] \\ \vdots \\ \mathbb{E}[x_n] \end{bmatrix} =: \boldsymbol{\mu}_x \quad (2-8)$$

gebildet. Die eindimensionale Varianz σ^2 wird im mehrdimensionalen Fall zur (Ko-)Varianzmatrix Σ der Dimension $n \times n$ und besitzt die Eigenschaften, positiv definit und symmetrisch zu sein. Die Berechnungsvorschrift lautet

$$\mathbb{V}[\mathbf{x}] = \mathbb{E}[(\mathbf{x} - \boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})^T] = \mathbb{E}[\mathbf{x}\mathbf{x}^T] - \boldsymbol{\mu}\boldsymbol{\mu}^T =: \boldsymbol{\Sigma}_x. \quad (2-9)$$

Ausgehend von zwei Zufallsvektoren $\mathbf{x} \in \mathbb{R}^{n_x}$, $\mathbf{y} \in \mathbb{R}^{n_y}$ ist zudem der Begriff der Kovarianz, bzw. Kreuz-Kovarianz $\mathbb{C}[\cdot, \cdot]$ von Bedeutung. Formal wird diese, ähnlich zur Varianz, angegeben mit

$$\mathbb{C}[\mathbf{x}, \mathbf{y}] = \mathbb{E}_{x,y}[(\mathbf{x} - \boldsymbol{\mu}_x)(\mathbf{y} - \boldsymbol{\mu}_y)^T] = \mathbb{E}_{x,y}[\mathbf{x}\mathbf{y}^T] - \boldsymbol{\mu}_x\boldsymbol{\mu}_y^T =: \boldsymbol{\Sigma}_{xy}.$$

Die Kreuz-Kovarianzmatrix $\boldsymbol{\Sigma}_{xy}$ hat dabei die Dimension $n_x \times n_y$. Aufgrund der Symmetrieeigenschaft gilt außerdem $\mathbb{C}[\mathbf{x}, \mathbf{y}] = \mathbb{C}[\mathbf{y}, \mathbf{x}]^T$. Bei alleiniger Betrachtung der Zufallsvariable \mathbf{x} kann die (Ko-)Varianzmatrix ebenfalls über den Kreuz-Kovarianz-Operator angegeben werden mit

$$\boldsymbol{\Sigma}_x = \mathbb{V}[\mathbf{x}] = \begin{bmatrix} \mathbb{V}[x_1] & \mathbb{C}[x_1, x_2] & \dots & \mathbb{C}[x_1, x_n] \\ \mathbb{C}[x_2, x_1] & \mathbb{V}[x_2] & \dots & \mathbb{C}[x_2, x_n] \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{C}[x_n, x_1] & \mathbb{C}[x_n, x_2] & \dots & \mathbb{V}[x_n] \end{bmatrix},$$

mit $\mathbb{V}[x_i] = \mathbb{C}[x_i, x_i]$, $i = 1, \dots, n$ als eindimensionale Varianzen.

Die eingeführten Operatoren bilden wichtige Werkzeuge innerhalb der Wahrscheinlichkeitstheorie. Neben ihrer Definition, ist es hilfreich zwei bestimmte Rechenregeln für die Umformung von wiederkehrenden Ausdrücken zu kennen. Zum einen liegen häufig zwei Zufallsvektoren \mathbf{x}, \mathbf{y} gleicher Dimension vor, an deren Summe $\mathbf{z} = \mathbf{x} + \mathbf{y}$ man interessiert

ist oder genauer nur an dem Erwartungsvektor μ_z und der Varianzmatrix Σ_z . In diesem Szenario können die Beziehungen

$$\begin{aligned}\mu_z &= \mathbb{E}[z] = \mathbb{E}[x + y] = \mathbb{E}[x] + \mathbb{E}[y] \\ &= \mu_x + \mu_y, \\ \Sigma_z &= \mathbb{V}[z] = \mathbb{V}[x + y] = \mathbb{V}[x] + \mathbb{V}[y] + \mathbb{C}[x, y] + \mathbb{C}[y, x] \\ &= \Sigma_x + \Sigma_y + \Sigma_{xy} + \Sigma_{yx}\end{aligned}\tag{2-10}$$

angegeben werden [DFO20]. Zum anderen kann oft ein linearer Zusammenhang zwischen den Zufallsvariablen x und y über $y = Ax + b$ mit $A \in \mathbb{R}^{n_y \times n_x}$ und $b \in \mathbb{R}^{n_y}$ zugrunde gelegt werden. Sind die Momente (μ_x, Σ_x) des Eingangs x bekannt, so lassen sich die Momente des Ausgangs y direkt angeben über [DFO20]

$$\begin{aligned}\mu_y &= \mathbb{E}[y] = \mathbb{E}[Ax + b] = A\mathbb{E}[x] + b = A\mu_x + b, \\ \Sigma_y &= \mathbb{V}[y] = \mathbb{V}[Ax + b] = A\mathbb{V}[x]A^T = A\Sigma_x A^T.\end{aligned}\tag{2-11}$$

2.2.3 Transformationssatz und Bayessche Regel

Die lineare Abhängigkeit zweier Zufallsvariablen x, y aus (2-11) stellt einen besonderen Spezialfall dar. Für den allgemeineren Fall eines nichtlinearen Zusammenhangs $x = g(y) : \mathbb{R}^{n_y} \rightarrow \mathbb{R}^{n_x}$ ist der Transformationssatz [Bis06] ein Hilfsmittel, um die Wahrscheinlichkeitsdichtefunktion einer Variablen durch die Wahrscheinlichkeitsdichtefunktion der anderen Variablen auszudrücken. Die zugehörige Vorschrift lautet

$$p_y(y) = p_x(x) \left| \frac{dx}{dy} \right| = p_x(g(y)) \left| \frac{dg(y)}{dy} \right|,\tag{2-12}$$

wobei $|\cdot|$ die Determinante und dx/dy die Jacobimatrix sind. Die Determinante stellt einen Skalierungsfaktor dar, welcher sicherstellt, dass bei der Veränderung der Variable die Axiome der Wahrscheinlichkeitstheorie (2-6) weiterhin eingehalten werden. Eine beispielhafte Anwendung des Transformationssatzes für die Bewegungsgleichung eines Pendels wurde dem Anhang hinzugefügt (s. (A1-1)).

Zuletzt wird die Bayessche Regel [Bis06], auch Satz von Bayes genannt, erläutert. Diese ist ein grundlegendes Werkzeug in der Stochastik und ermöglicht es, Wahrscheinlichkeiten auf der Basis von weiteren Informationen zu aktualisieren. Sie ist damit für zahlreiche Anwendungsfelder relevant, insbesondere wenn es um die kontinuierliche Datenverarbei-

tung von unsicheren Ereignissen geht. Formal ergibt sich der Satz von Bayes als Konsequenz aus den bereits eingeführten Produkt- und Summenregeln (2-7) zu:

$$p(\mathbf{x} | \mathbf{y})p(\mathbf{y}) = p(\mathbf{y} | \mathbf{x})p(\mathbf{x}) \Leftrightarrow \underbrace{p(\mathbf{x} | \mathbf{y})}_{\text{Posterior}} = \frac{\underbrace{p(\mathbf{y} | \mathbf{x})p(\mathbf{x})}_{\text{Likelihood Prior}}}{\underbrace{\int p(\mathbf{y} | \mathbf{x})p(\mathbf{x})d\mathbf{x}}_{\text{Evidence}}}. \quad (2-13)$$

Aufgrund ihres hohen Stellenwerts haben die einzelnen Komponenten ihre eigenen Bezeichnungen (*Prior*, *Likelihood*, *Posterior*, *Evidence*) erhalten. Im Grunde bietet die Bayessche Regel die Möglichkeit bedingte Wahrscheinlichkeitsverteilungen ineinander umzurechnen, weswegen in diesem Kontext häufig auch von einer Inversion der Verteilungen gesprochen wird. Eine besondere Bedeutung kommt den einzelnen Bestandteilen im Zusammenhang mit dem maschinellen Lernen zu. Der Prior $p(\mathbf{x})$ beinhaltet das Vorwissen über die Zufallsvariable \mathbf{x} , bevor etwaige Beobachtungen über die Zufallsvariable \mathbf{y} gemacht wurden. Beim maschinellen Lernen stellt \mathbf{x} bspw. den Parametervektor $\boldsymbol{\theta}$ eines künstlichen neuronalen Netzes dar. \mathbf{y} steht dann für die Trainingsdaten \mathcal{D} , mit deren Hilfe die Parameter bestimmt werden sollen. Der Likelihood-Ausdruck $p(\mathcal{D} | \boldsymbol{\theta})$ gibt somit indirekt an, wie hoch die Wahrscheinlichkeit für die gegebenen Daten unter der Voraussetzung für einen bestimmten Parametervektor ist. Zusammen mit dem Prior $p(\boldsymbol{\theta})$, welcher das Vorwissen über die Parameter des künstlichen neuronalen Netzes beschreibt, wird über die Bayessche Regel die Posterior-Verteilung $p(\boldsymbol{\theta} | \mathcal{D})$ berechnet. Die Evidenz bzw. der Nenner führt dabei zur Normalisierung, wodurch das erste Axiom eingehalten wird. Der Posterior spiegelt die Berücksichtigung der beobachteten Daten wider und wird daher als Update für den Prior angesehen.

2.2.4 Univariate und multivariate Normalverteilung

In der Wahrscheinlichkeitstheorie gibt es zahlreiche nennenswerte Wahrscheinlichkeitsdichtefunktionen $p(\mathbf{x})$, die von besonderer Bedeutung sind. Die mit Abstand bedeutendste Verteilung stellt die Normal- oder auch Gaußverteilung dar. Ihren hohen Stellenwert erhält sie dadurch, dass sie bei vielen natürlichen Prozessen auftritt, so unter anderem auch beim Messrauschen von Sensorsignalen eines mechatronischen Systems. Des Weiteren sind ihre charakteristischen Parameter, die ihre nominale Gestalt festlegen, für den Menschen intuitiv interpretierbar und damit leicht nachvollziehbar.

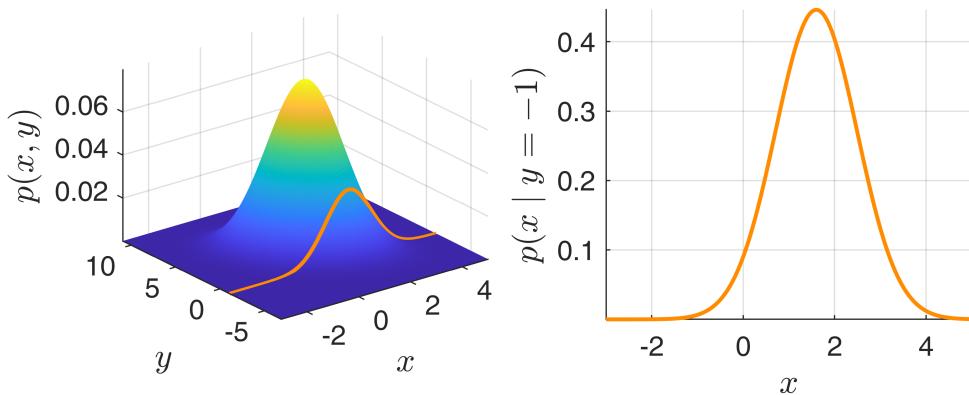


Bild 2-4: Multivariate (links) und (bedingte) univariate Normalverteilung (rechts).

Bild 2-4 zeigt beispielhaft die Form der PDF für den ein- und zweidimensionalen Fall. Im eindimensionalen Fall $x \in \mathbb{R}$ wird die univariate Normalverteilung über

$$x \sim \mathcal{N}_x(\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x - \mu)^2\right)$$

angegeben, wobei das "~-Symbol bedeutet, dass die Zufallsvariable der Verteilung auf der rechten Seite folgt. Alternativ wird auch geschrieben $p(x) = \mathcal{N}(\mu, \sigma^2)$. Die Verteilung wird eindeutig durch den Erwartungswert μ und die Varianz σ^2 charakterisiert, welche ihre Parameter darstellen. Ihren hohen Rang erhält die Normalverteilung unter anderem aufgrund der Tatsache und Besonderheit, dass die Parameter (μ, σ^2) mit den Berechnungen für die Momente ($\mathbb{E}[x], \mathbb{V}[x]$) übereinstimmen. Dies ist für andere bekannte Verteilungen nicht der Fall. Die Verallgemeinerung für den höher-/mehrdimensionalen Fall $\mathbf{x} \in \mathbb{R}^n$ lautet formal

$$\mathbf{x} \sim \mathcal{N}_x(\boldsymbol{\mu}, \Sigma) = \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu})\right), \quad (2-14)$$

wobei von der multivariaten Normalverteilung gesprochen wird. Der Exponent $(\mathbf{x} - \boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu})$ wird als quadratisches Mahalanobis Distanzmaß [DFO20] bezeichnet und definiert im zweidimensionalen Fall eine ellipsenförmige Darstellung über die Höhenlinien der Verteilungsfunktion. Mit den Dimensionen $\boldsymbol{\mu} \in \mathbb{R}^n$ und $\Sigma \in \mathbb{R}^{n \times n}$ ergibt sich die Anzahl der Parameter zu $\frac{2n+n(n+1)}{2}$, da die Kovarianzmatrix symmetrisch und positiv definit ist. Wie im eindimensionalen Fall stimmen die Parameter mit den ersten beiden Momenten überein. In Hinblick auf Bild 2-4 wird ersichtlich, dass sich das Maximum der PDF immer beim Erwartungswert befindet und sich die Dichte nach außen hin kontinuierlich gegen den Wert null verringert. Die Abfallrate wird dabei durch den Wert der Varianz bestimmt. Je größer die Varianz, desto weitreichender das Gebiet, welches durch einen signifikanten Wert der PDF abgedeckt wird. In Bild 2-4 (links) wird ein

formaler Zusammenhang zwischen der zweidimensionalen und eindimensionalen Normalverteilung angedeutet, welcher erst im Abschnitt 2.4.1 näher erläutert wird.

Für einige Anwendungen ist es notwendig, die Ableitung der stochastischen Variable \mathbf{x} nach der Parametrierung der zugrundeliegenden PDF, d. h. $\frac{d\mathbf{x}}{d\boldsymbol{\mu}}$, $\frac{d\mathbf{x}}{d\Sigma}$ für den Fall einer Normalverteilung, zu kennen. Hierfür ist der sogenannte Reparametrisierungstrick [DFO20] ein wichtiges Hilfsmittel. Ausgehend von einer Standardnormalverteilung $\mathbf{y} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_n)$, mit \mathbf{I}_n als Einheitsmatrix der Dimension n , gilt für die betrachtete Zufallsvariable \mathbf{x} der Zusammenhang

$$\mathbf{x} = \boldsymbol{\mu} + \mathbf{Ly}, \quad \mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma), \quad \text{mit} \quad \mathbf{L}\mathbf{L}^T = \Sigma. \quad (2-15)$$

Das Produkt $\mathbf{L}\mathbf{L}^T$ stellt die Cholesky-Zerlegung dar. Auf diese Weise lässt sich der stochastische Gradient (bzgl. der Parametrierung) auf eine deterministische Betrachtungsweise zurückführen und entsprechend einfach berechnen.

2.2.5 Numerische Approximation

Ebenso wie in anderen mathematischen Themengebieten, werden auch in der Wahrscheinlichkeitstheorie numerische Approximationsverfahren eingesetzt, um komplizierte Ausdrücke für die eine analytische Herleitung zu aufwendig oder gar unmöglich aufzustellen ist, zumindest näherungsweise berechenbar zu machen. Von der Vielzahl von Verfahren wird an dieser Stelle die *Monte-Carlo* (MC) Methode [DFO20] vorgestellt. Das Grundprinzip besteht darin, die Approximation auf der Grundlage von generierten Zufallszahlen vorzunehmen. Dabei gilt der Zusammenhang, dass die Anzahl der verwendeten Zufallszahlen die Güte der Approximation bestimmt. Zwei sehr häufig anfallende Ausdrücke involvieren die Berechnung des Erwartungswertoperators und die Integration über eine Zufallsvariable (bspw. zur Marginalisierung). Die statistische Schätzung für den Erwartungswert für eine Funktion $\mathbf{f}(\mathbf{x})$ und die Zufallsvariable \mathbf{x} lautet

$$\mathbb{E}[\mathbf{f}(\mathbf{x})] \approx \frac{1}{n_s} \sum_{i=1}^{n_s} \mathbf{f}(\mathbf{x}_i), \quad \mathbf{x}_i \sim p(\mathbf{x}),$$

wobei n_s die Anzahl der Zufallszahlen, welche auch *Partikel* oder *Sample* genannt werden, angibt und der zugrundeliegenden PDF $p(\mathbf{x})$ folgen. Die Schätzung der Approximation ist asymptotisch konsistent, da

$$\lim_{n_s \rightarrow \infty} \frac{1}{n_s} \sum_{i=1}^{n_s} \mathbf{f}(\mathbf{x}_i) = \int \mathbf{f}(\mathbf{x}) p(\mathbf{x}) d\mathbf{x},$$

gilt und sich die Summe dem Integral für zunehmende Partikelanzahl annähert [DFO20]. Außerdem ist der Schätzer unverzerrt (*unbiased*) von bspw. einem konstanten Störfaktor, der den Wert der Approximation negativ beeinflussen würde. Neben der Approximation des Erwartungsoperators, wird der MC Ansatz auch für die Integration über eine Zufallsvariable (Marginalisierung) über z. B.

$$p(\mathbf{x}) = \int p(\mathbf{x} \mid \mathbf{y}) p(\mathbf{y}) d\mathbf{y} \approx \frac{1}{n_s} \sum_{i=1}^{n_s} p(\mathbf{x} \mid \mathbf{y}_i), \quad \mathbf{y}_i \sim p(\mathbf{y}),$$

benutzt.

Offen bleibt die Frage nach der Berechnung der Zufallszahlen mittels eines deterministisch arbeitenden Rechners, welcher keinen Zugang zu einem stochastischen natürlichen Prozess, wie dem Rauschen eines Sensorsignals, hat. In diesem Fall generiert der Rechner *Pseudo-Zufallszahlen* auf der Basis einer geeigneten Iterationsvorschrift, deren Vorhersage anhand der Zahlen idealerweise schwer vorhersagbar ist. Die Simulation eines Doppelpendels bietet sich z. B. für eine solche Vorschrift an, da das chaotische Verhalten des Systems die Zufälligkeit der Zahlenfolge begünstigt. An dieser Stelle zeigt sich eine interessante Querverbindung von dynamischen Systemen und der Wahrscheinlichkeitstheorie, welche nicht nur bei der Erzeugung von Zufallszahlen auftritt, sondern auch bei anderen Fragestellungen von Bedeutung ist. Eine andere moderne Anwendung findet sich bei der Generierung von hoch dimensionalen Zufallszahlen, welche von der Posterior-Verteilung im Rahmen der Bayesschen Regel (vgl. (2-13)) stammen. Hierbei wird der sogenannte *Hamilton-Markov-Chain-Monte-Carlo* (HMC) Ansatz [Bet17] verwendet, welcher die gleichnamige hamiltonsche Physik ausnutzt, um sich zu stabilisierende Iterationsvorschriften für die Erzeugung von Partikeln der Posterior-Verteilung zu bestimmen. Dabei wird die Folge von Zufallszahlen wie ein dynamisches System mit einer bestimmten Trajektorie aufgefasst, welche sich auf einem konstanten Energieniveau aufhält. Im Rahmen dieser Arbeit wird dieser Ansatz nicht verwendet, bietet jedoch einige interessante weiterführende Fragestellungen und Forschungsansätze, um die hier behandelten Verfahren zu verbessern und weiterzuentwickeln.

Des Weiteren sind für die Praxis der empirische Erwartungsvektor $\hat{\mu}$ und die empirische Kovarianzmatrix $\hat{\Sigma}$ höchst relevant,

$$\begin{aligned} \mathbb{E}[\mathbf{x}] &\approx \frac{1}{n_s} \sum_{i=1}^{n_s} \mathbf{x}_i =: \hat{\mu}, \\ \mathbb{V}[\mathbf{x}] &\approx \frac{1}{n_s} \sum_{i=1}^{n_s} (\mathbf{x}_i - \mathbb{E}[\mathbf{x}]) (\mathbf{x}_i - \mathbb{E}[\mathbf{x}])^T =: \hat{\Sigma}, \end{aligned} \tag{2-16}$$

wobei für die Berechnung der empirischen Kovarianzmatrix in der Regel nicht der wahre Erwartungsvektor $\mathbb{E}[\mathbf{x}]$ zur Verfügung steht und durch den empirischen Erwartungsvektor $\hat{\mu}$ ausgetauscht werden muss. In diesem Fall ist die Schätzung allerdings verzerrt (*biased*) und muss für die Konsistenz durch den Vorfaktor $\frac{1}{n_s-1}$ anstelle von $\frac{1}{n_s}$ für $n_s > 1$ korrigiert werden [Bis06].

2.3 Anwendung der Wahrscheinlichkeitstheorie auf Pendeldynamik

Um die bereits eingeführten Begrifflichkeiten und Rechenregeln besser nachvollziehen zu können und einen Transfer zur Regelungstechnik herzustellen, wird als nächstes die Wahrscheinlichkeitstheorie auf bestimmte regelungstechnische Fragestellungen angewendet. Auf diese Weise können die entwickelten Verfahren zur Inbetriebnahme innerhalb der Hauptkapitel für den Leser besser nachvollzogen werden.

2.3.1 Parameteridentifikation mittels Bayesscher linearer Regression

Im Rahmen des modellbasierten Reglerentwurfs und der damit verbundenen physikalischen Modellbildung wird eine Parameteridentifikation für das Dynamikmodell durchgeführt. Dabei wird in ähnlicher Weise zum maschinellen Lernen anhand von gemessenen Zustandsdaten die Parametrierung über die Lösung eines Optimierungsproblems bestimmt. Die zugehörige Zielfunktion wird dabei so formuliert, dass eine robuste Parametrierung resultiert, die das dynamische Verhalten bestmöglich verallgemeinert und nicht nur für den Bereich, in welchem die Daten aufgenommen wurden, geeignet ist. Die Bayessche Regel (2-13) bietet in Verbindung mit der linearen Regression ein mächtiges Werkzeug für eine erweiterte Parameteridentifikation. Der Vorteil gegenüber einer regulären linearen Regression ist, dass sich die Unsicherheit über die nominalen Parameterwerte angeben lassen und ihr Einfluss auf Vorhersagen berücksichtigt werden kann. Illustrativ wird der Ansatz der Bayesschen linearen Regression [Bis06] zur Parameteridentifikation für ein unaktuiertes Pendel vorgestellt. Den Ausgangspunkt bildet die Bewegungsgleichung des Systems

$$\ddot{\varphi} = -\underbrace{\frac{g}{l}}_{=:w_1} \sin(\varphi) - \underbrace{\frac{d}{ml^2}}_{=:w_2} \dot{\varphi} = \underbrace{\begin{bmatrix} -\sin(\varphi) & -\dot{\varphi} \end{bmatrix}}_{=: \phi(\mathbf{x})^T} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}, \quad (2-17)$$

wobei vereinfachend die Substitutionen $w_1 := g/l$ und $w_2 := d/(ml^2)$ für die zu identifizierenden Parameter eingeführt wurde. Die Rückrechnung auf l, d mit bekanntem g, m ist über $l = g/w_1$ und $d = w_2/(ml^2) = (w_2 w_1^2)/(mg^2)$ möglich. Die Zielsetzung bei den nachfolgenden Ausführungen ist es, keine exakte Parametrierung für \mathbf{w} anhand von Zu-

standsmessungen zu identifizieren, sondern eine Wahrscheinlichkeitsverteilung $p(\mathbf{w})$ auf der Grundlage der Bayesschen Regel abzuleiten. Hiermit wird es dem Regelungstechniker ermöglicht eine Aussage darüber zu treffen, wie das Ergebnis der Identifikation zu bewerten ist und ob weitere Messungen notwendig sind, um die Genauigkeit zu erhöhen bzw. die Unsicherheit in Bezug auf die Parametrierung zu verringern.

Zunächst wird im Sinne des maschinellen Lernens der Eingang bzw. Zustand als $\mathbf{x} := [\varphi, \dot{\varphi}]^T$ und der Ausgang bzw. Beschleunigung mit $\ddot{\varphi}(\mathbf{x}; \mathbf{w}) = \phi(\mathbf{x})^T \mathbf{w}$ definiert. Dementsprechend werden die aufgenommenen Messdaten (n_d an der Zahl) über die Eingangsmatrix $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_{n_d}] \in \mathbb{R}^{n_x \times n_d}$ und den zugehörigen Ausgangsvektor $\mathbf{Y} = [y_1, \dots, y_{n_d}]^T \in \mathbb{R}^{n_d}$ angegeben. Aus der Definition des Ausgangs, bzw. der Bewegungsgleichung des Systems (2-17) folgt die Basisfunktion mit $\phi(\mathbf{x}) : \mathbb{R}^{n_x} \mapsto \mathbb{R}^{n_b}$, mit $\phi_i(\mathbf{x}) : \mathbb{R}^{n_x} \mapsto \mathbb{R}$, $i = 1, \dots, n_b$, wobei für den hier beschriebenen Fall $n_b = 2$ gilt. Da von nicht ideal aufgenommenen Messdaten ausgegangen werden soll, wird von einem normalverteilten mittelwertfreien Messrauschen mit einer Varianz von σ_n^2 am Ausgang ausgegangen. Damit lautet die weitere Annahme $y_i = \ddot{\varphi}(\mathbf{x}_i; \mathbf{w}) + \epsilon_i$, $\epsilon_i \sim \mathcal{N}(0, \sigma_n^2)$, bzw. unter Berücksichtigung von (2-15) die zugehörige PDF:

$$p(y | \mathbf{w}) = \mathcal{N}(\ddot{\varphi}(\mathbf{x}; \mathbf{w}), \sigma_n^2).$$

Für die Gesamtheit aller Ausgangsdaten \mathbf{Y} ergibt sich somit der Ansatz für die Likelihoodfunktion aus (2-13) zu

$$p(\mathbf{Y} | \mathbf{w}) = \prod_{i=1}^{n_d} \mathcal{N}(\phi(\mathbf{x}_i)^T \mathbf{w}, \sigma_n^2) = \mathcal{N}(\Phi \mathbf{w}, \sigma_n^2 \mathbf{I}_{n_d}), \quad (2-18)$$

mit der Gramschen Matrix $\Phi := [\phi(\mathbf{x}_1), \dots, \phi(\mathbf{x}_{n_d})]^T \in \mathbb{R}^{n_d \times n_b}$, welche die Transformation der Eingangsdaten über die Basisfunktionen beschreibt. Die Likelihood-Verteilung ordnet damit jeder Parametrisierung einen bestimmten Wahrscheinlichkeitsdichtewert zu. Da es sich bei (2-18) insbesondere für große n_d um ein Produkt von ggf. kleinen Zahlwerten nahe null handelt und dies zu numerischen Problemen auf einem Rechner führen kann, wird anstelle der Likelihood üblicherweise die logarithmische Likelihood

$$\log p(\mathbf{Y} | \mathbf{w}) = -\frac{1}{2\sigma_n^2} (\mathbf{Y} - \Phi \mathbf{w})^T (\mathbf{Y} - \Phi \mathbf{w}) - \frac{n_d}{2} \log(2\pi\sigma_n^2) \quad (2-19)$$

betrachtet. Auf diese Weise wird aus dem Produkt eine Summe, die vorteilhaft bei der numerischen Umsetzung ist.

Eine Möglichkeit die Parameter nun zu bestimmen, ist über die Maximierung der logarithmischen Likelihood (ML), wobei die analytische Lösung aus $\frac{d \log p(\mathbf{Y}|\mathbf{w})}{d\mathbf{w}} = 0$ zu dem Ergebnis

$$\mathbf{w}_{ML} = (\Phi^T \Phi)^{-1} \Phi^T \mathbf{Y} \quad (2-20)$$

führt. Dies entspricht der Lösung der kleinsten Fehlerquadrate, welche normalerweise beim klassischen modellbasierten Entwurf zum Einsatz kommt. Der hier vorgestellte Ansatz der Bayesschen linearen Regression ist damit konsistent zum herkömmlichen Vorgehen und erweitert ihn damit um probabilistische Überlegungen.

Für die Anwendung der Bayesschen Regel wird die Parametrierung \mathbf{w} als Zufallsvektor definiert und eine Prior-Verteilung $p(\mathbf{w}) = \mathcal{N}(\boldsymbol{\mu}_w, \Sigma_w)$ eingeführt. Die Prior-Verteilung beinhaltet das Vorwissen, welches vor der Aufnahme der Daten bekannt ist, und wird über die Festlegung des Erwartungsvektors $\boldsymbol{\mu}_w$, also der bestmöglichen Schätzung für die Parameter, und der Kovarianzmatrix Σ_w , womit die Unsicherheit über die Parametrisierung ausgedrückt wird, aufgestellt. Die Anwendung von (2-13) unter Vernachlässigung des Nenners $p(\mathbf{Y})$ führt anschließend auf

$$p(\mathbf{w} | \mathbf{Y}) \propto p(\mathbf{Y} | \mathbf{w})p(\mathbf{w}) = \mathcal{N}(\Phi\mathbf{w}, \sigma_n^2 \mathbf{I}_{n_d}) \mathcal{N}(\boldsymbol{\mu}_w, \Sigma_w).$$

Die Evidenz $p(\mathbf{Y})$ ist die Integrationskonstante zur Erfüllung des zweiten Axioms der Wahrscheinlichkeitstheorie (2-6) und erschließt sich aus der Tatsache, dass das Produkt zweier Normalverteilungen wiederum auf eine Normalverteilung führt (vgl. (A1-2) im Anhang), deren Normalisierungsfaktor aus der bekannten Form einer Normalverteilung (2-14) bestimmbar ist. Die analytische Lösung der Posterior-Verteilung kann dementsprechend kompakt mit

$$\begin{aligned} p(\mathbf{w} | \mathbf{Y}) &= \mathcal{N}(\boldsymbol{\mu}_{w|Y}, \Sigma_{w|Y}), \\ \boldsymbol{\mu}_{w|Y} &= \Sigma_{w|Y} (\Sigma_w^{-1} \boldsymbol{\mu}_w + \frac{1}{\sigma_n^2} \Phi^T \mathbf{Y}), \\ \Sigma_{w|Y} &= (\Sigma_w^{-1} + \frac{1}{\sigma_n^2} \Phi^T \Phi)^{-1}, \end{aligned} \quad (2-21)$$

angegeben werden [Bis06]. Für einen nicht informativen Prior, der eine unendliche hohe Unsicherheit aufweist $\Sigma_w \rightarrow \infty$, folgt aus der Posterior-Gleichung die Lösung des Maximum Likelihood Ansatzes $\boldsymbol{\mu}_{w|Y} \rightarrow \mathbf{w}_{ML}$, womit die Konsistenz gegeben ist.

Vor dem Hintergrund eines online lernfähigen Algorithmus oder dem Lernen auf kontinuierlichen Datenströmen bietet das Rahmenwerk der Bayesschen linearen Regression eine intuitive Interpretationsmöglichkeit. Ausgehend von einer aktuell bekannten Posterior-Verteilung kann diese für neue erfasste Daten als Prior-Verteilung für eine weitere An-

wendung der Bayesschen Regel verwendet werden. Auf diese Weise ergibt sich insgesamt ein iteratives Vorgehen, wobei neue Daten zu einem Update des Posteriors führen. Für die praktische Anwendung ist es notwendig, dass der Prior nicht zu restriktiv für zumindest theoretisch mögliche Werte von \mathbf{w} ist. Sind beispielsweise die Funktionswerte des Prior in einem bestimmten Gebiet, welches jedoch durch die Daten abgedeckt wird, zu gering, so wird das Update über die Bayessche Regel sehr viele Daten benötigen, um die fehlerhafte Annahme zu korrigieren. Im Extremfall weist der Prior für das Gebiet keine Wahrscheinlichkeitsdichte (entspricht theoretisch unmöglich) auf, sodass durch die Anwendung der Bayesschen Regel ausgeschlossen ist, dass der Posterior eine bestimmte Dichte in dem Gebiet erhält. Dies ist auf die multiplikative Vorschrift der Bayesschen Regel zurückzuführen. Grundsätzlich lässt sich zudem festhalten, dass je mehr Daten vorliegen, desto irrelevanter wird der Einfluss des Priors (und umgekehrt). Des Weiteren ist es entscheidend, welcher formale Ansatz für die Likelihoodfunktion und den Prior gemacht werden. Im obigen Beispiel wurde für beide eine Normalverteilung angenommen, was je nach Anwendungsfall und Datenbeschaffenheit nicht der Fall sein muss und ggf. durch andere Verteilungen besser abgebildet werden kann. Dementsprechend kann allerdings auch die analytische Berechenbarkeit verloren gehen, was zu aufwendigeren Berechnungen über numerische Approximationen führt.

Die Einführung des Parametervektors \mathbf{w} als Zufallsvariable bietet im Zusammenhang mit dem Posterior die Möglichkeit die Auswirkung auf die ursprüngliche Bewegungsgleichung (2-17) zu untersuchen. Hierzu wird die sogenannte prädiktive Verteilung [Bis06] der Ausgangsgröße bzw. Winkelbeschleunigung

$$\begin{aligned} p(y \mid \mathbf{Y}) &= \int p(y \mid \mathbf{Y}, \mathbf{w}) p(\mathbf{w} \mid \mathbf{Y}) d\mathbf{w} \\ &= \int \mathcal{N}(\ddot{\varphi}(\mathbf{x}; \mathbf{w}), \sigma_n^2) \mathcal{N}(\boldsymbol{\mu}_{w|Y}, \boldsymbol{\Sigma}_{w|Y}) d\mathbf{w} \\ &= \mathcal{N}(\boldsymbol{\phi}(\mathbf{x})^T \boldsymbol{\mu}_{w|Y}, \frac{1}{\sigma_n^2} + \boldsymbol{\phi}(\mathbf{x})^T \boldsymbol{\Sigma}_{w|Y} \boldsymbol{\phi}(\mathbf{x})) \end{aligned} \quad (2-22)$$

gebildet, wobei über die normalverteilte Parametrierung integriert wird (s. (A1-3)). Die Vereinfachung $p(y \mid \mathbf{Y}, \mathbf{w}) = p(y \mid \mathbf{w})$ ist dabei zulässig, da keine Abhängigkeit zu den Daten besteht. Zur Veranschaulichung der Zusammenhänge dient Bild 2-5. Im oberen, linken Teilbild ist die Phasenebene des Pendels abgebildet. Die Farbe des Hintergrunds gibt die Höhe der Winkelbeschleunigung an (vgl. (2-17) und Bild 2-2). Außerdem ist in schwarz eine Trajektorie des Systems zu sehen, von welcher die roten Messpunkte ausgewählt wurden. Anhand dieser Messpunkte wird die Bayessche lineare Regression durchgeführt. Hierzu wird der Prior im oberen, rechten Teilbild als Vorwissen eingeführt. Der rote Diamant kennzeichnet die wahre Parametrierung des Systems (beim maschinellen Lernen wird hierfür der Begriff *Ground Truth* benutzt). Es ist gut erkennbar, dass der

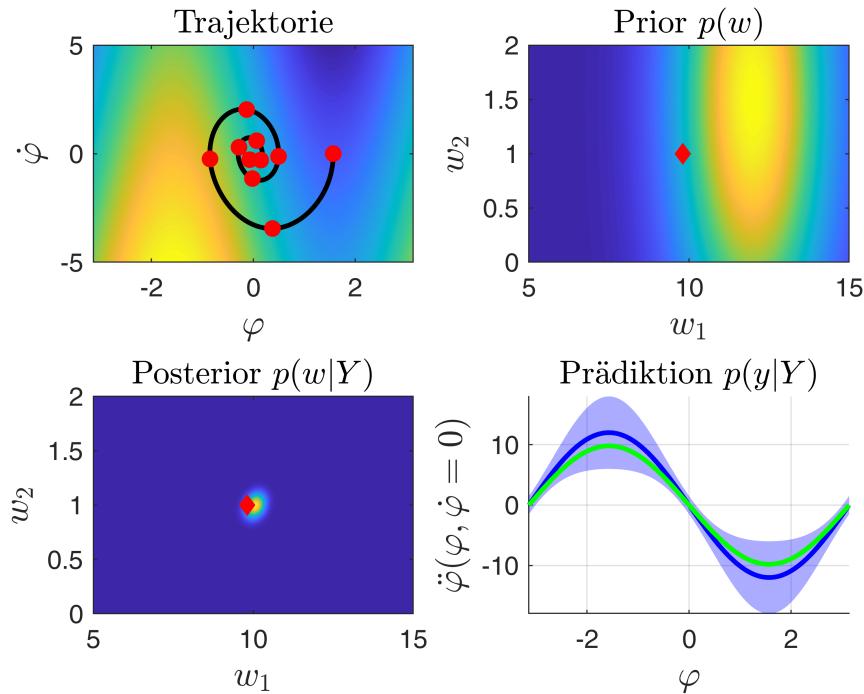


Bild 2-5: Probabilistische Parameteridentifikation für das Pendel mittels Bayesscher linearer Regression.

Mittelwert des Priors noch nicht mit dem Ground Truth übereinstimmt und die Varianz eine hohe Unsicherheit ausdrückt. Durch Anwendung von (2-21) und Berücksichtigung der Messdaten ergibt sich das untere, linke Teilbild für die Posterior-Verteilung. Der Erwartungswert stimmt nun fast mit dem Ground Truth überein und die Unsicherheit ist größtenteils verschwunden, sodass der Schätzung vertraut werden kann. Das untere, rechte Teilbild zeigt die Funktion der Bewegungsgleichung (2-17) für die Fixierung $\dot{\varphi} = 0$ (aus Darstellungsgründen) in grün, sowie ihre Prädiktion (2-22) in Form des Erwartungswertes (blaue Linie) und der dreifachen Standardabweichung (blaue Fläche). In Bezug auf die Identifikationsaufgabe steht dem Regelungstechniker nun nicht nur eine nominale Schätzung der Parametrierung (repräsentiert durch $\mu_{w|Y}$) zur Verfügung, sondern auch ein Maß (in Form der Varianz) zur Beurteilung der Aussagekraft. Damit kann entschieden werden, ob die Anzahl der Daten ausreichend ist oder ob weitere Daten für eine genauere Identifikation erforderlich sind. Des Weiteren kann über die prädiktive Verteilung direkt die Auswirkung der Parametrierung auf die Dynamik des Systems abgeschätzt werden und diese bei nachfolgenden Stabilitätsuntersuchungen und dem Regelungsentwurf einbezogen werden. Bisher wurde nur die Parametrierung eines dynamischen Systems als Zufallsvariable aufgefasst, allerdings zeigt sich durch die direkte Abhängigkeit, dass damit gleichermaßen auch der Zustand als Zufallsvariable betrachtet werden muss. In diesem Fall ist momentan unklar, wie eine Simulation bzw. Langzeitprädiktion mit dem probabilistischen Modell durchgeführt werden kann. Im nächsten Abschnitt wird diese

Fragestellung behandelt und der Einfachheit und Übersicht halber nur von einer Wahrscheinlichkeitsverteilung für den Zustand ausgegangen und die unsichere Parametrierung vernachlässigt.

2.3.2 Zustandspropagation mittels Moment Matching, Linearisierung und Unscented Transform

Die nachfolgenden Überlegungen werden weiterhin am Beispiel des Pendels mit dem nichtlinearen Zustandsraummodell

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}) = \begin{bmatrix} x_2 \\ -g_l \sin(x_1) - d_J x_2 + J_i u \end{bmatrix}, \quad (2-23)$$

vorgestellt, wobei zusätzlich eine Stellgröße u berücksichtigt wird. Zur besseren Übersicht wurden die Substitutionen $g_l := g/l$, $d_J := d/(ml^2)$ und $J_i := 1/(ml^2)$ eingeführt. Des Weiteren wird zur Simulation des Systems das Euler-Integrationsschema mit

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \Delta t \mathbf{f}(\mathbf{x}_k, \mathbf{u}_k) \quad (2-24)$$

vorausgesetzt, womit ausgehend von einem Anfangszustand und bekannten Steuerungsgrößen eine deterministische Langzeitprädiktion durchgeführt werden kann. Das Ziel dieses Abschnitts ist verschiedene Iterationsvorschriften zur probabilistischen Langzeitprädiktion kennen zu lernen. Genau genommen handelt es sich bei den Verfahren um eindeutige Berechnungsgleichungen, ähnlich zu (2-24), die damit als deterministisch bezeichnet werden müssten. Zur Kenntlichmachung des Kontextes und weil sie den wahrscheinlichkeitsbasierten Überlegungen zugeordnet sind, werden sie im Rahmen dieser Arbeit allerdings als probabilistisch bezeichnet. Die Ausgangssituation bildet ein normalverteilter Zustandsvektor (im Zeitschritt k)

$$\mathbf{x}_k \sim \mathcal{N}(\mathbf{m}_{x_k}, \mathbf{S}_{x_k}),$$

mit Zustandserwartungsvektor \mathbf{m}_{x_k} und Zustandskovarianzmatrix \mathbf{S}_{x_k} . Zur besseren Unterscheidung werden die Parameter der Normalverteilung, wenn sie sich auf einen Zustand beziehen, mit (\mathbf{m}, \mathbf{S}) anstelle von $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ beschrieben. Die Stellgröße \mathbf{u}_k bleibt weiterhin deterministisch, sodass für sie keine Verteilung angenommen wird. Eine äquivalente Betrachtung ist $\mathbf{u}_k \sim \mathcal{N}(\mathbf{m}_{u_k}, \mathbf{0})$. Der Grund hierfür ist, dass die Stellgröße durch die Steuereinheit dem System eindeutig vorgegeben werden kann und damit keine Unsicherheit über ihren Wert beinhaltet. In der Praxis könnten bspw. Fehler des Aktors zu einer Unsicherheit der Stellgröße führen, allerdings sollen solche Fälle in dieser Arbeit nicht betrachtet werden. Ausgehend von einem unsicheren aktuellen Zustand \mathbf{x}_k , stellt sich die

Frage nach der zugehörigen Folgezustandsverteilung $p(\mathbf{x}_{k+1})$. Da es sich bei (2-23) um einen nichtlinearen Ausdruck handelt, müsste für die Berechnung der Transformationssatz aus (2-12) angewendet werden. Aufgrund der beinhaltenden Terme ist die Berechnung allerdings analytisch als auch numerisch bereits für einfache Nichtlinearitäten aufwendig. Außerdem wird das Ergebnis der Folgezustandsverteilung keiner formal definierten Verteilung folgen, sondern eine beliebige Verteilung sein, welche schwierig für den Menschen interpretierbar ist. Es wird zudem kompliziert sein mit dieser Verteilung die darauffolgende Folgezustandsverteilung $p(\mathbf{x}_{k+2})$ zu berechnen und eine effiziente Langzeitprädiktion für weitere Zeitschritte aufzustellen. Eine gängige Alternative stellt daher der sogenannte (exakte) *Moment Matching* (MM) Ansatz dar [DFR15]. Dabei wird die Folgezustandsverteilung bestmöglich durch eine Normalverteilung approximiert

$$p(\mathbf{x}_{k+1}) \approx \mathcal{N}(\mathbf{m}_{x_{k+1}}, \mathbf{S}_{x_{k+1}}),$$

womit die Nachteile der exakten Berechnung behoben werden. Aufgrund der Tatsache, dass die Normalverteilung unimodal ist, bzw. nur ein einziges Maximum besitzt, kann durch diesen Approximationsansatz eine multimodale Zustandsverteilung nicht abgebildet werden, welche insbesondere bei probabilistischen Langzeitprädiktionen oft ausgeprägt ist (vgl. Abschnitt 2.3.3). In Bezug auf die Langzeitprädiktionen (in diesem Kontext auch Zustandspropagation genannt) bietet das MM jedoch einen wesentlichen Vorteil, da die Klasse der Normalverteilungen für alle Zeitschritte nie verlassen wird und damit der Berechnungsaufwand konstant bleibt. Speziell im Zusammenhang mit Optimierungsproblemen, welche direkt auf der Zustandspropagation aufbauen, spielt dies eine entscheidende Rolle bei der praktischen Realisierbarkeit auf einem Rechner. Des Weiteren wird sich der Regelungstechniker bei der Auswertung der probabilistischen Simulation in der Regel nur am Erwartungswert und der Varianz des Zustands orientieren, um eine Einschätzung des dynamischen (geregelten) Systemverhaltens zu erhalten. Multimodale Zustandsverteilung sind vor diesem Hintergrund nicht zweckmäßig und würden in den meisten Fällen auf eine nicht funktionierende Steuerung/Regelung hindeuten, welche auch durch die Größe der Zustandsvarianz alleine beschrieben wird. Eine tiefer gehende Auseinandersetzung mit diesem Thema wird in den Hauptkapiteln dieser Arbeit vorgenommen.

Nachfolgend werden drei Möglichkeiten beschrieben, um die Parametrierung der normalverteilten Folgezustandsverteilung $(\mathbf{m}_{x_{k+1}}, \mathbf{S}_{x_{k+1}})$ zu berechnen. Die erste Möglichkeit ist die (exakte) analytische Berechnung über die eingeführten Definitionen zum Erwar-

tungsvektor (2-8) und der Varianzmatrix (2-9). Beispielhaft für das betrachtete Pendelsystem (2-23) ergibt sich der Zustandserwartungsvektor zu

$$\begin{aligned} \mathbf{m}_{x_{k+1}}^{MM} &= \mathbb{E}[\mathbf{x}_{k+1}] \\ &= \mathbb{E}[\mathbf{x}_k] + \Delta t \begin{bmatrix} \mathbb{E}[x_k^{(2)}] \\ -g_l \mathbb{E}[\sin(x_k^{(1)})] - d_J \mathbb{E}[x_k^{(2)}] + J_i u_k \end{bmatrix} \\ &= \mathbf{m}_{x_k} + \Delta t \begin{bmatrix} m_{x_k}^{(2)} \\ -g_l \exp(-\frac{1}{2} S_{x_k}^{(1,1)}) \sin(m_{x_k}^{(1)}) - d_J m_{x_k}^{(2)} + J_i u_k \end{bmatrix}, \end{aligned} \quad (2-25)$$

wobei die Zahlen in den hochgestellten Klammern auf das Vektorelement referenzieren. Da es sich um das exakte Moment handelt, wird zudem das hochgestellte MM verwendet. Anhand der Gleichung ist erkennbar, dass die Entwicklung des Erwartungsvektors nicht nur vom vorherigen Erwartungsvektor \mathbf{m}_{x_k} abhängt, sondern auch von der Kovarianzmatrix \mathbf{S}_{x_k} . Insgesamt unterscheidet sich die Vorschrift nicht sonderlich von jener im deterministischen Fall. Im Grunde kommt lediglich der Vorfaktor $\exp(-\frac{1}{2} S_{x_k}^{(1,1)})$ hinzu. Eine Interpretation dieses Faktors lässt die Limes-Betrachtung zu. Für den Spezialfall $S_{x_k}^{(1,1)} \rightarrow 0$, in welchem es keinerlei Unsicherheit über den Zustand des Pendels gibt, strebt der Faktor $\exp(-\frac{1}{2} S_{x_k}^{(1,1)}) \rightarrow 1$, womit letztendlich die bekannte deterministische Euler-Integrationsgleichung zurückerhalten wird (vgl. (2-24)). (2-25) stellt damit eine Erweiterung des herkömmlichen Modellbegriffs um eine probabilistische Betrachtungsweise dar. Diese Ansicht lässt sich somit unmittelbar in die regelungstechnischen Zusammenhänge einbinden, welche in den Hauptkapiteln ausgeführt wird. Das andere Extrem der Limes-Untersuchung ist $S_{x_k}^{(1,1)} \rightarrow \infty$ mit $\exp(-\frac{1}{2} S_{x_k}^{(1,1)}) = 0$. Dies bedeutet, dass sich das Pendel überall im Zustandsraum gleichermaßen häufig aufhält und somit die Multiplikation mit dem Vorfaktor zur Vernachlässigung der Gravitationskraft führt. Anschaulich lässt sich dieser Fall interpretieren, wenn man sich unendlich viele Realisierungen des Zustands (und damit des Pendels) vorstellt. Aufgrund von $S_{x_k}^{(1,1)} \rightarrow \infty$ sind die Pendel über den ganzen Zustandsraum gleichmäßig verteilt, sodass sich zumindest in der Erwartungswertbetrachtung die Gravitationskraft aller Pendel gegenseitig aufheben würde, wodurch der entsprechende Term in (2-25) entfällt. Die Zustandskovarianzmatrix lässt sich über (2-10) und (2-11) folgendermaßen berechnen

$$\begin{aligned} \mathbf{S}_{x_{k+1}}^{MM} &= \mathbb{V}[\mathbf{x}_{k+1}] \\ &= \underbrace{\mathbb{V}[\mathbf{x}_k]}_{=\mathbf{S}_{x_k}} + \Delta t^2 \underbrace{\mathbb{V}[\mathbf{f}(\mathbf{x}_k, \mathbf{u}_k)]}_{(A1-4)} + \Delta t \underbrace{(\mathbb{C}[\mathbf{x}_k, \mathbf{f}(\mathbf{x}_k, \mathbf{u}_k)] + \mathbb{C}[\mathbf{x}_k, \mathbf{f}(\mathbf{x}_k, \mathbf{u}_k)]^T)}_{(A1-5)}. \end{aligned}$$

Die Ausdrücke wurden analytisch gelöst und aufgrund ihrer Länge dem Anhang hinzugefügt. An ihnen wird ersichtlich, dass die Berechnung bereits für das einfache Pendelbeispiel zu aufwendig ist und damit für die Praxis und beliebige mechatronische Systeme

ungeeignet ist. Die beiden nachfolgenden Ansätze bieten im Bezug auf diesen Aspekt eine effizientere Alternative.

Als Nächstes wird die Zustandspropagation mittels Linearisierung [Sim06] erläutert. Die Grundidee ist die bereits bekannte Gleichung (2-11) direkt anwendbar zu machen, wobei die lineare Transformation einer Normalverteilung wiederum auf eine Normalverteilung führt. Für den Folgezustand wird dafür eine Taylorreihen-Entwicklung erster Ordnung beim Erwartungswertvektor des aktuellen Zustands vorgenommen

$$\begin{aligned}\mathbf{x}_{k+1} &\approx \mathbf{m}_{x_k} + \Delta t \mathbf{f}(\mathbf{m}_{x_k}, \mathbf{u}_k) + \left(\mathbf{I}_{n_x} + \Delta t \frac{d\mathbf{f}}{d\mathbf{x}} \Big|_{\mathbf{x}=\mathbf{m}_{x_k}} \right) (\mathbf{x}_k - \mathbf{m}_{x_k}) \\ &= \underbrace{\left(\mathbf{I}_{n_x} + \Delta t \frac{d\mathbf{f}}{d\mathbf{x}} \Big|_{\mathbf{x}=\mathbf{m}_{x_k}} \right)}_{=:A} \mathbf{x}_k + \underbrace{\Delta t \left(\mathbf{f}(\mathbf{m}_{x_k}, \mathbf{u}_k) - \frac{d\mathbf{f}}{d\mathbf{x}} \Big|_{\mathbf{x}=\mathbf{m}_{x_k}} \mathbf{m}_{x_k} \right)}_{=:b}.\end{aligned}$$

Anschließend lassen sich die ersten beiden Momente kostengünstig über (2-11) berechnen:

$$\begin{aligned}\mathbf{m}_{x_{k+1}}^L &= A\mathbf{m}_{x_k} + b = \mathbf{m}_k + \Delta t \mathbf{f}(\mathbf{m}_{x_k}, \mathbf{u}_k), \\ \mathbf{S}_{x_{k+1}}^L &= A\mathbf{S}_{x_k} A^T.\end{aligned}$$

Im direkten Vergleich zur analytischen Berechnung aus (2-25) fällt auf, dass der besprochene Vorfaktor $\exp(-\frac{1}{2} S_{x_k}^{(1,1)})$ beim Ansatz der Linearisierung verloren geht. Die Approximationsgenauigkeit ist damit als schlechter einzustufen, wird jedoch durch die geringe Berechnungskomplexität ausgeglichen. Der Ansatz der Linearisierung ist in der Regelungstechnik nicht unbekannt, sondern erfährt beispielsweise weitverbreitete Anwendung beim Beobachterentwurf mittels des *Extended Kalman Filters* (EKF) [Sim06; Ada14]. Das Beobachtungsgesetz des EKF lässt sich in diesem Zusammenhang auch aus der Bayesschen Regel herleiten [Mur13], womit es eine interessante Querverbindung zwischen der Regelungstechnik, der Wahrscheinlichkeitstheorie und dem maschinellem Lernen gibt.

Das nächste Verfahren lehnt hieran an und bezieht sich auf den artverwandten *Unscented Kalman Filter* (UKF) [Sim06]. Dieser stellt einen Spezialfall des übergeordneten Partikelfilter (PF) dar, welcher sich auf den Überlegungen zur numerischen Approximation der Momente aus Abschnitt 2.2.5 stützt. Die Grundidee ist die Momente, ähnlich zu (2-16), über eine finite Anzahl von sogenannten Partikeln, d. h. wahrscheinlichen Realisierungen der Zufallsvariable, abzuschätzen. Der UKF nutzt vor diesem Hintergrund eine stark begrenzte Anzahl von Partikeln, deren Verortung im Zustandsraum anhand der resultierenden Approximationsgüte bestmöglich gewählt ist. Es handelt sich damit um eine ef-

fiziente Variante des PF. Die Partikel werden in diesem Zusammenhang *Sigma-Punkte* genannt und berechnen sich wie folgt:

$$\begin{aligned}\mathbf{x}_{s,k}^{(i)} &= \mathbf{m}_{x_k} + (\sqrt{(\kappa + n_x) \mathbf{S}_{x_k}})_{(i,:)}^T, \quad i = 1, \dots, n_x, \\ \mathbf{x}_{s,k}^{(i)} &= \mathbf{m}_{x_k} - (\sqrt{(\kappa + n_x) \mathbf{S}_{x_k}})_{(i,:)}^T, \quad i = n_x + 1, \dots, 2n_x, \\ \mathbf{x}_{s,k}^{(2n_x+1)} &= \mathbf{m}_{x_k},\end{aligned}\tag{2-26}$$

wobei der Ausdruck $(\sqrt{(\kappa + n_x) \mathbf{S}_{x_k}})_{(i,:)}^T$ die i -te Zeile der Matrixwurzel darstellt und über die Cholesky-Zerlegung berechnet wird. Die Variable κ ist ein Entwurfsparameter, welcher vom Anwender in Abhängigkeit vom System gewählt wird. Die $2n_x + 1$ Sigma Punkte werden anschließend über die Dynamikgleichung transformiert, bzw. auf den Folgezustand abgebildet

$$\mathbf{x}_{s,k+1}^{(i)} = \mathbf{x}_{s,k}^{(i)} + \Delta t \mathbf{f}(\mathbf{x}_{s,k}^{(i)}, \mathbf{u}_k), \quad i = 1, \dots, 2n_x + 1.\tag{2-27}$$

Danach werden die Momente über

$$\begin{aligned}\mathbf{m}_{x_{k+1}}^{UT} &= \sum_{i=1}^{2n_x+1} W^{(i)} \mathbf{x}_{s,k+1}^{(i)}, \\ \mathbf{S}_{x_{k+1}}^{UT} &= \sum_{i=1}^{2n_x+1} W^{(i)} (\mathbf{x}_{s,k+1}^{(i)} - \mathbf{m}_{x_{k+1}}^{UT}) (\mathbf{x}_{s,k+1}^{(i)} - \mathbf{m}_{x_{k+1}}^{UT})^T,\end{aligned}\tag{2-28}$$

mit Gewichtungsfaktoren $W^{(i)} = \begin{cases} \frac{1}{2(n_x+\kappa)}, & i = 1, \dots, 2n_x, \\ \frac{\kappa}{n_x+\kappa}, & i = 2n_x + 1, \end{cases}$

approximiert. Wobei die Herleitung der Gewichtungsfaktoren auf einer bestmöglichen Approximation der Momente basiert und die begrenzte Anzahl von Partikeln ausgleicht [Sim06]. Durch die symmetrische Anordnung der Sigma-Punkte um den Erwartungsvektor des aktuellen Zustands bzw. den wahrscheinlichsten Zustand werden die Folgemosmente effektiv erfasst, sodass sich insgesamt ein leicht berechenbares Verfahren mit einer hohen Güte ergibt. Das beschriebene Vorgehen wird als *Unscented Transform* (UT) bezeichnet und bietet sich vor allem bei stark nichtlinearen Systemen an, weswegen es in vielen Beobachtern umgesetzt wird. Im direkten Vergleich zum EKF wird beim UKF nicht die zugrundeliegende Dynamikfunktion linearisiert, sondern viel eher die Wahrscheinlichkeitsverteilung des Zustands selbst, wodurch die hohe Approximationsgüte nachvollzogen werden kann [Sim06].

Bild 2-6 zeigt beispielhaft zwei probabilistische Simulationen für das Pendel (2-23) (links: freie, unkontrollierte Schwingung und rechts: Bewegung mit konstant-wirkendem Dreh-

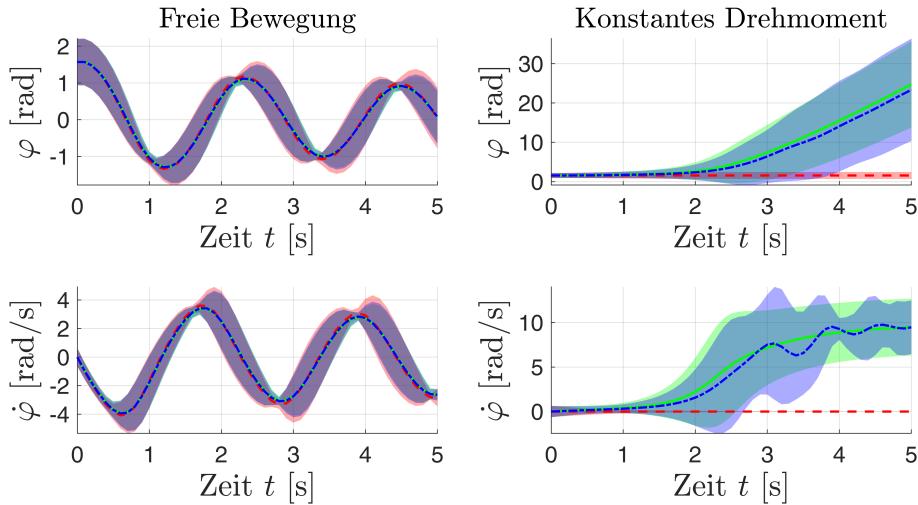


Bild 2-6: Probabilistische Simulation eines im Gelenk aktuierten Pendels. Gegeüberstellung der Verfahren zur Zustandspropagation über mehrere Zeitschritte (Grün: Exaktes Moment Matching, rot: Linearisierung und blau: Unscented Transform).

moment), welche mit den drei vorgestellten Verfahren berechnet wurden. Die Mittellinie ist jeweils der Erwartungswert des Zustands und das umliegende Band deutet die Unsicherheit in Form der zweifachen Standardabweichung, kurz 2σ , an. Die Kovarianz ist in dieser Darstellung nicht enthalten, sondern nur die Marginalverteilungen für die einzelnen Zustände. Im ersten Szenario wird für das Pendel eine Zustandsverteilung vorgegeben, deren Dichte sich größtenteils in der Nähe der oberen Ruhelage befindet. Gut erkennbar ist, wie sich nicht nur die Mittelwerte schwingend verhalten, sondern auch die zugehörige Varianz Zunahmen und Abnahmen erfährt. Die kleinste Unsicherheit tritt in der Regel kurz vor einem Umkehrpunkt des Pendels auf. Durch die dissipative Reibung im System fallen die Amplituden erwartungsgemäß mit der Zeit ab. Tatsächlich strebt die Verteilung für $t \rightarrow \infty$ auf die untere Ruhelage zu, sodass die Varianz komplett verschwindet $S_{x_k} \rightarrow 0$, da sich bildlich gesprochen alle möglichen Pendelrealisierungen aufgrund der Reibung nur in der unteren Ruhelage befinden und alle anderen Zustände komplett ausgeschlossen werden können. Des Weiteren ist gut erkennbar, dass die drei besprochenen Verfahren direkt hintereinander liegen und trotz ihrer verschiedenen Ansätze zum gleichen Ergebnis führen. Im zweiten Szenario wird ein konstantes Drehmoment von $u = 9.81 \text{ Nm}$ angelegt, so dass das Pendel anfängt, sich zu überschlagen. Nach einem kurzweiligen Beschleunigungsvorgang nimmt die Zustandsverteilung eine gleichbleibende Form an, wobei der Erwartungswert des Winkels zwar weiter anwächst; die zugehörige Varianz allerdings nicht weiter zunimmt. Offensichtlich unterscheiden sich die drei Verfahren für dieses Szenario in ihrer Qualität. Da der Moment Matching Ansatz auf einer exakten analytischen Berechnung der Momente basiert, weist er die best-mögliche Qualität auf und kann als Referenz für die beiden anderen Verfahren dienen. Die UT kommt sehr nah

an das Ergebnis des MM, unterliegt jedoch gewissen Schwankungen in Bezug auf die Winkelgeschwindigkeit. Dies kann auf die Nichtlinearität des Pendels zurückzuführen sein und macht sich aufgrund der Integratorstruktur weniger deutlich im Verlauf des Pendelwinkels bemerkbar. Die Linearisierung führt auf das schlechteste Ergebnis. Dies ist darauf zurückzuführen, dass sich das anliegende Drehmoment und die Gewichtskraft des Pendels ausgleichen und die Linearisierung ungefähr auf eine Einheitsmatrix führt, sodass es durch die Zustandspropagation zu einer gleichbleibenden Zustandsverteilung kommt. Bei der Anwendung der Verfahren sind solche Effekte zu berücksichtigen und müssen - je nach System und vorhandenen Nichtlinearitäten - gegeneinander abgewogen werden. Weiterführend wird die Zustandspropagation in Kapitel 4 für die hybride Optimalsteuerung verwendet, um möglicherweise vorhandene Modellfehler beim Entwurf zu berücksichtigen. Als Fazit der hier besprochenen Ergebnisse wird festgehalten, dass die UT der anderen Verfahren vorzuziehen ist, da sie einen guten Kompromiss zwischen Genauigkeit und Berechnungsaufwand darstellt.

2.3.3 Annäherung an exakte Folgezustandsverteilung

Im vorherigen Abschnitt wurde davon ausgegangen, dass die Folgezustandsverteilung als normalverteilt angenähert werden kann. Diese Annahme ist sinnvoll, da im Hinblick auf Langzeitprädiktionen die Klasse der Normalverteilung nie verlassen wird und kostengünstig berechenbar bleibt. Nachteilig ist, dass die Normalverteilung nur ein grobes Bild über die Positionierung des Zustands wiedergibt und beispielsweise die Aussagekraft einer möglicherweise multimodalen Verteilung verloren geht. Für die meisten Anwendungen ist eine genaue Berechnung der Zustandsverteilung nicht zweckmäßig und zu aufwendig. Der Vollständigkeit halber und vor dem Hintergrund besonderer Spezialfälle soll an dieser Stelle eine genauere numerische Approximation für die Folgezustandsverteilung vorgestellt werden. Die exakte Verteilung kann durch den MC Ansatz aus Abschnitt 2.2.5 angenähert werden. Dabei wird, wie beim PF, vorab eine finite Anzahl von Partikeln bestimmt und gleichermaßen zum deterministischen Fall eine Vorwärtsintegration für alle Partikel durchgeführt. Das Ergebnis dieses Ansatzes für das Pendel ist in Bild 2-7 zu sehen. Die Anfangsverteilung startet zu einem großen Teil in der Nähe der oberen instabilen Ruhelage. Von dieser Verteilung werden die Partikel durch zufälliges Sampling berechnet, wobei der Zusammenhang gilt: je mehr Partikel eingesetzt werden desto höher ist die Approximationsgüte. Anschließend werden die Partikel über die Systemdynamik transformiert. Nach 5 Zeitschritten ergibt sich das mittlere Teilbild. Es ist gut zu erkennen, dass die Verteilung stellvertretend für eine unendliche Anzahl von Pendeln steht, die sich aus der oberen Ruhelage entfernen und an Geschwindigkeit aufnehmen. Nach 100 Zeitschritten hat sich die anfängliche Normalverteilung in zwei separate Grenzschwingungen

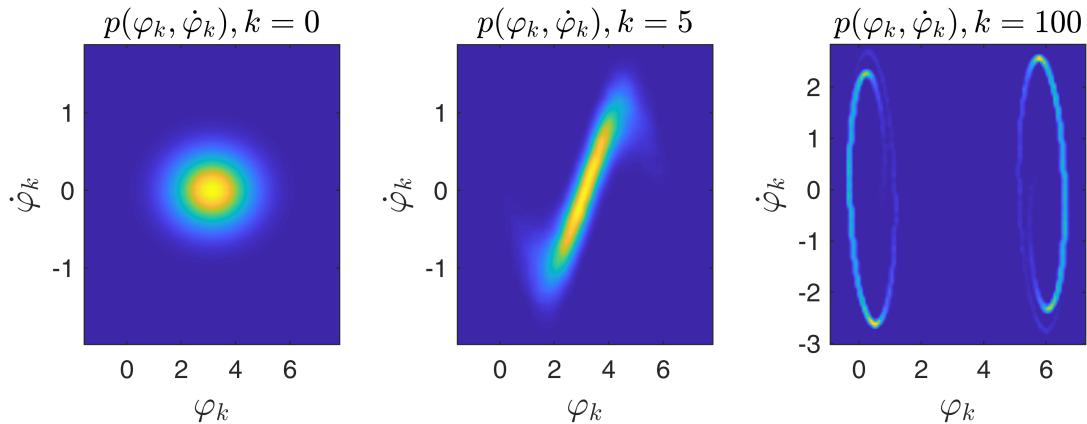


Bild 2-7: Zeitliche Entwicklung der Zustandsverteilung des frei schwingenden Pendels (2-23) (Links: normalverteilte Anfangsverteilung, mittig: Verteilung nach 5 Zeitschritten, rechts: Verteilung nach 100 Zeitschritten).

um die untere Ruhelage (um 2π versetzt) aufgespalten. Diese feine Zustandsverteilung wäre mit dem Moment Matching Ansatz nur grob abgebildet worden und könnte ohne vorhandenes Fachwissen falsch interpretiert werden. Der Erwartungsvektor der Verteilung würde sich beispielsweise in der Mitte der Schwingungen befinden bzw. genau in der oberen Ruhelage liegen. Tatsächlich befindet sich keines der Pendel zu diesem Zeitpunkt mehr in dieser Position. Hierbei wird eine besondere Eigenschaft des Erwartungswerts offensichtlich - er stellt lediglich ein Maß zur Beschreibung einer Verteilung dar und muss zwangsläufig nicht Teil einer Menge von plausiblen Realisierungen sein. Des Weiteren würde sich die Varianz über den gesamten Bereich der Schwingungen erstrecken und keine genaue Verortung der Pendel zulassen. Je nach Anwendungsfall ist in der Praxis also abzuwägen, ob die Approximation mittels einer Normalverteilung ausreichend oder die komplexere MC Simulation notwendig ist.

2.4 Maschinelles Lernen mittels Gauß-Prozess-Regression

Das maschinelle Lernen befasst sich im Allgemeinen mit der Gewinnung von Erkenntnissen und Mustern, die auf der Grundlage von datenbasierten Lernalgorithmen und nicht durch eine explizite menschliche Programmierung erlangt werden. In diesem Rahmen wird auch der Begriff einer KI verwendet. Für die in dieser Arbeit besprochene Aufgabenstellung wird das maschinelle Lernen dazu eingesetzt, um den Regelungstechnischen Entwicklungsprozess unterstützend zu erweitern und zu vereinfachen. Insbesondere geht es dabei darum, aus vorhandenen gemessenen Systemdaten, welche beispielsweise den Zustand betreffen, gewinnbringende Erkenntnisse auszunutzen. Hierfür wurde innerhalb der Einleitung der Begriff der hybriden Modellierung eingeführt, welche einen anpassungsfähigen, flexiblen und datengetriebenen Teil besitzt, der auf der Basis des maschinellen Lernens bestimmt wird. Das ML bildet insgesamt betrachtet ein breites Spektrum

von Aufgaben und Lösungsmöglichkeiten ab. Es eröffnet damit eine Vielzahl von Anknüpfungspunkten zur Regelungstechnik, um den methodischen Werkzeugkasten nützlich zu erweitern und zu verbessern. In Hinblick auf die Inbetriebnahme wird sich im Rahmen dieser Arbeit nur auf einige Anknüpfungspunkte fokussiert, die den höchsten Stellenwert für die Thematik haben und daher intensiv aufgearbeitet werden. Hierzu werden die Grundideen des ML, wie der datenzentrierte Ansatz, die Rückkopplung von Informationen (aus dem RL) und das Konzept der Generalisierung, um neuen unvorhergesehenen Situationen zu begegnen, im Regelungstechnischen Kontext betrachtet. Ein maßgebendes Instrument ist hierfür die bereits eingeführte Wahrscheinlichkeitstheorie, die die Unsicherheit über bestimmte Vorgänge quantifizierbar macht, so unter anderem die Variabilität der Trainingsdaten als auch die Bestimmung von Parametern einer ML-Ansatzfunktion.

Eine beliebte Klasse von Funktionen, die im ML häufig eingesetzt werden, stellen die künstlichen neuronalen Netze dar. Der Grund hierfür liegt in ihrer ausgezeichneten Verallgemeinerungsfähigkeit und ihrem leichten Aufbau, der je nach Anwendungsfall flexibel anpassbar ist. Mittlerweile gibt es daher eine schiere Menge von Netzvariationen, die bildlich gesprochen häufig auch als "Zoo" bezeichnet wird. Für einen Laien ist es oftmals schwer, das richtige Netz für seinen Anwendungsfall auszuwählen und anschließend geeignet zu trainieren. Dies stellt einen Nachteil bei dem Einsatz von künstlichen neuronalen Netzen dar und wird durch die Wahl der passenden Netzgröße weiter verstärkt. Die Größe geht einher mit der Anzahl von Parametern des Netzes. Ist die Anzahl zu gering, kann die Aufgabenstellung womöglich nicht gelöst werden und das Netz neigt zur Unteranpassung, wodurch bestimmte Nuancen nicht abgebildet werden können. Ist die Anzahl auf der anderen Seite zu groß, ist das zugrundeliegende Optimierungsproblem zur Bestimmung der Parameter unter Umständen nur schwer lösbar und es besteht die Gefahr der Überanpassung, womit die Generalisierung verloren geht und ggf. Rauscheffekte unbeabsichtigt nachgebildet werden. Aus diesem Grund wird beim ML zwischen aleatorischer und epistemischer Unsicherheit unterschieden [HW21]. Die aleatorische Unsicherheit wird auch als Zufallsunsicherheit bezeichnet und resultiert aus den natürlichen Schwankungen des betrachteten Prozesses. Diese sind im Allgemeinen unvermeidlich und dem Prozess intrinsisch. Ein Beispiel ist das Messrauschen eines Sensors. Epistemische Unsicherheit ist hingegen auf das begrenzte und unvollständige Wissen über ein System zurückzuführen, welches bspw. durch diverse Modellannahmen verursacht wird. Im Gegensatz zur aleatorischen Unsicherheit kann die epistemische Unsicherheit durch zusätzliches Wissen, bspw. in Form einer größeren Datenmenge, verringert werden. Eine Möglichkeit den angesprochenen Gegebenheiten und Problematiken zu begegnen, bietet die Klasse der nicht-parametrischen ML-Modelle. Hierzu zählt das Verfahren der Gauß-Prozess-Regression, welches in diesem Unterabschnitt detailliert eingeführt wird und die Basis für alle entwickelten Verfahren dieser Arbeit darstellt. Die Besonderheit ist, dass

vom Anwender keine feste Anzahl von Parametern vorgegeben werden muss, sondern das Modell auf natürliche Weise mit der Anzahl der Trainingsdaten verknüpft ist und entsprechend mit ihnen anwächst und somit an Komplexität zunimmt. Damit geht die Einschränkung einher, dass ohne besondere Anpassung, das Verfahren nur für relativ kleine Datenmengen geeignet ist. Vor dem Anwendungsgebiet der Regelungstechnik, in welcher einfache handhabbare Modelle erstrebenswert und sogar vorteilhaft in Hinblick auf bspw. Echtzeitfähigkeit und Interpretierbarkeit sind, stellt dies keinen erheblichen Nachteil dar. Weiterführend handelt es sich um ein probabilistisches Verfahren, welches insbesondere darauf ausgelegt ist, die Unsicherheit über die gesuchte Funktion in Form einer Varianzangabe auszudrücken. Zusammen mit dem bereits eingeführten erweiterten (probabilistischen) Modellbegriff stellt die Gauß-Prozess-Regression ein ideales Hilfsmittel dar, um innerhalb eines hybriden Modells, die a-priori unbekannten Modellfehler zu identifizieren und zu kompensieren. Die Schnittstelle des physikalischen und datenbasierten Modellteils stellt somit also die Wahrscheinlichkeitstheorie dar.

2.4.1 Partitionierung einer multivariaten Gaußverteilung

Im Folgenden wird die Herleitung der Gauß-Prozess-Regression nach einer Kombination von [Ras06; Bis06; Mur13] besprochen. Den Ausgangspunkt stellt eine allgemeine multivariate Normal- bzw. Gaußverteilung nach $\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ mit $\mathbf{x} \in \mathbb{R}^n$ (vgl. Abschnitt 2.2.4) dar. Außerdem wird eine Partitionierung der Form $\mathbf{x} = [\mathbf{x}_1^T, \mathbf{x}_2^T]^T$, $\mathbf{x}_1 \in \mathbb{R}^{n_1}$, $\mathbf{x}_2 \in \mathbb{R}^{n_2}$ mit $n_2 = n - n_1$ und

$$\boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}, \quad \boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{1,1} & \boldsymbol{\Sigma}_{1,2} \\ \boldsymbol{\Sigma}_{2,1} & \boldsymbol{\Sigma}_{2,2} \end{bmatrix},$$

angenommen. Zur besseren Übersicht wird die Präzisionsmatrix

$$\boldsymbol{\Lambda} := \boldsymbol{\Sigma}^{-1} = \begin{bmatrix} \boldsymbol{\Lambda}_{1,1} & \boldsymbol{\Lambda}_{1,2} \\ \boldsymbol{\Lambda}_{2,1} & \boldsymbol{\Lambda}_{2,2} \end{bmatrix}$$

definiert, welche die Elemente der inversen Kovarianzmatrix besitzt, die im Folgenden häufig verwendet werden. Von dieser Situation ausgehend, wird ein Ausdruck für die bedingte Wahrscheinlichkeitsverteilung $p(\mathbf{x}_1 | \mathbf{x}_2)$ gesucht, welche die Grundlage der Gauß-Prozess-Regression darstellt. Hierfür wird die Produktregel aus (2-7) angewandt und führt als Zwischenergebnis auf

$$p(\mathbf{x}_1, \mathbf{x}_2) = p(\mathbf{x}_1 | \mathbf{x}_2)p(\mathbf{x}_2). \tag{2-29}$$

Da es sich bei der linken Seite um eine Normalverteilung handelt und $p(\mathbf{x}_2)$ die Marginalverteilung bzgl. der Variable \mathbf{x}_2 darstellt, liegt die Schlussfolgerung nahe, dass es sich bei $p(\mathbf{x}_1 | \mathbf{x}_2)$ ebenfalls um eine Normalverteilung handeln muss (vgl. (A1-2)). Für die gesuchte bedingte Verteilung wird daher der Ansatz $p(\mathbf{x}_1 | \mathbf{x}_2) = \mathcal{N}(\boldsymbol{\mu}_{1|2}, \boldsymbol{\Sigma}_{1|2})$ getroffen. Durch einen Koeffizientenvergleich der quadratischen Terme innerhalb der Exponentialfunktionen folgt für die Kovarianzmatrix zunächst

$$\mathbf{x}_1^T \boldsymbol{\Lambda}_{1,1} \mathbf{x}_1 \stackrel{!}{=} \mathbf{x}_1^T \boldsymbol{\Sigma}_{1|2}^{-1} \mathbf{x}_1 \quad \Rightarrow \quad \boldsymbol{\Sigma}_{1|2} = \boldsymbol{\Lambda}_{1,1}^{-1}.$$

Weiterführend werden die linearen Terme ausgewertet, womit der Erwartungsvektor zu

$$\begin{aligned} -2\mathbf{x}_1^T (\boldsymbol{\Lambda}_{1,1} \boldsymbol{\mu}_1 - \boldsymbol{\Lambda}_{1,2} (\mathbf{x}_2 - \boldsymbol{\mu}_2)) &\stackrel{!}{=} -2\mathbf{x}_1^T \boldsymbol{\Sigma}_{1|2}^{-1} \boldsymbol{\mu}_{1|2}, \\ \Rightarrow \quad \boldsymbol{\mu}_{1|2} &= \boldsymbol{\mu}_1 - \boldsymbol{\Lambda}_{1,1}^{-1} \boldsymbol{\Lambda}_{1,2} (\mathbf{x}_2 - \boldsymbol{\mu}_2) \end{aligned}$$

bestimmt ist. Es bleibt die Frage nach den Elementen der Präzisionsmatrix $\boldsymbol{\Lambda}$ offen. Hierfür wird das Theorem zur Bestimmung der *Inverse einer partitionierten Matrix* (A1-6) genutzt. Damit folgen für die gesuchten Elemente die Zusammenhänge

$$\begin{aligned} \boldsymbol{\Lambda}_{1,1} &= (\boldsymbol{\Sigma}_{1,1} - \boldsymbol{\Sigma}_{1,2} \boldsymbol{\Sigma}_{2,2}^{-1} \boldsymbol{\Sigma}_{1,2}^T)^{-1}, \\ \boldsymbol{\Lambda}_{1,2} &= -(\boldsymbol{\Sigma}_{1,1} - \boldsymbol{\Sigma}_{1,2} \boldsymbol{\Sigma}_{2,2}^{-1} \boldsymbol{\Sigma}_{1,2}^T)^{-1} \boldsymbol{\Sigma}_{1,2} \boldsymbol{\Sigma}_{2,2}^{-1}, \end{aligned}$$

sodass sich die Elemente der Inversen auf die Elemente der ursprünglichen Kovarianzmatrix zurückführen lassen. Als resultierendes Endergebnis leiten sich die geschlossenen Ausdrücke

$$\begin{aligned} \boldsymbol{\mu}_{1|2} &= \boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{1,2} \boldsymbol{\Sigma}_{2,2}^{-1} (\mathbf{x}_2 - \boldsymbol{\mu}_2), \\ \boldsymbol{\Sigma}_{1|2} &= \boldsymbol{\Sigma}_{1,1} - \boldsymbol{\Sigma}_{1,2} \boldsymbol{\Sigma}_{2,2}^{-1} \boldsymbol{\Sigma}_{1,2}^T, \end{aligned} \tag{2-30}$$

ab. Der Vollständigkeit halber führt dieses Ergebnis außerdem in Bezug auf (2-29) auf die Marginalverteilung $p(\mathbf{x}_2) = \mathcal{N}(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_{2,2})$. (2-30) kann so interpretiert werden, dass die Kenntnis eines Teils eines normalverteilten Zufallsvektors zu einer bedingten Wahrscheinlichkeitsverteilung für den anderen Teil führt, welcher wiederum normalverteilt ist. Bedingungen und auch Marginalisierungen führen somit für den Spezialfall einer Normalverteilung wieder auf die selbe Klasse zurück. Diese Tatsache stellt einen bedeutenden Aspekt der Gauß-Prozess-Regression dar, wobei die Klasse der Normalverteilungen im gesamten Verfahren nie verlassen wird. Eine visuelle Erklärung liefert zudem Bild 2-4, in welchem der Wert für $Y = -1$ als bekannt angenommen wird und sich ausgehend von der zwei- wiederum eine (bedingte) eindimensionale Normalverteilung für $p(x | y = -1)$ ergibt, welche mit Hilfe von (2-30) berechnet wird.

2.4.2 Lösung der Regressionsaufgabe

Bei einer Regressionsaufgabe geht es darum, eine Funktion $f(\mathbf{x}) : \mathbb{R}^{n_x} \mapsto \mathbb{R}$ zu finden, bzw. diese bestmöglich anzunähern [Bis06]. Hierfür liegen Ein- und Ausgangsdaten der Form $\mathbf{x}_i, y_i = f(\mathbf{x}_i) + \epsilon_i, i = 1, \dots, n_d, \epsilon_i \sim \mathcal{N}(0, \sigma_n^2)$ vor, die durch ein gewisses Rauschen verfälscht wurden. Dabei wird in den meisten Fällen vereinfacht angenommen, dass es sich um ein normalverteiltes Rauschen handelt. Die Gesamtheit aller Daten wird verkürzt durch die beiden Matrizen $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_{n_d}]^T$ und $\mathbf{Y} = [y_1, \dots, y_{n_d}]^T$ beschrieben. Eine Lösungsmöglichkeit kann mit Hilfe von Gauß-Prozessen (GPs) [Ras06] erzielt werden. Ein GP ist die Verallgemeinerung der bekannten multivariaten Normalverteilung auf unendlich viele Zufallsvariablen, deren gesamtheitliche Betrachtung auch als Funktion oder parameterabhängige Zufallsvariable interpretiert werden kann. Ein GP ist vollständig und eindeutig durch eine Mittelwertfunktion $m(\mathbf{x}) := \mathbb{E}[f(\mathbf{x})]$ und eine Kovarianzfunktion (auch Kernel genannt) $k(\mathbf{x}, \mathbf{x}') := \mathbb{C}[f(\mathbf{x}), f(\mathbf{x}')]$, die symmetrisch und positiv semi-definit ist, definiert. Formal wird geschrieben

$$f(\mathbf{x}) \sim \mathcal{GP}(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')),$$

womit \mathbf{x} den kontinuierlichen Parameter und f die parameterabhängige (ebenfalls kontinuierliche) Zufallsvariable darstellen. Nach der Definition kann der GP nun an beliebigen Stellen \mathbf{x} ausgewertet werden. Eine Auswertung an den Stellen der Eingangsdaten führt beispielsweise auf die multivariate Normalverteilung

$$\begin{aligned} \mathbf{F} &\sim \mathcal{N}(\boldsymbol{\mu}_F, \boldsymbol{\Sigma}_F), \\ \boldsymbol{\mu}_F &= [m(\mathbf{x}_1), \dots, m(\mathbf{x}_{n_d})]^T, \\ \boldsymbol{\Sigma}_F &= \begin{bmatrix} k(\mathbf{x}_1, \mathbf{x}_1) & \dots & k(\mathbf{x}_1, \mathbf{x}_{n_d}) \\ \vdots & \ddots & \vdots \\ k(\mathbf{x}_{n_d}, \mathbf{x}_1) & \dots & k(\mathbf{x}_{n_d}, \mathbf{x}_{n_d}) \end{bmatrix}, \end{aligned}$$

mit $f_i = f(\mathbf{x}_i)$, $\mathbf{F} = [f_1, \dots, f_{n_d}]^T$. Die Verteilung ist als Marginalverteilung für die beschriebenen Funktionswerte anzusehen, wobei über die restlichen (unendlichen vielen) anderen Funktionswerte ($f(\mathbf{x}) \mid \mathbf{x} \neq \mathbf{x}_i, i = 1, \dots, n_d$) mit Bezug zu (2-7) integriert worden ist [Ras06]. Das Rauschen der Ausgangsdaten kann als übergeordneter Prozess verstanden werden, der einer Integration über die wahren Funktionswerte über

$$\begin{aligned} p(\mathbf{Y}) &= \int p(\mathbf{Y} \mid \mathbf{F}) p(\mathbf{F}) d\mathbf{F} \\ &= \int \mathcal{N}(\mathbf{F}, \sigma_n^2 \mathbf{I}_{n_d}) \mathcal{N}(\boldsymbol{\mu}_F, \boldsymbol{\Sigma}_F) d\mathbf{F} \\ &= \mathcal{N}(\boldsymbol{\mu}_F, \boldsymbol{\Sigma}_F + \sigma_n^2 \mathbf{I}_{n_d}), \end{aligned}$$

mit $\boldsymbol{\mu}_Y = \boldsymbol{\mu}_F$ und $\Sigma_Y = \Sigma_F + \sigma_n^2 \mathbf{I}_{n_d}$, bedarf (s. (A1-3)). Zur Lösung der Regressionsaufgabe wird nun ein fiktiver (frei wählbarer) Funktionswert $\hat{f}(\mathbf{x})$ der GP-Auswertung hinzugefügt. Insgesamt ergibt sich somit die erweiterte multivariate Normalverteilung

$$\begin{bmatrix} \mathbf{Y} \\ \hat{f}(\mathbf{x}) \end{bmatrix} \sim \mathcal{N}(\boldsymbol{\mu}_{\hat{f}}(\mathbf{x}), \Sigma_{\hat{f}}(\mathbf{x})),$$

$$\boldsymbol{\mu}_{\hat{f}}(\mathbf{x}) = [\boldsymbol{\mu}_Y^T, m(\mathbf{x})]^T,$$

$$\Sigma_{\hat{f}}(\mathbf{x}) = \begin{bmatrix} & & k(\mathbf{x}_1, \mathbf{x}) & \\ & \Sigma_Y & \vdots & \\ & & k(\mathbf{x}_{n_d}, \mathbf{x}) & \\ k(\mathbf{x}, \mathbf{x}_1) & \cdots & k(\mathbf{x}, \mathbf{x}_{n_d}) & k(\mathbf{x}, \mathbf{x}) \end{bmatrix}.$$

Ersichtlich ist die partitionierte Struktur, womit die bekannten Gleichungen aus Abschnitt 2.4.1 zur Anwendung gebracht werden dürfen, um die bekannten Ausgangsdaten der gesuchten (verrauschten) Funktion als bekannt vorauszusetzen. Für die bedingte Verteilung ergeben sich somit die zentralen Gleichungen der Gauß-Prozess-Regression zu

$$\begin{aligned} p(\hat{f}(\mathbf{x}) | \mathbf{Y}) &= \mathcal{N}(\mu_{\hat{f}|Y}(\mathbf{x}), \Sigma_{\hat{f}|Y}(\mathbf{x})), \\ \mu_{\hat{f}|Y}(\mathbf{x}) &= m(\mathbf{x}) + \mathbf{k}(\mathbf{x}, \mathbf{X}) \Sigma_Y^{-1} (\mathbf{Y} - \boldsymbol{\mu}_Y), \\ \Sigma_{\hat{f}|Y}(\mathbf{x}) &= k(\mathbf{x}, \mathbf{x}) - \mathbf{k}(\mathbf{x}, \mathbf{X}) \Sigma_Y^{-1} \mathbf{k}(\mathbf{x}, \mathbf{X})^T, \end{aligned} \quad (2-31)$$

worin die verkürzende Definition $\mathbf{k}(\mathbf{x}, \mathbf{X}) := [k(\mathbf{x}, \mathbf{x}_1), \dots, k(\mathbf{x}, \mathbf{x}_{n_d})]$ eingeführt wurde. Diese Gleichungen können alternativ auch über die Bayessche Regel (2-13) hergeleitet werden, wobei sie dann entsprechend den Posterior darstellen. Aus der Mittelwertgleichung wird ersichtlich, dass dieser sich aus dem ursprünglich angenommenen Mittelwert $m(\mathbf{x})$ und einem additiven Korrekturterm $\mathbf{k}(\mathbf{x}, \mathbf{X}) \Sigma_Y^{-1} (\mathbf{Y} - \boldsymbol{\mu}_Y)$ zusammensetzt. Die Korrektur hängt wiederum von der Abweichung zwischen gemessenen Daten und der Annahme ab $(\mathbf{Y} - \boldsymbol{\mu}_Y)$ und wird über die sogenannte Kalman-Verstärkung $\mathbf{k}(\mathbf{x}, \mathbf{X}) \Sigma_Y^{-1}$ (gleichermaßen zum regelungstechnischen Beobachterentwurf [Ada14]) passend skaliert. Die Varianzgleichung ist ebenfalls gut nachvollziehbar. Sie setzt sich aus einer Grundsicherheit $k(\mathbf{x}, \mathbf{x})$ und einem mindernden (negativen) Term zusammen. Aufgrund der positiven Definitheit von Σ_Y , gilt für den hinteren Term $\mathbf{k}(\mathbf{x}, \mathbf{X}) \Sigma_Y^{-1} \mathbf{k}(\mathbf{x}, \mathbf{X})^T \geq 0$, wodurch die Unsicherheit lediglich sinken kann. Tatsächlich kann der Beweis geführt werden, dass mehr Daten die Unsicherheit immer weiter reduzieren und niemals steigern können [Viv98]. Dies ist zu den Überlegungen zur Bayesschen Regel aus Abschnitt 2.3.1 konsistent (vgl. auch Bild 2-5).

Bild 2-8 zeigt die Zusammenhänge bei der Gauß-Prozess-Regression für ein einfaches eindimensionales Beispiel. Auf der linken Seite ist der Fall dargestellt, in welchem noch

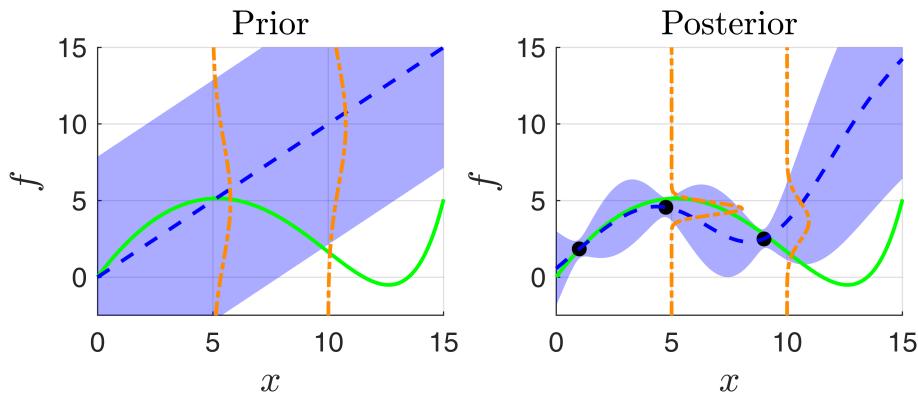


Bild 2-8: Visuelle Darstellung der Gauß-Prozess-Regression für ein eindimensionales Beispiel. Die grüne Linie zeigt die gesuchte Funktion (Ground Truth), welche anhand der schwarzen Datenpunkte durch den GP (blau) nachgebildet werden soll.

keine Daten über die gesuchte Funktion vorliegen. Als Annahmen für den GP wurde eine lineare (steigende) Mittelwertfunktion (blaue, gestrichelte Linie) ausgewählt. Die exakte Kovarianzfunktion ist für das betrachtete Beispiel nicht relevant, bewirkt jedoch ein konstantes Unsicherheitsband um den sich verändernden Mittelwert. Die induzierte Normalverteilung wird beispielhaft an zwei unterschiedlichen Stellen ($x = 5$ und $x = 10$) ausgewertet und ist durch die orangen Linien repräsentiert. Da es sich um Verteilungen bzgl. der Funktion f handelt, sind diese nicht horizontal, sondern vertikal abgebildet. Die hohe Unsicherheit aufgrund des Nichtvorhandenseins von Daten spiegelt sich durch ihre Flachheit (hohe Varianz) wider. Im rechten Teilbild wird nun davon ausgegangen, dass drei Messpunkte (symbolisiert durch die schwarzen Punkte) aufgezeichnet wurden. Die Daten liegen nicht exakt auf der grünen Linie, da ein gewisses Messrauschen bei diesem Beispiel berücksichtigt wurde. Auf Grundlage der Messdaten wird (2-31) ausgewertet und entsprechend die a-priori Verteilung der linken Seite zum Posterior aktualisiert. Der Posterior ist an die Messdaten angepasst und spiegelt diese durch das Durchlaufen dieser mit dem verschobenen Mittelwert wider. Des Weiteren variiert die vorher konstante Varianz nun in Abhängigkeit der Entfernung zu den Messdaten. Je näher die Messdaten, desto niedriger die Varianz bzw. Unsicherheit. Die Unsicherheit nimmt bei den Messpunkten jedoch nicht vollständig ab, denn das Rauschen verbleibt an diesen Stellen. Da die Daten sich größtenteils auf der linken Seite befinden, wird hier die gesuchte zugrundeliegende Funktion bereits sehr gut durch die Posterior-Verteilung wiedergegeben. Die rechte Seite ist weiterhin durch ein hohes Maß an Unsicherheit geprägt und strebt aufgrund der Konstruktion der Gauß-Prozess-Regression im Mittel zurück auf die angenommene a-priori Mittelwertfunktion. Weiterführend sind außerdem die beiden bedingten Wahrscheinlichkeitsverteilungen an den Stellen ($x = 5$ und $x = 10$) explizit gezeigt. Es ist gut erkennbar,

wie sich die ursprüngliche Flachheit verändert und nun durch eine engere Verteilungen repräsentiert ist.

2.4.3 Kovarianzfunktion und Hyperparameter

Der angesetzte GP hängt von einer Mittelwert- $m(\mathbf{x})$ und einer Kovarianzfunktion $k(\mathbf{x}, \mathbf{x}')$ ab. Hierin wird das Vorwissen, welches über die Regressionsaufgabe vorhanden ist, in das Verfahren eingefügt. Die Bedeutung der Mittelwertfunktion ist intuitiv leicht nachvollziehbar. Durch den Bayesschen Charakter des Verfahrens können möglicherweise unpassende Funktionen durch die beobachteten Daten ausgeglichen werden (vgl. Bild 2-8) und stellen somit kein sonderlich großes Problem dar. Ein legitimer (oft verwendeter) Ansatz ist daher für die Mittelfunktion $m(\mathbf{x}) \equiv 0$ den konstanten Wert null anzusetzen, was keine Einschränkung im Bezug auf die Regressionsfähigkeit des GPs darstellt. In den meisten Fällen wird darüber das nicht vorhandene Vorwissen ausgedrückt. Des Weiteren ist eine geeignete Kovarianzfunktion $k(\mathbf{x}, \mathbf{x}')$ für die vollständige Definition des GPs auszuwählen. In diesem Zusammenhang beschreibt die Kovarianzfunktion bzw. der Kernel, die Eigenschaften der gesuchten Funktion in Bezug auf ihre Glattheit und den Grad ihrer Differenzierbarkeit. Außerdem kann über sie eine bestimmte Periodizität ausgedrückt werden [Ras06]. Entsprechend lässt sich über den Kernel erweitertes Vorwissen im Vergleich zur rudimentären Mittelwertfunktion einbeziehen.

Der gängigste Ansatz ist der sogenannte *Squared Exponential* (SE) Kernel, welcher die Form

$$k_{SE}(\mathbf{x}, \mathbf{x}'; \boldsymbol{\theta}) = \sigma_f^2 \exp(-(\mathbf{x} - \mathbf{x}')^T \mathbf{W}^{-1} (\mathbf{x} - \mathbf{x}')), \quad (2-32)$$

mit Gewichtungsmatrix $\mathbf{W} = \text{diag}(l_1^2, \dots, l_{n_x}^2)$, welche als Elemente die sogenannten *Lengthscale*-Parameter besitzt, sowie der Signalvarianz σ_f^2 , hat. Der Kernel ist durch physikalische Prozesse motiviert und bedeutet im Grunde, dass zwei nahe beieinanderliegende Eingangspaare $(\mathbf{x}, \mathbf{x}')$ stark korrelieren, wohingegen weit auseinanderliegende Paare keine Korrelation aufweisen. Dieser Ansatz spiegelt sich in vielen physikalischen Abläufen wider, und besagt, dass kleine Änderungen auch nur zu kleinen Wirkungsänderungen führen. Dementsprechend werden durch die Kovarianzfunktion glatte, nicht sprungfähige Funktionen beschrieben [Ras06]. Die parametrische Abhängigkeit ist gewollt, um dem GP eine gewisse Flexibilität einzuräumen und damit er sich auf unterschiedliche Rahmenbedingungen anpassen kann. Vor diesem Hintergrund werden die sogenannten *Hyperparameter* über $\boldsymbol{\theta} := [l_1, \dots, l_{n_x}, \sigma_f, \sigma_n]^T \in \mathbb{R}^{n_x+2}$ eingeführt, die für die vollständige Definition des GPs notwendig sind. Die eingangs erwähnte nicht-parametrische Eigenschaft der Gauß-Prozess-Regression mag aufgrund der eingeführten

Hyperparameter irreführend erscheinen, hat jedoch weiterhin ihre Berechtigung. Im Vergleich zu einem künstlichen neuronalen Netz, welches je nach Anwendung eine Vielzahl von Parametern benötigt, ist die Anzahl der Hyperparameter stark begrenzt und skaliert nur mit der Dimension der Eingangsdaten und nicht mit der Anzahl von Trainingsdaten bzw. der Komplexität der gesuchten Funktion.

Eine Möglichkeit die Hyperparameter automatisiert zu bestimmen, besteht in der Optimierung anhand einer geeigneten Gütfunktion. Hierfür kann ähnlich zu (2-19) die logarithmische Likelihood maximiert werden

$$\begin{aligned} \boldsymbol{\theta}_* &= \arg \max_{\boldsymbol{\theta}} \log p(\mathbf{Y} \mid \boldsymbol{\theta}), \quad \theta_i > 0, \quad i = 1, \dots, n_x + 2, \\ \log p(\mathbf{Y} \mid \boldsymbol{\theta}) &= -\frac{1}{2}(\mathbf{Y} - \boldsymbol{\mu}_Y)^T \boldsymbol{\Sigma}_Y^{-1}(\mathbf{Y} - \boldsymbol{\mu}_Y) - \frac{1}{2} \log |\boldsymbol{\Sigma}_Y| - \frac{n_d}{2} \log 2\pi, \end{aligned} \quad (2-33)$$

die auf der Basis der gemessenen Daten ausgewertet wird und die Berechnung der Determinante $|\boldsymbol{\Sigma}_Y|$ der Ausgangsdatenmatrix beinhaltet [Ras06]. Die Einführung der log-Funktion ist für die numerische Stabilität notwendig und verändert den nominalen Wert der optimierten Hyperparameter $\boldsymbol{\theta}_*$ nicht. In der Praxis wird die Optimierung mittels eines gradientenbasierten Verfahrens durchgeführt, wofür der Gradient über

$$\frac{\partial \log p(\mathbf{Y} \mid \boldsymbol{\theta})}{\partial \theta_i} = \frac{1}{2} \text{Spur} \left((\boldsymbol{\alpha} \boldsymbol{\alpha}^T - \boldsymbol{\Sigma}_Y^{-1}) \frac{d\boldsymbol{\Sigma}_Y}{d\theta_i} \right), \quad \boldsymbol{\alpha} := \boldsymbol{\Sigma}_Y^{-1}(\mathbf{Y} - \boldsymbol{\mu}_Y) \quad (2-34)$$

ausgewertet wird. Hierin zeigt sich, dass der Gradient auf die Ableitung der Kovarianzmatrix und damit auf die des Kernels zurückzuführen ist. Eine beispielhafte Berechnung für den SE-Kernel ist dem Anhang beigefügt (s. (A1-7)).

2.5 Ergänzende theoretische und praktische Aspekte

Neben den eingeführten grundlegenden Gleichungen der Gauß-Prozess-Regression sollen weiterführend einige ergänzende theoretische und praktische Aspekte vorgestellt werden. Diese sollen sowohl das Verständnis für den Umgang mit GPs vertiefen als auch vor dem Hintergrund der Regelungstechnik fachspezifische Überlegungen aufgreifen und erläutern.

2.5.1 Matérn und Piecewise Squared Exponential Kernel

Die Kovarianzfunktion stellt einen entscheidenden Einflussfaktor bei der Arbeit mit einem GP dar. Neben dem bereits eingeführten SE-Kernel aus (2-32) gibt es eine ganze Reihe weiterer Kernelfunktionen [Ras06], die je nach Anwendungsgebiet ausgewählt werden. Der SE-Kernel repräsentiert besondere Glattheitsanforderungen. Dies wird dadurch ausgedrückt, dass er unendlich mal differenzierbar $k_{SE}(\cdot, \cdot) \in C^\infty$ ist. Für viele physikalische

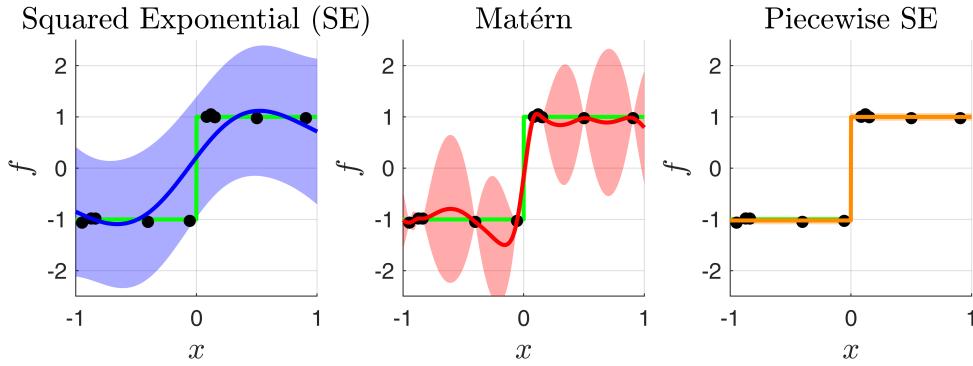


Bild 2-9: Gauß-Prozess-Regression mit verschiedenen Kovarianzfunktionen zur Abbildung einer Sprungfunktion (grün).

Vorgänge, die über einen GP beschrieben werden sollen, ist der SE-Kernel eine geeignete Wahl. Um eine breite Klasse von Funktionen abbilden zu können, hat sich neben dem SE-Kernel, der sogenannte Matérn-Kernel als beliebter Ansatz bewährt. In seiner 32-Variante lautet er folgendermaßen:

$$\begin{aligned} k_M(\mathbf{x}, \mathbf{x}'; \boldsymbol{\theta}) &= \sigma_f^2 (1 + \sqrt{3d}) \exp(-\sqrt{3d}), \\ d(\mathbf{x}, \mathbf{x}'; \boldsymbol{\theta}) &:= (\mathbf{x} - \mathbf{x}')^T \mathbf{W}^{-1} (\mathbf{x} - \mathbf{x}'), \end{aligned} \quad (2-35)$$

und stellt mit $k_M(\cdot, \cdot) \in C^2$ weniger restriktive Glattheitsanforderungen dar. Aufgrund dieser Tatsache wird im Vergleich zum SE-Kernel allerdings auch eine größere Menge an Daten benötigt, um die Unsicherheit über die Funktion gleichermaßen zu reduzieren. Für die Praxis ebenfalls relevant ist eine Spezialvariante des SE-Kernel - der sogenannte *Piecewise-SE*-Kernel [Bij18]. Für die Anwendung dieses Kernels ist es notwendig, den Eingangsraum \mathbb{R}^{n_x} in disjunkte Mengen $D_i, i = 1, \dots, n_d$ zu unterteilen, z. B. über

$$\mathbf{x} \in \bigcup D_i, \quad \text{mit} \quad D_i \cap D_j = \emptyset, j = 1, \dots, n_d, i \neq j.$$

Den Mengen entsprechend lässt sich anschließend der Kernel stückweise über den Zusammenhang

$$k_{PSE}(\mathbf{x}, \mathbf{x}'; \boldsymbol{\theta}) = \begin{cases} \sigma_f^2 \exp(-(\mathbf{x} - \mathbf{x}')^T \mathbf{W}^{-1} (\mathbf{x} - \mathbf{x}')), & \mathbf{x}, \mathbf{x}' \in D_i \\ 0, & \text{sonst,} \end{cases}$$

beschreiben, wobei die beiden Eingangspaare aus der gleichen Teilmenge stammen müssen, um einen Effekt auf die Korrelation zu haben. Der Vorteil dieses Ansatzes wird bei der Betrachtung des nachfolgenden Beispiels deutlich.

Bild 2-9 zeigt die Gauß-Prozess-Regression für den Fall einer unbekannten Sprungfunktion (Ground Truth). In der Regelungstechnik sind Sprungfunktionen häufig im Zusam-

menhang mit Systemen mit Haftreibung anzutreffen, deren Verhalten sich abrupt bei einem Vorzeichenwechsel der Geschwindigkeit ändert. Der grüne Verlauf zeigt die gesuchte Sprungfunktion. Die schwarzen Punkte stellen die gemessenen Daten dar. Die Gauß-Prozess-Regression wurde mit drei verschiedenen Kerneln durchgeführt. Das linke Teilbild zeigt das Ergebnis für den SE-Kernel, wobei gut zu erkennen ist, dass der Sprung aufgrund der hohen Glattheitsanforderungen nicht zufriedenstellend wiedergegeben werden kann. Der zugrundeliegende GP weist ein zu hohes Rauschen auf, sodass der Verlauf insgesamt eine hohe Varianz hat. Der Sprung wird bei diesem Kernel fälschlicherweise als Messrauschen interpretiert. Die Anwendung des Matérn-Kernels aus dem mittleren Teilbild kann zumindest an den Datenpunkten eine niedrigere Unsicherheit kennzeichnen. An den anderen Stellen ist die Varianz allerdings weiterhin zu groß, was unter anderem auf ungeeignete Hyperparameter zurückzuführen ist. Das beste Ergebnis wird im rechten Teilbild erzielt, in dem der PSE-Kernel mit den disjunkten Mengen $D_1 = (-\infty, 0)$ und $D_2 = [0, \infty)$ verwendet wird. In diesem Fall wird die gesuchte Funktion mit äußerst geringer Unsicherheit wiedergegeben und das Messrauschen wird richtig identifiziert. In Bezug auf haftende Systeme ist der PSE-Kernel somit sehr hilfreich und einfach anzuwenden, da die Sprungstelle a-priori bereits bekannt ist und nicht aufwendig bestimmt werden muss.

2.5.2 Interpretation als radiales Basisfunktionen-Netz

Die Gauß-Prozess-Regression gehört zu der Klasse der nicht-parametrischen Verfahren, da sie, mit Ausnahme der wenigen Hyperparameter, keine aufwendige Parameteridentifikation benötigt, sondern mit der Anzahl der Trainingsdaten auf natürliche Weise mitwächst (vgl. (2-31)). Damit hat sie zunächst einen bedeutenden Vorteil bspw. gegenüber parametrischen künstlichen neuronalen Netzen, da diese bei einer vergleichbaren Größe deutlich mehr zu bestimmende Parameter besitzen. Des Weiteren können bspw. durch den SE-Kernel beliebig oft differenzierbare Funktionen beschrieben werden, was bei einem künstlichen neuronalen Netz mit einer begrenzten Größe nicht der Fall ist. Um diese Tatsache besser einsehen zu können, besteht die Möglichkeit den GP in ein radiales Basisfunktionen-Netz (RBF) umzuformen, wodurch eine bessere Vergleichbarkeit zwischen einem nicht-parametrischen und parametrischen Ansatz entsteht.

Zunächst wird die Mittelwertfunktion des Posterior (2-31) wie folgt betrachtet:

$$\begin{aligned} \mu_{\hat{f}|Y}(\mathbf{x}) &= m(\mathbf{x}) + \mathbf{k}_{SE}(\mathbf{x}, \mathbf{X}) \underbrace{\Sigma_Y^{-1}(\mathbf{Y} - \boldsymbol{\mu}_Y)}_{=: \boldsymbol{\alpha}} \\ &= m(\mathbf{x}) + \sum_{i=1}^{n_d} \alpha_i k_{SE}(\|\mathbf{x} - \mathbf{X}_i\|_{W^{-1}}^2). \end{aligned} \tag{2-36}$$

Diese wird als additive Kombination des Prior Mittelwerts und einer gewichteten Summe umgeschrieben. Bei der gewichteten Summe zeigt sich die Struktur eines RBF, durch die konstanten Gewichtungsfaktoren α_i und die durch den SE-Kernel induzierten Basisfunktionen. Es wird deutlich, dass sich die Zentren des RBF bei einem jeden Datenpunkt befinden und sich das Netz mit jedem weiteren Datenpunkt um ein zusätzliches Zentrum automatisch vergrößert [Ras06].

Um die Beschreibungsfähigkeit des GPs zu bestimmen, wird dieser in seine parametrische Form überführt. Hierzu wird die Mittelwertfunktion durch den gängigen parametrischen Ansatz $y(\mathbf{x}) = \phi(\mathbf{x})^T \mathbf{w}$ (vgl. Ausführungen zur Bayesschen linearen Regression in Abschnitt 2.3.1) ausgedrückt, indem der Verbindungsansatz $\mathbf{w} = \Phi^T \boldsymbol{\alpha}$ eingeführt wird. Aus der Rückrechnung folgt der bekannte Zusammenhang aus (2-36):

$$\begin{aligned} y(\mathbf{x}) &= m(\mathbf{x}) + \phi(\mathbf{x})^T \mathbf{w} \\ &= m(\mathbf{x}) + \phi(\mathbf{x})^T \Phi^T \boldsymbol{\alpha} \\ &= m(\mathbf{x}) + \underbrace{\phi(\mathbf{x})^T [\phi(\mathbf{X}_1), \dots, \phi(\mathbf{X}_{n_d})]}_{=:k(\mathbf{x}, \mathbf{X})} \Sigma_Y^{-1} (\mathbf{Y} - \boldsymbol{\mu}_Y) \\ &= m(\mathbf{x}) + k(\mathbf{x}, \mathbf{X}) \Sigma_Y^{-1} (\mathbf{Y} - \boldsymbol{\mu}_Y), \end{aligned} \tag{2-37}$$

worin die Basisfunktion $\phi(\mathbf{x})$ noch als unbekannt gilt. Zunächst ist jedoch eine Anmerkung bezüglich der Berechnungskomplexität der beiden Darstellungen anzuführen. Das Training des GPs ist auf die Bestimmung der Hyperparameter aufgrund der Maximierung der Likelihoodfunktion zurückzuführen. Dafür muss in (2-33) eine $n_d \times n_d$ große Kovarianzmatrix Σ_Y invertiert werden. Für den parametrischen Ansatz und die Bestimmung der Gewichte kann (2-20) als Lösung herangezogen werden, woraus ersichtlich wird, dass für die Lösung eine Matrix mit Dimension $n_w \times n_w$ invertiert werden muss. Je nach Anwendungsfall und zur Verfügung stehender Rechenkapazität kann daher eine passende Auswahl zwischen parametrischen und nicht-parametrischen Ansätzen getroffen werden. In Bezug auf die effiziente Inbetriebnahme und die formulierten Anforderungen soll im Rahmen dieser Arbeit von einer geringen Datenmenge und einer hohen Beschreibungsfähigkeit zur Abbildung einer breiten Klasse von unbekannten Funktionen ausgegangen werden. Für die Themenstellung dieser Arbeit gilt somit $n_w \gg n_d$, womit die Gauß-Prozess-Regression vor dem Hintergrund der Berechnungskomplexität ein geeignetes ML-Werkzeug für die Inbetriebnahme darstellt.

Als Nächstes wird die in (2-37) beinhaltete Basisfunktion $\phi(\mathbf{x})$ näher untersucht. Es liegt nahe, hierfür die Kovarianzfunktion als inneres Produkt [Ras06] über den Ansatz

$$k(\mathbf{x}, \mathbf{x}') = \phi(\mathbf{x})^T \phi(\mathbf{x}'), \tag{2-38}$$

zu definieren, wodurch die beiden Eingangspaare funktional voneinander getrennt werden. Insbesondere der SE-Kernel weist in diesem Zusammenhang eine interessante Eigenschaft auf, auf die im Speziellen weiter eingegangen werden soll. Hierfür wird vereinfachend $\sigma_f^2 = 1$ und $\mathbf{W}^{-1} = \mathbf{I}_{n_x}$ gesetzt. Durch die Umformung der beinhaltenden Exponentialfunktion als unendliche Summe folgt

$$\begin{aligned} k_{SE}(\mathbf{x}, \mathbf{x}') &= \exp\left(-\frac{1}{2}(\mathbf{x} - \mathbf{x}')^T(\mathbf{x} - \mathbf{x}')\right) \\ &= \exp\left(-\frac{\mathbf{x}^T \mathbf{x}}{2}\right) \exp(\mathbf{x}^T \mathbf{x}') \exp\left(-\frac{\mathbf{x}'^T \mathbf{x}'}{2}\right) \\ &= \exp\left(-\frac{\mathbf{x}^T \mathbf{x}}{2}\right) \sum_{n=0}^{\infty} \frac{(\mathbf{x}^T \mathbf{x}')^n}{n!} \exp\left(-\frac{\mathbf{x}'^T \mathbf{x}'}{2}\right). \end{aligned} \quad (2-39)$$

Die äußeren Faktoren weisen somit bereits die gewünschte Struktur von (2-38) auf. Für die mittlere Summe gilt, dass sie einen weiteren bekannten Kernel darstellt. Es handelt sich dabei um die monomiale Kovarianzfunktion [Ras06], welche mit

$$k_{Mn}(\mathbf{x}, \mathbf{x}') = (\mathbf{x}^T \mathbf{x}')^n = \boldsymbol{\phi}_{Mn}(\mathbf{x})^T \boldsymbol{\phi}_{Mn}(\mathbf{x}')$$

angegeben wird. Der Exponent $n \in \mathbb{N}_0$ gibt hierbei die Ordnung an. Für $n = 2$ und $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^2$ folgt beispielsweise

$$k_{M2}(\mathbf{x}, \mathbf{x}') = (\mathbf{x}^T \mathbf{x}')^2 = [x_1^2, \sqrt{2}x_1x_2, x_2^2] [x_1'^2, \sqrt{2}x_1'x_2', x_2'^2]^T = \boldsymbol{\phi}_{M2}(\mathbf{x})^T \boldsymbol{\phi}_{M2}(\mathbf{x}').$$

Die unendliche Summe aus (2-39) lässt sich somit weiter umformulieren zu

$$\begin{aligned} k_{SE}(\mathbf{x}, \mathbf{x}') &= \exp\left(-\frac{\mathbf{x}^T \mathbf{x}}{2}\right) \left(\sum_{n=0}^{\infty} \frac{1}{\sqrt{n!}} \boldsymbol{\phi}_{Mn}(\mathbf{x})^T \boldsymbol{\phi}_{Mn}(\mathbf{x}') \frac{1}{\sqrt{n!}} \right) \exp\left(-\frac{\mathbf{x}'^T \mathbf{x}'}{2}\right) \\ &= \boldsymbol{\phi}_{SE}(\mathbf{x})^T \boldsymbol{\phi}_{SE}(\mathbf{x}'), \\ \boldsymbol{\phi}_{SE}(\mathbf{x}) &:= \exp\left(-\frac{\mathbf{x}^T \mathbf{x}}{2}\right) [\boldsymbol{\phi}_{M0}(\mathbf{x})^T, \dots, \frac{1}{\sqrt{\infty!}} \boldsymbol{\phi}_{M\infty}(\mathbf{x})^T]^T, \end{aligned}$$

womit der SE-Kernel in der gesuchten Form eines inneren Produkts dargestellt ist. Anhang der letzten Gleichung wird ersichtlich, dass der SE-Kernel unendlich viele Basisfunktionen impliziert und damit in Bezug auf ein parametrisches Modell eine Identifikation von unendlich vielen Parametern benötigen würde. Diese Eigenschaft spiegelt sich bei der Anwendung des GP so wider, indem er bezüglich seiner Beschreibungsfähigkeit nicht begrenzt ist und damit beliebige Funktionen aus C^∞ (in Abhängigkeit der Datenlage) genau beschreiben kann. In dieser Hinsicht ist er damit vergleichbaren parametrischen Ansätzen überlegen und für die Anwendungen dieser Arbeit besonders geeignet.

2.5.3 Details zur Anwendung und Implementierung

Abschließend sollen noch einige weitere Details angeführt werden, welche die grundlegenden Prinzipien der Gauß-Prozess-Regression umfassender erläutern. Dabei geht es sowohl um die praktische Anwendung als auch um einige ergänzende Einzelheiten zur programmiertechnischen Implementierung des Verfahrens. Diese Aspekte sind relevant, da sie die Zusammenhänge und ggf. Problematiken in den Hauptkapiteln verständlicher machen, womit die gefundenen Lösungsvorschläge nachvollziehbarer werden.

Über- und Unteranpassung

Der erste Aspekt umfasst das allumfassende maschinelle Lernproblem der Über- und Unteranpassung des ausgewählten Modellansatzes. Dieses Problem wird häufig mit dem Begriff *Ockhams Rasiermesser* beschrieben und besagt, dass einfache gegenüber komplexeren Theorien oder Modellen vorzuziehen sind, sofern sie denselben Sachverhalt gleichermaßen gut beschreiben. Die Einfachheit bezieht sich hierbei auf die Anzahl von Variablen und möglichen Modellhypthesen, die zusammen gesehen in einer klaren und logischen Beziehung zueinander stehen. Für das maschinelle Lernen bedeutet dies, dass die Komplexität der Ansatzfunktion geeignet zur Problemstellung gewählt werden muss und keine Über- bzw. Unteranpassung an die Trainingsdaten aufweisen darf. In Bezug zur Gauß-Prozess-Regression wird dieser maßgeblich durch die Wahl der Hyperparameter bestimmt. Diese werden wiederum anhand der Likelihoodfunktion aus (2-33) identifiziert. Unter der Vernachlässigung der Prior Mittelwertfunktion $m(\mathbf{x}) = 0$, bzw. des Ausgangsdatenvektors $\boldsymbol{\mu}_Y = 0$ lautet das zugrundeliegende Optimierungsproblem

$$\min_{\boldsymbol{\theta}} \log p(\mathbf{Y} | \boldsymbol{\theta}) \propto \underbrace{\mathbf{Y}^T \boldsymbol{\Sigma}_Y^{-1} \mathbf{Y}}_{\text{Datenanpassung}} + \underbrace{\log |\boldsymbol{\Sigma}_Y|}_{\text{Modellkomplexität}} + \underbrace{n_d \log 2\pi}_{\text{Konstante}}, \quad (2-40)$$

worin bereits auf natürliche Weise eine Lösung für das Problem der Über- und Unteranpassung in Form einer Mehrzieloptimierung auftaucht. Der erste Term steht dabei für die Anpassung an die Datenlage, wohingegen der zweite Term für die Modellkomplexität steht. Ein zu stark angepasster GP kann zwar die Daten gut nachbilden, weist ggf. aber auch eine zu hohe Komplexität auf und versagt daher bei der Generalisierung in Bezug auf unvorhergesehene Eingangsdaten. Entgegengesetzt ist eine zu schwache Datenanpassung (bspw. in Form einer linearen Ausgleichsgerade) nicht aussagekräftig genug. Hieraus folgt die Forderung nach einem Kompromiss zwischen diesen beiden Zielen.

Nachfolgend wird der Fokus auf eine eindimensionale SE-Kovarianzfunktion mit $k_{SE}(x, x'; l) = \exp(-(x - x')^2/l^2)$, bzw. $\theta = l$, gelegt, um die Wirkungsweise der beiden Terme zur Datenanpassung und Modellkomplexität besser einordnen zu können. Für $l \rightarrow 0$ strebt $\boldsymbol{\Sigma}_Y \rightarrow \mathbf{I}_{n_d}$ und der Datenterm nimmt sein Minimum $\mathbf{Y}^T \mathbf{Y}$ an, gleichzeitig nimmt $\log |\boldsymbol{\Sigma}_Y| = \log \prod \lambda_i = \log 1 = 0$ jedoch ein Maximum an, wobei $\lambda_i, i = 1, \dots, n_d$

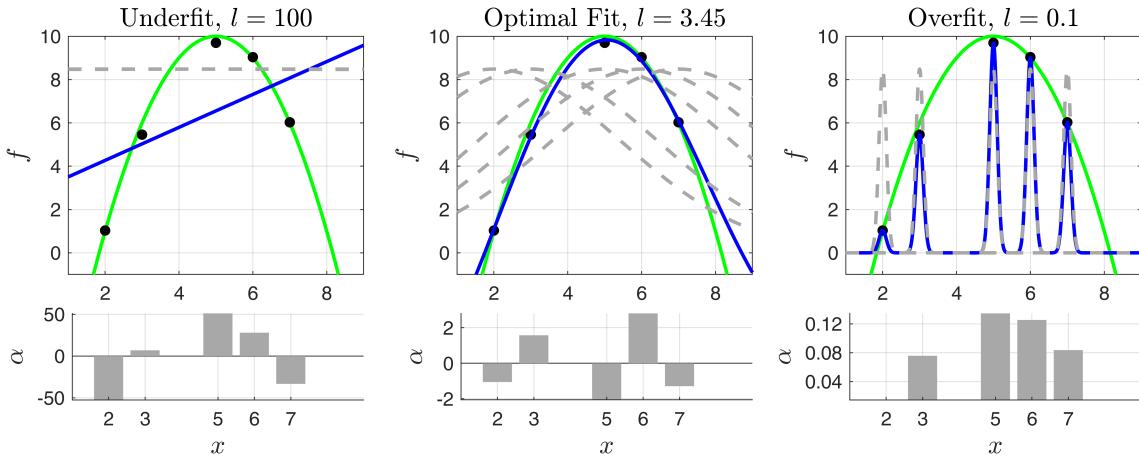


Bild 2-10: Auswirkung des Lengthscale-Hyperparameters inkl. der Darstellung der grundeliegenden Basisfunktionen (gestrichelt grau) und Gewichtungsfaktoren im unteren Balkendiagramm.

die Eigenwerte der Datenmatrix sind. Hieraus folgt eine Überanpassung, da die Funktionswerte weitestgehend unkorreliert sind und sich insgesamt eine flexible, zu stark angepasste Regressionsfunktion ergibt. Für den entgegengesetzten Fall von $l \rightarrow \infty$ strebt $\Sigma_Y \rightarrow 1_{n_d}$, wobei genau ein Eigenwert $\lambda_1 = n_d$ und die übrigen Eigenwerte gleich 0 sind. Dementsprechend strebt $\mathbf{Y}^T \Sigma_Y^{-1} \mathbf{Y}$ gegen ∞ (das Maximum des Terms) und $\log |\Sigma_Y| = \log 0 = -\infty$ stellt wiederum ein Minimum dar. Es resultiert die Unteranpassung, denn die Funktionswerte sind stark korreliert (zusammenhängend) und es folgt eine nahezu lineare Funktion als Lösung der Regressionsaufgabe. Die beschriebene Wirkungsweise ist in Bild 2-10 anhand eines eindimensionalen Beispiels visualisiert. Die grüne Linie zeigt den Ground Truth, von welchem die 5 schwarzen Datenpunkte als Grundlage für die Regression genutzt wurden. Vereinfachend wird von den GPs nur die Posterior Mittelwertfunktion durch die blauen Linien gezeigt. Des Weiteren zeigen die grauen gestrichelten Linien die Basisfunktionen (normiert auf σ_f) aus der RBF-Darstellung. Entsprechend zeigen die unteren Diagramme die dazugehörigen Gewichtungsfaktoren $\alpha_i, i = 1, \dots, 5$ (vgl. (2-36)). Der Lengthscale-Parameter nimmt von links nach rechts ab, sodass der Effekt dieses Parameters gut erkennbar ist. Das linke Bild zeigt die Unteranpassung (durch einen nahezu linearen Funktionsverlauf) und das rechte Bild die Überanpassung, wobei die Basisfunktionen eine impulsförmige Gestalt annehmen. Im mittleren Bild ist das Ergebnis der Optimierung mit einem passenden Hyperparameter zu sehen. Dabei wird ein guter Kompromiss zwischen den beiden Extremfällen gefunden und der Ground Truth durch den GP genau nachgebildet.

Für den Fall einer höherdimensionalen Regression bieten die dimensionsabhängigen Lengthscale-Parameter (s. (2-32)) in diesem Zusammenhang eine praktische Möglichkeit, die Sensitivität bzgl. der einzelnen Eingangsdimensionen auszuwerten. Dies ist unter dem

Begriff *Automatic Relevance Determination* (ARD) in der Literatur [Ras06] bekannt und verweist auf die Größenordnung der optimierten Hyperparameter, welche eine nichtlineare Sensitivitätsanalyse ermöglichen. Je höher der bestimmte Lengthscale-Parameter, desto irrelevanter ist die zugehörige Eingangsdimension für die Regression. Niedrige Werte weisen hingegen auf eine starke Abhängigkeit und hohe Relevanz hin. Für die Praxis bedeutet dies, dass die Eingangsdimension sukzessive verringert werden kann, sobald die Hyperparameter einen Anlass dazu geben, wodurch die Gauß-Prozess-Regression insgesamt an Effektivität dazugewinnt. Diesen Unterabschnitt abschließend, sollen neben der Lengthscale auch die zwei weiteren Hyperparameter kurz erläutert werden. Die Varianz der Ausgangsgröße $\mathbb{V}[y(\mathbf{x})] = k(\mathbf{x}, \mathbf{x}) = \sigma_f^2 + \sigma_n^2$ setzt sich aus der Summe der Signalvarianz und der Varianz des Rauschprozesses zusammen, womit der GP als ganzes durch zwei Unsicherheitsquellen (aleatorische und epistemische Unsicherheit [HW21]) beeinflusst wird. Im Allgemeinen stellt sich daher durch die Bestimmung der Hyperparameter ein Kompromiss zwischen diesen beiden Quellen ein, der die Rauschvarianz nicht überbewertet und die Signalvarianz korrekt abschätzt. Zur Orientierung kann hierbei festgehalten werden, dass 95% der Funktionswerte sich in einem $\pm 2\sigma_f$ -Band um den Mittelwert herum befinden.

Unbeschränkte Optimierung der Hyperparameter

Der nächste Punkt betrifft das Optimierungsproblem aus (2-33), wobei es sich um eine Optimierung unter Nebenbedingungen handelt. Für die Hyperparameter $\theta_i > 0$ mit $i = 1, \dots, n_x + 2$ gilt, dass sie keine negativen Werte annehmen dürfen, um konsistent zu sein. Um die Optimierung zu vereinfachen und die Beschränkungen zu umgehen, wird in der Praxis eine Transformation der Hyperparameter angestrebt. Genauer lautet die Transformationsvorschrift $\tilde{\boldsymbol{\theta}} := \log \boldsymbol{\theta}$, wobei der monotone Anstieg und der positive Definitionsbereich der log-Funktion zum Tragen kommt. Das Optimierungsproblem ändert sich dadurch folgendermaßen:

$$\log \boldsymbol{\theta}_* = \arg \max_{\tilde{\boldsymbol{\theta}}} \log p(\mathbf{Y} \mid \exp(\tilde{\boldsymbol{\theta}})),$$

wobei der Gradient aus (2-34) durch den Zusammenhang

$$\frac{\partial \log p(\mathbf{Y} \mid \exp(\tilde{\boldsymbol{\theta}}))}{\partial \log \theta_i} = \frac{\partial \log p(\mathbf{Y} \mid \boldsymbol{\theta})}{\partial \theta_i} \theta_i,$$

angepasst werden muss. Auf diesem Weg können die einfacheren unbeschränkten Optimierungsverfahren im Rahmen der Hyperparameterbestimmung genutzt werden.

Numerische Stabilität

Die numerische Stabilität spielt eine wichtige Rolle bei der Realisierung eines Verfahrens innerhalb eines Computerprogramms. Die Gauß-Prozess-Regression weist vor die-

sem Hintergrund einige Besonderheiten auf. Maßgeblich für das Verfahren ist die Inversion der Datenmatrix Σ_Y , die sich unter bestimmten Voraussetzungen als problematisch herausstellen kann. Als Maß für die numerische Instabilität wird in diesem Zusammenhang die Konditionszahl $\kappa = \lambda_{\max}/\lambda_{\min}$ [Han09] herangezogen, welche sich aus dem größten und kleinsten Eigenwert der betrachteten Matrix zusammensetzt. Optimalerweise liegt die Konditionsmatrix bei 1. Eine unzureichende Stabilität ist für den Fall $\kappa \gg 1$ gegeben. Vor diesem Hintergrund ist zunächst im Zusammenhang mit der Gauß-Prozess-Regression eine auf die Signalvarianz normierte Schreibweise der inversen Datenmatrix über

$$\Sigma_Y^{-1} = (\Sigma_F + \sigma_n^2 \mathbf{I}_{n_d})^{-1} = \frac{1}{\sigma_n^2} \left(\frac{\sigma_f^2}{\sigma_n^2} \tilde{\Sigma}_F + \mathbf{I}_{n_d} \right)^{-1},$$

sinnvoll. Von dieser Form ausgehend, bietet der Satz von Gershgorin [Han09] die Möglichkeit, den Bereich, in welchem sich die Eigenwerte befinden, zumindest ansatzweise abzuschätzen. Durch die Anwendung des Satzes folgt für den Wertebereich der Eigenwerte $\lambda \in [1, \sigma_r^2 n_d + 1]$ mit $\sigma_r := \frac{\sigma_f}{\sigma_n}$ als Signal-Rausch-Verhältnis. Im schlechtesten Fall bedeutet dies einen kleinsten Eigenwert von $\lambda_{\min} = 1$, sowie einen größten Eigenwert von $\lambda_{\max} = \sigma_r^2 n_d + 1$. Hieraus folgt wiederum eine Konditionszahl von

$$\kappa = \frac{\lambda_{\max}}{\lambda_{\min}} = \sigma_r^2 n_d + 1,$$

die eine lineare Abhängigkeit bzgl. der Anzahl der Daten und eine quadratisch Abhängigkeit bzgl. des Signal-Rausch-Verhältnisses aufweist. Unter der Voraussetzung einer konstanten Signalvarianz ist somit ein gewisses Rauschen innerhalb der Daten hilfreich und trägt zur numerischen Stabilität des Verfahrens bei. Aus diesem Grund wird in der Praxis die Unsicherheit künstlich durch einen geringen additiven Term erhöht

$$\hat{\Sigma}_Y^{-1} = (\Sigma_F + (\sigma_n^2 + \sigma_\epsilon^2) \mathbf{I}_{n_d})^{-1}, \quad \sigma_\epsilon \approx \frac{\sigma_n}{100},$$

sodass die numerische Stabilität gefördert wird. Des Weiteren zeigt die Gleichung der Konditionszahl, dass sich eine geringe Anzahl von Daten positiv auswirken kann. Dies geht mit der allgemeinen Betrachtungsweise einher, dass sich die Gauß-Prozess-Regression besonders für kleine Datenmengen auszeichnet und für diese ein geeignetes maschinelles Lernverfahren darstellt. Außerdem kann festgestellt werden, dass $\tilde{\Sigma}_F \rightarrow \mathbf{I}_{n_d}$ und $\lambda_{\max}/\lambda_{\min} \rightarrow 1$ gilt, wenn die Eingangsdaten der Matrix \mathbf{X} weit auseinander liegen und/oder der Lengthscale-Parameter besonders klein ist. Diese Ausgangssituation ist somit besonders wünschenswert. Auf der anderen Seite kann $\tilde{\Sigma}_F \rightarrow \mathbf{1}_{n_d}$ zu $\lambda_{\max}/\lambda_{\min} \rightarrow \infty$ führen, wenn die Eingangsdaten nah beieinander liegen und/oder der Lengthscalepara-

meter besonders groß ist. Dieser Fall sollte nach Möglichkeit verhindert werden, um die Leistungsfähigkeit des Verfahrens nicht zu beeinträchtigen. Praktischerweise ist eine Anpassung des Optimierungsproblems (2-40) hierzu hilfreich, wobei Strafterme eingeführt werden, die bspw. zu hohen Lengthscale-Parameter oder einem zu hohen Signal-Rausch-Verhältnis entgegenwirken, jedoch gleichzeitig das Mehrzieloptimierungsproblem schwerer lösbar machen.

3 Steuerungsentwurf mittels Bayesscher Optimierung

Der Entwurf einer Steuerung für ein dynamisches System spielt eine zentrale Rolle innerhalb der Regelungstechnik. Die Anwendungsgebiete umfassen die Automatisierungs-, sowie die Luft-, Raum- und Fahrzeugtechnik. Des Weiteren werden Steuerungen im industriellen Umfeld und insbesondere in der Verfahrenstechnik eingesetzt. Das übergeordnete Ziel einer Steuerung ist, das Systemverhalten über die Stellgröße so zu beeinflussen, dass sich daraus eine gewünschte Systemantwort ergibt. Je nach Komplexität des dynamischen Systems kann sich der Entwurf einer Steuerung unterschiedlich schwer gestalten. Liegt ein linearer Zusammenhang zwischen der Stellgröße und den Zuständen vor, lässt sich eine geeignete Steuerung rasch ermitteln. Gibt es hingegen nichtlineare Abhängigkeiten oder komplizierte Wechselwirkungen, gestaltet sich ein Steuerungsentwurf in der Regel schwierig. Typischerweise erfolgt der Entwurf auf der Basis eines physikalisch geprägten Simulationsmodells, welches durch den Ingenieur im Rahmen eines Modellbildungsprozesses aufgestellt wird. Die Grundidee der modellbasierten Vorgehensweise ist, eine geeignete Steuerung für das Simulationsmodell zu entwerfen und diese dann gleichermaßen für das reale System zu verwenden. Ein wesentliches Qualitätsmerkmal ist dabei die Prädiktionsgenauigkeit des aufgestellten Modells. Bildet das Modell die Realität nicht genau genug ab, funktioniert die entwickelte Steuerung für das reale System nicht, so dass üblicherweise in erster Linie eine Verbesserung des Modells angestrebt wird. Vor dem Hintergrund komplexer mechatronischer Systeme, welche schwierig zu beschreibende physikalische Effekte beinhalten, ist diese herkömmliche Herangehensweise allerdings kompliziert umzusetzen und damit für die Praxis in der Regel kein möglicher Lösungsweg. Aus diesem Grund wird in diesem Kapitel der Einsatz eines maschinellen Lernverfahrens - der sogenannten Bayesschen Optimierung (BO) - für den Steuerungsentwurf vorgestellt. Die BO basiert zu einem großen Teil auf der eingeführten Gauß-Prozess-Regression und nutzt deren probabilistischen Aspekte für eine gezielte bzw. effiziente Identifikation einer parametrisierten Steuerung.

Das Kapitel ist folgendermaßen aufgebaut: Zunächst wird in Abschnitt 3.1 die Aufgabe des Steuerungsentwurfs erläutert und den Einsatz der BO weiter motiviert. Danach wird in 3.2 die BO grundlegend eingeführt und vor dem Hintergrund des regelungstechnischen Entwurfs betrachtet. Im Rahmen dieses Abschnitts wird zudem der Stand der Forschung ausführlich erläutert. Anschließend wird in 3.3 ein komplexes mechatronisches System - das Ultraschalldrahtbonden - als Anwendungsbeispiel vorgestellt und die Problematik beim zugehörigen Steuerungsentwurf besprochen. Dies steht in enger Verbindung zum Folgeabschnitt 3.4, in welchem die physikalische Modellbildung mit verschiedenen Tieften thematisiert wird. In 3.5 erfolgt die Verknüpfung von der BO und dem Ultraschall-

drahtbonden, wobei eine speziell auf den Prozess angepasste Variante der BO entwickelt wird. Den Abschluss dieses Kapitels bildet 3.6, worin die Anwendung und die Ergebnisanalyse erläutert werden. In 3.7 wird abschließend ein ausführlicher Ausblick zu weiterführenden Forschungsthemen gegeben.

3.1 Aufgabenstellung des parametrisierten Steuerungsentwurfs

Den Ausgangspunkt der Betrachtung stellt ein beliebiges diskretes dynamisches System mit

$$\mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k, \mathbf{u}_k), \quad (3-1)$$

mit einem bekannten Anfangszustand \mathbf{x}_0 und einer finiten Anzahl von Zeitschritten $k = 0, \dots, H$ dar. Für das System soll eine Steuerung ausgelegt werden, wobei noch zu klären ist, welche Ziele dabei zu erfüllen sind. Zunächst wird weiterführend davon ausgegangen, dass für die Steuerung eine passende Ansatzfunktion mit

$$\mathbf{u}_k = \mathbf{u}_k(t_k; \boldsymbol{\theta}) \quad (3-2)$$

zur Verfügung steht, welche von einer bestimmten Parametrisierung $\boldsymbol{\theta} \in \mathbb{R}^{n_\theta}$ abhängt. In diesem Zusammenhang wird in der Regelungstechnik der Begriff einer parametrisierten Steuerung verwendet. Grundsätzlich kann an dieser Stelle auch von einer parametrisierten Regelung mit $\mathbf{u}_k = \mathbf{u}_k(\mathbf{x}_k; \boldsymbol{\theta})$ ausgegangen werden, allerdings wird vor dem Hintergrund des Anwendungsbeispiels weiterhin ausschließlich eine parametrisierte Steuerung betrachtet. Die Aufgabe besteht nun darin, eine passende Parametrisierung auszuwählen, so dass das Steuerungsziel erreicht wird. Bisher noch unbekannt ist, nach welchem Kriterium eine Bewertung von unterschiedlichen Parametrisierungen vorgenommen werden kann. Hierzu wird beispielhaft das quadratisch gewichtete Gütekriterium $J(\boldsymbol{\theta}) : \mathbb{R}^{n_\theta} \mapsto \mathbb{R}_+$ über

$$\min_{\boldsymbol{\theta}} J(\boldsymbol{\theta}) = \Delta t \sum_{k=0}^H (\mathbf{x}_k - \mathbf{x}_G)^T \mathbf{W}_x (\mathbf{x}_k - \mathbf{x}_G) + \mathbf{u}_k^T \mathbf{W}_u \mathbf{u}_k \quad (3-3)$$

eingeführt. Hierin ist \mathbf{x}_G der geforderte Zielzustand, welcher über die parametrisierte Steuerung erreicht werden soll, Δt ist die zeitliche Schrittweite und $\mathbf{W}_x \in \mathbb{R}^{n_x \times n_x}$, $\mathbf{W}_u \in \mathbb{R}^{n_u \times n_u}$ sind vorgebbare Gewichtungsmatrizen. Anhand des Gütemaßes können verschiedene Parametrisierungen ihrer Wertigkeit nach geordnet und miteinander verglichen werden. Das Ziel besteht darin, das Optimierungsproblem (3-3) unter den Nebenbedingungen von (3-1) und (3-2) zu lösen und diejenige Parametrisierung $\boldsymbol{\theta}^*$ zu bestimmen, welche den kleinsten Gütfunktionswert liefert. Als Beispiel wird in [GTZ07] der Auf-

schwung eines Doppelpendels auf einem Wagen (s. Bild 1-4) realisiert, indem als Ansatz für die Steuerung des Wagens eine Cosinusreihe mit vier Parametern verwendet wird. Der Einsatz der Cosinusreihe wird damit begründet, dass sie die natürliche Bewegung des Doppelpendels berücksichtigt und sie gleichzeitig dafür sorgt, dass der Wagen wieder an seine Anfangsposition in der Mitte der Wagenstrecke zurückkehrt. Die Wahl der Ansatzfunktion ist daher für diese Form des Steuerungsentwurfs entscheidend und erfordert möglicherweise besonderes Expertenwissen über das betrachtete System.

Um (3-3) zu lösen, wird beim modellbasierten Entwurf (vgl. Bild 1-2) so vorgegangen, dass zunächst ein physikalisches Dynamikmodell zum realen System (3-1) aufgestellt wird. Da es sich um eine möglicherweise unvollständige bzw. fehlerbehaftete Beschreibung handelt, wird für das Modell die Schreibweise \hat{f} (ähnlich zu einer Zustandsschätzung eines Beobachters) benutzt. Anschließend kann durch eine Simulation weiterführend auch eine Schätzung der Gütfunktion \hat{J} umgesetzt werden, womit eine Lösung des Optimierungsproblems ohne Verwendung des realen Systems berechnet werden kann. Das Problem bei diesem Ansatz ist, dass dabei angenommen wird, dass das aufgestellte Dynamikmodell keine zu große Abweichung zum realen System aufweist und dieses genau genug beschreibt. Ist die Abweichung zu groß, so kann die gefundene Parametrisierung der Steuerung zwar für das Modell qualitativ hochwertige Ergebnisse liefern, aber am realen System unter Umständen versagen und unbefriedigende Resultate erzeugen. Besonders ausgeprägt ist dieses Phänomen für komplexe, schwierig zu modellierende Systeme, für die gegebenenfalls nur eine simplifizierte Modellbildung angesetzt werden kann. In diesem Fall stößt der beschriebene modellbasierte Entwurfsprozess an seine Grenzen und kann möglicherweise nicht weiterverfolgt werden. Von so einem Szenario soll an dieser Stelle weiterführend ausgegangen werden. Als Ansatz für eine erfolgreiche Realisierung wird die Verwendung der BO vorgeschlagen, da sie als maschinelles Lernverfahren direkt mit dem realen System in Interaktion tritt und eine erweiterte Auslegung auf der Basis von Messdaten erfolgt. Im nächsten Abschnitt wird dieser Lösungsweg näher erläutert.

3.2 Bayessche Optimierung als Lösungsansatz

Die BO ist ein ML-Verfahren zur effizienten und automatisierten Suche von optimalen Parametern. Ihr Ansatz basiert auf dem Bayesschen Theorem (2-13), um Unsicherheiten zu beschreiben und Rückschlüsse aus bereits gesammelten Daten zu ziehen. Das Verfahren ist darauf ausgelegt, einen Kompromiss zum Exploration-Exploitation-Dilemma zu finden und wird damit den globalen Optimierungsverfahren zugeordnet. Es ist somit, im Gegensatz zu gradientenbasierten Verfahren, in der Lage, die Konvergenz zu lokalen Optima zu unterbinden. Ein weiterer Vorteil ist, dass keine Kenntnis über den Gradien-

ten bzgl. der Gütfunktion vorhanden sein muss. Lediglich die punktuelle Auswertung der Gütfunktion, welche ggf. durch einen Rauschprozess verfälscht sein kann, ist eine Voraussetzung für die Anwendbarkeit.

Die Grundidee des Verfahrens ist, die Zielfunktion als datenbasierten GP zu modellieren und anhand seiner probabilistischen Vorhersagen eine Auswahl geeigneter Gebiete im Parameterraum zu bestimmen, an welchen die optimalen Parameter zu finden sind. Der GP wird iterativ durch die Hinzunahme von neuen Beobachtungen aktualisiert, die durch die Auswertung der Gütfunktion an den erfolgversprechendsten Stellen gewonnen werden. Die BO hat sich insgesamt als nützliches Werkzeug in Situationen erwiesen, in denen die Evaluierung der Zielfunktion teuer und/oder zeitaufwendig ist, wie z. B. bei der Optimierung von Hyperparametern eines tiefen künstlichen neuronalen Netzes. Durch ihre Fähigkeit, ressourceneffizient zu sein und sich auf die vielversprechendsten Bereiche des Parameterraums zu konzentrieren, hat die BO in verschiedenen praktischen Anwendungsbereichen Einzug gehalten. Darunter fallen die explorativen Bohrungen, um neue Ölreserven zu erschließen, sowie die aufwendigen Experimente und Studien, die im Zusammenhang mit der Entwicklung von neuartigen Medikamenten im Bereich der Pharmazie notwendig sind [SSW⁺16].

Die BO wird häufig mit anderen alternativen Lösungsmöglichkeiten verglichen. Dazu zählen die zufällige, die gitterbasierte und die manuelle Suche. Die zufällige Suche wählt ihrem Namen nach in jedem Iterationsschritt eine zufällige Parametrierung aus, wodurch ihre Effizienz als besonders schlecht einzustufen ist. Die manuelle Suche, welche vom menschlichen Anwender vorgenommen wird, stellt im Allgemeinen eine bessere Alternative als die zufällige Suche dar. Diese wird daher häufig eingesetzt, ist jedoch stark von den Fähigkeiten des Anwenders und dessen Expertise in Bezug auf das Optimierungsproblem abhängig. Außerdem gibt es deutliche Einschränkungen, welche die Anzahl der Parameter betreffen. Für niedrigdimensionale Probleme ist es für den Menschen machbar, einen Überblick über die bereits ausprobierten und möglicherweise geeigneten Parameter-Kandidaten zu behalten, wohingegen höherdimensionale Probleme im Allgemeinen eine Herausforderung darstellen und eine vorher festgelegte Systematik möglicherweise nur mit Mühe eingehalten werden kann. Des Weiteren stellt die gitterbasierte Suche ein gängiges Lösungsverfahren dar. Hierbei wird der Suchraum in einer geeigneten Weise diskretisiert und dann nach und nach ausgewertet. Unter Umständen sind in bestimmten Gebieten feinere Diskretisierungsschritte notwendig, um eine gewünschte Genauigkeit zu erreichen. Ebenso wie bei der manuellen Suche, hat dieser Ansatz jedoch das Problem, für höherdimensionale Optimierungen schlechte Skalierungseigenschaften aufzuweisen. In diesem Zusammenhang wird auch vom *Fluch der Dimensionalität* [Bis06] gesprochen, welcher den zugrundeliegenden Effekt beschreibt. Bei einer gleichmäßigen Diskretisie-

rung des Suchraums nimmt mit der Dimension die Anzahl der benötigten Diskretisierungspunkte exponentiell zu, wodurch der Speicherbedarf und der Rechenaufwand immens ansteigen. Im direkten Vergleich stellt die BO im Allgemeinen, aufgrund ihres sich sukzessiv verbessernden Charakters, die Beste der genannten Lösungsmöglichkeiten dar und ist diesen daher vorzuziehen.

Im Nachfolgenden wird die BO [SSW⁺16] formal bzw. mathematisch beschrieben. Die Basis bildet die Gauß-Prozess-Regression (s. 2.4), wobei die Schätzung der Gütfunktion (3-3) über eine parameterabhängige Zufallsvariable nach

$$\hat{J}(\boldsymbol{\theta}) \sim \mathcal{GP}(m(\boldsymbol{\theta}), k(\boldsymbol{\theta}, \boldsymbol{\theta}'; \boldsymbol{\eta})), \quad (3-4)$$

definiert wird. Unter der Voraussetzung eines vorhandenen Datensatzes, bestehend aus $\mathbf{D}_\theta = [\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_{n_d}]^T \in \mathbb{R}^{n_d \times n_\theta}$ und $\mathbf{D}_J = [J_1, \dots, J_{n_d}]^T \in \mathbb{R}^{n_d}$, werden die Hyperparameter $\boldsymbol{\eta} \in \mathbb{R}^{n_\theta+2}$ der Kovarianzfunktion $k(\cdot, \cdot)$ standardmäßig über die Maximierung der logarithmischen Likelihoodfunktion (vgl. (2-33))

$$\begin{aligned} \boldsymbol{\eta}_* &= \arg \max_{\boldsymbol{\eta}} \log(p(\mathbf{D}_J)) \\ &= \arg \max_{\boldsymbol{\eta}} -\frac{1}{2}(\mathbf{D}_J - \mathbf{m}_D)^T \mathbf{K}(\boldsymbol{\eta})^{-1} (\mathbf{D}_J - \mathbf{m}_D) - \frac{1}{2} \log(\det \mathbf{K}(\boldsymbol{\eta})), \end{aligned} \quad (3-5)$$

bestimmt, mit $\mathbf{m}_D = [m(\boldsymbol{\theta}_1), \dots, m(\boldsymbol{\theta}_{n_d})]^T$ und symmetrischer und positiv definiter Gram-Matrix $\mathbf{K}(\boldsymbol{\eta}) \in \mathbb{R}^{n_d \times n_d}$ mit den Elementen $K_{i,j} = k(\boldsymbol{\theta}_i, \boldsymbol{\theta}_j; \boldsymbol{\eta})$, $i, j = 1, \dots, n_d$. Die Maximierung wird auch als Training des GP bezeichnet. Nach der Identifikation der Hyperparameter liefern die Posteriorgleichungen des GP (vgl. (2-31))

$$\begin{aligned} p(\hat{J}(\boldsymbol{\theta}) | \mathbf{D}_y) &= \mathcal{N}(\mu(\boldsymbol{\theta}), \sigma^2(\boldsymbol{\theta})), \\ \mu(\boldsymbol{\theta}) &= m(\boldsymbol{\theta}) + \mathbf{k}_D^T(\boldsymbol{\theta}) \mathbf{K}^{-1} (\mathbf{D}_y - \mathbf{m}_D), \\ \sigma^2(\boldsymbol{\theta}) &= k(\boldsymbol{\theta}, \boldsymbol{\theta}) - \mathbf{k}_D^T(\boldsymbol{\theta}) \mathbf{K}^{-1} \mathbf{k}_D(\boldsymbol{\theta}), \end{aligned} \quad (3-6)$$

mit $\mathbf{k}_D(\boldsymbol{\theta}) = [k(\boldsymbol{\theta}, \boldsymbol{\theta}_1; \boldsymbol{\eta}_*), \dots, k(\boldsymbol{\theta}, \boldsymbol{\theta}_{n_d}; \boldsymbol{\eta}_*)]^T$, eine probabilistische Schätzung über die Gestalt der Gütfunktion, wobei die Schätzung alle bisher bekannten Datenpunkte berücksichtigt. Das iterative Vorgehen der BO ergibt sich danach durch eine kontinuierliche Erweiterung des Datensatzes $\mathbf{D}_\theta \leftarrow \{\mathbf{D}_\theta, \boldsymbol{\theta}_{n_d+1}\}$, $\mathbf{D}_J \leftarrow \{\mathbf{D}_J, J_{n_d+1}\}$, wodurch die Schätzung in jeder Iteration weiter verbessert wird.

Bisher unklar ist, wie anhand des GP die nächste Parametrierung $\boldsymbol{\theta}_{n_d+1}$ bestimmt bzw. nach welchem Kriterium das nächste Experiment am realen System festgelegt wird. Hierfür wird im Rahmen der BO eine sogenannte Akquisitionsfunktion $\alpha(\boldsymbol{\theta}) : \mathbb{R}^{n_\theta} \mapsto \mathbb{R}$ [SSW⁺16],

welche von der aktuellen Schätzung bzgl. der Gütfunktion des GP (3-6) abhängig ist, benutzt und ein unterlagertes Optimierungsproblem

$$\boldsymbol{\theta}_{n_d+1} = \arg \max_{\boldsymbol{\theta}} \alpha(\boldsymbol{\theta}), \quad (3-7)$$

automatisiert gelöst. Mittlerweile gibt es ein breites Spektrum von möglichen Funktionen, die je nach Anwendungsfall besser oder schlechter geeignet sind. Einige wichtige Akquisitionsfunktionen werden für einen Überblick im Nachfolgenden genauer vorgestellt.

Das erste Kriterium lautet *Probability of Improvement* (PI) [Kus64] und ist über

$$\alpha_{PI}(\boldsymbol{\theta}) = P(\hat{J}(\boldsymbol{\theta}) \leq \xi_J) = \Phi(\gamma(\boldsymbol{\theta})), \quad \gamma(\boldsymbol{\theta}) = \frac{\xi_J - \mu(\boldsymbol{\theta})}{\sigma(\boldsymbol{\theta})}$$

definiert. Hierbei ist ξ_J ein a-priori festgelegter Grenzwert, der erreicht werden soll, oder alternativ der aktuell beste Gütfunktionswert aus den bestehenden Beobachtungen, welcher bei jeder auftretenden Verbesserung entsprechend ersetzt wird. Die Wahrscheinlichkeit einer Verbesserung kann besonders vorteilhaft durch die Verwendung eines GP analytisch bestimmt werden, sodass $\Phi(\cdot)$ in diesem Zusammenhang die Standardverteilungsfunktion einer Normalverteilung darstellt. Ein Problem dieses Ansatzes ist, dass sich die Optimierung ggf. zu stark auf einen bestimmten Bereich im Suchraum fokussiert und diesen nicht mehr verlässt bzw. die Suche gewissermaßen stagniert.

Das nächste Kriterium lautet *Expected Improvement* (EI) [Moc74] und folgt der Grundidee, die Größe der Verbesserung weiter zu quantifizieren. Hierfür wird der Erwartungswert im Bezug auf die Dichte, die unter dem Grenzwert liegt, mit

$$\begin{aligned} \alpha_{EI}(\boldsymbol{\theta}) &= \mathbb{E}[\max(0, \xi_J - \hat{J}(\boldsymbol{\theta}))] \\ &= \sigma(\boldsymbol{\theta})(\gamma(\boldsymbol{\theta})\Phi(\gamma(\boldsymbol{\theta}))) + \phi(\gamma(\boldsymbol{\theta})) \end{aligned} \quad (3-8)$$

herangezogen. Hierin ist $\phi(\cdot)$ die Standardnormalverteilung. Innerhalb vieler BO Implementierungen stellt diese Akquisitionsfunktion den Standard dar, da sie sich in zahlreichen Veröffentlichungen als robust herausgestellt hat. In [NH18] wird eine geringfügig abgewandelte Variante vorgeschlagen, welche eine Skalierung mit Hilfe der Varianz bzw. Standardabweichung vornimmt und über

$$\begin{aligned} \alpha_{SEI}(\boldsymbol{\theta}) &= \mathbb{E}[\max(0, \xi_J - \hat{J}(\boldsymbol{\theta}))]\mathbb{V}[\max(0, \xi_J - \hat{J}(\boldsymbol{\theta}))]^{-\frac{1}{2}} \\ &= \alpha_{EI}(\boldsymbol{\theta})(\sigma^2(\boldsymbol{\theta})(\gamma(\boldsymbol{\theta})^2 + 1)\Phi(\boldsymbol{\theta}) + \gamma(\boldsymbol{\theta})\phi(\gamma(\boldsymbol{\theta}))) - \alpha_{EI}(\boldsymbol{\theta})^2)^{-\frac{1}{2}}, \end{aligned}$$

beschrieben wird. Aufgrund der zusätzlichen Skalierung wird sie als *Scaled Expected Improvement* (SEI) bezeichnet. Im Rahmen der Veröffentlichung wurden mit ihr gute Er-

gebnisse bzgl. verschiedener Testfunktionen erzielt. In Verbindung mit dem Anwendungsbeispiel der vorliegenden Arbeit konnte jedoch keine Überlegenheit zu der unskalierten Variante festgestellt werden, sodass sich ihre Anwendbarkeit zunächst noch durch weitere unabhängige Veröffentlichungen herausstellen muss.

Der Vollständigkeit halber wird als letzte Akquisitionsfunktion der *Lower Confidence Bound* (LCB) [SKKS12] Ansatz genannt, welcher formal

$$\alpha_{LCB}(\boldsymbol{\theta}) = -\mu(\boldsymbol{\theta}) + \kappa\sigma(\boldsymbol{\theta}),$$

lautet. Dabei stellt $\kappa > 0$ einen Explorationsparameter dar, der nicht notwendigerweise konstant sein muss, sondern sich über die Iterationen hinweg verändern kann. Sinnvoll ist bspw. ein Ansteigen des Wertes, um den Suchraum weiter zu erforschen und lokalen Optima zu entkommen. Sofern eine rasche Konvergenz beabsichtigt ist, kann es jedoch auch vorteilhaft sein den Wert langsam gegen null konvergieren zu lassen, um ein bestimmtes Optimum zu fokussieren. Vorteilhaft ist, dass gegenüber den anderen Akquisitionsfunktionen kein Grenzwert benötigt wird und für bestimmte iterationsabhängige κ -Funktionen obere Schranken, die den Worst-Case beschreiben, hergeleitet werden können [SKKS12].

Algorithmus 1 Bayessche Optimierung

- 1: **Eingabe:** Initiale Datenbasis $(\mathbf{D}_\theta, \mathbf{D}_J)$ aus n_{init} Experimenten, Iterationsbudget n_{budget} , GP Mittelwert- und Kovarianzfunktion $(m(\boldsymbol{\theta}), k(\boldsymbol{\theta}, \boldsymbol{\theta}'))$, ggf. Grenzwert ξ_J bzw. Explorationsfaktor κ .
 - 2: **Wiederhole** bis das Iterationsbudget n_{budget} aufgebraucht ist:
 - 3: Aktualisiere Gauß-Prozess bzw. Hyperparameter η nach (3-5).
 - 4: Evaluiere bei $\boldsymbol{\theta}_q = \arg \max_{\boldsymbol{\theta}} \alpha(\boldsymbol{\theta})$ und erhalte J_q .
 - 5: Erweitere Datenmenge $(\mathbf{D}_\theta, \mathbf{D}_J) \leftarrow (\mathbf{D}_\theta \cup \boldsymbol{\theta}_q, \mathbf{D}_J \cup J_q)$.
 - 6: **Ausgabe:** Optimale Parameter $\boldsymbol{\theta}_i, i = \text{index } \min_i \mathbf{D}_J^{(i)}, i = 1, \dots, n_{init} + n_{budget}$.
-

Algorithmus 1 fasst die wesentlichen Schritte der BO übersichtlich zusammen. In der 1. Zeile erfolgt die Initialisierung, wobei die initiale Datenmenge in der Regel durch zufällige Experimente bestimmt wird. Das Iterationsbudget legt die Anzahl der zur Verfügung stehenden Experimente fest und ist üblicherweise auf zeitliche Rahmenbedingungen zurückzuführen. Alternativ kann auch ein Konvergenzkriterium angesetzt werden. Die Mittelwert- und Kovarianzfunktion bestimmen die grundsätzlichen Annahmen, die über den GP getroffen werden. Hinzu kommen gegebenenfalls übergeordnete Parameter, die die Akquisitionsfunktion betreffen. In jeder Iteration wird zunächst der GP durch die bestehende Datenmenge aktualisiert und eine Anpassung der Hyperparameter anhand von (3-5) vorgenommen (Zeile 3). Anschließend erfolgt die Lösung des unterlagerten Optimierungsproblems (3-7), welches die gewählte Akquisitionsfunktion umfasst (Zeile 4).

Abschließend wird in der 5. Zeile ein neues Experiment ausgeführt und das gesammelte Datenpaar den bestehenden Daten hinzugefügt. Nach dem Erreichen des Iterationsbudgets wird diejenige Parametrisierung ausgewählt, die dem kleinsten Gütefunktionswert zugeordnet ist (Zeile 6). Um eine mögliche Verfälschung durch etwaiges Rauschen in den Daten zu berücksichtigen, besteht auch die Möglichkeit den Mittelwert des Posteriors $\mu(\theta)$ (3-6) im Rahmen eines weiteren Optimierungsproblems zu minimieren und die dabei gefundene Lösung als optimale Parametrisierung auszugeben.

Stand der Forschung

Dieser Abschnitt beschäftigt sich mit allgemeinen und regelungstechnisch spezifischen Veröffentlichungen, in denen die BO zur Lösung genutzt wird. Bei den Veröffentlichungen mit einem regelungstechnischen Hintergrund wird zwischen Arbeiten unterschieden, die eine bestimmte Form von Vorwissen über das technische System oder die Steuerung bzw. Regelung beinhalten, und solchen Arbeiten, die keine besondere Vorkenntnis ausnutzen. Weiterführend wird hierauf der Beitrag der vorliegenden Arbeit erläutert und eingeordnet.

Einen gesamtheitlichen Überblick über die Verwendung der BO anstelle des manuellen Einstellens bietet [SSW⁺16], worin insbesondere noch einmal die Unterschiede der vorgestellten Akquisitionsfunktionen behandelt und verschiedene Optimierungsrouterien zur Lösung der unterlagerten Probleme vorgestellt werden. In [SLA12] werden hierzu passend weiterführende Analysen vorgenommen, die sich auf praktische Aspekte fokussieren. Unter anderem werden der SE- (2-32) und der Matern-Kernel (2-35) gegenübergestellt und die zu hohen Glattheitsanforderungen des SE-Kernels für praktische Anwendungen kritisiert. Bei den Anwendungsbeispielen dieser Arbeit konnte die Kritik jedoch nicht bestätigt werden, sodass der SE-Kernel in allen Fällen zufriedenstellende Ergebnisse lieferte. Die hohen Glattheitsanforderungen müssen daher grundsätzlich nicht als negativ erachtet werden und schaffen ggf. eine weitere Ebene, um Vorwissen einzubringen.

Der Einsatz der BO für regelungstechnische Fragestellungen wird vielseitig in der Literatur untersucht. In [CVS⁺20] wird eine umfassende Zusammenfassung zu den möglichen Schnittstellen zwischen der Regelungstechnik und verschiedenen maschinellen Lernverfahren gegeben. Die Schnittstellen werden bei dem Dynamikmodell, der Steuerung oder Regelung und der Gütfunktion gesehen. Innerhalb einer hierarchisch aufgestellten Struktur, wird die BO dabei auf der Ebene der Gütfunktion eingeordnet, wodurch eine untergeordnete Korrektur eines fehlerbehafteten Dynamikmodells umgangen wird. Dies ist unter anderem besonders vorteilhaft, wenn unklar ist, wie die Zustände eines Systems genau definiert sind oder sie schwierig, durch entsprechende Sensorik, erfasst werden können. Die Autoren führen zudem verschiedene Wege zur Einbringung von Vorwissen an, so

kann bspw. eine gewisse Struktur oder Ansatzfunktion für die Steuerung bekannt sein oder eine möglicherweise vorhandene Parametrisierung bereits eingegrenzt werden.

Spezifische Anwendungsfälle, in denen die BO zur Bestimmung von Parametern in einem Regelungstechnischen Kontext eingesetzt wird, jedoch kein besonderes Vorwissen über die Aufgabe ausgenutzt wird, sind in [SKOT13; CHJA18; CSPD16; NMST20; MHB⁺16; RWL19; Bij18; ZBP21; SAA20; FKDZ19] zu finden. Die erstgenannten Publikationen [SKOT13] und [CHJA18] konzentrieren sich auf den Beobachterentwurf und die automatisierte Bestimmung der beinhalteten Parameter mittels BO. Im Falle eines UKF, welcher auf den Transformationsgleichungen der UT (2-26)-(2-28) basiert, wird bspw. der Entwurfsparameter κ , der die Entfernung der Sigma-Punkte um den Erwartungsvektor steuert, optimal in Bezug auf die Nichtlinearitäten des Systems ausgelegt. Auf diese Weise kann die Qualität der Verteilungsapproximation gesteigert werden. Als Motivation für die Verwendung der BO wird in diesen Arbeiten angeführt, dass kein spezifisches Fachwissen über die Beobachterstruktur, das dynamische System und dessen zugrundeliegenden Terme benötigt und so die aufwendige manuelle Suche umgangen wird.

In den anderen aufgeführten Veröffentlichungen wird Bezug zum Steuerungs- bzw. Regelungsentwurf genommen. In [CSPD16] wird ein Regler, welcher als Zustandsautomat realisiert ist und gewisse Umschaltzeiten als Parameter besitzt, mit Hilfe der BO für einen Laufroboter ausgelegt. Im direkten Vergleich mit anderen Lösungsmöglichkeiten zeichnet sich die BO hierbei als äußerst effizient aus, was in Hinblick auf mögliche Beschädigungen aufgrund eines instabilen Laufvorgangs ein besonders wichtiges Kriterium darstellt. Hervorzuheben ist, dass dabei kein Vorwissen über die Laufdynamik des Roboters genutzt wurde. Die Autoren begründen dies mit der schwierigen Modellbildung aufgrund der niedrigen Qualität der verbauten Komponenten, die das Gesamtsystem sehr schwingungsanfällig machen. In [MHB⁺16] wird die BO für den Entwurf eines linear-quadratischen Reglers verwendet. Allerdings werden nicht direkt die Parameter des linearen Reglers als Optimierungsvariablen definiert, sondern die Elemente der Gewichtungsmatrizen des riccatischen Gütemaßes (vgl. (3-3)). Der Grund hierfür ist, dass die BO somit nur im Raum stabiler Regler arbeitet und mögliche Selektierungen, die zu einem instabilen Verhalten führen würden, ausgeschlossen sind. Ähnlich zum modellbasierten Entwurf ist diese Schlussfolgerung allerdings nur so gut wie die Prädiktionsqualität des linearen Dynamikmodells auf dessen Berechnung das riccatische Gütemaß basiert. In den Arbeiten [NMST20], [Bij18] und [RWL19] wird die BO tiefergehend im industriellen Kontext und für die Auslegung von einfachen Reglern untersucht. Die betrachteten Systeme sind dabei ein Steuerungsventil, welches zur Komprimierung von Gasen eingesetzt wird, eine Windrand zur Stromerzeugung und ein mechanisches Positionierungssystem. In allen Fällen wird höchstens ein lineares Modell als Vorwissen verwendet. Ein komple-

xerer Regelungsansatz wird in den Arbeiten [SAA20] und [ZBP21], repräsentiert durch eine modellprädiktive Regelung (MPC), verwendet. Dabei wird bspw. der Stellhorizont der MPC für das betrachtete System durch die BO optimiert. Zuletzt sei auch die Arbeit von [FKDZ19] zu erwähnen, worin die Parameter eines künstlichen neuronalen Netzes, welches den Regler darstellt, mittels BO bestimmt werden. Da die Anzahl der Parameter zu den vorherigen Anwendungsfällen vergleichsweise hoch ist, wird allerdings eine Dimensionsreduktion mit Hilfe eines *Principle Component Analysis* (PCA) Ansatzes dem BO Algorithmus hinzugefügt. Auf diese Weise kann der Suchraum deutlich reduziert und der Einsatz der BO, welche grundsätzlich besser für niedrig dimensionale Probleme geeignet ist, ermöglicht werden.

Die Veröffentlichungen des vorhergehenden Abschnitts erzielen insgesamt gute Ergebnisse bei der Verwendung der BO im regelungstechnischen Kontext. Etwaiges Vorwissen wird jedoch nur rudimentär genutzt, größtenteils in Form von linearen Dynamikmodellen. Unklar bleibt dabei, wie sich komplexeres Vorwissen, möglicherweise vorliegend als nichtlineares analytisches oder Black-Box Modell, im Rahmen der BO einfügen lässt. Hierfür bieten sich unter anderem die angenommene Mittelwert- und Kovarianzfunktion des GPs (s. (3-4)) an. Üblicherweise werden hierfür aufgrund eines nicht vorhandenen oder unklaren Vorwissens die Ansätze $m(\boldsymbol{\theta}) = 0$ und $k(\cdot, \cdot) = k_{SE}(\cdot, \cdot)$ getroffen. Da sich die BO an die Datenbasis anpasst, stellen diese Ansätze keine Einschränkung an das Verfahren dar, bieten jedoch ein Verbesserungspotential in Bezug auf die Effizienz bei der Inbetriebnahme. Komplexere Funktionen, die ein bestimmtes Fachwissen ausnutzen, wirken sich im Rahmen der BO so aus, dass bestimmte Gebiete im Suchraum ggf. ausgeschlossen werden können und durch die BO nicht aufwendig exploriert werden müssen. Die BO kann sich somit direkt auf die Gebiete fokussieren, die die besten Gütekennwerte generieren und das zur Verfügung stehende Iterationsbudget dort zweckmäßiger nutzen.

Zwei Veröffentlichungen, die die Schnittstelle der Kovarianzfunktion zur Einbringung von Fachwissen explizit ausnutzen, sind [RAMA19] und [MBH⁺17]. Die Schnittstelle wird dabei sehr unterschiedlich verwendet. In [RAMA19] wird ein im Kreis laufender Roboter betrachtet, dessen Steuerungsparameter ausgelegt werden sollen. Die Autoren nutzen für diesen Fall eine eigene Variante des SE-Kernel, welche über

$$k_{ISE}(\boldsymbol{\theta}, \boldsymbol{\theta}'; \boldsymbol{\eta}) = k_{SE}(\phi(\boldsymbol{\theta}), \phi(\boldsymbol{\theta}'); \boldsymbol{\eta}),$$

beschrieben wird und als Zusatz den Buchstaben *I* für *Informed* trägt. Hierin ist $\phi(\boldsymbol{\theta})$ eine Transformation der Parametrisierung, die als Basis das vorhandene Vorwissen eines mechanischen Dynamikmodells nutzt. Es werden verschiedene Möglichkeiten für die Transformation vorgeschlagen. Eine Möglichkeit ist die simulierte Trajektorie, welche durch die Verwendung eines fixierten $\boldsymbol{\theta}$ entsteht, zu verwenden. Im Zuge dessen

wird bzgl. der SE-Berechnungsvorschrift kein Vergleich zwischen Parameterwerten vorgenommen, sondern ein Vergleich zwischen Trajektorien. Für besonders sensitive Systeme wie den Laufroboter ist dieser Ansatz sinnvoll, da zwei ähnliche Parametrisierungen ohne die Transformation zu einer hohen Kovarianz führen würden, allerdings nicht die möglicherweise gänzlich unterschiedlichen Trajektorien widerspiegeln. Mit Hilfe der Transformation gelingt es somit, die Parametrisierungen in einen geeigneten Raum für einen Vergleich zu überführen und das vorhandene Vorwissen zu berücksichtigen. Um die Dimension dieses Raums deutlich zu reduzieren und damit Berechnungszeit einzusparen, schlagen die Autoren vor, vorab (offline) eine ausreichend große Datenbank von simulierten Trajektorien zu erstellen und eine niedrig dimensionale Transformation über eine lernfähige Autoencoder-Struktur zu erlernen. Anschließend wird online die BO unter Verwendung des Autoencoders zur Identifikation der Parametrisierung ausgeführt.

In [MBH⁺17] wird die Schnittstelle zur Kovarianzfunktion ebenfalls genutzt, um ein dynamisches Modell einfließen zu lassen. Dafür wird ein sogenannter *Multisource* GP definiert, welcher mehrere verschiedene Informationsquellen vorsieht. Zunächst wird ein erweiterter Parametervektor mit $\tilde{\boldsymbol{\theta}} := [\boldsymbol{\theta}^T, \delta]^T$ eingeführt, wobei $\delta \in \{0, 1\}$ eine binäre Variable darstellt, die angibt, ob es sich um einen simulierten oder real gemessenen Datenpunkt handelt. Der Kernel wird demnach über den Ausdruck

$$k(\tilde{\boldsymbol{\theta}}, \tilde{\boldsymbol{\theta}}') = k_{sim}(\boldsymbol{\theta}, \boldsymbol{\theta}') + k_\delta(\delta, \delta') k_{err}(\tilde{\boldsymbol{\theta}}, \tilde{\boldsymbol{\theta}}')$$

angesetzt und setzt sich somit additiv aus einem Term für die Simulation und einem Korrekturterm zusammen. Für $k_{sim}(\cdot, \cdot)$ und $k_{err}(\cdot, \cdot)$ kann bspw. der SE-Kernel verwendet werden. Das Verbindungselement stellt $k_\delta(\cdot, \cdot) \in \{0, 1\}$ dar, womit überprüft wird, ob beide Eingaben vom realen System stammen. Nur in diesem Fall ist der korrektive Fehlerterm aktiv und wird im Rahmen der Gauß-Prozess-Regression berücksichtigt. Der Ansatz verfolgt die Grundidee, dass simulierte Daten nur einen Teil der wahren Gütfunktion erklären können. Dementsprechend sinkt die Unsicherheit bei den ausgewerteten Stellen des Parameterraums nur geringfügig im Vergleich zu einer Auswertung am realen System, wobei die Unsicherheit (abgesehen vom Messrauschen) vollständig verschwindet.

Bild 3-1 zeigt beispielhaft die Wirkungsweise des eingesetzten GPs in [MBH⁺17]. Die grüne durchgezogene Linie zeigt den Ground Truth bzw. die unbekannte wahre Gütfunktion. Der linke (schwarze) Datenpunkt stellt ein Experiment am realen System dar und stammt somit vom Ground Truth. Der rechte Datenpunkt wurde anhand des Simulationsmodells erzeugt und befindet sich abseits des Ground Truth, um die Auswirkung von vorhandenen Modellfehlern anzudeuten. Die grüne gestrichelte Linie wird somit dem Simulationsmodell zugeordnet. In der linken Bildhälfte stimmen Realität und Modell gut

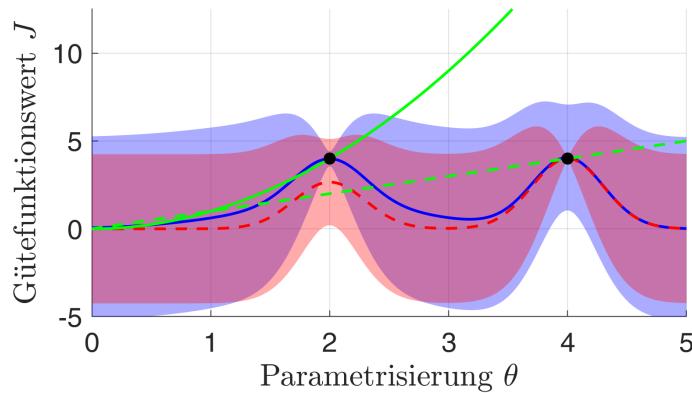


Bild 3-1: Visuelle Darstellung eines Multisource Gauß-Prozesses (blau, rot), welcher auf zwei verschiedenen Informationsquellen basiert. Die grüne durchgezogene Linie zeigt die wahre Gütfunktion und die grüne gestrichelte Linie eine modellbasierte Schätzung.

miteinander überein, wohingegen es auf der rechten Seite deutliche Unterschiede gibt, der (steigende) Trend tendenziell aber stimmt. Die blauen Verläufe gehören zur vollständigen Kovarianzfunktion, also inkl. dem Korrekturterm, und kennzeichnen den Mittelwert und die zweifache Standardabweichung. Gut erkennbar ist, dass die Varianz beim realen Datenpunkt stark reduziert ist, wohingegen der fiktive Datenpunkt der Simulation die Unsicherheit nur geringfügig beeinflusst. Die roten Verläufe werden der Simulation zugeordnet und weisen das gegenteilige Verhalten auf. Hierbei führt der fiktive Datenpunkt zu einer deutlichen Reduktion der Varianz und der reale Datenpunkt hat nur eine geringe Auswirkung. Die BO-Routine enthält aufgrund des erweiterten GP-Ansatzes eine zusätzliche Abfrage, ob ein reales oder fiktives (simuliertes) Experiment ausgeführt werden soll. Dies geschieht anhand eines weiteren Kriteriums, welches durch bestimmte vorab festgelegte Grenzwerte bestimmt ist. In dieses Kriterium fließen die Grundsätze ein, dass ein simuliertes Experiment zeitlich deutlich schneller durchgeführt werden kann als ein reales, dafür jedoch weniger Informationen beinhaltet und damit vergleichsweise als minderwertiger anzusehen ist. Aufgrund dieser Erweiterung wählt die BO automatisch zwischen realen und simulativen Auswertungen und wechselt, je nach festgelegten Kriterien, zwischen diesen hin und her. Im Kern berücksichtigt [MBH⁺17] somit zwar a-priori-Wissen über die Regelungsaufgabe, macht allerdings auch darauf aufmerksam, dass dieses Wissen unvollständig bzw. nicht perfekt sein kann und daher nur bedingt und mit einer gewissen Unsicherheit, ausgenutzt werden sollte.

Die genannten Arbeiten haben gemeinsam, dass die Mittelwertfunktion $m(\cdot)$ des GP als Schnittstelle vernachlässigt wird, obwohl diese den intuitiv einfachsten Weg darstellt, um Expertenwissen einzubinden. In [MBH⁺17] wird dieser Aspekt innerhalb einer kritischen Analyse ihres Ansatzes aufgegriffen. Der Umweg über den Multisource-GP ist dann besonders nützlich, wenn das Vorwissen relativ ungenau ist und nur eine grundsätzliche

Tendenz in Hinblick auf die Gütfunktion beinhaltet. Andernfalls sehen die Autoren die Mittelwertfunktion als besser geeignete Schnittstelle an, da über sie keine Generierung von Datenpunkten notwendig ist. Das gesamte Simulationsmodell ist demnach direkt bei der Auswertung einer bestimmten Stelle θ vorhanden und nicht nur partiell an den Stellen, die durch die BO ausgewählt wurden. Vor dem Hintergrund der Berechnungskomplexität ist die Verwendung der Mittelwertfunktion ebenfalls vorzuziehen, da die fiktiven Daten der Simulation die Datenmenge rasch ansteigen lassen und damit gleichermaßen die Berechnungsdauer zunimmt. Problematisch an der Verwendung der Mittelwertfunktion als Schnittstelle ist lediglich die Einbindung von zu stark fehlerbehafteten Annahmen. In [Sch19] wurde hierzu bspw. ein einfaches Pendelsystem betrachtet und die Effizienz der BO bzgl. verschiedener mechanischer Parametrierungen untersucht. Als Ergebnis wurde folgender Zusammenhang herausgearbeitet: Je größer die Abweichung der Parametrierung zum realen System, desto niedriger die Effizienz des Verfahrens. Ab einer gewissen Grenze konnte sogar eine Verschlechterung gegenüber dem Fall ohne Vorwissen, d. h. $m(\cdot) = 0$, beobachtet werden. Anschaulich kann dieser Fall so interpretiert werden, dass die wahre Gütfunktion eine Sinusschwingung darstellt und als Mittelwertfunktion der Cosinus benutzt wird. Dabei benötigt die BO zahlreiche (zusätzliche) Auswertungen, um den Fehler, welcher durch die Mittelwertfunktion induziert wird, auszugleichen, sodass insgesamt die Effizienz deutlich sinkt. Letztendlich steht der Anwender der BO vor der Frage, in wie weit das vorhandene Vorwissen eingebunden werden soll bzw. welche Prädiktionsfähigkeit es besitzt. Insgesamt betrachtet ist die Mittelwertfunktion in den überwiegenden Fällen vorzuziehen. Dies wird bspw. durch die Arbeit in [DFE20] motiviert. Hierin wird explizit die Rolle der Mittelwertfunktion im Rahmen der BO umfassend durch ein breites Spektrum von gängigen Testfunktionen untersucht und ein positives Fazit für die Verwendung ausgesprochen.

Das Ziel dieses vorliegenden Kapitels ist, die BO für die Inbetriebnahme eines komplexen Anwendungsbeispiels - dem Ultraschalldrahtbonden - auszulegen. Sowohl eine theoretische als auch praktische Umsetzung dieses Konzepts wurde in der Literatur bisher nicht beschrieben und stellt daher eine neuartige Entwicklung dar. Es soll eine optimale Parametrisierung einer vorab festgelegten Ansatzfunktion für die Steuerung direkt am realen System identifiziert werden. Der Ultraschalldrahtbondprozess zeichnet sich dadurch aus, dass ein modellbasierter Entwurf nach momentanen Stand der Forschung aufgrund der schwierigen Modellbildung nicht realisierbar ist. Die vorliegende Arbeit setzt hier mit der Verwendung der BO an und ermöglicht somit eine effiziente und praktische Auslegung der Steuerung. Um dieses Ziel zu erreichen, ist es notwendig eine an das System angepasste Variante der BO zu entwickeln, die die Besonderheiten in Hinblick auf die Entwurfskriterien adäquat berücksichtigt. Zudem wird zur Effizienzsteigerung das vorhandene Expertenwissen bestmöglich ausgenutzt. Um das entwickelte Verfahren zu va-

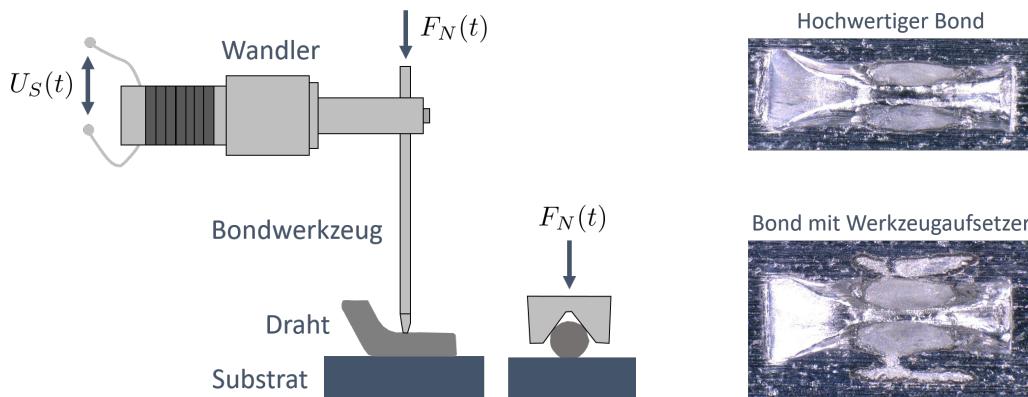


Bild 3-2: Schematischer Aufbau und Komponenten [SKHS21] des Ultraschalldrahtbondprozesses (links) und zwei exemplarische Aufnahmen einer Aluminium Bondverbindung aus der Draufsicht mittels Lichtmikroskops (rechts).

lidieren, werden verschiedene Versuchsreihen am Bondautomaten durchgeführt und die allgemein vorteilhaften Eigenschaften der BO bewertet. Nachfolgend wird zunächst eine vollständige Einführung in den Ultraschalldrahtbondprozess gegeben.

3.3 Anwendungsbeispiel: Ultraschalldrahtbondprozess

Ultraschalldrahtboden ist ein Festkörperfügeprozess [Har10; HSB⁺20]. Dabei handelt es sich um eine Standardtechnologie für die Herstellung elektrischer Verbindungen in der Mikro- und Leistungselektronik und wird zudem auch in der Batterieproduktion im Rahmen der Elektromobilität verwendet. Bild 3-2 zeigt die Hauptkomponenten eines Ultraschalldrahtbondprozesses. Eine oszillierende Relativbewegung zwischen Draht und Substrat (Untergrund) wird durch eine Wechselspannung $U_S(t)$ mit Ultraschallfrequenzen, üblicherweise im Bereich von 40 bis 150 kHz, induziert, die auf einen piezoelektrischen Wandler (*Transducer*) angelegt ist. Der Transducer wandelt die elektrische Anregung in mechanische Vibrationen um, die über ein Bondwerkzeug in die Verbindungszone übertragen wird. Das Bondwerkzeug drückt den Draht mit einer Normalkraft $F_N(t)$ auf das Substrat. Die beiden metallischen Partner, z. B. Aluminiumdraht auf einem goldbeschichteten Substrat, werden durch Interdiffusion und Bildung von intermetallischen Verbindungen kalt zusammengeschweißt. Die Vibration wird für eine Prozesszeit von ungefähr 300 ms in Abhängigkeit des Drahtdurchmessers erzeugt. In der Regel wird der Transducer während der Prozessdauer durch einen zugrundeliegenden Frequenzregler in Resonanz gehalten, um die maximale Amplitude des Werkzeugs zu erhalten. Im Rahmen dieser Arbeit wird davon ausgegangen, dass dieser Frequenzregler vorhanden und passend eingestellt ist bzw. keiner weiteren Optimierung bedarf. Nach der Erzeugung der ersten Bondverbindung wird in einem herkömmlichen Prozess eine Drahtschleife zu einem zweiten Ort gezogen. An der zweiten Stelle wird der Draht ebenfalls mit dem Un-

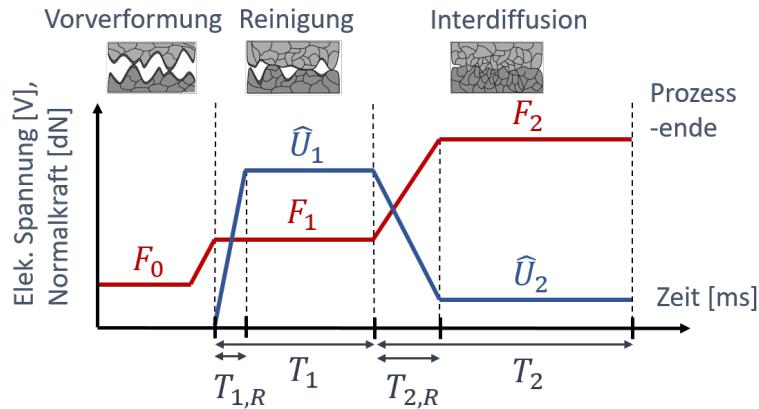


Bild 3-3: Exemplarischer Verlauf der parametrisierten Steuerung für den Ultraschalldrahtbondprozess.

tergrund verbunden und daraufhin durch ein Messer abgeschnitten. Für diese Arbeit ist die Betrachtung der ersten Verbindung ausreichend, sodass der Draht direkt im Anschluss an die erste Verbindungsstelle abgeschnitten wird. Bild 3-2 zeigt hierzu auf der rechten Seite zwei verschiedene Bondverbindungen an deren rechter Seite der scharfe Schnitt des Messers erkennbar ist.

Die Steuerung des Prozesses ist durch die Normalkraft $F_N(t)$ und die Spannungsamplitude $\hat{U}_S(t)$ definiert. In Bild 3-3 ist die vorgeschlagene parametrisierte Steuerfunktion $\mathbf{u}(t; \boldsymbol{\theta}) = [F_N(t; \boldsymbol{\theta}), \hat{U}_S(t; \boldsymbol{\theta})]^T$ für die beiden Eingänge dargestellt. Die genaue Form der Steuerung wird durch $\boldsymbol{\theta} = [F_0, F_1, F_2, \hat{U}_1, \hat{U}_2, T_1, T_2]^T \in \mathbb{R}^7$ charakterisiert. Die Übergangszeiten (Rampenlängen) werden auf 25% der jeweiligen Gesamtphasenzeit T_1, T_2 eingestellt. Der formelmäßige Zusammenhang der Steuerung ist im Anhang unter (A3-1) angegeben. Die Steuerung basiert auf dem Expertenwissen verschiedener Einrichter des Prozesses und hat sich bereits über mehrere Jahre hinweg als besonders zuverlässig und geeignet erwiesen, um qualitativ hochwertige Bondverbindungen zu erzeugen.

Physikalisch gesehen, besteht der Bondprozess aus vier Phasen, wobei eine kontinuierliche Überlappung zwischen den letzten drei stattfindet [Gei09; LTW17]: In der ersten Vordeformations-Phase wird der Draht zunächst ohne Vibrationseinbringung auf das Substrat gedrückt und verformt. Wenn das Bondwerkzeug zu vibrieren beginnt, wird eine Relativbewegung zwischen Draht und Substrat verursacht. Dies entfernt Verunreinigungen wie Staubpartikel oder Oxidschichten und reduziert zudem die Rauheit beider Oberflächen. Die Phase dient somit der Reinigung und Aktivierung. Danach folgt eine Phase der größeren plastischen Verformung des Drahtes, unterstützt durch den ultraschallbedingten Erweichungseffekt [USA⁺14]. In der abschließenden Diffusions-Phase nimmt die Kontaktfläche zu und beide Kontaktpartner diffundieren ineinander bis zum Ende des Prozesses, was letztendlich zu einer soliden (Bond-)Verbindung führt. Eine geeignete Steuer-

funktion zur Erzeugung stabiler Bondverbindungen unterstützt die Bildung dieser physikalischen Phasen, muss sie jedoch nicht direkt in ihren Parametern widerspiegeln. Basierend auf Expertenerfahrung hat die Steuerfunktion daher drei Phasen. Die erste Phase ist die Vordeformationsphase mit der initialen Kraft F_0 . Die Dauer dieser Phase ist irrelevant, sollte allerdings ausreichend sein, um eine anfängliche Kontaktfläche hervorzurufen. Die zweite Steuerphase, mit den Parametern (T_1, F_1, \hat{U}_1) , deckt die physikalischen Prozesse der Reinigung und Aktivierung, sowie die erste Verformung und Interdiffusion grob ab. Die letzte Steuerphase enthält die Parameter (T_2, F_2, \hat{U}_2) und ist für den größten Teil der Verformung und für die finale Verschweißung verantwortlich.

Es gibt mehrere Kriterien, die die Qualität einer Verbindung definieren. Gemäß [DVS17] werden insbesondere die Scherfestigkeit und optische Kriterien, wie die schädliche Kollision des Werkzeugs mit dem Substrat, berücksichtigt. Die Scherfestigkeit einer Verbindung wird mit einer Kraftmessdose in Kombination mit einem Schermeißel gemessen. Der Meißel bewegt sich über die Verbindung in einer bestimmten geringen Höhe über dem Substrat, während gleichzeitig die maximale Kraft F_S bis zum Bruch der Verbindung gemessen wird. Es handelt sich dabei also um eine zerstörende Messmethode, die pro Bondverbindung nur einmal durchgeführt werden kann. Ein Kontakt zwischen Werkzeug und Substrat kann auftreten, wenn die verwendete Steuerung zu viel Energie in das System einführt. In diesem Fall verformt sich der Draht zu stark und die Kanten des Bondwerkzeugs kollidieren mit dem Substrat, wodurch beide Komponenten möglicherweise beschädigt werden. Dieses Szenario muss, insbesondere beim Bonden auf empfindlichen rissanfälligen Substraten wie Halbleiterplatten, während des Betriebs vermieden werden. Beim Steuerungsentwurf ist es unerwünscht, aber zur Exploration zulässig. Aufgetretene Kollisionen können mit einem optischen Mikroskop nach dem Prozess erkannt werden. Bild 3-2 zeigt hierzu auf der rechten Seite eine Verbindung ohne (oben) und mit (unten) einem Werkzeugaufsetzer.

Der Prozessfähigkeitsindex C_{pK} ist ein wichtiges statistisches Maß in der Industrie und wird in diesem Kontext verwendet, um die Qualität des Bondprozesses zu quantifizieren. Er hängt vom Mittelwert und der Varianz der Scherkraft ab, die durch Prozess- und Messrauschen negativ beeinflusst wird. Daher ist die Scherkraft als eine Zufallsvariable mit dem Mittelwert $\mathbb{E}[F_S(\theta)]$ und der Varianz $\mathbb{V}[F_S(\theta)]$ aufzufassen. Der Prozessfähigkeitsindex [DVS17] wird dann definiert durch

$$C_{pK}(\theta) = \frac{\mathbb{E}[F_S(\theta)] - LSL}{3\sqrt{\mathbb{V}[F_S(\theta)]}}, \quad (3-9)$$

wobei LSL (*Lower Specification Limit*) die untere Spezifikationsgrenze ist. Sie bestimmt die minimale Scherkraft, die erreicht werden soll, und wird je nach Anwendungsfall und

Materialpaarung von einem Experten ausgewählt. Um den C_{pK} -Wert zu berechnen, wird der Mittelwert und die Varianz empirisch durch

$$\begin{aligned}\mathbb{E}[F_S(\boldsymbol{\theta})] &\approx \frac{1}{n_{rep}} \sum_{i=1}^{n_{rep}} F_S^{(i)}(\boldsymbol{\theta}) =: \mu_{F_S}(\boldsymbol{\theta}), \\ \mathbb{V}[F_S(\boldsymbol{\theta})] &\approx \frac{1}{n_{rep}-1} \sum_{i=1}^{n_{rep}} (F_S^{(i)}(\boldsymbol{\theta}) - \mu_{F_S}(\boldsymbol{\theta}))^2 =: \sigma_{F_S}^2(\boldsymbol{\theta}),\end{aligned}$$

approximiert, wobei $n_{rep} > 1$ (Repetition) separate Bonds mit derselben zugrundeliegenden Steuerung erzeugt werden, was zu den Messdaten $F_S^{(i)}(\boldsymbol{\theta}), i = 1, \dots, n_{rep}$ führt. Werkzeugkollisionen und andere optische Kriterien werden durch die binäre Variable $g(\boldsymbol{\theta})$ erfasst, wobei 0 eine gute Verbindung und 1 eine Verbindung mit einem optischen Mangel repräsentieren. Genauer gesagt, wenn mindestens eine der n_{rep} Verbindungen einen optischen Mangel aufweist, wird $g(\boldsymbol{\theta})$ auf 1 gesetzt.

Der Entwurf der parametrisierten Steuerung für den Ultraschalldrahtbondprozess wird dann mathematisch als das folgende restriktierte Optimierungsproblem formuliert:

$$\boldsymbol{\theta}_* = \arg \max_{\boldsymbol{\theta}} C_{pK}(\boldsymbol{\theta}), \quad \text{u.B.v.} \quad g(\boldsymbol{\theta}) = 0. \quad (3-10)$$

Dieses Problem wird in der Praxis in der Regel durch manuelles Ausprobieren gelöst, da es bisher keine automatisierte Lösungsstrategie in der Industrie gibt. Das Optimierungsproblem hängt von vielfältigen Rahmenbedingungen ab, womit es nicht ausreicht es einmalig zu lösen. In der Praxis treten bspw. unterschiedliche Materialpaarungen auf, die eine Anpassung der Steuerungsparameter erfordern. Außerdem kann die genaue Geometrie des Substrats oder der Drahtdurchmesser einen Einfluss auf die optimale Parametrisierung haben. Aus diesen Gründen muss das Optimierungsproblem je nach Anwendungsfall aufwendig neu durch einen Einrichter gelöst werden. Das Ziel dieser Arbeit ist, das Problem mittels BO zu lösen und damit vom menschlichen Anwender größtenteils zu entkoppeln. Die Vorteile dieser Vorgehensweise sind vielfältig, werden aber vor allem in einer Zeit- und Kostenersparnis gesehen. Des Weiteren wird durch den Einsatz der BO die Möglichkeit geschaffen, einen Steuerungsentwurf im Sinne der Regelungstechnik durchzuführen. Dieser Entwurf ist gegenüber der manuellen Auslegung von Vorteil, da er nicht auf einer subjektiven Einschätzung des menschlichen Einrichters basiert, sondern nach objektiven Kriterien der BO bewertet wird. Der Ansatz dieser Arbeit ist es jedoch nicht das Expertenwissen unberücksichtigt zu lassen, sondern über die erwähnten GP-Schnittstellen anderweitig einzubeziehen. In diesem Zuge ist die bereits erläuterte Wahl der Steuerungsfunktion aus Bild 3-3 eine Form der Einbringung von Fachwissen. Durch

ihre Wahl wird der Suchraum an möglichen Steuerungsverläufen drastisch eingeschränkt und trägt somit zu einer Effizienzsteigerung bei.

Den Schlussfolgerungen aus dem Abschnitt zum Stand der Forschung folgend, ist die Mittelwertfunktion $m(\theta)$ des GP eine einfache und nachvollziehbare Möglichkeit, um Vorwissen über die Steuerungsaufgabe (3-10) innerhalb der BO aufzugreifen. Mit der Kenntnis eines simplen Simulationsmodells könnte hierbei eine Auswertung des C_{pK} -Werts vorgenommen werden, welche wiederum als geeignete Mittelwertfunktion eingesetzt werden könnte. Im Grunde ist bereits eine Schätzung der Scherfestigkeit F_S ausreichend, da der C_{pK} -Wert im wesentlichen von diesem abhängt und für die Varianz $\mathbb{V}[F_S]$ der Einfachheit halber ein konstanter Faktor angesetzt werden kann (s. (3-9)). Das nächste Kapitel widmet sich daher der Aufstellung eines geeigneten Simulationsmodells für den Ultraschalldrahtbondprozess, welches auf der Grundlage von physikalischen Gesetzmäßigkeiten hergeleitet wird.

3.4 Physikalische Modellbildung für das Ultraschalldrahtboden

Die physikalische Modellbildung des Ultraschalldrahtbondprozesses wird innerhalb der Literatur vielseitig diskutiert. Die Modelltiefen unterscheiden sich dabei deutlich und reichen von einfachen Einmassenschwingern mit besonderen Reibcharakteristiken bis hin zu komplexen (FEM-)Modellen. Im Rahmen dieses Unterkapitels sollen einige der Modelle vorgestellt und ihre Vor- bzw. Nachteile erläutert werden. Des Weiteren soll ihre Eignung als Vorwissen im Rahmen der BO besprochen werden. Bild 3-4 zeigt auf der linken Seite einen klassischen mechanischen Modellansatz, welcher sich aus einer Verkettung von Massen und Feder-Dämpfer-Elementen zusammensetzt, und auf der rechte Seite ein hoch diskretisiertes FEM-Modell, wobei die Druckverteilung in der Kontaktfläche hervorgehoben ist [Ung17]. Im Nachfolgenden werden diese und verwandte Modellierungsarten näher beschrieben. Dabei verfolgen alle Modelle das Ziel, die zeitliche Entwicklung der Scherfestigkeit $F_S(t)$ genau wiederzugeben bzw. vorherzusagen.

Ein erster einfacher Modellierungsansatz besteht darin, nur die Drahtmasse mit einer gefederten Verbindung zum Untergrund zu betrachten. Das Masse-Feder-System auf der linken Seite von Bild 3-4 wird hierzu als Anschauung verwendet. Genauer genommen, wird nur die Masse (Draht) auf der ganz rechten Seite betrachtet, welche durch eine parallele Feder-Anordnung mit dem Substrat verbunden ist. Der gesamte linke Teil wird also

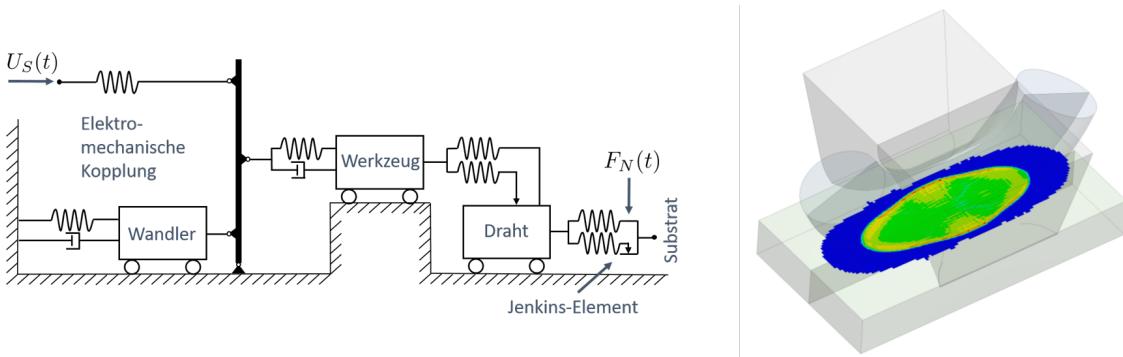


Bild 3-4: Ansätze zur physikalischen Modellierung des Ultraschalldrahtbondprozesses [Ung17]. Mehrmassenschwinger mit Jenkins-Element (links) und Finite-Elemente-Methode zur Beschreibung der Kontaktfläche (rechts).

zunächst vernachlässigt. Das beschriebene Modell lässt sich durch die Differentialgleichungen (angelehnt an [Ung17; SKHS21])

$$\begin{aligned}\dot{x}_1(t) &= \dot{s}_D(t), \\ \dot{x}_2(t) &= \ddot{s}_D(t) = \frac{1}{m_D}(F_T(t) - c_D s_D(t) - F_J(t)), \\ \dot{x}_3(t) &= \dot{F}_J(t) = c_J \dot{s}_D(t)(1 - \frac{1}{2}(1 + \text{sign}(\dot{s}_D(t))F_J(t))) \left| \frac{F_J(t)}{\mu(W_\mu(t))F_N(t)} \right|, \\ \dot{x}_4(t) &= \dot{W}_\mu(t) = \frac{1}{2}F_J(t)\dot{s}_D(t)(1 + \text{sign}(\dot{s}_D(t))F_J(t)) \left| \frac{F_J(t)}{\mu(W_\mu(t))F_N(t)} \right|,\end{aligned}\quad (3-11)$$

mathematisch beschreiben. Darin repräsentiert $\boldsymbol{x} = [s_D, \dot{s}_D, F_J, W_\mu]^T$ den Zustand des Systems, welcher wiederum aus der Wagenposition s_D des Drahts, seiner Geschwindigkeit \dot{s}_D , einer internen Kraftkomponente F_J und der geleisteten Reibungsenergie W_μ besteht. Die Eingänge in das System werden mit Hilfe der wirkenden Normalkraft F_N und der Tangentialkraft F_T , welche horizontal am Wagen bzw. Draht wirkt, beschrieben. Zur Vereinfachung wird zunächst die Tangentialkraft als Ursache für die Vibration der Drahtmasse angesehen. Um einen Zusammenhang zur angelegten elektrischen Spannung U_S , welche den eigentlichen Eingang darstellt, herzustellen, kann eine Proportionalität mit $F_T \propto U_S$ angenommen werden. Des Weiteren enthalten die Differentialgleichungen spezifische mechanische Parameter: c_D, c_J sind Federsteifigkeiten und $\mu(W_\mu(t))$ ein energieabhängiger Reibungswert, welcher die Stärke der Anbindung repräsentiert. (3-11) beschreibt auf einfache Weise die Relativbewegung zwischen dem Draht und dem Untergrund. Die Besonderheit des Modells ist die Verwendung des sogenannten Jenkins-Elements (s. Bild 3-4) in Kombination mit einer Evolutionsgleichung [Ung17], welches den Haft- und Gleitreibungszustand charakterisiert und den Übergang kontinuierlich fließend abbildet. Die Differentialgleichung bzgl. \dot{F}_J beschreibt in diesem Zusammenhang verschiedene Fälle, die das Element annehmen kann. Unter anderem muss die Normal-

kraft F_N einen passenden Wert annehmen, da ansonsten keine Reibungsenergie erzeugt wird. Ist die Kraft zu gering, gleitet der Draht wirkungslos über den Untergrund. Ist die Kraft hingegen zu groß, wird der Draht so stark auf den Untergrund gedrückt, dass er sich elastisch mit dem Untergrund bewegt bzw. an diesem haftet und keine wirksame Reibung entstehen kann. Neben der Normalkraft ist die Tangentialkraft F_T ein weiterer Steuereingang, durch die eine sinusförmige Schwingung des Drahtes vorgegeben wird. Eine optimale Ansteuerung wird im Rahmen des Modells durch das Integratorverhalten der gesamten eingetragenen Reibungsenergie W_μ beschrieben, welche sich proportional zur Scherkraft F_S verhält und somit im Idealfall ein kontinuierliches Anwachsen widerspiegeln sollte.

Das auf diesem Wege hergeleitete Modell (3-11) beschreibt den Bondprozess physikalisch und für den Menschen nachvollziehbar. In Hinblick auf das reale System und die schematische Darstellung in Bild 3-2 lässt sich das Modell auf der ausgewählten Modellierungstiefe folgendermaßen erweitern: In Bild 3-4 sind weitere Massen für das Bondwerkzeug und den Wandler vorgesehen. Die Verbindung der Massen erfolgt durch Feder-Dämpfer-Elemente, die insgesamt betrachtet zu eigenen (Ein-)Schwingvorgängen führen. Des Weiteren deutet der vertikale Balken die elektro-mechanische Verbindung der Piezo-elektronik an, die durch ein Hebelgesetz definiert wird. Somit lässt sich anstelle der eingeführten Tangentialkraft F_T die ursprüngliche elektrische Spannung U_S am Transducer als Stellgröße für das erweiterte Modell verwenden. Das erweiterte Modell zeichnet sich dadurch aus, dass es neben der Drahdynamik die Bewegungen der anderen Komponenten miteinbezieht und deren Schwingungs- bzw. Resonanzverhalten. Vor dem Hintergrund des Steuerungsentwurfs ist dieser jedoch komplizierter geworden, da nun die komplexere Wirkungskette vom Wandler bis hin zum Substrat zu berücksichtigen ist.

Das komplexe FEM-Modell [Ung17], welches auf der rechten Seite von Bild 3-4 zu sehen ist, kann auf der Basis des bereits erläuterten Modellansatzes nachvollzogen werden. Die Grundidee ist, die Kontaktfläche zwischen Draht und Untergrund gleichmäßig zu diskretisieren und für jedes finite Element die Differentialgleichungen aus (3-11) bzw. die des Jenkins-Elements anzusetzen. Unter der Annahme einer parabelförmigen Druckverteilung, die durch die wirkende Normalkraft und die elastische Verformung des Drahtes entlang der vertikalen Richtung hervorgerufen wird, werden die einzelnen Diskretisierungspunkte unterschiedlich angeregt. Die elastische Verformung wird dabei über ein unterlagertes FEM-Modell für das Materialverhalten des Drahtes verwendet. Aufgrund der feinen Diskretisierung wächst in jedem Element die Reibungsenergie auf unterschiedliche Weise an, sodass sich lokal voneinander differenzierbare Scherkräfte beobachten lassen. Repräsentativ ist dies durch die Farben für die Druckverteilung in Bild 3-4 dargestellt. Demnach wachsen die lokalen Scherkräfte ellipsenförmig von innen nach außen an.

Die Scherkraft des Gesamtmodells wird nach Prozessende über eine gewichtete Summe über alle Elemente gebildet. Zusammengenommen ist das komplexe FEM-Modell dafür geeignet, um die lokalen Effekte innerhalb der Kontaktfläche wiederzugeben. Dieser Genauigkeitszuwachs wird jedoch durch eine deutlich erhöhte Berechnungszeit im Vergleich zum einfacheren Masse-Feder-System kompensiert und bietet vor dem Hintergrund des Gütemaßes bzw. der resultierenden Scherfestigkeit keine signifikanten neuen Informationen. Grundsätzlich lässt sich das Simulationsergebnis des Masse-Feder-Systems als die gewichtete Summe über die lokalen Scherfestigkeiten des FEM-Modells auffassen. Die Tauglichkeit des FEM-Modells für die Verwendung im Rahmen der BO als Vorwissen über den Prozess ist damit bereits in Frage zu stellen. Die lokalen Effekte sind im Rahmen des Steuerungsentwurfs nicht von Bedeutung, sodass das einfachere Modell vorzuziehen ist.

Ein wesentlicher Nachteil der bisher vorgestellten Modellierungsansätze stellt zudem die Berechnungszeit dar. Das FEM-Modell benötigt beispielsweise eine Berechnungsdauer in der Größenordnung von einer Stunde für die Simulation von den erwähnten 300 ms, die der Prozess für die Erzeugung einer Bondverbindung braucht. Die Ursache hierfür kann bei dem Diskretisierungsgrad des FEM-Modells gesehen werden, ist jedoch nicht maßgeblich. Das Hauptproblem liegt vor allem bei der zeitlichen Schrittweite des zugrundeliegenden Integrationsverfahrens. Das aufgeschaltete Steuersignal für die elektrische Spannung besitzt eine Frequenz von mehreren Kilohertz, womit die zeitliche Schrittweite in einem ähnlichen Bereich für eine stabile numerische Simulation liegen muss. Die notwendige und sehr geringe Schrittweite ist daher der Hauptgrund für die lange Berechnungszeit der bisher vorgestellten Modelle. In Hinblick auf den BO-Algorithmus 1 und die 4. Zeile, in welcher die Akquisitionsfunktion optimiert wird, ist eine lange Berechnungszeit der a-Priori-Mittelwertfunktion besonders unpraktikabel. Die Akquisitionsfunktion besitzt in der Regel viele lokale Optima, die vor dem Hintergrund einer globalen Optimierungsstrategie eine übermäßig hohe Anzahl von Funktionsauswertungen erfordert. Diese Anforderung können die bisherigen Ansätze nicht erfüllen.

Die angesprochene Problematik motiviert den Einsatz von einfacheren Modellen, die nicht auf der Ebene der Drahtbewegung agieren, sondern auf einer höheren Modellierungstiefe angesiedelt sind. Eine Variante besteht darin, nicht die exakte hochfrequente Schwingung zu betrachten, sondern sich nur auf ihre Amplitude zu fokussieren. Dadurch ist für die Simulation eine zeitliche Diskretisierung der Amplitudenverläufe vorzunehmen für die vergleichsweise eine beträchtlich größere Schrittweite ausreichend und damit

zielführender ist (vgl. Bild 3-3). In [SAS⁺18; SSH⁺20; SKHS21] schlagen die Autoren zwei verknüpfte und fallabhängige Differentialgleichungen über

$$\dot{\gamma}(t) = \begin{cases} c_\gamma P(t) & \text{Für } 0 < \gamma(t) \leq 1, \\ 0 & \text{sonst,} \end{cases} \quad \dot{\tau}(t) = \begin{cases} c_\tau \gamma(t) P(t) & \text{Für } P(t) \geq P_0, \\ 0 & \text{sonst,} \end{cases}$$

vor. Hierin stellt γ den Reinigungsgrad innerhalb der Kontaktfläche dar und τ ist mit der Scherfestigkeit gleichzusetzen. Des Weiteren ist P die eingetragene Reibleistung, welche durch die Steuerungsgrößen hervorgerufen wird. Der Reinigungsgrad γ verläuft zwischen den Grenzen 0 (ungereinigt) und 1 (vollständig gereinigt), kann aufgrund der Konstruktion nur monoton steigen und beinhaltet einen prozessabhängigen Skalierungsfaktor c_γ . Um einen Einfluss auf die Scherkraft zu haben, wird ein bestimmtes Mindestmaß an Leistung, durch die Einführung eines Schwellwerts P_0 , gefordert. Wird diese Grenze überschritten, nimmt der Scherkraftzuwachs τ proportional zum Reinigungsgrad γ über den Parameter c_τ zu. Dieser Ansatz zeichnet sich durch seinen einfachen und leicht nachvollziehbaren Aufbau aus und beschreibt sehr gut das Sättigungsverhalten der Reinigung und Scherkraft am Ende der Prozessdauer. Zudem werden nur wenige Parameter benötigt, um das Modell vollständig zu definieren. Ein Nachteil des Modells ist, dass es nur bedingt aussagekräftig und in seiner Prädiktionsfähigkeit bzgl. der Scherkraft stark eingeschränkt ist.

Eine etwas detailliertere Sichtweise auf der Beschreibungsebene der Amplituden wird in [Gau09] vorgeschlagen. Der Bondvorgang wird darin über die Gleichungen

$$\begin{aligned} \dot{\gamma}(t) &= k_1(t) + \left(k_2(t) - k_1(t) - \frac{dA(t)}{dt} \frac{1}{A(t)} \right) \gamma(t) - k_2(t) \gamma(t)^2, \\ k_1(t) &= \frac{4f\mu_{ox}F_N(t)}{A(t)} (\hat{s}_0 - c\mu_{ox}F_N(t)), \\ k_2(t) &= \frac{4f}{A(t)} (c\mu_{ox}F_N(t))(\mu_{ox} - \mu_{met})F_N(t), \end{aligned}$$

ausgedrückt. Die resultierende Scherkraft lässt sich bei diesem Modell über den Zusammenhang $F_S = \gamma A$ bestimmen, wobei γ wiederum den Reinigungsgrad darstellt und A die Anbindungsfläche im Kontakt wiedergibt. Das Modell hängt von etwaigen (mechanischen) Parametern ab. So gehen neben der Schwingungsfrequenz f , auch die Steifigkeit der Verbindung c und spezifische Reibungswerte für die Oxid- μ_{ox} und metallische μ_{met} Schicht ein. Des Weiteren stellt \hat{s}_0 die Amplitude des Drahtes im unbelasteten frei-schwingenden Fall dar, welche durch die zunehmende Belastung durch die Anbindung vermindert wird. Eine besondere Charakteristik des Modells ist die quadratische Abhängigkeit bzgl. γ . Dieser Ansatz wird in [Gau09] durch verschiedene experimentelle Messreihen begründet und basiert daher nicht direkt auf physikalischen Ge-

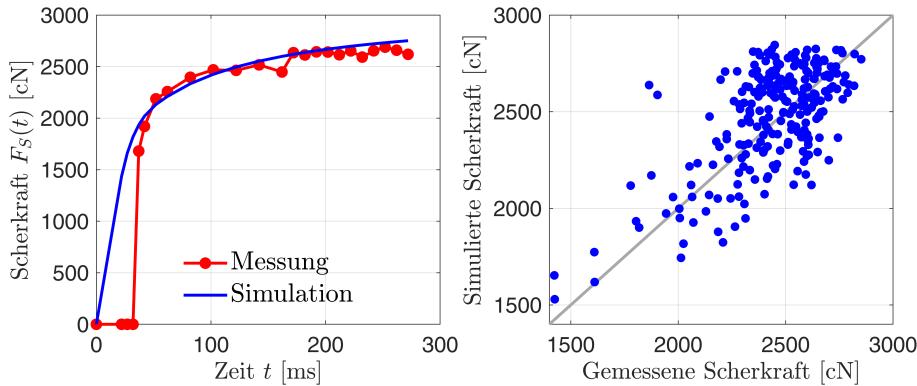


Bild 3-5: Zusammenfassung der Validierungsergebnisse aus [Her20]. Gegenüberstellung von Messung und Simulation in Bezug auf eine zeitliche Entwicklung des Scherkraft für eine bestimmte Steuerparametrisierung (rechts) und gesamtheitliche Betrachtung der gemessenen und simulierten Scherkräfte am Prozessende (links).

setzmäßigkeiten. Der große Vorteil dieses Modells ist, dass es eine Balance zwischen Genauigkeit und Berechnungsdauer herstellt. Unter der Voraussetzung eines konstanten Wertes für A muss lediglich eine Differentialgleichung erster Ordnung ausgewertet werden, wodurch die Auswertung für eine bestimmte Parametrisierung sehr schnell vorgenommen werden kann. An dieser Stelle ist außerdem die aktuelle Arbeit aus [vBS⁺20] zu erwähnen, in welcher der Ansatz aus [Gau09] aufgegriffen und durch weitere Effekte und Annahmen ergänzt wird. Darunter fallen bspw. die Drahthöhenabnahme, die während des Schweißvorgangs durch das Einsinken des Drahtes in das Substrat stattfindet und eine parabelförmige Druckverteilung in der Kontaktfläche. Durch den zusätzlichen Detaillierungsgrad steigt allerdings auch die Anzahl der Differentialgleichungen des Zustandsraummodells auf drei, wodurch der Einsatz des Modells innerhalb der BO, welche eine schnelle Auswertung des Modells erfordert, wiederum in Frage gestellt wird.

Einen detaillierteren Überblick und eine vollständige Analyse des Modells aus [Gau09] wurde in der studentischen Arbeit [Her20], welche zu einem großen Anteil auf den Vorarbeiten des Autoren dieser Arbeit basiert, ausgearbeitet. Auf dieser Grundlage wurde ein eigenes Simulationsmodell hergeleitet, das den Anforderungen für den Einsatz der BO bestmöglich begegnet. Unter anderem wurde darauf geachtet, dass das Modell wenige Parameter besitzt, um den Identifikationsaufwand auf ein Minimum zu reduzieren und vergleichsweise schnell bzgl. verschiedener Parametrisierungen ausgewertet werden kann. Die Ergebnisse der Studienarbeit sind in Bild 3-5 dargestellt. Das linke Teilbild visualisiert die Scherkraft-Trajektorien einer Messung und der zugehörigen Simulation. Die zeitliche Messung der Scherkraft ist in der Praxis äußerst aufwendig, da für jeden Diskretisierungspunkt ein eigener Bond erzeugt werden muss und der Prozess an dem passenden Zeitpunkt abgebrochen werden muss, um die aktuelle Scherkraft an dieser Stelle über das

zerstörerische Abscheren ermitteln zu können. Mit Hilfe des Modells wird der gemessene Verlauf sehr gut wiedergegeben, wobei zu berücksichtigen ist, dass dem Verlauf lediglich eine gewisse Parametrisierung zugrunde liegt. Da im Rahmen der BO die Scherfestigkeit am Ende des Prozesses bzw. der Trajektorie von entscheidender Bedeutung ist, reicht es aus, die Endergebnisse aus Messung und Simulation miteinander zu vergleichen. Das Teilbild auf der rechten Seite zeigt hierzu die Auswertung eines Datensatzes, welcher 240 separate Einzelmessungen enthält. Das bedeutet, dass hinter jedem blauen Datenpunkt eine eigene Bonderzeugung mit einer bestimmten Parametrisierung steht. Für die Simulation wurde die selbe Parametrisierung benutzt, sodass das Diagramm insgesamt die Prädiktionsgenauigkeit des entwickelten Modells widerspiegelt. Im Idealfall sollten sich die Punkte genau auf der diagonalen (grauen) Linie befinden, um eine hohe Genauigkeit des Modells anzugeben. Tatsächlich reflektiert die hohe Streuung der Daten das genaue Gegenteil. Das Modell ist demnach zwar in der Lage, eine einzige Scherkraft-Trajektorie (vgl. Bild 3-5, links) gut wiederzugeben, versagt jedoch bei der Generalisierung für unterschiedliche Steuerungen (rechts). In den erwähnten Veröffentlichungen findet diesbezüglich keine Auswertung auf einen vergleichbar großen Datensatz statt, sodass nach aktuellem Stand der Forschung kein Modellansatz in der Lage ist, das Generalisierungsproblem geeignet zu adressieren.

Aufgrund dieser Tatsache ist die Einbindung von Vorwissen in Form eines Dynamikmodells des Ultraschalldrahtbondprozesses im Rahmen der BO als problematisch einzustufen. Die Anforderungen, um für die BO gewinnbringend eingesetzt werden zu können, sind nach aktuellem Stand nicht mit der komplizierten und mühsamen Modellbildung in Einklang zu bringen. Etwaige Voruntersuchungen haben in diesem Zuge gezeigt, dass das aufgestellte Modell aus Bild 3-5 zu keiner Effizienzsteigerung der BO beitragen konnte. Abseits der Genauigkeit des Modells gibt es weitere Aspekte, die nur schwerlich über physikalische Gesetzmäßigkeiten abzubilden sind. Zum einen kann in Bezug zu (3-9) festgehalten werden, dass es sich bei dem Bondprozess um einen sensitiven und hochgradig stochastischen Ablauf handelt. Ein Simulationsmodell ist vor diesem Hintergrund zwar in der Lage eine Aussage über die resultierende Scherkraft zu machen, jedoch nicht über die zugehörige Streuung bzw. Varianz. Um diese Größe abbilden zu können, müsste eine weitaus umfassendere Modellierung auf der Grundlage von stochastischen Differentialgleichungen unternommen werden. Des Weiteren tritt in der Praxis unter Umständen das sogenannte Überbonden auf, wobei zu viel Energie in die Bondverbindung eingebracht wird. Das monotone Anwachsen der Scherkraft muss daher relativ zu den Modellannahmen betrachtet werden, denn beim Überbonden kann die Kraft mit zunehmender Zeit auch wieder abnehmen und die Bondqualität mindern. Keines der Literatur bekannten physikalischen Modelle kann diesen Effekt im Ansatz beschreiben. Gesamtheitlich betrachtet ist das Ultraschalldrahtbonden daher ein Beispiel für ein schwierig zu modellierendes

System, wobei selbst mit erheblichen Aufwand kein ausreichend gutes Modell aufgestellt werden kann, welches einen modellbasierten Steuerungsentwurf ermöglichen würde. Die BO stellt in diesem Zusammenhang eine ideale Möglichkeit dar, um die komplizierte Modellbildung zu umgehen und dennoch einen adäquaten Steuerungsentwurf durchführen zu können. Ein wesentlicher Vorteil ist dabei, dass dem stochastischen Charakter des Systems bereits durch den eingesetzten GP und der Konstruktion des Verfahrens begegnet wird. Dessen ungeachtet, benötigt die Standardroutine der BO (s. Algorithmus 1) diverse Anpassungen, um für den Steuerungsentwurf beim Ultraschallbonden eingesetzt zu werden. Im nachfolgenden Abschnitt werden diese Anpassungen erläutert, wobei darauf geachtet wird, möglichst viele Vorkenntnisse über den Prozess in das Verfahren einzubinden.

3.5 Angepasste Bayessche Optimierung für den Steuerungsentwurf beim Ultraschalldrahtbonden

In diesem Unterkapitel wird die angepasste Variante der BO vorgestellt. Das Entwurfsziel besteht darin, dass Optimierungsproblem (3-10) am realen System durch direkte Interaktion zu lösen. Dabei wird die Parametrisierung bzw. Steuerung identifiziert, welche den höchsten Prozessfähigkeitsindex liefert und gleichzeitig keine Schädigung in Form von Werkzeugaufsetzern herbeiführt. Die Grundidee ist, drei GPs sowohl für den Mittelwert, als auch für die Standardabweichung der Scherkraft und für die Nebenbedingung bzgl. Werkzeugaufsetzer einzusetzen (vgl. (3-10)), wobei die nachfolgenden a-priori und a-posteriori Annahmen getroffen werden:

$$\begin{aligned} \hat{\mu}_{F_S}(\boldsymbol{\theta}) &\sim \mathcal{GP}(m_\mu(\boldsymbol{\theta}), k_\mu(\boldsymbol{\theta}, \boldsymbol{\theta}')), & p(\hat{\mu}_{F_S}(\boldsymbol{\theta}) \mid \mathbf{D}_\mu) &= \mathcal{N}(\mu_\mu(\boldsymbol{\theta}), \sigma_\mu^2(\boldsymbol{\theta})), \\ \hat{\sigma}_{F_S}(\boldsymbol{\theta}) &\sim \mathcal{GP}(m_\sigma(\boldsymbol{\theta}), k_\sigma(\boldsymbol{\theta}, \boldsymbol{\theta}')), & p(\hat{\sigma}_{F_S}(\boldsymbol{\theta}) \mid \mathbf{D}_\sigma) &= \mathcal{N}(\mu_\sigma(\boldsymbol{\theta}), \sigma_\sigma^2(\boldsymbol{\theta})), \quad (3-12) \\ \hat{g}(\boldsymbol{\theta}) &\sim \mathcal{GP}(m_g(\boldsymbol{\theta}), k_g(\boldsymbol{\theta}, \boldsymbol{\theta}')), & p(\hat{g}(\boldsymbol{\theta}) \mid \mathbf{D}_g) &= \mathcal{N}(\mu_g(\boldsymbol{\theta}), \sigma_g^2(\boldsymbol{\theta})). \end{aligned}$$

Für die Versuchsreihen der eigenen vorhergehenden Veröffentlichung [HHTT22] und dieser Arbeit wurde der Matérn-Kernel aus (2-35) für alle Kovarianzfunktionen verwendet. Obwohl derselbe Kernel für alle Kovarianzfunktionen verwendet wurde, unterscheiden sie sich dennoch in ihrer exakten Gestalt, da andere Hyperparameter gemäß (3-5) für jede Kovarianzfunktion festgelegt werden. Die Mittelfunktionsfunktionen in Bezug auf die Varianz der Scherkraft σ_{F_S} und die Einschränkung g wurden auf konstante Werte gesetzt: $m_\sigma = 60$ cN, $m_g = 0$. Im Fall der Einschränkung entspricht diese Annahme einer optimistischen Initialisierung, da angenommen wird, dass die Einschränkung im gesamten Parameterbereich nicht verletzt wird. Für die Mittelfunktion in Bezug auf μ_{F_S} werden zwei Fälle betrachtet. Im ersten Fall wird $m_\mu = \text{LSL} = 2500$ cN ebenfalls konstant

gesetzt, und es wird angenommen, dass kein besonderes weiteres Vorwissen vorhanden ist. Im zweiten Fall wird anstelle einer konstanten, eine quadratische Funktion mit $m_\mu(\boldsymbol{\theta}) = \boldsymbol{\theta}^T \mathbf{A} \boldsymbol{\theta} + \mathbf{b}^T \boldsymbol{\theta} + c$ verwendet, wobei die Größen $\mathbf{A}, \mathbf{b}, c$ über eine *Least-Squares*-Regression an einen bestehenden Datensatz angepasst wurden. Dieser Ansatz wird im Laufe dieses Unterkapitels noch näher erläutert und repräsentiert eine *Reverse Engineering* Strategie und soll stellvertretend für ein physikalisches Ersatzmodell stehen.

Um den Prozessfähigkeitsindex zu optimieren, wird eine Schätzung zu (3-9) mit

$$\hat{C}_{pK}(\boldsymbol{\theta}) = \frac{\hat{\mu}_{F_S}(\boldsymbol{\theta}) - \text{LSL}}{3\hat{\sigma}_{F_S}(\boldsymbol{\theta})},$$

vorgenommen, wofür die zugehörigen GPs für $\hat{\mu}_{F_S}$ und $\hat{\sigma}_{F_S}$ miteinander kombiniert werden müssen. Da \hat{C}_{pK} nichtlinear von diesen Größen abhängt, ist die zugehörige WahrscheinlichkeitsdichteVerteilung $p(\hat{C}_{pK})$ nicht mehr gaußförmig. Dennoch kann die genaue Verteilung analytisch berechnet werden [DR13]. Im Allgemeinen ist sie stark gewölbt und weist keine Momente auf. Die Form kann unimodal, bimodal, symmetrisch oder asymmetrisch sein. Die Autoren von [DR13] schlagen jedoch eine Normalapproximation über die Beziehung

$$\begin{aligned} p(\hat{C}_{pK}) &\approx \mathcal{N}(\mu_C(\boldsymbol{\theta}), \sigma_C^2(\boldsymbol{\theta})), \\ \mu_C(\boldsymbol{\theta}) &= \frac{\mu_\mu - \text{LSL}}{3\mu_\sigma}, \\ \sigma_C^2(\boldsymbol{\theta}) &= \left(\frac{\sigma_\sigma}{\sigma_\mu} \right)^2 \left(\left(\frac{\sigma_\mu}{3\sigma_\sigma} \right)^2 + \left(\frac{\mu_\mu - \text{LSL}}{3\mu_\sigma} \right)^2 \right) \end{aligned} \quad (3-13)$$

vor, welche für $\sigma_\sigma/\mu_\sigma \leq 0.1$ gültig ist und eine hohe Approximationsgüte aufweist. Für das betrachtete Anwendungsbeispiel wurde in vorläufigen Untersuchungen festgestellt, dass der Grenzwert im schlechtesten Fall nicht signifikant überschritten wird und damit die Normalapproximation angenommen werden darf.

Zur besseren Verständlichkeit der nachfolgenden Erklärungen wird auf Bild 3-6 und Algorithmus 2 verwiesen, welche die Schritte der angepassten BO-Implementierung zusammenfassen. Der Begriff *Batch-Element* wird später an entsprechender Stelle erläutert und kann für den Moment mit der identifizierten Parametrisierung gleichgesetzt werden, wo bei $n_{batch} = 1$ gilt. Bild 3-6 zeigt die Beziehungen der involvierten Funktionen während einer Iteration anhand eines fiktiven eindimensionalen Beispiels ($\theta \in \mathbb{R}$). Es wird angenommen, dass bereits fünf Auswertungen der realen unbekannten Funktionen vorliegen (initiale Datenbasis aus Zeile 1), welche durch die blauen und roten Kreise im oberen Bild dargestellt sind. Zusätzlich sind die wahre $C_{pK}(\theta)$ -Funktion und die Nebenbedingung $g(\theta)$ bzgl. Werkzeugaufsetzern durch die durchgezogene und gestrichelte grüne Linie re-

Algorithmus 2 Bayessche Optimierung für Ultraschalldrahtbonden

- 1: **Eingabe:** Initiale Datenbasis $(\mathbf{D}_\theta, \mathbf{D}_\mu, \mathbf{D}_\sigma, \mathbf{D}_g)$ aus n_{init} Experimenten, Iterationsbudget n_{budget} , Anzahl der Batch-Elemente n_{batch} , GP Mittelwert- und Kovarianzfunktionen $(m_\mu, m_\sigma, m_g, k_\mu, k_\sigma, k_g)$, Grenzwerte ξ_C, ξ_g , LSL.
 - 2: **Wiederhole** bis das Iterationsbudget n_{budget} aufgebraucht ist:
 - 3: Aktualisiere Gauß-Prozesse bzw. Hyperparameter $(\boldsymbol{\eta}_\mu, \boldsymbol{\eta}_\sigma, \boldsymbol{\eta}_g)$ nach (3-5).
 - 4: Berechne Batch-Elemente $\boldsymbol{\theta}_q^{(b)} = \arg \max_{\boldsymbol{\theta}} \alpha_{EI,g}(\boldsymbol{\theta}; \boldsymbol{\theta}_q), b = 1, \dots, n_{batch}$ (3-17).
 - 5: Evaluiere bei $\boldsymbol{\theta}_q^{(b)}$ und erhalte $(\mu_{F_s,q}^{(b)}, \sigma_{F_s,q}^{(b)}, g_q^{(b)})$ \triangleright Experiment(e) reales System
 - 6: Erweitere Datenmenge
 $(\mathbf{D}_\theta, \mathbf{D}_\mu, \mathbf{D}_\sigma, \mathbf{D}_g) \leftarrow (\mathbf{D}_\theta \cup \boldsymbol{\theta}_q^{(b)}, \mathbf{D}_\mu \cup \mu_{F_s,q}^{(b)}, \mathbf{D}_\sigma \cup \sigma_{F_s,q}^{(b)}, \mathbf{D}_g \cup g_q^{(b)})$.
 - 7: Berechne Prozessfähigkeitsindex $\mathbf{D}_C^{(i)} = \frac{\mathbf{D}_\mu^{(i)} - \text{LSL}}{3\mathbf{D}_\theta^{(i)}}, i = 1, \dots, n_{init} + n_{budget}$.
 - 8: **Ausgabe:** Optimale Parametrisierung $\boldsymbol{\theta}_j$,
 $j = \text{index } \max_{j, \mathbf{D}_{g,j}=0} \mathbf{D}_{C,j}, j = 1, \dots, n_{init} + n_{budget}$.
-

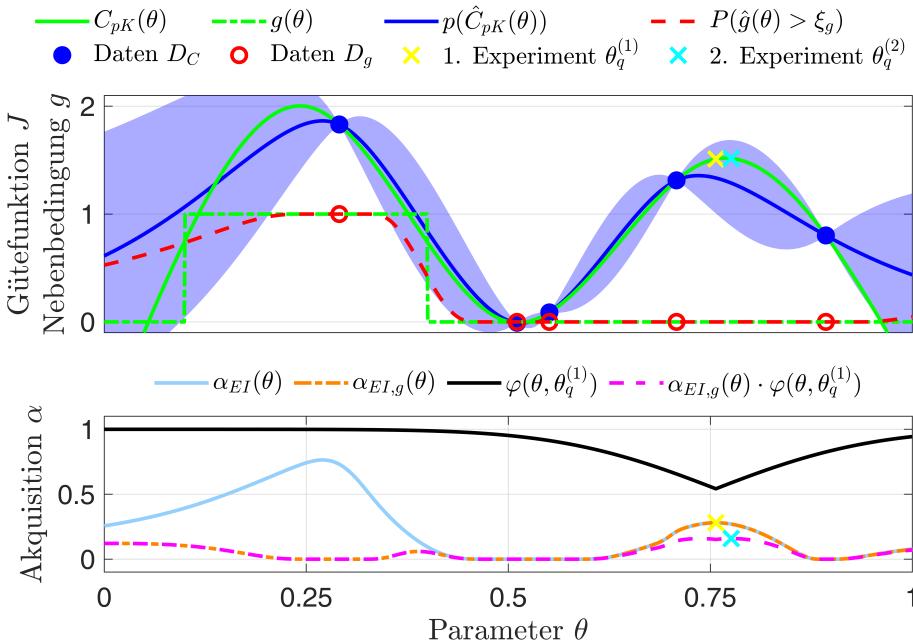


Bild 3-6: Anschauliche Darstellung für eine Iteration der entwickelten BO-Methode im eindimensionalen Fall mit einem Steuerungsparameter. Eine detaillierte Erklärung der Zusammenhänge wird im Fließtext gegeben.

präsentiert (Ground Truth). Bezüglich der Zielgröße ist zudem die aktuelle Schätzung des GP in Blau dargestellt. Hierbei stellt die durchgezogene Linie den Mittelwert $\mu_C(\theta)$ dar und der schattierte Bereich entspricht der Standardabweichung $\sigma_C(\theta)$ (s. (3-12) und (3-13)). Es wird davon ausgegangen, dass die Aktualisierung der GPs (Zeile 3) stattgefunden hat.

Ein wesentlicher Bestandteil aus Zeile 4 ist die Akquisitionsfunktion $\alpha(\theta)$, die das unterlagerte Optimierungsproblem definiert. Die Akquisitionsfunktion wird benötigt, um die Parametrisierung des nächsten Experiments aus den verfügbaren Informationsquellen der involvierten GPs abzuleiten. Es wird das EI-Kriterium $\alpha_{EI}(\theta) = \mathbb{E}[\max(0, \hat{C}_{pK}(\theta) - \xi_C)]$ nach der Berechnungsvorschrift aus (3-8) eingesetzt. Im unteren Teil von Abbildung 3-6 erfolgt die Auswertung des GPs bzgl. der Gütefunktion. Die durchgezogene, hellblaue Linie repräsentiert die Akquisitionsfunktion für die erwartete Verbesserung gemäß (3-8). Der zugehörige Schwellenwert ξ_C ist auf 1 festgelegt. Daher zeigt die Funktion in einem Bereich um den Parameterwert von 0.5 einen Wert von nahezu 0. Das Maximieren dieser Funktion würde dazu führen, dass die realen Funktionen aus der oberen Abbildung in der Nähe des ganz linken Datenpunkts ausgewertet würden. Dieser Punkt entspricht dem globalen Maximum. Jedoch wird die einzuhaltende Nebenbedingung in diesem Bereich nicht erfüllt ($g(\theta) = 1$ für $\theta \in (0.1, 0.4)$). Daher ist eine Auswertung an dieser Stelle nicht gewünscht und sollte im Rahmen der BO automatisiert als ungeeignet identifiziert werden. Im Nachfolgenden wird eine solche Automatisierung besprochen und anhand des eindimensionalen Beispiels weiter erläutert.

Durch die Berücksichtigung der Werkzeugkollisionen und die Einführung eines eigenen GPs (vgl. (3-12)) gehört das entwickelte BO-Verfahren zu der übergeordneten Kategorie des *Safe Reinforcement Learnings*. Hierbei handelt es sich um spezielle Verfahren, die eine gewisse Form von Sicherheit beim Entwurf fordern. In der Kombination mit der BO und regelungstechnischen Anwendungen wird der Sicherheitsaspekt in den Veröffentlichungen rund um [BKS23; BG20; MBK⁺21; WHB⁺21] besprochen. Dabei ist ein Ansatz, von einem sicheren bzw. stabilen Gebiet zu starten und dieses allmählich zu vergrößern. Die Exploration findet daher nur an Rändern des zuverlässigen Gebietes statt. Für Anwendungen für welche die Einhaltung von sicherheitskritischen Aspekten äußerst wichtig und als strikt aufzufassen ist, ist der Ansatz des sukzessiven Explorierens überaus geeignet. Vor dem Hintergrund des Ultraschalldrahtbondens ist eine Verletzung der Nebenbedingung während der Lernphase nicht als sonderlich kritisch zu beurteilen. Abgesehen von einer Zunahme des Verschleißes wird der Prozess nicht merklich negativ durch Kollisionen beeinflusst. Entscheidend ist das Endresultat und die identifizierte Steuerung, welche am Ende der Optimierung durch den BO-Algorithmus ausgegeben wird. Von ihr wird eingefordert, dass sie den bestmöglichen Prozessfähigkeitsindex unter

der zuverlässigen Vermeidung von Kollisionen erzeugt. Aus diesem Grund sind auftretende Kollisionen während der Optimierung zwangsläufig nicht hinderlich, sondern tragen zum Informationszuwachs über ungeeignete Gebiete bei. Dementsprechend muss die Nebenbedingung nicht hart eingehalten und mit dem oberen Ansatz einer Gebietserweiterung berücksichtigt werden. Zumal die Gebietserweiterung die grundsätzliche Effizienz der BO relativiert und als eher langsam aufgrund der beschränkten Explorationsfähigkeit einzustufen ist.

Aus diesem Grund wird in dieser Arbeit der Ansatz verfolgt, die Nebenbedingung durch einen Strafterm innerhalb der Gütfunktion bzw. Akquisitionsfunktion zu berücksichtigen. Hierfür wurde in (3-12) die Annahme getroffen, dass es sich bei der binären Markierung g um eine Variable handelt, die kontinuierliche Werte annehmen kann. Formal gesehen, ist dies eine Ungenauigkeit, da korrekterweise ein klassifizierender GP [Ras06] verwendet werden müsste. Im Gegensatz zu der kontinuierlichen GP-Variante sind hierfür jedoch diverse Approximationen notwendig (bspw. *Laplace*-Approximation oder Erwartungswertpropagationsverfahren, welche in [Ras06] ausführlich behandelt werden), um die zugrundeliegende Berechnungskomplexität zu reduzieren. Bei den Untersuchungen dieser Arbeit hat sich gezeigt, dass der Mehraufwand durch die Einführung eines solchen GPs nicht gerechtfertigt ist und zu keinerlei Verbesserung führt. Zudem kam es bei manchen Experimenten zu numerischen Problemen, die auf die notwendigen Approximationen zurückzuführen sind. Insgesamt betrachtet, unterscheidet sich die Ausgabe des kontinuierlichen nicht besonders von der des klassifizierenden GPs. In beiden Fällen liefert der GP eine Wahrscheinlichkeitsaussage über den Wert der Nebenbedingung. Diesbezüglich liefern beide Ansätze keine scharfe Gebietsgrenze, an welcher die Nebenbedingung ihren Wert abrupt ändert, sondern berechnen stattdessen einen glatten und kontinuierlichen Übergang. Dieser Übergang stellt die Unsicherheit des maschinellen Lernverfahrens über den Wert dar und spiegelt das stochastische Verhalten des Ultraschalldrahtbondens in geeigneter Weise wider.

Innerhalb dieses Kontexts ist die Wahrscheinlichkeitsdichtefunktion für die Nebenbedingung definiert als $p(\hat{g}(\boldsymbol{\theta}) \mid \mathcal{D}_g) = \mathcal{N}(\mu_g(\boldsymbol{\theta}), \sigma_g^2(\boldsymbol{\theta}))$ (vgl. (3-12)). Im Gegensatz zur Zielgröße liegt das Interesse hier nicht am konkreten Wert der Einschränkung, sondern vielmehr an der Wahrscheinlichkeit des Auftretens einer Werkzeugkollision. Daher erfolgt eine Integration über die Wahrscheinlichkeitsdichtefunktion ab einem bestimmten Grenzwert ξ_g :

$$P(\hat{g}(\boldsymbol{\theta}) > \xi_g) = \int_{\xi_g}^{\infty} \mathcal{N}(\mu_g(\boldsymbol{\theta}), \sigma_g^2(\boldsymbol{\theta})) d\hat{g} = \frac{1}{2} \left(1 + \operatorname{erf} \left(\frac{\xi_g - \mu_g(\boldsymbol{\theta})}{\sqrt{2\sigma_g^2(\boldsymbol{\theta})}} \right) \right), \quad (3-14)$$

mit $\text{erf}(\cdot)$ als Fehlerfunktion. Nachfolgend wird $g(\theta) = 0$ als eine weiche Einschränkung behandelt und dementsprechend mit der Akquisitionsfunktion (3-8) multipliziert, wobei die Gegenwahrscheinlichkeitswert von (3-14) zu verwenden ist:

$$\alpha_{EI,g}(\theta) = \alpha_{EI}(\theta) \left(1 - P(\hat{g}(\theta) > \xi_g) \right). \quad (3-15)$$

Auf diese Weise lassen sich Bereiche im Parameterraum berücksichtigen, in denen die Wahrscheinlichkeit für eine Werkzeugkollision hoch ist, jedoch auch hohe Werte für den Prozessfähigkeitsindex zu erwarten sind. Diese Explorationseigenschaft ist besonders wichtig für die frühen Iterationen der BO, in denen nur eine geringe Datenmenge verfügbar ist und der GP bezüglich der Nebenbedingung ungenaue Vorhersagen liefert.

Die gestrichelte rote Linie im oberen Bild von Abbildung 3-6 zeigt die Wahrscheinlichkeit einer Werkzeugkollision nach (3-14). Der Schwellenwert ξ_g wurde auf 0.5 festgelegt. Es ist zu erkennen, dass die Wahrscheinlichkeit einer Werkzeugkollision auf der rechten Seite nahezu null ist, während die Wahrscheinlichkeit auf der linken Seite glatt zu einem Wert von eins für die einzige beobachtete Werkzeugkollision übergeht. Für abnehmende Werte von θ in Richtung null fällt die Wahrscheinlichkeit erneut ab, da hier extrapoliert wird und eine optimistische Mittelwertfunktion von null gewählt wurde. Die gestrichelte orangefarbene Linie im unteren Bild zeigt die Gewichtung der Akquisitionsfunktion für die erwartete Verbesserung mit der Gegenwahrscheinlichkeit für eine Werkzeugkollision (vgl. (3-15)). Der linke Bereich wird abgewertet, sodass das neue Maximum auf der rechten Seite angenommen wird und an der Stelle des gelben Kreuzes als nächstes ausgewertet wird.

Bis zu diesem Punkt wurde in jeder Iteration lediglich ein Experiment angenommen. Im Rahmen des Ultraschallbondens besteht eine Auswertung darin, eine leere Substratplatte in die Bondmaschine einzuführen und diese automatisch mit der ausgewählten Parameterisierung mit Bondverbindungen zu bestücken. Anschließend wird die Substratplatte aus der Bondmaschine entfernt und in den Scherkraftprüfer platziert, wo der Scherwiderstand gemessen wird. Die resultierenden Werte werden dann manuell in die Datenbank einge tragen, auf der der BO-Algorithmus operiert. Aufgrund der vergleichsweise schnellen und automatisierten Durchführung von Bonden und Scheren ist es sinnvoll, mehrere Parameterisierungssätze direkt in einer BO-Iteration zu berechnen und parallel zu bewerten. Diese simultanen Bewertungen werden als *Batch* bezeichnet, wobei der Ansatz aus [GDHL15] für die Implementierung aufgegriffen wird. Um die Anwendung einer Batch-Berechnung weiter zu begründen, erfolgt eine Betrachtung der spezifischen Zeit, die für eine Iteration benötigt wird. Diese Zeit setzt sich aus der BO-Berechnung t_{calc} , der Vorbereitungszeit des Experiments $t_{prepare}$ und der eigentlichen Auswertungszeit t_{eval} zusammen. Die Gesamtzeit summiert sich dann zu $T_{single} = (t_{calc} + t_{prepare} + t_{eval}) \cdot n_{iter}$, wobei n_{iter} die

Anzahl der Iterationen ist. Bei n_{batch} Batchelementen ergibt sich eine Gesamtzeit von $T_{batch} = (n_{batch}t_{calc} + t_{prepare} + n_{batch}t_{eval}) \cdot \frac{n_{iter}}{n_{batch}}$. Diese Berechnung geht von einer konstanten Vorbereitungszeit und linear skalierenden Zeiten für Berechnung und Auswertung aus, was in den meisten Fällen eine eher pessimistische Annahme ist. Für die experimentelle Konfiguration dieser Arbeit lassen sich die nachfolgenden Zeiten grob abschätzen mit: $t_{calc} = 10$ s, $t_{prepare} = 100$ s, $t_{eval} = 50$ s. Bei $n_{iter} = 100$ und $n_{batch} = 6$ ergeben sich Gesamtzeiten von $T_{single} = 4$ Stunden und $T_{batch} = 1.9$ Stunden, was einer Reduzierung von mehr als 50% entspricht. Diese Reduzierung geht jedoch mit einer verringerten Effizienz einher, da alle identifizierten Batchelemente auf demselben GP basieren und daher erst nach der Bewertung aller Batchelemente und nicht einzeln aktualisiert wird.

Um mehrere Parametrisierungen bzw. Experimente zu berücksichtigen, muss die gewichtete Akquisitionsfunktion $\alpha_{EI,g}$ weiter angepasst werden. Der Ansatz in [GDHL15] sieht vor, eine lokale Straffunktion einzuführen mit

$$\varphi(\boldsymbol{\theta}, \boldsymbol{\theta}_q) = \frac{1}{2} \operatorname{erf} \left(- (2\sigma_C^2(\boldsymbol{\theta}_q))^{-\frac{1}{2}} (\|\mathbf{G}\|_2 \|\boldsymbol{\theta} - \boldsymbol{\theta}_q\|_2 - \xi_C + \mu_C(\boldsymbol{\theta}_q)) \right), \quad (3-16)$$

wobei $\mathbf{G} := \frac{d\mu_C(\boldsymbol{\theta})}{d\boldsymbol{\theta}} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_q}$ angesetzt wird.

Diese Straffunktion hängt von der Parametrisierung $\boldsymbol{\theta}$ und einer spezifischen Stelle $\boldsymbol{\theta}_q$ ab, welche das zuvor ausgewählte Batchelement repräsentiert. Die Stärke der Strafe hängt von der Norm des Gradienten der Posterior-Mittelwertfunktion $\|\mathbf{G}\|_2$ multipliziert mit dem Abstand zwischen den betrachteten Parametern und dem vorherigen Batchelement $\|\boldsymbol{\theta} - \boldsymbol{\theta}_q\|_2$ ab. Weitere Komponenten sind der Posterior-Mittelwert $\mu_C(\boldsymbol{\theta}_q)$ und die Varianz $\sigma_C^2(\boldsymbol{\theta}_q)$, die am vorherigen Batchelement ausgewertet werden, sowie der Schwellenwert ξ_C . Die Grundidee besteht darin, für jedes berechnete Batchelement einen Strafbetrag zu verwenden und das nächste Batchelement durch das Produkt der eingeschränkten Akquisitionsfunktion und aller lokalen Strafterme zu bestimmen:

$$\boldsymbol{\theta}_q^{(b)} = \arg \max_{\boldsymbol{\theta}} \begin{cases} \alpha_{EI,g}(\boldsymbol{\theta}) & b = 1, \\ \alpha_{EI,g}(\boldsymbol{\theta}) \prod_{i=1}^{b-1} \varphi(\boldsymbol{\theta}, \boldsymbol{\theta}_q^{(i)}) & b = 2, \dots, n_{batch}. \end{cases} \quad (3-17)$$

Zusammengefasst werden die einzelnen unterlagerten Optimierungsprobleme (3-17) nacheinander gelöst, bis alle n_{batch} Batchelemente berechnet wurden. Anschließend werden alle Batchelemente am realen System getestet. Bild 3-6 veranschaulicht dieses Konzept durch die schwarze durchgezogene Linie, die den lokalen Strafbetrag $\varphi(\cdot, \cdot)$ (3-16) am Ort des gelben Kreuzes (1. Batchelement $\boldsymbol{\theta}_q^{(1)}$) repräsentiert. Das Produkt der schwarzen Linie und der gestrichelten orangefarbenen Linie ergibt die gestrichelte magentafarbene Linie, deren maximaler Wert sich an der Position des türkisfarbenen Kreuzes (2. Batchelement

$\theta_q^{(2)}$) befindet (vgl. (3-17)). Für dieses anschauliche Beispiel beträgt die Gesamtanzahl der Batchelemente zwei, sodass als nächstes die Auswertung an den berechneten Stellen stattfinden würde. In diesem Beispiel liegen die Batchelemente relativ nah beieinander, was im Allgemeinen nicht der Fall ist, sodass sich die Batchelemente oftmals im Parameterraum auf unterschiedliche Optima verteilen, um diese zu explorieren.

Der entwickelte BO-Algorithmus 2, welcher speziell für den Entwurf einer optimalen Steuerung für das Ultraschalldrahtbonden aufgestellt wurde, ist an dieser Stelle vollständig erläutert worden. In einem direkten Vergleich zum Standardverfahren der BO (s. Algorithmus 1) zeigt sich, dass er um zusätzliche GPs für unterschiedliche Zielsetzungen erweitert wurde. So werden für den Prozessfähigkeitsindex zwei und für die Nebenbedingung ein GP(s) eingesetzt. Des Weiteren wird innerhalb einer Iteration nicht nur ein Experiment durchgeführt, sondern direkt mehrere gleichzeitig, um die automatisierte Bonderzeugung auszunutzen und eine zeitliche Effizienzsteigerung zu erreichen. Damit erfüllt die BO die praktischen Anforderungen und bietet eine alternative Lösungsmöglichkeit zum herkömmlichen manuellen Einstellen. Im folgenden Abschnitt wird die BO am realen System validiert und ihre Ergebnisse ausgewertet.

3.6 Anwendung und Ergebnisanalyse

In den durchgeführten Experimenten wurde ein automatischer Drahtbonder (Typ *BJ955*) der Firma Hesse Mechatronics verwendet. Als Substratmaterial dienten Aluminium-Dibond-Platten. Der Draht bestand ebenfalls aus Aluminium und wies einen Durchmesser von 500 µm auf. Die Scherfestigkeiten wurden mithilfe eines Xyztec Sigma-Prüfgeräts gemessen. Dieses Setup stellt einen typischen Aufbau aus der industriellen Praxis dar. Im Hinblick auf den in Algorithmus 2 dargestellten Ablauf wurde die Anzahl der initialen Experimente n_{init} auf 10, die Anzahl der Batchelemente n_{batch} auf 6, das Iterationsbudget n_{budget} auf 15 (was einer Begrenzung der Experimente auf $n_{init} + n_{batch}n_{budget} = 100$ entspricht), sowie die Grenzwerte ξ_C auf 2 und ξ_g auf 0.5 festgelegt. Die untere Spezifikationsgrenze LSL beträgt 2500 cN und die Anzahl der Verbindungen pro Parametrisierung n_{rep} beläuft sich auf 10. Der Suchraum für die Parametrisierung ist auf der Grundlage von Expertenwissen eingeschränkt, wobei die untere Grenze mit θ_{lb} (*Lower Bound*) und die obere Grenze mit θ_{ub} (*Upper Bound*) definiert und Tabelle 3-1 zu entnehmen sind.

Tabelle 3-1: Beschränkungen für die Parametrisierung

	F_0 [cN]	F_1 [cN]	F_2 [cN]	\hat{U}_1 [V]	\hat{U}_2 [V]	T_1 [ms]	T_2 [ms]
θ_{lb}	300	375	375	43.3	7.75	5.5	29.5
θ_{ub}	900	1125	1125	80.6	54.25	38.5	206.5

Alle eingesetzten GPs gehen von der Matérn-Kovarianzfunktion (2-35) aus. Um die bestimmten Hyperparameter zu vergleichen, wurden die Eingabewerte über die unteren und oberen Grenzen aus Tabelle 3-1 auf das Einheitsintervall normiert. Die damit verbundene Auswertung wird am Ende dieses Unterkapitels besprochen. Der entwickelte BO-Algorithmus 2 wurde in der Programmierumgebung MATLAB [Mat24] implementiert. Die Optimierung der Akquisitionsfunktion erfolgt in zwei Schritten mittels einer Zufallsuche mit einer Million Kandidaten, um dem erwartbaren Auftreten von mehreren lokalen Optima zu begegnen. Der beste Kandidat wird als initiale Ausgangsschätzung (*Initial Guess*) für eine anschließende verfeinernde Optimierung mit der integrierten Routine *fminsearch*, welche auf dem gradientenfreien Nelder-Mead-Verfahren [GK02] basiert, verwendet. Die enorme Anzahl von Auswertungen der Akquisitionsfunktion ist unproblematisch, da ihre Rechenkomplexität nach Konstruktion gering ist und sie vergleichsweise schnell berechnet werden kann.

Es erfolgt ein Vergleich von vier verschiedenen Lösungsansätzen für (3-10):

- **Zufallssuche (*Random Search*):** Für jedes Experiment wird eine zufällige Parametrisierung aus der Gleichverteilung $\theta \sim \mathcal{U}(\theta_{lb}, \theta_{ub})$ ausgewählt. Dieser Fall dient als Referenz für die anderen Verfahren und wird aufgrund der Dimension des Problems als ungeeignete Lösungsstrategie für die Praxis angesehen.
- **Manuelle Einstellung (*Manual Tuning*):** Ein Nicht-Experte, der mit dem Prozess und seinen physikalischen Effekten vertraut ist, bestimmt die optimale Parametrisierung manuell. Die Person muss in jeder Iteration 6 Parametrisierungen auswählen, die dann parallel bewertet werden. Dies entspricht dem Vorgehen des BO-Algorithmus 2. Alle bisherigen Experimente können jederzeit eingesehen werden, um Schlussfolgerungen für die nächsten Parametrisierungen zu ziehen.
- **BO mit konstanter Prior-Mittelwertfunktion:** Dies ist das Basisszenario, in welchem keine spezifischen a-priori Kenntnisse über die Prozessdynamik verfügbar sind. Daher werden die Prior-Mittelwertfunktionen für die mittlere Scherkraft und die Standardabweichung auf konstante Werte gesetzt, also $m_\mu = 2500$ cN und $m_\sigma = 60$ cN, was zu einer eher pessimistischen Initialisierung mit einem Prozessfähigkeitsindexwert von 0 führt.
- **BO mit quadratischer Prior-Mittelwertfunktion:** Dies ist ein Referenzszenario, bei dem eine quadratische anstelle einer konstanten Prior-Mittelwertfunktion $m_\mu(\theta) = \theta^T A \theta + b^T \theta + c$ für die mittlere Scherkraft verwendet wird. Die mittlere Funktion für die Standardabweichung bleibt konstant (s. o.). Die Größen A, b, c werden durch die Methode der kleinsten Quadrate an alle Daten angepasst, die aus den anderen drei Lösungsansätzen gesammelt wurden. Die Wahl dieser Ansatzfunktion wird durch den physikalischen Effekt des Überbonden [Har10] begründet,

welcher dafür verantwortlich ist, dass die Scherkraft bei einem zu hohen Energieeintrag mit der Zeit nach dem Erreichen des Maximums wieder absinken kann.

In Abbildung 3-7 (links) ist der Verlauf des Prozessfähigkeitsindex über die Anzahl der Experimente/Iterationen dargestellt. Konstante Plateaus deuten darauf hin, dass in den entsprechenden Experimenten keine Verbesserung erfolgt ist. Es ist zu beachten, dass ein Anstieg im Zielwert nur auftritt, wenn der C_{pK} -Wert höher ist als in allen vorherigen Experimenten und die Nebenbedingung bezüglich optischer Kriterien erfüllt ist. Allen Ansätzen wurden dieselben 10 anfänglichen Experimente zugewiesen, sodass sich die Verläufe anfangs nicht voneinander unterscheiden.

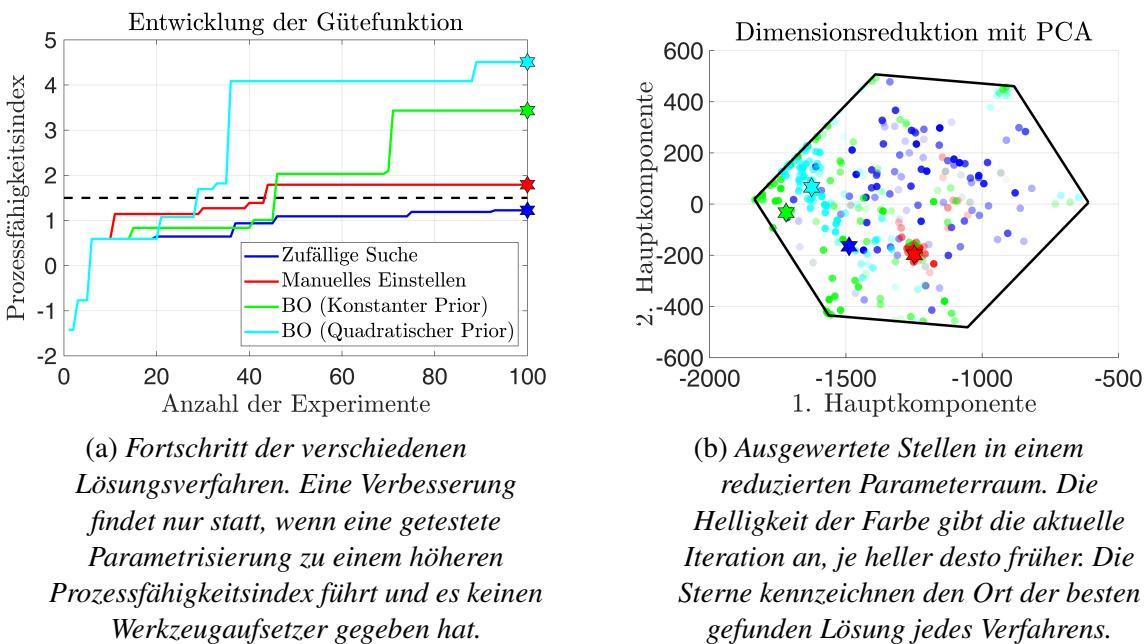


Bild 3-7: Ergebnisse der experimentellen Validierung der Bayesschen Optimierung am Bondautomaten

Für industrielle Prozesse ist üblicherweise ein Mindest- C_{pK} -Wert im Bereich von 1.33 bis 1.67 erforderlich [DVS17] (schwarze gestrichelte Linie). Die Zufallssuche (blau) verharrt schnell nahe dem Wert 1, ohne signifikante Verbesserungen zu erzielen. Basierend auf diesen Beobachtungen, deutet sich an, dass die Zielfunktion in einem großen Teil des Suchraums relativ flach verläuft. Die manuelle Einstellung (rot) führt zu einer gewissen Effizienzsteigerung. Nach etwa 50 Experimenten kann jedoch keine weitere Verbesserung festgestellt werden. Im Gegensatz dazu findet die BO mit einem konstanten Prior (grün) eine deutlich bessere Parametrisierung, obwohl sie zuvor auf dem gleichen Niveau wie die manuelle Einstellung verharret. Das anfängliche, zufalls-ähnliche Verhalten wird auf die ursprüngliche Exploration des Suchraums zurückgeführt, wobei sich zeigt, dass dieses Verhalten für die späteren Iterationen von Nutzen ist. Die BO mit einem quadratischen Prior (türkis) liefert die besten Ergebnisse im Vergleich zu allen anderen Ansätzen.

Mehrere Experimente sind erforderlich, um diese Parametrisierung zu identifizieren, was darauf hindeutet, dass der quadratische Prior trotz seiner breiten Datenbasis noch einige Abweichungen zu der realen Zielgröße aufweist. Dennoch lenkt die Exploration in eine Region potenziell hochwertiger Parametrisierungen, was insgesamt betrachtet zu einer hohen Effizienz führt. Dieser Fall verdeutlicht, wie eine geeignete, wenn auch nicht zwangsläufig perfekte Prior-Mittelwertfunktion den Fortschritt der BO verbessert. Anstelle des hier verwendeten datengetriebenen ML-Modells könnte ein physikalisches Modell des Prozesses in Betracht gezogen werden, um die Leistungsfähigkeit zu steigern. Wie jedoch in Abschnitt 3.4 gezeigt wurde, ist bisher kein für die BO geeignetes physikalisches Modell bekannt. Die identifizierte quadratische Ansatzfunktion könnte hier dazu eingesetzt werden, um im Sinne einer Reverse Engineering Strategie, ein geeignetes Modell aufzustellen, welches im Rahmen der BO gewinnbringend eingesetzt werden kann. Ein entscheidender Effekt könnte dabei das Überbonden [Har10] sein, welches für eine Abnahme der Scherkraft im Zusammenhang mit falsch gewählten Eingangsgrößen und einem zu hohen Energieeintrag steht. Eine Weiterentwicklung dieses Ansatzes wird als Ausblick auf zukünftige Forschung offen gelassen und bietet ein großes Potential, um den Bondprozess physikalisch besser beschreiben zu können. Ergänzend wurde der BO-Algorithmus neben Aluminium auch mit den Materialien Kupfer und Nickel validiert. Im Bild A3-1 des Anhangs sind hierzu die Ergebnisse gezeigt. Ein rascher Anstieg des Prozessfähigkeitsindex lässt sich auch hier beobachten, sodass hieraus geschlossen werden kann, dass keine besondere Abhängigkeit zwischen der BO-Performanz und der Materialpaarung besteht.

In der folgenden Analyse wird der Fortschritt im Parameterraum behandelt, wie in Bild 3-7 (rechts) dargestellt. Eine Visualisierung aller 7 Parameter durch mehrere Schnittebenen des 7-dimensionalen Raums ist aufgrund der Komplexität nicht übersichtlich, daher wurden die Daten mittels einer Hauptkomponentenanalyse (PCA) in einen 2-dimensionalen Raum transformiert. Hierzu wurde eine Singulärwertzerlegung (SVD) der Datenmatrix durchgeführt, die aus sämtlichen Parameterwerten aller Experimente besteht. Anschließend wurden die Parameterwerte in einen 2-dimensionalen Raum transformiert, indem die ersten beiden Spalten der Matrix mit den linken Singulärvektoren verwendet wurden, die den beiden größten Singulärwerten zugeordnet werden. Die Parameterwerte im Kontext der Zufallssuche zeigen, wie erwartet, eine breite Diversifikation. Auffällig ist die Konzentration roten Punkte in einem bestimmten Gebiet. Dieser Bereich wurde vom manuellen Einsteller lokal erkundet, da die resultierenden Werte der Zielfunktion relativ gut waren und die Einschränkung erfüllt wurde. Dieses Vorgehen kann als sichere Exploration betrachtet werden, bei der der Bereich getester Parameter iterativ erweitert wird. Dieses Verhalten ähnelt dem Safe Reinforcement Learning Ansatz aus dem Abschnitt zum Stand der Forschung und kann daher auf diese Weise interpretiert werden. Es wurde

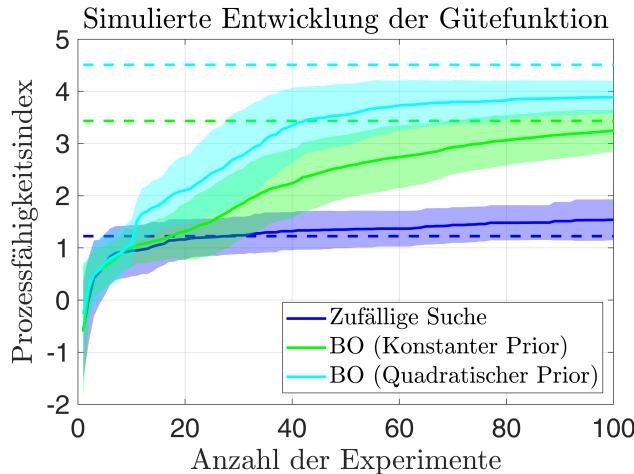


Bild 3-8: Fortschritt des Prozessfähigkeitsindex in der Simulationsumgebung. Die durchgezogenen Linien zeigen den Mittelwert und die transparenten Flächen die Standardabweichung für mehrere verschiedene Identifikationen. Die gestrichelten Linien deuten die Endergebnisse aus der Validierung am realen Prozess an.

zudem festgestellt, dass der Überblick des Anwenders über die vorherigen Experimente bei etwa 30 bis 40 Experimenten verloren ging, was den mangelnden Fortschritt bei der Zielfunktion in Bild 3-7 (links) erklärt. Im Gegensatz dazu, erkundet die BO besonders die Grenzen des Suchraums und konzentriert sich relativ rasch auf die Region, in der die höchsten C_{pK} -Werte gefunden wurden. Aus Gründen der Vertraulichkeit gegenüber dem Industriepartner dürfen die optimalen Parameter im Rahmen dieser Arbeit nicht offenlegt werden.

Simulationsbasierte Ergebnisse zur Untersuchung der Robustheit

Aufgrund der zeit- und kostenintensiven Experimente am realen System war es nicht möglich, den Einfluss der initialen Experimente und somit die Robustheit jeder Methode umfassend zu untersuchen. Aus diesem Grund wurden GPs auf der Basis von sämtlichen bisher gesammelten Daten trainiert, um das reale System realitätsnah simuieren zu können. Die Vorhersagen dieser virtuellen Umgebung (im nachfolgenden auch Referenz-GPs genannt) sind daher vergleichsweise präzise. Dies ermöglicht es, die verwendeten Methoden - mit Ausnahme der manuellen Abstimmung - für verschiedene Anfangsexperimente zu testen. Bild 3-8 zeigt in diesem Zusammenhang die zugehörigen Ergebnisse für 50 verschiedene separate Durchläufe. Zunächst ist zu erkennen, dass die Durchläufe bezüglich des realen Systems vernünftig sind, da sie mit den simulierten Durchläufen größtenteils übereinstimmen (vgl. zu Bild 3-7). Darüber hinaus übertrifft die verwendete BO-Methode robust die Zufallssuche und konvergiert zu hohen Werten der Zielgröße. Die Tatsache, dass das globale Maximum aus den Messungen (gestrichelte türkisfarbene Linie) von der quadratischen Prior-BO nicht erreicht wird, könnte mit der Regularisierung der Referenz-GPs zusammenhängen. Die Mittelwertfunktion des GP

Tabelle 3-2: Identifizierte Hyperparameter der trainierten Gauß-Prozesse. Hervorzuhebende hohe bzw. niedrige Werte sind rot und blau markiert.

\mathcal{GP}	σ_f^2	l_{F_0}	l_{F_1}	l_{F_2}	l_{U_1}	l_{U_2}	l_{T_1}	l_{T_2}	σ_n^2
$\hat{\mu}_{F_S}(\boldsymbol{\theta})$	$4.4 \cdot 10^5$	13.7	2	1.9	1.7	1.4	2.9	2.6	$1.6 \cdot 10^3$
$\hat{\sigma}_{F_S}(\boldsymbol{\theta})$	$1.7 \cdot 10^3$	1.3	1.9	0.2	0.9	0.9	2.3	0.8	335
$\hat{g}(\boldsymbol{\theta})$	0.3	3	0.7	1.1	0.5	0.5	1	1.3	0.05

kann in diesem Fall die Extremwerte der Messungen nicht genau abbilden. Die gefundenen optimalen Parameter sind jedoch identisch zu denen der Messung. Zusammengefasst zeigt sich anhand dieser Ergebnisse, dass die BO eine hohe Robustheit gegenüber verschiedenen Anfangsexperimenten aufweist. Als Fazit zu den realen und simulierten Resultaten ist festzustellen, dass der Einsatz der BO für den Steuerungsentwurf am Ultraschalldrahtbondprozess äußerst geeignet und eine Alternative zum herkömmlichen modellbasierten Steuerungsentwurf auf der Basis eines physikalischen Modells ist. Sofern ein hinreichend präzises physikalisches Modell bekannt ist, muss dieses zwangsläufig bei der Anwendung der BO nicht verworfen werden, sondern kann über die Prior-Mittelwertfunktion eingebunden werden. Hierzu hat der Reverse Engineering Grundgedanke über die quadratische Ansatzfunktion gezeigt, dass die Berücksichtigung von Vorwissen eine deutliche Effizienzsteigerung herbeiführt. Neben dem modellbasierten Entwurf stellt die BO zudem eine echte Alternative zum praktischen manuellen Einstellen dar, wobei die Suche nach den optimalen Parametern teil-automatisiert wird und anhand von objektiven Gütekriterien interaktiv vorgenommen wird. Auf diese Weise wird für die Einrichtung des Prozesses kein Expertenwissen mehr benötigt, sondern kann auch von Nicht-Experten vorgenommen werden. Diverse Fehlerquellen, wie die subjektive Wahrnehmung des Einrichters oder bspw. die beobachtete zu rasche Fokussierung auf ein lokales Optima (vgl. Bild 3-7), können damit ausgeschlossen werden.

Ergänzende Auswertung der Hyperparameter

Neben der Erprobung der Algorithmen innerhalb einer simulierten Umgebung, bieten die trainierten GPs die Möglichkeit, die erlernten Hyperparameter η_* genauer zu untersuchen und damit weitere Einsichten in den Prozess zu erhalten. Insbesondere die erlernten Längenskalen l , die den einzelnen Eingangsdimensionen zugeordnet werden, zeigen die Relevanz einer bestimmten Parameterdimension für die Ausgabe. Dies wird in der Literatur als automatische Relevanzbestimmung (*Automatic Relevance Determination (ARD)*) bezeichnet [Ras06]. Die Hyperparameter sind in Tabelle 3-2 aufgeführt. Eine direkte Vergleichbarkeit ist möglich, da die Parameterdimensionen auf das Einheitsintervall standardisiert wurden. Der Einfluss eines bestimmten Parameters wird durch den Wert der Skala l bestimmt, wobei höhere Werte auf einen geringeren Einfluss hindeuten (vgl. hierzu auch Bild 2-10). In Bezug auf die mittlere Scherkraft $\hat{\mu}_{F_S}(\boldsymbol{\theta})$ und das Label $\hat{g}(\boldsymbol{\theta})$ zeigt sich,

dass die Normalkraft der Vorverformungsphase F_0 im Vergleich zur Standardabweichung $\hat{\sigma}_{F_S}(\boldsymbol{\theta})$ weniger relevant ist. Dies ist plausibel, da eine ausreichende Anfangskontaktfläche durch einen breiten Bereich von Normalkraftwerten gebildet wird. Andererseits haben die Werte der Ultraschallspannungen (\hat{U}_1, \hat{U}_2) einen vergleichsweise hohen Einfluss auf die mittlere Scherkraft. Gleches gilt für die Standardabweichung der Scherkraft $\hat{\sigma}_{F_S}(\boldsymbol{\theta})$. Jedoch zeigt sich, dass die Kraft F_2 und die Zeit T_2 der zweiten Phase den größten Einfluss haben, was bedeutet, dass diese Werte mit hoher Genauigkeit korrekt gewählt werden müssen, um eine geringe Standardabweichung zu erzielen. Die Werte für die Signal- und Rauschvarianz (σ_f^2, σ_n^2) stimmen mit den Beobachtungen während der Experimente überein.

Fazit der Untersuchungen

Die Ergebnisse aus der realen Welt und den Simulationen zeigen, dass die Anwendung der BO für den Steuerungsentwurf des Ultraschalldrahtbondens sehr geeignet ist. Der vorgeschlagene Ansatz übertrifft die zufällige Suche und die manuelle Abstimmung hinsichtlich Effizienz und Robustheit. Des Weiteren wurde gezeigt, dass die Einbeziehung einer quadratischen Prior-Mittelfunktion vorteilhaft ist und die Leistung verbessert. Dieses Ergebnis wird als Leitfaden für weitere Forschung betrachtet, bei der die quadratische Prior-Mittelfunktion durch ein physikalisches Simulationsmodell des Bondprozesses ersetzt wird. Die vorliegende Arbeit legt dafür die Grundlage, da sie auf zahlreichen Messungen mit einer Vielzahl von Steuersignalen bzw. Parametrisierungen basiert und somit gut validiert ist. Ergänzend wurden die Hyperparameter der trainierten GPs untersucht, wobei festgestellt wurde, dass die Aufsetzkraft während der Vorverformungsphase wenig Einfluss auf die Bindungsqualität hat. Die mittlere Scherkraft reagiert empfindlich auf die Amplitudenwerte der Spannung, während die Varianz am stärksten auf die Normalkraft in der zweiten Phase und deren Dauer reagiert.

3.7 Weiterführender Entwurf mit Bayesscher Optimierung

Im Nachfolgenden soll ein Überblick über mögliche Erweiterungen und Verbesserungen für die BO im Regelungstechnischen Kontext gegeben werden. Ein erster Ansatzpunkt wird bei der Wahl der Akquisitionsfunktion gesehen. Hierfür wurde der Ansatz des EI (3-8) genutzt, welcher innerhalb der Literatur als eine Standardwahl angesehen wird. Neben ihm existieren allerdings auch komplexere Ansätze, die sich nicht an dem nominalen Wert der Gütfunktion orientieren, sondern sich auf den Informationsgehalt eines jeden Experiments fokussieren. Das nächste Experiment soll bei dieser Grundidee an einer Stelle im Parameterraum durchgeführt werden, die einen verhältnismäßig hohen Informationszuwachs in Bezug zum gesuchten globalen Optimum darstellt. In [HS12] wird, dieser Idee folgend, die sogenannte *Entropy Search* (ES) als Akquisitionsfunktion vorge-

schlagen. Als Maß für den Informationsgehalt einer kontinuierlichen Verteilungsfunktion $p(\hat{J}(\boldsymbol{\theta}))$ wird die differentielle Entropie über

$$H(p) = \mathbb{E}[-\log(p)] = - \int p \log(p) d\hat{J},$$

angegeben, wobei sich bspw. für den Fall einer eindimensionalen Normalverteilung die analytische Form $H(\mathcal{N}(\mu, \sigma)) = \frac{1}{2}(1 + \log(2\pi\sigma^2))$ ergibt. Eine Verteilung mit einer hohen Varianz weist dementsprechend eine hohe Entropie auf, sodass eine Auswertung einen gleichermaßen hohen Informationsgewinn verspricht. Entgegengesetzt weisen Verteilungen mit einer niedrigen Varianz keinen allzu großen Informationszuwachs auf, da der Wert der zugrundeliegenden Zufallsvariable bereits gut abgeschätzt werden kann. Der Grundidee von [HS12] weiter folgend, wird die ES-Akquisitionsfunktion über den Ausdruck

$$\alpha_{ES}(\boldsymbol{\theta}) = H(p(\hat{\boldsymbol{\theta}}_* | \mathbf{D}_J)) - \mathbb{E}_{p(\hat{J}(\boldsymbol{\theta}) | \mathbf{D}_J)}[H(p(\hat{\boldsymbol{\theta}}_* | \mathbf{D}_J \cup \{\boldsymbol{\theta}, \hat{J}(\boldsymbol{\theta})\}))],$$

angegeben. Hierin stellt $p(\hat{\boldsymbol{\theta}}_* | \mathbf{D}_J)$ eine Verteilung für die optimale Parametrisierung dar, welche auf der Grundlage des GP bzgl. $p(\hat{J}(\boldsymbol{\theta}) | \mathbf{D}_J)$ abgeleitet wird. Zur Bestimmung sind aufwendige Approximationen durch MC-Verfahren notwendig. In diesem Zusammenhang gibt der vordere Term $H(p(\hat{\boldsymbol{\theta}}_* | \mathbf{D}_J))$ den Informationsgehalt dieser Verteilung an und wird mit dem hinteren Term $\mathbb{E}_{p(\hat{J}(\boldsymbol{\theta}) | \mathbf{D}_J)}[H(p(\hat{\boldsymbol{\theta}}_* | \mathbf{D}_J \cup \{\boldsymbol{\theta}, \hat{J}(\boldsymbol{\theta})\}))]$ in Relation gesetzt. Dieser gibt an, wie stark sich der Informationsgehalt ändern würde, wenn an der Stelle $\boldsymbol{\theta}$ ein Experiment durchgeführt werden würde. Anstelle des wahren Funktionswertes wird die Wahrscheinlichkeitsverteilung des GP $p(\hat{J}(\boldsymbol{\theta}) | \mathbf{D}_J)$ verwendet, wodurch der Einsatz des Erwartungswertoperators notwendig wird. Für niedrig dimensionale Probleme, z. B. $n_\theta \leq 3$, weist die ES-Akquisitionsfunktion vorteilhafte Eigenschaften auf und ist den anderen Funktionen im Allgemeinen überlegen. Ein großes Problem ist allerdings ihre schlechtere Skalierbarkeit und damit schwere Anwendbarkeit auf höher dimensionale Probleme. Der Grund hierfür sind die numerisch rechenintensiven Approximationen, die bspw. für den hinteren Term und den Erwartungswertoperator aufeinander aufbauen und damit bei der Berechnungskomplexität besonders zum Tragen kommen.

Eine Folgearbeit ist [WJ17], worin das grundsätzliche Problem darin gesehen wird, dass die Suche bzw. Approximation im Raum der Parametrisierung stattfindet. Es wird daher vorgeschlagen, nicht den Informationsinhalt bzgl. der optimalen Parametrisierung $\boldsymbol{\theta}_*$ zu erkunden, sondern bzgl. des optimalen Funktionswertes J_* , da dieser eine skalare Größe darstellt. Die umgeformte Akquisitionsfunktion heißt *Maximum-Value Entropy Search* (MES) und wird mit

$$\alpha_{MES}(\boldsymbol{\theta}) = H(p(\hat{J}(\boldsymbol{\theta}) | \mathbf{D}_J)) - \mathbb{E}_{p(\hat{J}_* | \mathbf{D}_J)}[H(p(\hat{J}(\boldsymbol{\theta}) | \mathbf{D}_J \cup \hat{J}_*))]$$

ausgedrückt, wobei die Verteilung der optimalen Gütfunktionswerte $p(\hat{J}_* \mid \mathbf{D}_J)$ eine zentrale Rolle spielt. Diese Verteilung muss gleichermaßen zu $p(\hat{\theta}_* \mid \mathbf{D}_J)$ anhand des GP approximiert werden, ist allerdings deutlich günstiger bzgl. der Berechnung, da $\dim(J) = 1 \leq n_\theta$ gilt. In diesem Zusammenhang stellt die hintere Verteilung $p(\hat{J}(\boldsymbol{\theta}) \mid \mathbf{D}_J \cup \hat{J}_*)$ keine herkömmliche Normalverteilung mehr dar, sondern ist durch die angenommene Kenntnis über das globale Optimum \hat{J}_* nach oben hin abgeschnitten, woraus sich eine sogenannte *truncated* (gekürzte) Normalverteilung ergibt. Eine analytische Berechnung der differentiellen Entropie ist mit dieser Verteilung möglich, sodass sich hieraus insgesamt eine berechnungseffiziente Akquisitionsfunktion ergibt. Im direkten Vergleich zum EI (3-8) ist ein wesentlicher Vorteil, dass kein Grenzwert ξ_J für die Berechnung von $\alpha_{MES}(\cdot)$ notwendig ist, sondern eine automatisierte Abschätzung auf der Basis des GPs getroffen wird. Die Anwendung dieses Ansatzes ist daher besonders geeignet, wenn a-priori kein Grenzwert bekannt ist und über den zugrundeliegenden Steuerungsentwurf kaum Vorwissen vorhanden ist. Im Falle des Ultraschalldrahtbondens und der durchgeföhrten Experimente konnte jedoch zuverlässig ein relativ genauer Grenzwert angegeben werden, sodass ein Einsatz der MES-Akquisitionsfunktion nicht notwendig war. Als Ausblick für weitere regelungstechnische Anwendungsgebiete der BO stellt die MES-Akquisitionsfunktion allerdings einen klaren Mehrwert dar.

Einen weiteren Punkt zur Verbesserung bietet die Bestimmung der Hyperparameter. Dafür wurde bisher die Maximierung der Log-Likelihood (s. (2-33)) vorgeschlagen, bei deren Optimierung ein sogenannter *Point Estimate* der Hyperparameter identifiziert wird. Damit ist gemeint, dass lediglich ein bestimmter zahlenmäßiger Vektor zur Definition des GP bzw. der Kernelfunktion verwendet wird. Ein präziserer Ansatz wäre es, die Unsicherheit über die Hyperparameter explizit zu berücksichtigen, indem sie als Zufallsvariablen definiert und mit Hilfe der Bayesschen Regel (2-13) angegeben werden. Für die Prior-Verteilung ließe sich wiederum eine Normalverteilung verwenden und ein gewisses Vorwissen im Sinne der ARD durch ihre Parametrierung vorgeben. Als Folge für die BO ist die Akquisitionsfunktion nun von der Hyperparametrierung abhängig, wodurch der Erwartungswert über

$$\alpha(\boldsymbol{\theta}) = \mathbb{E}_{p(\boldsymbol{\eta} \mid \mathbf{D}_J)}[\alpha(\boldsymbol{\theta}; \boldsymbol{\eta})] = \int \alpha(\boldsymbol{\theta}; \boldsymbol{\eta}) p(\boldsymbol{\eta} \mid \mathbf{D}_J) d\boldsymbol{\eta} \approx \frac{1}{n_s} \sum_{i=1}^{n_s} \alpha(\boldsymbol{\theta}; \boldsymbol{\eta}_i), \boldsymbol{\eta}_i \sim p(\boldsymbol{\eta} \mid \mathbf{D}_J)$$

einzuführen ist. Hierbei wird der Erwartungswert über n_s Realisierungen der Posterior-Verteilung der Hyperparameter $p(\boldsymbol{\eta} \mid \mathbf{D}_J)$ approximiert, wobei der *Markov-Chain Monte Carlo* (MCMC) aus z. B. [Kuß06] verwendet werden kann. Die Einbeziehung der Unsicherheit über die Hyperparameter ist dann besonders sinnvoll, wenn die Datenlage noch

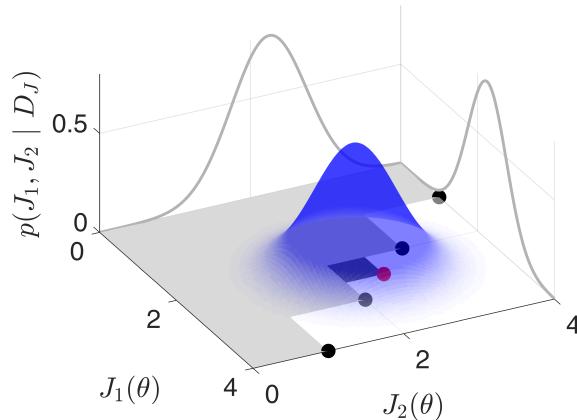


Bild 3-9: Identifikation einer Paretofront mit Expected Hypervolume Improvement

sehr gering ausfällt, also für die frühen Iterationen der BO. Bei einer ausreichend großen Datenmenge kann dieser Freiheitsgrad entfallen und auf die herkömmliche Maximierung der logarithmischen Likelihood zurückgegriffen werden.

Als letzter Punkt ist eine Erweiterung auf eine Mehrzieloptimierung zu nennen. Im Zusammenhang mit dem Ultraschalldrahtbonden gibt es dazu diverse Anknüpfungspunkte. So setzt sich bspw. der Prozessfähigkeitsindex bereits aus dem Mittelwert und der Standardabweichung der Scherkraft zusammen, wofür separate GPs angesetzt worden sind, die zwei bestimmte Zielwerte darstellen. Bei der Mehrzieloptimierung steht nicht mehr eine gewisse optimale Parametrisierung im Vordergrund, sondern die Paretofront \mathcal{P} , die durch mehrere verschiedene optimale Parametrisierungen aufgespannt wird. Der Einrichter des Bondprozesses erhält nach der Ausführung der BO damit die Möglichkeit, einen für ihn passenden Punkt auf der Paretofront auszuwählen. Das Ziel der BO ist, in jeder Iteration die Paretofront nach und nach zu erweitern und neue dominante Punkte hinzuzufügen. Ein Ansatz hierzu wird in [ED18; YEDB19] vorgestellt, wobei das sogenannte *Hypervolume* (HV), welches durch die Paretofront und einen Referenzpunkt aufgespannt wird, betrachtet wird. Die zugehörige Akquisitionsfunktion für die BO lautet *Expected Hypervolume Improvement* (EHVI) mit

$$\alpha_{EHVI}(\boldsymbol{\theta}) = \int (HV(\mathcal{P} \cup \hat{\mathbf{J}}(\boldsymbol{\theta})) - HV(\mathcal{P})) p(\hat{\mathbf{J}}(\boldsymbol{\theta}) | D_J) d\hat{\mathbf{J}}, \quad (3-18)$$

und berücksichtigt die Unsicherheit über die verschiedenen Gütfunktionswerte, welche durch die gelernten GPs ausgedrückt wird. Bild 3-9 verdeutlicht hierzu visuell die grundsätzlichen Zusammenhänge anhand eines Beispiels mit den Gütfunktionswerten $J_1(\boldsymbol{\theta})$ und $J_2(\boldsymbol{\theta})$. Der hellgraue Bereich zeigt die aktuell bekannte Paretofront bzw. das Hypervolumen in Form einer Fläche im zweidimensionalen Fall, welche durch die bisherigen dominanten Messpunkte (schwarz) aufgespannt wird. In blau ist die multivariate Normalverteilung der zugrundeliegenden GPs für eine bestimmte Parametrisierung darge-

stellt. Zudem sind die marginalen Verteilungen an den Seiten in hellgrau visualisiert. Auf der Basis von (3-18) wird entsprechend die Parametrisierung als nächstes durch den BO-Algorithmus ausgewählt, die zu einer Maximierung des Hypervolumens im Erwartungswert führt. Die Verbesserung in dieser Metrik ist durch das dunkelgraue Gebiet angedeutet, sodass eine Auswertung bei der Parametrisierung des roten Punktes durchgeführt wird und die Paretofront somit an dieser Stelle ein Update erhält. In der studentischen Arbeit [Rei23] wurde der Ansatz des EHVI für den Ultraschalldrahtbondprozess erprobt. Dabei wurden der Prozessfähigkeitsindex und die Prozesszeit als Optimierungsziele ausgewählt. Erwartungsgemäß verschlechtert sich der Prozessfähigkeitsindex für geringer werdende Prozesszeiten, wodurch sich der typische Verlauf einer Paretofront ausbildet. Insgesamt wurden auch im Rahmen dieser Arbeit gute Ergebnisse mit dem BO-Ansatz erzielt.

4 Hybride Optimalsteuerung dynamischer Systeme

Die Optimalsteuerungstheorie ist ein Zweig der mathematischen Optimierung, der sich damit befasst, eine Steuerung für ein dynamisches System über einen bestimmten Zeitraum zu finden, sodass ein Funktional unter Berücksichtigung von gewissen Nebenbedingungen optimiert wird [Bet10; PLB12]. Sie hat zahlreiche Anwendungen im wissenschaftlichen und industriellen Umfeld. Zum Beispiel kann das dynamische System ein Roboterarm sein, dessen Steuerung der Vorgabe der passenden Momente in seinen Gelenken entspricht. Das Ziel könnte hierbei sein, eine Position so schnell wie möglich mit minimalem Energieaufwand zu erreichen. Diese Aufgabe wird in der Regel durch die Entwicklung eines mathematischen Modells des realen Systems, basierend auf der Anwendung physikalischer Gesetze, gelöst und einem anschließenden Entwurfsprozess, der optimale Steuersignale im Hinblick auf das Modell angibt. Die resultierende Steuerung wird dann auf das reale System angewendet, wobei eine wesentliche Annahme ist, dass das etablierte Modell das reale System ausreichend genau beschreibt. Ist diese Annahme verletzt, ist der Optimalsteuerungsentwurf hinfällig und muss auf einem anderen Wege durchgeführt werden. Der nachfolgende Abschnitt erläutert diesen Zusammenhang mathematisch detaillierter und ordnet die hybride Optimalsteuerung in den Gesamtkontext ein. Dabei wird Bezug zur modellbasierten Inbetriebnahme genommen und eine andere Ausgangssituation als im letzten Kapitel, welche den Einsatz der BO motivierte, zugrunde gelegt.

4.1 Klassischer modellbasierter Optimalsteuerungsentwurf

Den Ausgangspunkt der Betrachtung stellt ein nichtlineares diskretes dynamisches System in der Zustandsraumdarstellung

$$\mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k, \mathbf{u}_k),$$

mit dem Zustand $\mathbf{x} \in \mathbb{R}^{n_x}$ und der Stellgröße $\mathbf{u} \in \mathbb{R}^{n_u}$ dar. Die Zeit wird durch eine äquidistante Diskretisierung der Form $\tau := \{t_k = k\Delta t \mid k = 0, \dots, H \in \mathbb{N}, \Delta t \in \mathbb{R}_+\}$ ausgedrückt. Die Aufgabe des Optimalsteuerungsentwurfs besteht nun darin, das System aus einem (initialen) Anfangszustand \mathbf{x}_I zum Zeitpunkt $t = 0$ s in der Zeit $T = \Delta t H$ in einen gewünschten Zielzustand \mathbf{x}_G durch die passende Wahl des Stellgrößenverlaufs $\mathbf{u}_k, k = 0, \dots, H$ zu überführen. Eine Bewertung der aufgeschalteten Stellgrößen findet

dabei anhand einer Gütfunktion J statt. Beispielhaft wird hierfür wiederum das riccati-sche Gütemaß mit

$$J(\mathbf{x}_k, \mathbf{u}_k) = \Delta t \sum_{k=0}^H (\mathbf{x}_k - \mathbf{x}_G)^T \mathbf{W}_x (\mathbf{x}_k - \mathbf{x}_G) + \mathbf{u}_k^T \mathbf{W}_u \mathbf{u}_k$$

angenommen. Diese Wahl entspricht einer Mehrzieloptimierung, wobei der Abstand zum Zielzustand $(\mathbf{x}_k - \mathbf{x}_G)$ so schnell wie möglich verkleinert und gleichzeitig der quadratische Energieaufwand $\mathbf{u}_k^T \mathbf{u}_k$ klein gehalten werden soll. Das Gütemaß ist so konstruiert, dass es über einen Skalarisierungsansatz mit den Gewichtungsmatrizen \mathbf{W}_x und \mathbf{W}_u einen Kompromiss zwischen den beiden Zielen darstellt. Neben der Optimierung der Gütfunktion, ist ein weiteres Entwurfsziel die Einhaltung von Nebenbedingungen. Diese hängen mit dem Anwendungsfall und dem zugrundeliegenden System bzw. dessen Zuständen und Stellgrößen zusammen. Im einfachsten Fall müssen dafür lediglich Beschränkungen der Form $\|\mathbf{x}_k\| \leq \mathbf{x}_{max}$ und $\|\mathbf{u}_k\| \leq \mathbf{u}_{max}$ eingehalten werden. Zusammengefasst kann hierfür eine Funktion mit $\mathbf{g}(\cdot) \leq \mathbf{0}$ angesetzt werden, die alle Bedingungen berücksichtigt. Weiterführend wird davon ausgegangen, dass eine physikalische Modellbildung für das dynamische System durchgeführt worden ist und eine Approximation $\hat{\mathbf{f}} \approx \mathbf{f}$ vorhanden ist. Anhand des Modells wird nun ein Optimierungsproblem formuliert, welches vollkommen entkoppelt vom realen System zu betrachten ist. Für die Lösung des Optimierungsproblems gibt es verschiedene Ansätze [Bet10; PLB12]. Im Rahmen dieser Arbeit sind insbesondere die sogenannten Schießverfahren (*Shooting Methods*) [Kel17] von besonderer Bedeutung. Dabei wird zwischen dem *Single-* und *Multiple-Shooting* unterschieden.

Das zum Single-Shooting zugehörige Optimierungsproblem lautet

$$\min_{\mathbf{u}_k} J(\mathbf{u}_k), \quad \text{u.B.v.} \quad \mathbf{g}(\mathbf{u}_k) \leq \mathbf{0}, \quad k = 1, \dots, H, \quad (4-1)$$

und beinhaltet die Minimierung der Gütfunktion durch die richtige Wahl des Stellgrößenverlaufs unter der Einhaltung der erwähnten technischen Nebenbedingungen. Die Stellgröße jedes einzelnen Zeitschritts stellt im Verbund den Optimierungsvektor dar. In dieser Darstellung steckt das aufgestellte Modell implizit in der Berechnung der Gütfunktion und Gleichheitsnebenbedingung, da für ihre Auswertungen eine Integration über das Zeitgitter τ mit dem ausgewählten Stellgrößenverlauf erforderlich ist. Die exakte Einhaltung der Modelldynamik ist somit bei jeder Iteration des Optimierungssolvers sichergestellt. Die Evaluierung eines bestimmten Stellgrößenverlaufs gleicht, bildlich gesprochen, einem Schießvorgang, welcher den Zielzustand \mathbf{x}_G erreichen soll. Hiernach leitet sich der entsprechende Name des Verfahrens ab. Ein ähnlicher Ansatz zum Single- wird beim

Multiple-Shooting gewählt. Das zugehörige Optimierungsproblem lautet in seiner voll-diskretisierten Form

$$\begin{aligned} \min_{\boldsymbol{x}_k, \boldsymbol{u}_k} J(\boldsymbol{x}_k, \boldsymbol{u}_k), \quad \text{u.B.v.} \quad & \boldsymbol{g}(\boldsymbol{x}_k, \boldsymbol{u}_k) \leq \mathbf{0}, & k = 1, \dots, H, \\ & \boldsymbol{h}(\boldsymbol{x}_k, \boldsymbol{u}_k) := \boldsymbol{x}_{k+1} - \hat{\boldsymbol{f}}(\boldsymbol{x}_k, \boldsymbol{u}_k) = \mathbf{0}, & k = 1, \dots, H-1, \end{aligned} \quad (4-2)$$

wobei die Modelldynamik nun explizit als Gleichheitsnebenbedingung $\boldsymbol{h}(\cdot) = \mathbf{0}$ auftaucht und neben den Stellgrößen nun auch die Zustandsvektoren eines jeden Schrittes als OptimierungsvARIABLEN vertreten sind. Im Gegensatz zum Single- wird beim Multiple-Shooting jeder Zeitschritt als ein Schießvorgang angesehen. Es gibt allerdings auch Abstufungen des Verfahrens, die nicht jeden Zeitschritt als Schuss annehmen, sondern mehrere Zeitschritte zusammenfassen und damit eine geringere Komplexität aufweisen bzw. sich eher in Richtung des Single-Shootings orientieren. Im Allgemeinen sind die Konvergenzeigenschaften des Multiple-Shooting als besser zu bewerten, wodurch es häufiger für schwierigere Optimalsteuerungsprobleme bzw. komplexere Systeme angewandt wird [Bet98].

Durch eine computergestützte Lösung der Optimierungsprobleme erhält der Anwender den zur Aufgabe passenden Steuerungsverlauf. Dieser ist jedoch relativ zum aufgestellten Modell zu sehen und hängt daher stark von dessen Prädiktionsqualität ab. Der Steuerungsverlauf kann daher bei der Anwendung am realen System aufgrund von Modellfehlern ein anderes Verhalten hervorrufen, als es durch die Simulation zu erwarten gewesen wäre (vgl. Bild 1-2). Um dieser Problematik bei der Inbetriebnahme zu begegnen, wird ein hybrider Modellierungsansatz vorgeschlagen. Hierbei werden Messdaten, die vom realen System stammen, als Grundlage für eine Korrektur der Dynamikgleichungen verwendet. Als maschinelles Lernverfahren werden wiederum GPs eingesetzt, die die Unsicherheit über die Dynamik quantifizieren und im Entwurfsprozess berücksichtigen. Das Optimalsteuerungsproblem (4-2) wird dahingehend mit dem hybriden Modell formuliert und die Modellunsicherheiten explizit einbezogen. Das Ziel dieses Hauptkapitels ist es ein hybrides Optimalsteuerungsverfahren aufzustellen, welches eine erweiterte Inbetriebnahme für partiell bekannte Systemdynamiken ermöglicht und das Problem der herkömmlichen Vorgehensweise adressiert.

Im Gegensatz zum vorherigen Kapitel 3 bzw. dem Entwurf mittels der BO, wird im Rahmen dieses Kapitels davon ausgegangen, dass das System soweit identifizierbar ist, dass eine Klarheit darüber herrscht, was die Zustände des Systems genau sind und wie sie mathematisch definiert sind. Außerdem soll vorausgesetzt werden, dass die Zustände in Form einer Messung, Berechnung oder Beobachtung zugänglich sind. Beide Vorausset-

zungen waren beim Ultraschalldrahtbondprozess nicht gegeben, sodass eine Korrektur auf der Ebene der Dynamikgleichung praktisch nicht umsetzbar war. In diesem Zusammenhang stellte die BO eine Möglichkeit dar, um einen Steuerungsentwurf auf der Basis der Gütfunktion durchzuführen, wobei die genutzten GPs nicht als Korrektur für die Dynamik, sondern als Anpassung auf der Ebene der Gütfunktion fungierten.

4.2 Mehrfachpendel auf einem Wagen als Anwendungsbeispiel

Eine interessante Systemklasse stellen Mehrfachpendel auf einem Wagen dar. Ihre Eigenschaften in einem wissenschaftlichen Kontext innerhalb der Regelungstechnik sind für die Erforschung und Entwicklung neuartiger Steuerungs- und Regelungskonzepte vorteilhaft. Sie sind durch einen einfachen Aufbau gekennzeichnet und lassen sich somit vergleichsweise einfach als Prüfstand realisieren. So steht dieser Arbeit ebenfalls ein Prüfstand (s. Bild 1-4) zur Verfügung, an welchem das entwickelte hybride Optimalsteuerungsverfahren und Erweiterungen getestet werden. Das akademisch geprägte System besteht aus einem Wagen, welcher sich ausschließlich horizontal bewegen kann, und einem oder mehreren nacheinander angeordneten Pendelarmen. Die Pendelarme sind dabei rotatorisch miteinander verbunden. Der Fußpunkt dieser Konfiguration wird an dem Wagen, ebenfalls rotatorisch, befestigt. Mögliche Aufgabenstellungen, die an Pendelsysteme geknüpft sind, sind der Aufschwung aus der unteren stabilen in die obere instabile Ruhelage, sowie das Balancieren in der oberen Ruhelage. Um den Zielzustand (die obere Ruhelage) durch die Bewegung des Wagens zu erreichen und beizubehalten, ist eine komplizierte Steuerabfolge erforderlich, die selbst für einen Menschen im Sinne des manuellen Einstellens nur äußerst schwer zu identifizieren ist. Dies liegt unter anderem an den Eigenschaften, die diese Systeme charakterisieren, bspw. den trigonometrischen Zusammenhängen, welche die Bewegung beschreiben. Eine genaue Darstellung wird in diesem Unterkapitel durch eine physikalische Modellbildung gegeben. Das System zählt außerdem zu der Klasse der uneraktuierten Systeme, da jeweils die Anzahl der Freiheitsgrade größer als die Anzahl der Steuereingänge ist. Die Pendelarme können nur indirekt über den Wagen beeinflusst werden. Das Doppelpendel weist zudem ein chaotisches, sensitivs Bewegungsverhalten auf. Kleine Änderungen der Ausgangssituation bzw. des Steuerungssignals führen dabei rasch zu größer werdenden Abweichungen des Bewegungsablaufs. Insbesondere für die praktische Umsetzung müssen beim Steuerungs- und Regelungsentwurf Zustandsbeschränkungen, in Form von Weg-, Geschwindigkeit- und Beschleunigungs-Begrenzungen, berücksichtigt werden.

Am Lehrstuhl für Regelungstechnik und Mechatronik der Universität Paderborn wurde ein Doppelpendel-Prüfstand konzipiert und realisiert (s. Bild 1-4). Die Pendelarme wurden dabei modular ausgelegt, sodass ein schneller Umbau zum Einfach- oder Dreि-

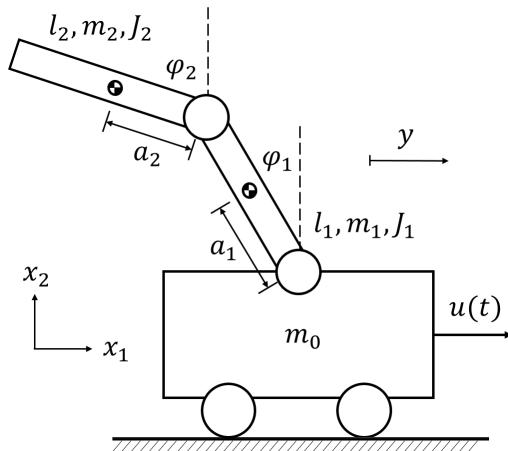


Bild 4-1: Prinzipskizze des Doppelpendels auf einem Wagen (vgl. Bild 1-4).

fachpendel vorgenommen werden kann. Die Antriebseinheit besteht aus dem Linearmotor LKL 20-85 Bosch Rexroth und einem Frequenzumrichter, der in einem kraft- oder geschwindigkeitsgeregelten Modus betrieben werden kann. Aufgrund der Rahmenkonstruktion ist der Weg des Linearantriebs auf ± 0.6 m begrenzt. Des Weiteren können eine maximale Geschwindigkeit von $\pm 5 \frac{\text{m}}{\text{s}}$ und eine maximale Beschleunigung von $\pm 100 \frac{\text{m}}{\text{s}^2}$ gestellt werden. Die Zustandserfassung erfolgt durch optische, hochauflösende Winkelencoder, die 20.000 Striche pro Umdrehung besitzen. Die Winkelgeschwindigkeiten der Pendelarme werden auf der Recheneinheit näherungsweise über eine Finite-Differenzen-Approximation bestimmt. Die digitale Signalverarbeitung wird von einer dSPACE Echtzeithardware durchgeführt, die mit einer Abtastfrequenz von 1kHz arbeitet.

Modellbildung für das Doppelpendel auf einem Wagen

Die Modellbildung für das Doppelpendel auf einem Wagen orientiert sich an [TKOT11; Tim13] und wird nach dem *Lagrange-Formalismus* durchgeführt. Bild 4-1 zeigt das physikalische Ersatzbild des Prüfstands mit den für die Dynamik relevanten mechanischen Parametern und Koordinaten. Das Mehrkörpersystem besteht aus drei starren Körpern. Der Wagen besitzt die Masse m_0 und die Pendelarme die Massen m_i mit $i = 1, 2$. Zudem weisen die Pendelarme eine Länge von l_i und ein Trägheitsmoment von J_i auf – ihr Abstand vom Fußpunkt zum Schwerpunkt beträgt jeweils a_i . Die horizontal verlaufende Koordinate $y(t)$ wird dem Wagen zugeordnet und die beiden Winkelkoordinaten $\varphi_i(t)$, $i = 1, 2$ den Pendelarmen. Die Winkel werden, ausgehend von der vertikalen Nulllage, positiv gegen den Uhrzeigersinn gezählt. Zunächst soll die am Wagen wirkende Kraft als Systemeingang $u_F(t)$ betrachtet werden. Etwaige Reibungseffekte zwischen dem Linearantrieb und der Wagenstrecke werden vernachlässigt, da sie durch die interne Geschwindigkeitsregelung des Linearmotors kompensiert werden. Die Reibung der rotatorischen Gelenke wird jedoch über lineare Ansätze mit den Dämpfungskonstanten d_i

berücksichtigt. Weitere Störeinflüsse, wie bspw. Schwingungen der Rahmenkonstruktion, werden nicht modelliert.

Der Lagrange-Formalismus setzt zunächst die Bestimmung der Orts- bzw. Schwerpunktskoordinaten $\mathbf{x}^{(i)} = [x_1^{(i)}, x_2^{(i)}]^T, i = 0, 1, 2$, durch die generalisierten Koordinaten $(\varphi_1, \varphi_2, y)$ voraus. Hierzu werden die Vektoren

$$\mathbf{x}^{(0)} = \begin{bmatrix} y \\ 0 \end{bmatrix}, \quad \mathbf{x}^{(1)} = \begin{bmatrix} y - a_1 \sin(\varphi_1) \\ a_1 \cos(\varphi_1) \end{bmatrix}, \quad \mathbf{x}^{(2)} = \begin{bmatrix} y - l_1 \sin(\varphi_1) - a_2 \sin(\varphi_2) \\ l_1 \cos(\varphi_1) + a_2 \cos(\varphi_2) \end{bmatrix},$$

definiert. Mit dieser Charakterisierung werden die gesamte kinetische Energie V und die gesamte potentielle Energie U des Systems aufgestellt.

$$V = \frac{1}{2} \left(\sum_{i=0}^2 m_i \left\| \dot{\mathbf{x}}^{(i)} \right\|_2^2 + \sum_{i=1}^2 J_i \dot{\varphi}_i^2 \right), \quad U = \sum_{i=0}^2 m_i g x_2^{(i)}. \quad (4-3)$$

Dabei bezeichnet $\|\cdot\|_2$ die euklidische Norm und g die Gravitationskonstante mit dem Wert $9.81 \frac{\text{m}}{\text{s}^2}$. Für die dissipativen Reibmomente in den Pendelgelenken werden lineare Ansätze mit $F_1 = d_2(\dot{\varphi}_2 - \dot{\varphi}_1) - d_1 \dot{\varphi}_1$ und $F_2 = d_2(\dot{\varphi}_1 - \dot{\varphi}_2)$ verwendet. Zu den Reibmomenten tritt die am Wagen angreifende Kraft F_0 hinzu, welche gleichzeitig die einzige Eingangsgröße $u_F(t)$ in das System darstellt. Die Lagrangeschen Gleichungen zweiter Art [Sko18] für das betrachtete System in Indexschreibweise lauten:

$$\frac{d}{dt} \frac{\partial L(\mathbf{q}, \dot{\mathbf{q}})}{\partial \dot{q}_i} - \frac{\partial L(\mathbf{q}, \dot{\mathbf{q}})}{\partial q_i} = F_i, \quad i = 0, 1, 2. \quad (4-4)$$

Darin ist $L(\mathbf{q}, \dot{\mathbf{q}}) = V(\mathbf{q}, \dot{\mathbf{q}}) - U(\mathbf{q})$ die sogenannte Lagrangefunktion, die sich aus der Differenz aus kinetischer Energie V und potentieller Energie U (s. (4-3)) zusammensetzt. In dem Vektor $\mathbf{q} = [q_1, q_2, q_0]^T = [\varphi_1, \varphi_2, y]^T$ werden die verallgemeinerten Lagekoordinaten und in dem Vektor $\dot{\mathbf{q}} = [\dot{\varphi}_1, \dot{\varphi}_2, \dot{y}]^T$ die verallgemeinerten Geschwindigkeiten zusammengefasst. Die einzelnen Bestandteile der Gleichungen (4-4) führen auf das nachfolgende nichtlineare System von Differentialgleichungen (DGL) zweiter Ordnung:

$$\mathcal{M}(\mathbf{q}) \ddot{\mathbf{q}} = \mathcal{F}(\mathbf{q}, \dot{\mathbf{q}}) + \mathcal{B} u_F, \quad (4-5)$$

mit der Matrix

$$\begin{aligned}\mathcal{M}^{(3 \times 3)} &= \begin{bmatrix} \mathcal{M}_{1,1}^{(2 \times 2)} & \mathcal{M}_{1,2}^{(2 \times 1)} \\ \mathcal{M}_{2,1}^{(1 \times 2)} & \mathcal{M}_{2,2}^{(1 \times 1)} \end{bmatrix} \\ &= \begin{bmatrix} J_1 + a_1^2 m_1 + l_1^2 m_2 & a_2 l_1 m_2 \cos(\varphi_1 - \varphi_2) & (l_1 m_2 - a_1 m_1) \cos(\varphi_1) \\ a_2 l_1 m_2 \cos(\varphi_1 - \varphi_2) & J_2 + a_2^2 m_2 & -a_2 m_2 \cos(\varphi_2) \\ (l_1 m_2 - a_1 m_1) \cos(\varphi_1) & -a_2 m_2 \cos(\varphi_2) & m_1 + m_2 + m_0 \end{bmatrix},\end{aligned}\quad (4-6)$$

und den Vektoren

$$\begin{aligned}\mathcal{F}^{(3 \times 1)} &= \begin{bmatrix} \mathcal{F}_1^{(2 \times 1)} \\ \mathcal{F}_2^{(1 \times 1)} \end{bmatrix} \\ &= \begin{bmatrix} d_2(\dot{\varphi}_2 - \dot{\varphi}_1) - d_1 \dot{\varphi}_1 - a_2 l_1 m_2 \sin(\varphi_1 - \varphi_2) \dot{\varphi}_2^2 + g(a_1 m_1 + l_1 m_2) \sin(\varphi_1) \\ d_2(\dot{\varphi}_1 - \dot{\varphi}_2) - a_2 l_1 m_2 \sin(\varphi_1 - \varphi_2) \dot{\varphi}_1^2 + g a_2 m_2 \sin(\varphi_2) \\ (l_1 m_2 - a_1 m_1) \sin(\varphi_1) \dot{\varphi}_1^2 - a_2 m_2 \sin(\varphi_2) \dot{\varphi}_2^2 \end{bmatrix}, \\ \mathcal{B}^{(3 \times 1)} &= [0, 0, 1]^T.\end{aligned}\quad (4-7)$$

Die symmetrische Matrix \mathcal{M} wird Massenmatrix genannt und der Vektor \mathcal{F} fasst Coriolis-Zentrifugal-, sowie dissipative Kräfte zusammen. Durch den Eingangsvektor \mathcal{B} ist gut zu erkennen, dass sich nur die dritte DGL direkt durch die Aktorkraft $u_F(t)$ beeinflussen lässt. Aufgrund der Geschwindigkeitsregelung des Linearantriebs bietet es sich an, eine partielle Zustandslinearisierung [Tim13] durchzuführen, sodass anstelle der Aktorkraft F_0 die Wagenbeschleunigung \ddot{y} als Eingangsgröße $u_y(t)$ betrachtet wird. Die generalisierten Koordinaten $\mathbf{q} = [\mathbf{q}_1, \mathbf{q}_2]^T$ werden zunächst in die passiven ($\mathbf{q}_1 = [\varphi_1, \varphi_2]^T$) Freiheitsgrade und den aktiven ($q_2 = y$) Freiheitsgrad unterteilt. Diese Art der Partitionierung wurde bereits in (4-6), (4-7) angedeutet und führt auf die Darstellung

$$\begin{aligned}\mathcal{M}_{1,1} \ddot{\mathbf{q}}_1 + \mathcal{M}_{1,2} \ddot{\mathbf{q}}_2 &= \mathcal{F}_1, \\ \mathcal{M}_{2,1} \ddot{\mathbf{q}}_1 + \mathcal{M}_{2,2} \ddot{\mathbf{q}}_2 &= \mathcal{F}_2 + u_F.\end{aligned}\quad (4-8)$$

Die obere Gleichung wird nun nach dem passiven Anteil umgestellt. Diese Umformung ist möglich, da \mathcal{M} positiv definit ist, und die Inverse $\mathcal{M}_{1,1}^{-1}$ somit existiert. Anschließend wird die umgeformte obere in die untere Gleichung eingesetzt und nach der Eingangskraft aufgelöst:

$$u_F = \mathcal{M}_{2,1} \mathcal{M}_{1,1}^{-1} \mathcal{F}_1 + (\mathcal{M}_{2,2} - \mathcal{M}_{2,1} \mathcal{M}_{1,1}^{-1} \mathcal{M}_{1,2}) \ddot{\mathbf{q}}_2 - \mathcal{F}_2. \quad (4-9)$$

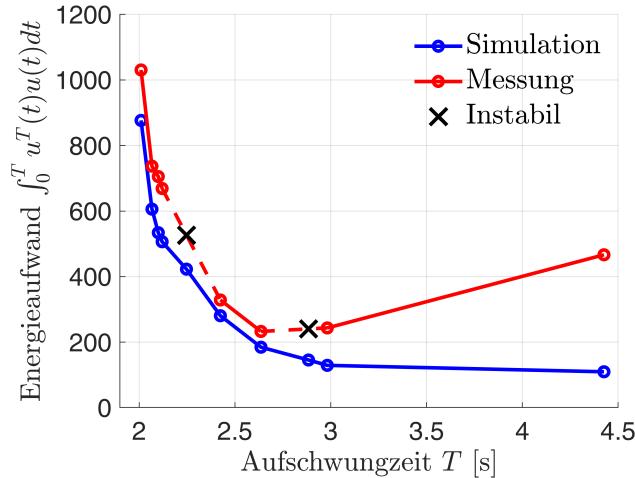


Bild 4-2: Simulierte Paretofront und zugehörige Prüfstandsmessungen für das Dreifachpendel auf einem Wagen.

Die entstandene Gleichung (4-9) für $u_F(\ddot{q}_2)$ motiviert die Einführung einer fiktiven Rückführung der Wagenbeschleunigung \ddot{y} , sodass die untere Gleichung in (4-8) vollständig kompensiert wird. Übrig bleibt ein reduziertes System mit der Beschleunigung des Wagens als neuen Systemeingang

$$\begin{aligned} \mathcal{M}_{1,1}\ddot{q}_1 &= \mathcal{F}_1 - \mathcal{M}_{1,2}u_y, \\ \ddot{q}_2 &= u_y, \end{aligned} \tag{4-10}$$

welches eine geringere Berechnungszeit im Vergleich zum Originalsystem (4-5) benötigt.

Abweichung zwischen Simulation und Realität am Beispiel der Paretofront eines Dreifachpendels auf einem Wagen

Eine Erweiterung des Doppelpendels stellt das Dreifachpendel auf einem Wagen dar. Für die Realisierung wird am Prüfstand lediglich ein weiterer Pendelarm an das Ende des äußeren Pendelarms montiert. Die grundsätzlichen Eigenschaften ändern sich durch diesen Aufbau nicht, sondern werden in gewisser Weise verstärkt. Das chaotische und sensitive Bewegungsverhalten bleibt erhalten, allerdings wird es durch den zusätzlichen Pendelarm erhöht, sodass die Prädiktionsfähigkeit eines zugehörigen Modells generell abnimmt. Um die Effekte eines nicht idealen Modells zu verdeutlichen, wurden am Dreifachpendel verschiedene Studien durchgeführt. Hierzu ist insbesondere die studentische Arbeit [Ism18] zu nennen, in welcher der Aufschwung und die Stabilisierung nach dem Vorgehen in [Tim13] erfolgreich umgesetzt werden konnte. Die Inbetriebnahme gestaltete sich jedoch äußerst schwierig aufgrund der erwähnten Problematik in Bezug auf die genaue Modellbildung des Systems. Ein besonders aussagekräftiges Ergebnis ist im Zusammenhang mit der Paretofront des Optimalsteuerungsentwurfs entstanden. In Bild 4-2 ist hierzu die Auswertung der simulierten Paretofront dargestellt. Als Ziele wurden hier-

bei der Energieaufwand und die Aufschwungzeit betrachtet. Durch die Veränderung der jeweiligen Gewichtung kann die in blau dargestellte Paretofront abgetastet werden. Die Abtastpunkte sind durch die blauen Punkte visualisiert. In der Simulation prägt sich der typische Verlauf einer Paretofront aus, wobei entlang der Kurve nur dominante Punkte miteinander verbunden werden. Die den blauen Punkten zugehörige Optimalsteuerung wurde gleichermaßen am Prüfstand erprobt. Die rote Kurve stellt die gemessenen Ergebnisse dar. Dabei wird deutlich, dass der Energieaufwand insgesamt höher ausfällt als es durch die Simulation prognostiziert wurde. Der Grund hierfür ist die Verwendung der Zwei-Freiheitsgrade-Regelungsstruktur (s. Bild 2-1), wobei die eingesetzte Optimalsteuerung um einen Reglereingriff erweitert wird, um etwaige Modellfehler bzw. äußere Störungen zu kompensieren. Der Reglereingriff ist notwendig, da sich das System ansonsten zu stark von der Soll-Trajektorie entfernt und den geforderten Zielzustand nicht erreichen kann. Der zusätzliche Energieaufwand ist dabei für die einzelnen Punkte der Paretofront nicht konstant, sondern hängt von der Aufschwungzeit ab. Je länger der Aufschwung, desto mehr Energie wird für die Einhaltung der Trajektorie gebraucht. Dies ist nachvollziehbar, da die Prädiktionsfähigkeit mit der Zeit abnimmt und sich bestehende Modellfehler entsprechend zeitlich aufsummieren. Neben diesem Sachverhalt sind außerdem zu einigen blauen keine zugehörigen roten Punkte zu beobachten (so bei $T = 2.25$ s und $T = 2.8$ s). Dies liegt daran, dass zu diesen Punkten kein erfolgreicher Aufschwung am Prüfstand umgesetzt werden konnte. In diesen Fällen war der Reglereingriff zu groß, sodass das System instabil wurde und die technischen Rahmenbedingungen verletzt wurden. Dieses Ergebnis ist dahingehend interessant, da sich die benachbarten Paretopunkte bzw. ihre Trajektorien nicht allzu stark von den instabilen Fällen unterscheiden und daher kein besonderer Grund angegeben werden kann, warum eine Realisierung nicht möglich war. Insgesamt ergibt sich damit ein Bild, welches eine ideale Paretofront aus der Simulation zeigt, welche allerdings in der Realität nicht gleichermaßen umgesetzt werden kann. Die in rot dargestellte reale Paretofront weicht dabei von der Simulation ab, und kann daher im Grunde nicht mehr als Paretofront bzgl. des realen System bezeichnet werden, da sie einige Punkte (am rechten Rand) enthält, die von anderen Punkten dominiert werden. Dieser Sachverhalt verdeutlicht die Schwierigkeiten, welche sich bei der Inbetriebnahme durch einen modellbasierten Entwurf ergeben und motiviert den Einsatz von hybriden Verfahren, welche die Informationsquelle der Messdaten nutzen, um den modellbasierten Optimalsteuerungsentwurf zu verbessern. Bevor der hybride Ansatz näher besprochen wird, soll zunächst ein rein datengetriebenes Verfahren vorgestellt werden und dessen Vor- und Nachteile erläutert werden. Auf diese Weise lässt sich die hybride Optimalsteuerung besser in den Gesamtkontext einordnen und im weiteren Verlauf des Kapitels ein Vergleich zwischen einem Ansatz ohne und mit Vorwissen gezogen werden. Im Rahmen des modellbasierten RLs und im Zusammenhang mit der Regelungstechnik

ist hierzu der PILCO-Ansatz zu nennen, welcher im nächsten Abschnitt detailliert mit der Anwendung auf das Doppelpendel auf einem Wagen erklärt wird.

4.3 Probabilistic Inference for Learning Control

In diesem Unterkapitel wird der PILCO-Algorithmus nach [Dei10; DRF12; DEPF14; DFR15] vorgestellt. PILCO ist ein modellbasiertes RL-Verfahren zur Lösung von Regelungsaufgaben. Das Verfahren setzt dabei wenig Vorwissen über die Regelungsaufgabe voraus und beinhaltet das Erlernen eines dynamischen Modells auf der Grundlage von Messungen am betrachteten realen System. Zudem zeichnet sich PILCO durch seine hohe Effizienz im Hinblick auf die Auswertung vorhandener Messdaten und die Nutzung von Rechenkapazitäten aus. Aus diesem Grund wird PILCO in der Literatur als ein *State-of-the-Art* RL-Algorithmus für die kontinuierliche Regelung von realen dynamischen Systemen angesehen.

Eine physikalische Modellbildung wird im Kontext von RL als Expertenwissen bezeichnet, welches für die Anwendung von PILCO nicht vorausgesetzt wird. Aus diesem Grund wird bei PILCO die Systemdynamik mittels Gauß-Prozess-Regression nachgebildet. Genauer betrachtet, wird für jede Zustandsdimension ein eigener GP angesetzt. Außerdem wird die Annahme getroffen, dass die GPs voneinander unabhängig sind, sich also gegenseitig nicht beeinflussen können. Diese Annahme stellt keine Einschränkung dar, da jede Lernaufgabe in Hinblick auf ihre Zustandsdimension separat voneinander betrachtet werden kann. Eine Verknüpfung der einzelnen Gleichungen kommt nur durch die zeitliche Entwicklung bzw. Integration des Zustandsvektors (Zustandspropagation, s. Abschnitt 2.3.2) zustande, jedoch nicht durch den zugehörigen GP selbst. Durch den probabilistischen Charakter eines GP wird über ihn die Unsicherheit über die Dynamikfunktion ausgedrückt. Erst durch die Aufnahme von Messdaten und die Bildung des Posterior für jeden GP wird diese Unsicherheit in bestimmten Bereichen des Zustandsraums reduziert, wodurch ein kontinuierlicher Lernvorgang der Dynamik des Systems angestoßen wird. Der probabilistische Charakter steht daher im engen Zusammenhang mit dem Kompromiss zwischen einem explorativen und exploitativen Verhalten des PILCO-Algorithmus und ist essentiell wichtig für den Erfolg des Verfahrens. Unter anderem wurde in [DFR15] durch diverse Versuche gezeigt, dass die Einbringung der Wahrscheinlichkeitstheorie für das erfolgreiche Lernen notwendig ist, denn der Austausch der probabilistischen GPs mit einem vergleichbaren deterministischen Modellansatz führte für die betrachteten Anwendungsbeispiele zum Misserfolg. Ein Einfluss einer stochastischen Komponente, welche das Ergebnis negativ beeinflussen könnte, wird über die mehrfache Wiederholung der Experimente ausgeschlossen.

Algorithmus 3 Probabilistic Inference for Learning Control (PILCO)

- 1: **Initialisierung:** Wähle eine zufällige Reglerparametrisierung θ . Schalte dem System die Regelungsstrategie $r(\mathbf{x}; \theta)$ auf und sammle erste initiale Messdaten.
 - 2: **Wiederhole** bis die Regelungsaufgabe erfüllt ist:
 - 3: Lerne die Systemdynamik mittels GP anhand der vorhandenen Messdaten.
 - 4: **Wiederhole** bis ein Konvergenzkriterium erreicht ist:
 - 5: Berechne die Gesamtkosten $J(\theta)$ auf Basis des gelernten GP-Modells.
 - 6: Führe eine gradientenbasierte Verbesserung der Reglerparametrisierung durch und erhalte θ^* .
 - 7: Schalte dem System die Regelungsstrategie $r(\mathbf{x}; \theta^*)$ auf und sammle parallel weitere Messdaten, die den bestehenden Messdaten hinzugefügt werden.
-

Das übergeordnete Ziel bei PILCO ist es, einen parametrisierten Regler $\mathbf{u} = r(\mathbf{x}; \theta)$ zu finden, der die erwarteten Langzeitkosten bzw. das Gütemaß

$$J(\theta) = \sum_{k=0}^H \mathbb{E}[c(\mathbf{x}_k)] \in \mathbb{R}, \quad \mathbf{x}_0 \sim \mathcal{N}(\boldsymbol{\mu}_0, \Sigma_0), \quad (4-11)$$

über $H+1$ Zeitschritte minimiert. Dabei ist $c(\mathbf{x}_k)$ eine momentane Kostenfunktion, die jeden Zustand \mathbf{x}_k anhand eines Zielzustandes \mathbf{x}_G bewertet. Zudem wird der Erwartungswert der Kosten betrachtet, da die Zustände \mathbf{x}_k aufgrund der GP-Modellierung nur mit einer bestimmten Wahrscheinlichkeitsverteilung vorliegen. Die Minimierung der Gesamtkosten (4-11) wird mittels gradientenbasierten Optimierungssolvern durchgeführt. Die Optimierungsvariablen sind dabei die Elemente der Parametrisierung θ , durch die der Regler und dessen Qualität beeinflusst wird. Während der Optimierung wird mehrfach eine Langzeitprädiktion der Form $p(\mathbf{x}_0 | \theta) \mapsto p(\mathbf{x}_1 | \theta) \mapsto \dots \mapsto p(\mathbf{x}_H | \theta)$ mit Hilfe des GP-Modells berechnet, um die Langzeitkosten (4-11) auszuwerten und eine kontinuierliche Verbesserung des Reglers vorzunehmen. Eine Besonderheit dabei ist, dass die Zustandsverteilungen über den Moment Matching Ansatz als Normalverteilungen approximiert werden. Bei der Verwendung des SE-Kernels für die einzelnen GPs kann in [DFR15] gezeigt werden, dass die Berechnung für den Erwartungsvektor und die Kovarianzmatrix des Zustands analytisch aufgestellt werden kann. Das hat den Vorteil, dass die zugehörige Berechnung relativ günstig ist und gleichermaßen der Gradient für die Optimierung analytisch abgeleitet werden kann.

Algorithmus 3 fasst die wesentlichen Schritte zusammen, welche im Nachfolgenden näher erläutert werden. PILCO beinhaltet eine Initialisierungsphase (Zeile 1), sowie eine äußere (Zeile 2-7) und eine innere (Zeile 4-6) Schleife. In der Initialisierungsphase wird dem System eine zufällige Steuerfolge aufgeschaltet. Dabei werden erste Messdaten bzw. Zustandsübergänge des Systems aufgenommen. Anschließend wird die äußere Schleife von PILCO ausgeführt. Die Schleife beginnt mit dem Training eines GP-Modells, welches die

Systemdynamik abbilden soll. Training bedeutet in diesem Zusammenhang, dass die Hyperparameter der zugehörigen GPs über einen Optimierungssolver variiert werden, sodass die Likelihoodfunktion maximiert wird. Das Training wird anhand der bestehenden Messdaten durchgeführt. Im Anschluss erfolgt die Optimierung der Reglerparametrisierung θ (innere Schleife). Auf Grundlage des gelernten GP-Modells wird hierbei das Systemverhalten laufend mit unterschiedlichen Reglerparametrisierungen simuliert, bis eine optimale Parameterkonfiguration θ_* gefunden wurde und die Optimierung konvergiert. Im Anschluss an die Reglerparameteroptimierung wird dem System die aktuelle Regelungsstrategie $r(\mathbf{x}; \theta_*)$ aufgeschaltet und parallel neue Messdaten aufgenommen. Die neuen Messdaten werden den bestehenden Messdaten hinzugefügt und die äußere Schleife erneut ausgeführt. Dem GP-Training stehen dabei weitere Messdaten zur Verfügung, wodurch sich ein genaueres Modell des Systems ergibt. Durch die Veränderung des Modells führt die Optimierung innerhalb der inneren Schleife zu einer besseren Lösung θ_* . Die äußere Schleife wird solange ausgeführt, bis die Regelungsaufgabe erfüllt wurde. Das Ergebnis von PILCO ist ein optimierter Regler, der einen gewünschten Zielzustand erreicht und stabilisiert. Zusätzlich steht am Ende des Verfahrens das trainierte ML-Modell zur Verfügung, welches bspw. für einen weiterführenden Beobachterentwurf verwendet werden kann [DHH09].

PILCO setzt eine parametrisierbare Ansatzfunktion mit $r(\mathbf{x}; \theta) : \mathbb{R}^{n_x} \mapsto \mathbb{R}^{n_u}$ für die Darstellung des Reglers voraus. Dabei steht n_x für die Anzahl der Zustandsgrößen und n_u für die Anzahl der Stellgrößen. Da in dieser Arbeit Systeme mit einer Stellgröße betrachtet werden, gelten die nachfolgenden Betrachtungen für den Spezialfall $n_u = 1$. Bei PILCO wird für jede Stellgröße der Ansatz

$$r(\mathbf{x}; \theta) = \mathbf{k}(\mathbf{M}, \mathbf{x})^T \underbrace{\mathbf{K}(\mathbf{M}, \mathbf{M})^{-1} \mathbf{v}}_{:=\alpha} = \sum_{i=1}^{n_m} k(\mathbf{m}_i, \mathbf{x}) \alpha_i, \quad (4-12)$$

mit $\mathbf{M} = [\mathbf{m}_1, \dots, \mathbf{m}_{n_m}] \in \mathbb{R}^{n_x \times n_m}$ gewählt, wobei n_m für die Anzahl der Basisfunktionen steht. Bei formaler Betrachtung von (4-12) fällt auf, dass es sich um die Posterior-Gleichung für den Mittelwert eines GP handelt (vgl. (2-31)). Die Ansatzfunktion für die Stellgröße entspricht daher einem GP mit der SE-Kovarianzfunktion. Für die Posterior-Varianz wurde keine Gleichung angegeben, da diese im Kontext einer Regelung identisch null gesetzt wird. Es handelt sich daher um einen deterministischen GP, der keine Unsicherheiten aufweist. Jedem Zustand wird somit eine eindeutige Stellgröße zugeordnet. Um den GP-Regler vollständig zu definieren, müssen Ein-, Ausgangsdaten und Hyperparameter festgelegt werden. Anders als bei der Modellierung mittels GP, handelt es sich bei den Ein- und Ausgangsdaten des GP-Reglers um fiktive Daten, die unter anderem die Optimierungsvariablen des PILCO-Algorithmus darstellen. In der Matrix \mathbf{M} aus (4-12)

werden die Eingangsdaten $\mathbf{m}_i \in \mathbb{R}^{n_x}$, $i = 1, \dots, n_m$ und in dem Vektor $\mathbf{v} \in \mathbb{R}^{n_m}$ die Ausgangsdaten zusammengefasst. Mit der Variable α wird das Produkt aus $\mathbf{K}(\mathbf{M}, \mathbf{M})^{-1}$ und \mathbf{v} abgekürzt. Die Hyperparameter der SE-Kovarianzfunktion lauten σ_f^2 , σ_w^2 und $l_1^2, \dots, l_{n_x}^2$. Die Parameter σ_f^2 und σ_w^2 sind nach [DFR15] mit den Werten 1 und 0.01 fest vorgegeben. Die restlichen Hyperparameter l_i^2 werden den Optimierungsvariablen zugeordnet. Insgesamt ergibt sich somit der Parametervektor $\boldsymbol{\theta} := [\mathbf{m}_1^T, \dots, \mathbf{m}_{n_m}^T, \mathbf{v}^T, l_1, \dots, l_{n_x}]^T$ aus dem Raum $\mathbb{R}^{n_x n_m + n_m + n_x}$, der sämtliche Optimierungsvariablen des GP-Reglers beinhaltet. Als momentane Kostenfunktion wird bei PILCO ein begrenzt quadratischer Ansatz mit

$$c_q(\mathbf{x}) = 1 - \exp(-\frac{1}{2}(\mathbf{x} - \mathbf{x}_G)^T \mathbf{W}_x^{-1} (\mathbf{x} - \mathbf{x}_G)) \in [0, 1], \quad (4-13)$$

eingeführt. Über die symmetrische und positiv definite Matrix \mathbf{W}_x kann eine Gewichtung der unterschiedlichen Zustandsdimensionen eingestellt werden. Ist das System noch sehr weit von seinem Zielzustand entfernt, werden die maximalen Kosten von 1 zugewiesen, wohingegen das Erreichen und die Stabilisierung des Zielzustands mit den minimalen momentanen Kosten von 0 belohnt werden.

In der eigenen Veröffentlichung [HTHT18] wurde der PILCO-Algorithmus zum ersten Mal auf ein reales Doppelpendel auf einem Wagen angewandt und der Aufschwung, sowie die Stabilisierung der oberen Ruhelage erfolgreich umgesetzt. Eine wesentliche Voraussetzung dabei war, dass die technischen Rahmenbedingungen, die durch die Prüfstandskonstruktion (s. Bild 1-4) vorgegeben waren, eingehalten wurden. Hierfür musste die bestehende Kostenfunktion (4-13) geeignet erweitert werden. Um die Nebenbedingungen des Wagens zu berücksichtigen wurde die sogenannte doppelseitige Hingefunktion $c_h(\cdot)$ eingeführt und additiv an die bestehenden Kosten bzgl. des Zielzustands geknüpft. Insgesamt ergibt sich somit eine Gesamtkostenfunktion von

$$J(\boldsymbol{\theta}) = \sum_{k=0}^H \mathbb{E}[c_q(\mathbf{x}_k) + \sum_{i=1}^{n_x} c_h(x_k^{(i)})], \quad \text{mit} \quad c_h(x^{(i)}) = \begin{cases} x_{min}^{(i)} - x^{(i)} & , x^{(i)} < x_{min}^{(i)}, \\ x^{(i)} - x_{max}^{(i)} & , x^{(i)} > x_{max}^{(i)}, \\ 0 & , \text{sonst}, \end{cases} \quad (4-14)$$

wodurch Zustandsüberschreitungen der i -ten Zustandsdimension durch eine lineare Zunahme der Kosten bestraft werden.

Bild 4-3 zeigt auszugsweise die Zustandsverläufe aus der 10. und 27. (finalen) Lerniteration des Verfahrens². Im Bild sind in blau die probabilistischen Prädikationen des aktuellen

²Ein Video der experimentellen Validierung ist unter dem nachfolgenden Link https://youtu.be/N-yrQu9zuOI?si=zyQjmHCXLdfv9a7_ zu finden.

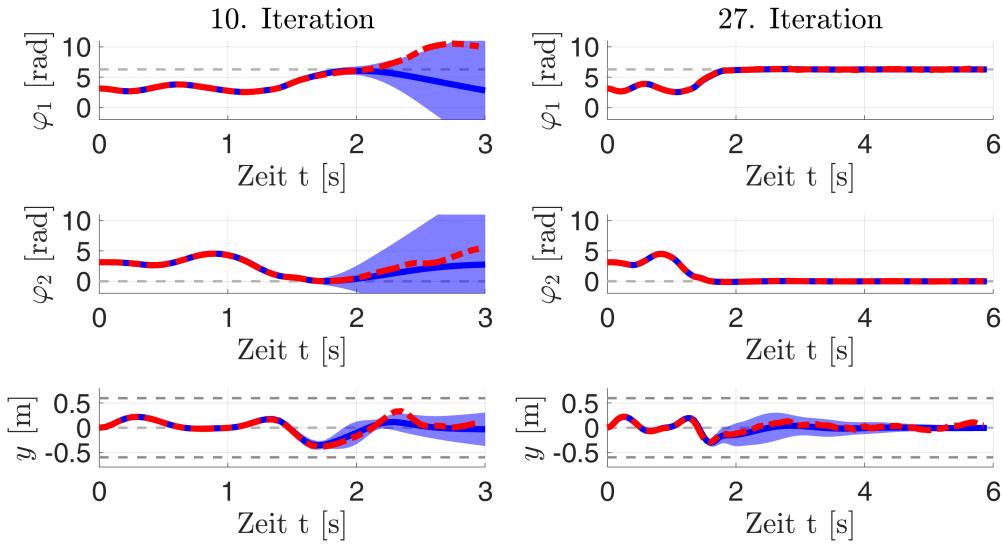


Bild 4-3: Anwendung von PILCO auf ein reales Doppelpendel auf einem Wagen. Darstellung der 10. und 27. Lerniteration des Verfahrens.

GP-Modells dargestellt, sowie die real gemessenen Verläufe (rot) bei der Anwendung des gleichen Reglers. Die ersten Iterationen sind durch eine hohe Unsicherheit bzw. Zustandsvarianz gekennzeichnet. In den ersten 2 s im 10. Versuch ist das Verhalten des Systems gut vorhergesagt, wird jedoch zum Ende der Trajektorie, wegen der Annäherung an die instabile Ruhelage, schlechter, sodass es noch eine deutliche Abweichung zum Zielzustand gibt. Auch wenn das System in dieser Iteration noch nicht im Zielzustand stabilisiert werden kann, ist bereits die Erreichung des Zustands bei etwa 2 s ein wesentlicher Schritt zur Erfüllung des Regelungsziels, denn die dabei aufgenommenen Messdaten haben einen bedeutenden Wert für das zugrundeliegende ML-Modell. Anhand dieser Daten kann das Verhalten in der Nähe des Zielzustands besser vorhergesagt werden, wohingegen die ersten Iterationen durch eine fehleranfällige Extrapolation des Modells geprägt war. Im Laufe der nächsten Iterationen wird der Regler bzw. die Parametrisierung weiter verbessert und der Stellhorizont bei der Einhaltung gewisser Grenzwerte sukzessive erweitert. Nach 27 Lerniterationen wurde das Ziel durch PILCO erreicht, wobei die gesamte Systemdynamik alleine auf der Grundlage von gemessenen Zustandsdaten abgebildet wurde und keinerlei physikalische Zusammenhänge ausgenutzt wurden.

Die Arbeit in [HTHT18] verdeutlicht daher die praktische Anwendbarkeit von Reinforcement Learning Algorithmen in einem regelungstechnischen Kontext und adressiert die Problematik einer fehlerbehafteten Modellbildung durch ein rein datengetriebenes Vorgehen. Die gelöste Aufgabe stellt dabei eine anspruchsvolle Herausforderung dar, welche selbst mit den klassischen bewährten regelungstechnischen Herangehensweisen nur schwerlich erfolgreich umgesetzt werden kann (s. Ausführungen zu Bild 4-2). Trotz der sehr guten Ergebnisse in Bezug auf das komplexe Anwendungsbeispiel muss der PILCO-

Algorithmus allerdings auch in bestimmten Punkten kritisch beurteilt werden. Ein Punkt ist die Dauer des Lernvorgangs. Zwar ist die reine Interaktionszeit, welche PILCO für die Erprobung einer neu gefundenen Regelung benötigt, äußerst gering mit etwa 2 s pro Iteration, dennoch ist auch die Offline-Berechnung bzw. die Lösung des Optimierungsproblems zu berücksichtigen. Aufgrund der hohen Komplexität des Reglers (die Anzahl der Parameter liegt bei etwa 1.000), welcher für das Doppelpendel erforderlich ist, beträgt die Lösung eines Optimierungsproblems innerhalb einer Iteration etwa 2 Stunden. Bei den 27 durchgeführten Iterationen beträgt die gesamte Berechnungszeit damit in etwa zwei Tage. Im Vergleich zu einem herkömmlichen physikalischen Ansatz, welcher idealerweise direkt nach einer adäquaten Identifikation der Modellparameter funktioniert, ist der Lernvorgang damit als äußerst lang zu bezeichnen. Ein Grund hierfür ist, dass das Dynamikmodell von Grund auf gelernt werden muss und dafür erst eine entsprechende Datenlage geschaffen werden muss. Selbst ein einfaches Modell in Form eines linearen Zustandsraummodells um die untere Ruhelage, welches relativ einfach aufgestellt werden könnte, findet bei der Standardvariante des PILCO-Verfahrens keine Verwendung.

Stand der Forschung

In den Folgearbeiten zu [DFR15] bzw. PILCO wurde daher die Einbindung von Vorwissen über die Systemdynamik bearbeitet. Hierbei sind insbesondere [BNv⁺14] und [CH15] zu nennen. In [BNv⁺14] wird die Einbindung eines linearen Modells behandelt und als Anwendungsbeispiel ein Greifroboter, welcher einen Becher aufheben soll, angeführt. Die Einbindung erfolgt dabei über die Mittelwertfunktion des GP-Modells, wobei wiederum eine analytische Berechnung erhalten bleibt. Die Autoren in [CH15] sind daran interessiert, nichtlineares Vorwissen in den PILCO-Algorithmus zu integrieren. Es wird ebenfalls die Prior-Mittelwertfunktion benutzt, allerdings nicht direkt die nichtlineare Dynamikfunktion darüber eingebunden. Die Begründung hierfür ist, dass die analytische Berechenbarkeit verloren geht. In [CH15] wird deswegen ein Vorgehen auf der Grundlage von simulierten Daten des Vorwissens vorgeschlagen. Die simulierten Messdaten werden dann im Rahmen von weiteren zusätzlichen GPs genutzt und zum Training dieser verwendet. Die Posterior-Mittelwertfunktionen der trainierten GPs werden dann über den Prior der ursprünglichen PILCO-GPs eingebunden. Als Vorteil bleibt somit zwar die analytische Berechenbarkeit, insbesondere des Gradienten bzgl. der Gütfunktion, erhalten, allerdings wird das Vorwissen auch nur indirekt über die simulierten Daten eingebunden. Auf diese Weise geht der größte Vorteil des Vorwissens verloren - die Extrapolationseigenschaft - in Zustandsgebieten, in denen (noch) keine Messdaten vorliegen. Ergänzend sind an dieser Stelle auch die Folgearbeiten zu PILCO von [DNM⁺17; EPDP13; GMR16; ADA⁺22] und [MR17] zu nennen. In [DNM⁺17] wird nicht die übliche deterministische GP-Regelung (4-12) eingesetzt, sondern ein in der Regelungstechnik gängiger PID-Regler optimiert. [EPDP13] erweitert PILCO für den Fall des sogenannten *Imitation Learning*, wobei der

menschliche Anwender eine Soll-Trajektorie manuell vorgibt und die Regelungsaufgabe darin besteht dieser Trajektorie bestmöglich zu folgen. Als Anwendungsbeispiel wird ein Roboterarm zum Spielen von Tischtennis betrachtet. Des Weiteren wird in [MR17] die Nutzung des GP-Dynamikmodells im Rahmen eines parallelen Beobachterentwurfs erforscht. Bei der Regleroptimierung wird dabei die Filterung durch den Beobachter explizit berücksichtigt. Durch dieses Vorgehen lassen sich selbst Systeme, die ein starkes Messrauschen aufweisen, durch die gelernte Regelung stabilisieren. Die weiteren genannten Arbeiten untersuchen PILCO im Zusammenhang mit anderen probabilistischen maschinellen Lernverfahren und tauschen den GP-Ansatz entsprechend aus. Im Grundsatz unterscheiden sich die entwickelten Verfahren daher nicht zum ursprünglichen PILCO-Verfahren, weisen jedoch im Bezug auf die Skalierbarkeit und die Verallgemeinerung auf komplexere Zustandsverteilungen bessere Eigenschaften auf.

Neben den bereits erwähnten Kritikpunkten, ist ein weiterer Punkt in Bezug zum Regelungsansatz (4-12) anzumerken. Es handelt sich hierbei um eine komplexe Ansatzfunktion, die sich durch eine gewisse Anzahl von Parametern kennzeichnet. Die Anzahl der Parameter und damit auch die Komplexität der Regelung wird vom Anwender a-priori festgelegt. Während des Verfahrens ist nicht vorgesehen, dass sich der Regler diesbezüglich verändert. Das Problem besteht darin, dass die Anzahl passend zur Regelungsaufgabe bzw. dem System ausgewählt werden muss. Ist die Komplexität des Reglers zu gering, so kann die Aufgabe nicht gelöst werden und es stellt sich eine Stagnation im Laufe der Lerniterationen ein. Auf der anderen Seite ist eine zu hohe Komplexität ebenfalls kontraproduktiv, da die Lösung der Regelungsaufgabe damit grundsätzlich ermöglicht wird, andererseits aber das zugrundeliegende Optimierungsproblem immer schwieriger zu bewältigen ist und suboptimale Parametrisierungen wahrscheinlicher werden. Hinzu kommt, dass die Dauer des Lernvorgangs insgesamt unter Umständen wesentlich zunimmt. Des Weiteren kann die GP-Regelung vor dem Hintergrund der Zwei-Freiheitsgrade-Regelungsstruktur (s. Bild 2-1) betrachtet werden. Während die klassische Regelungsstruktur eine eindeutige Unterscheidung zwischen dem Steuerungs- und Regelungsanteil des Stellsignals liefert, ist die GP-Regelung (4-12) dahingehend nicht analysierbar. In Bezug auf das Doppelpendel-Anwendungsbeispiel ist die gelernte Regelung lediglich für die Aufschwung-Trajektorie zulässig und enthält damit den Charakter einer Steuerung. Dennoch gleicht sie kleinere Störungen, die entlang der Aufschwung-Trajektorie auftreten können, aus, sodass sie von der Wirkungsweise ähnlich zum Ansatz der Zwei-Freiheitsgrade-Regelungsstruktur stabilisierend funktioniert. Die Aufteilung des Stellsignals und eine damit verbundene Interpretierbarkeit der gelernten Regelung ist jedoch nicht möglich.

Daher passen [KD18] und [HLZ18; HKZ17] PILCO in die Richtung eines Optimalsteuerungsproblems (4-1) an und behandeln die Stellgrößen eines jeden Zeitschritts direkt als Optimierungsvariablen. Die Autoren aus [KD18] konzentrieren sich dabei auf ein autonomes Lernszenario (keine vorherige Annahme über Wissen) und können Optimalitätsbedingungen ableiten, indem sie das Maximum-Prinzip von Pontryagin [PLB12] unter der einschränkenden Annahme, dass der GP einen SE-Kernel hat, anwenden. Sie erzielen gute Ergebnisse in zustands- und stellgrößenbeschränkten simulationsgestützten Testexperimenten im Vergleich zu PILCO, da diese Einschränkungen innerhalb des Optimalsteuerungsproblems natürlich durch fest einzuhaltende Nebenbedingungen behandelt werden können und nicht über einen Strafterm innerhalb der Gütfunktion berücksichtigt werden müssen (vgl. (4-14)). [HLZ18; HKZ17] sind ähnlich zu [KD18] im Sinne der Verwendung von GPs, diskutieren jedoch zusätzlich verschiedene Approximationen für die Zustandspropagation unter Unsicherheit, mit Ausnahme der UT (s. Abschnitt 2.3.2). Da beide Veröffentlichungen die Steuerungen als einzige Optimierungsvariablen definieren, wird ihr Ansatz als Single-Shooting klassifiziert. Als Regelungskonzept wird in beiden Fällen ein modellprädiktiver Ansatz (MPC) vorgeschlagen, der jedoch in der Praxis schwer umgesetzt werden kann aufgrund der hohen Echtzeitanforderungen, die sich durch die Lösung des komplexen Optimierungsproblems in jedem Zeitschritt ergeben. Diese Anmerkung gilt insbesondere für das Anwendungsbeispiel dieses Kapitels aufgrund seiner hohen Dynamik und notwendigerweise niedrigen Abtastzeit.

Beitrag und Prinzipien der hybriden Optimalsteuerung

Der Beitrag der vorliegenden Arbeit vor dem Hintergrund des Stands der Forschung ist vielseitig. Zunächst kann anhand der aufgelisteten Veröffentlichungen und den eigenen Ergebnissen im Zusammenhang mit PILCO geschlussfolgert werden, dass sich das probabilistische Rahmenwerk und der iterative Lernvorgang erfolgversprechend in die Regelungstechnik und die Inbetriebnahme integrieren lassen. Durch das direkte Lernen auf der Basis von Messdaten vom realen System wird die Problematik der unzureichenden Modellbildung adressiert. Dabei bleibt eine gewisse Flexibilität im Bezug auf das betrachtete System erhalten und geht mit einer zeitlichen Einsparung hinsichtlich einer aufwendigeren Modellierungsarbeit einher. Als Grundprinzip wird das Vorgehen von PILCO und die Verwendung von probabilistischen GPs daher zur Durchführung einer besseren und effizienteren Inbetriebnahme einer Optimalsteuerung übernommen. Die wahrscheinlichkeitsbasierte Komponente ist dabei vorteilhaft, um gewisse Unsicherheiten über die Dynamikfunktion einzubeziehen und so Vorkommnisse am realen Prozess, bspw. die Verletzung von technischen Rahmenbedingungen oder ein instabiles Systemverhalten, besser vorhersagen zu können. Im Vergleich zu den vorhergehenden Arbeiten wird dabei die Einbindung von Vorwissen explizit bei der Entwicklung des damit hybriden Optimalsteuerungsverfahrens berücksichtigt. Das Ziel ist es, beliebige nichtlineare Dynamikfunk-

tionen als Annahme zuzulassen und diese durch die aufgestellten GPs und den PILCO-ähnlichen Lernvorgang korrigieren zu lassen. Die UT stellt hierbei ein wesentliches Hilfsmittel dar, um die Zustandspropagation des hybriden Modells abzubilden. Ein weiterer Vorteil der Nutzung der UT ist, dass sie nicht nur den Einsatz einer nichtlinearen Dynamikfunktion ermöglicht, sondern auch eine freie Wahl der Kovarianzfunktion des GPs beinhaltet. Vergleichbare Arbeiten sind nur auf den SE-Kernel beschränkt. Ein weiterer entscheidender Punkt ist in Bezug auf das Lösungsverfahren des Optimalsteuerungsproblems zu nennen. Anstelle des Single- wird in dieser Arbeit das Multiple-Shooting eingesetzt, da es für komplexere Anwendungsfälle, wie dem Mehrfachpendel auf einem Wagen, robustere bzw. zuverlässigere Ergebnisse liefert. Neben dem reinen Optimalsteuerungsentwurf wird in diesem Kapitel auch die Entwicklung einer zugehörigen Regelung besprochen. Hierzu wurde bei PILCO das Problem der fehlenden Interpretierbarkeit angeführt. Anstelle eines modellprädiktiven Ansatzes, welcher in der Praxis nur schwerlich umgesetzt werden kann, insbesondere für ein hochdynamisches Doppelpendel, wird auf die Zwei-Freiheitsgrade-Regelungsstruktur zurückgegriffen und ein zeitvarianter Riccati-Regler auf der Basis des hybriden Modells aufgestellt. Eine Validierung des entwickelten Verfahrens findet an einem simulierten vollaktuierten Doppelpendel und einem realen Doppelpendel auf einem Wagen statt.

4.4 Entwicklung der hybriden Optimalsteuerung und Erprobung an einem simulierten voll-aktuierten Doppelpendel

Den Ausgangspunkt bilden Messdaten in Form eines Tripels mit $(\mathbf{x}_k, \mathbf{u}_k, \mathbf{x}_{k+1})$ nach jedem Versuch bzw. Experiment am realen System, die in den nachfolgenden Matrizen gesammelt werden:

$$\mathbf{X}_k, \mathbf{X}_{k+1} \in \mathbb{R}^{n_x \times n_d}, \mathbf{U}_k \in \mathbb{R}^{n_u \times n_d}, \quad (4-15)$$

wobei n_d die Anzahl der vorhandenen Datenpunkte ist. Die Anzahl der Datenpunkte wächst mit jeder Lerniteration und somit auch diese Matrizen. Darüber hinaus können sie durch Messrauschen verfälscht sein. Für jeden Datenpunkt wird der Fehler $\Delta \mathbf{x}_{k+1}$ im Hinblick auf das physikalische Modell \mathbf{f}_{phy} über

$$\Delta \mathbf{X}_{k+1}^{(i)} = \mathbf{X}_{k+1}^{(i)} - \mathbf{f}_{phy}(\mathbf{X}_k^{(i)}, \mathbf{U}_k^{(i)}), \quad i = 1, \dots, n_d, \quad (4-16)$$

berechnet. Diese Modellfehler werden im Rahmen des Verfahrens als Zufallsvariablen behandelt und für jede Zustandsdimension mit einem separaten GP modelliert:

$$\Delta x_{k+1}^{(j)} \sim \mathcal{GP}_j(0, k(\mathbf{x}_k^u, \mathbf{x}_k^{u'}; \boldsymbol{\eta}_j)), \quad j = 1, \dots, n_x,$$

mit Nullmittelwert und Kovarianzfunktion $k(\cdot, \cdot; \boldsymbol{\eta})$, Hyperparametern $\boldsymbol{\eta}_j$ und abkürzender Schreibweise $\mathbf{x}_k^u := [\mathbf{x}_k^T, \mathbf{u}_k^T]^T$. Der Mittelwert ist null, da die Modellfehler unbekannt sind und bereits alles Vorwissen über die Dynamik vom physikalischen Modell erfasst wird. Die Kovarianzfunktion kodiert allgemeine Annahmen, wie den Grad der Glättigkeit oder Periodizität der zugehörigen Modellfehlerfunktion. Optimale Hyperparameter $\boldsymbol{\eta}_j^*$ werden durch den Ansatz der maximalen marginalen logarithmischen Likelihood bestimmt.

Unter Verwendung der Daten in (4-15), (4-16) und der optimalen Hyperparameter wird die (prädiktive) Posterior-Verteilung über den Satz von Bayes berechnet, welche per Definition erneut eine Normalverteilung ist:

$$\begin{aligned} p(\Delta x_{k+1}^{(j)} | \Delta \mathbf{X}_{k+1}^{(j,:)}) &= \mathcal{N}(\Delta \mu_j(\mathbf{x}_k^u), \Delta \sigma_j^2(\mathbf{x}_k^u)), \\ \Delta \mu_j(\mathbf{x}_k^u) &= \mathbf{k}_{X_k^U}(\mathbf{x}_k^u)^T \mathbf{K}_{\eta_j^*}^{-1} \Delta \mathbf{X}_{k+1}^{(j,:)^T}, \\ \Delta \sigma_j^2(\mathbf{x}_k^u) &= k(\mathbf{x}_k^u, \mathbf{x}_k^u; \boldsymbol{\eta}_j^*) - \mathbf{k}_{X_k^U}(\mathbf{x}_k^u)^T \mathbf{K}_{\eta_j^*}^{-1} \mathbf{k}_{X_k^U}(\mathbf{x}_k^u), \end{aligned} \quad (4-17)$$

wobei $\Delta \mathbf{X}_{k+1}^{(j,:)}$ die j -te Zeile von $\Delta \mathbf{X}_{k+1}$ ist, sowie die vektorielle Definition gilt $\mathbf{k}_{X_k^U}(\mathbf{x}_k^u) = [k(\mathbf{x}_k^u, \mathbf{X}_k^{U(1)}; \boldsymbol{\eta}_j^*), \dots, k(\mathbf{x}_k^u, \mathbf{X}_k^{U(n_d)}; \boldsymbol{\eta}_j^*)]^T$. $\mathbf{K}_{\eta_j^*} \in \mathbb{R}^{n_d \times n_d}$ ist die symmetrische und positiv definite Gram-Matrix mit den Einträgen

$$K_{\eta_j^*}^{(r,c)} = k(\mathbf{X}_k^{U(r)}, \mathbf{X}_k^{U(c)}; \boldsymbol{\eta}_j^*), \quad r, c = 1, \dots, n_d,$$

mit $\mathbf{X}_k^U := [\mathbf{X}_k^T, \mathbf{U}_k^T]^T$, wobei der zusätzliche Subindex die verwendete Spalte angibt. Unter Verwendung der Standardannahme [DFR15], dass die dimensionszugehörigen GPs unabhängig von einander sind, lautet das gesamte datengetriebene Modell damit formal

$$\begin{aligned} p(\Delta \mathbf{x}_{k+1} | \Delta \mathbf{X}_{k+1}) &= \mathcal{N}(\Delta \boldsymbol{\mu}(\mathbf{x}_k^u), \Delta \boldsymbol{\Sigma}(\mathbf{x}_k^u)), \\ \Delta \boldsymbol{\mu}(\mathbf{x}_k^u) &= [\Delta \mu_1(\mathbf{x}_k^u), \dots, \Delta \mu_{n_x}(\mathbf{x}_k^u)]^T, \\ \Delta \boldsymbol{\Sigma}(\mathbf{x}_k^u) &= \text{diag}([\Delta \sigma_1^2(\mathbf{x}_k^u), \dots, \Delta \sigma_{n_x}^2(\mathbf{x}_k^u)]). \end{aligned} \quad (4-18)$$

Entsprechend zu [KD18] wird für das gesamte datengetriebene Modell eine Kurzschreibweise durch

$$\mathbf{f}_{gp}(\mathbf{x}_k, \mathbf{u}_k) := [\Delta \boldsymbol{\mu}, \Delta \boldsymbol{\Sigma}], \quad (4-19)$$

mit der Abbildung $\mathbf{f}_{gp} : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \mapsto \mathbb{R}^{n_x} \times \mathbb{R}^{n_x \times n_x}$ eingeführt.

Bis zu diesem Punkt wurden deterministische Größen als Eingang für den GP angenommen. Für die Berücksichtigung der Zustandspropagation wird nun ein normalverteilter

Zustand $\mathbf{x}_k \sim \mathcal{N}(\mathbf{m}_k, \mathbf{S}_k)$ angenommen. Gemäß der Definition in (4-16) lautet der Ansatz des Hybridmodells

$$\mathbf{x}_{k+1} = \mathbf{f}_{phy}(\mathbf{x}_k, \mathbf{u}_k) + \Delta\mathbf{x}_{k+1}(\mathbf{x}_k, \mathbf{u}_k). \quad (4-20)$$

Die Momente der Verteilung des Nachfolgezustands werden berechnet, indem das Gesetz der totalen Erwartung

$$\mathbf{m}_{k+1} = \mathbb{E}_{x_k}[\mathbb{E}_{\Delta x_{k+1}}[\mathbf{x}_{k+1}]] = \mathbb{E}_{x_k}[\mathbf{f}_{phy}] + \mathbb{E}_{x_k}[\Delta\boldsymbol{\mu}], \quad (4-21)$$

und das Gesetz der totalen Varianz

$$\begin{aligned} \mathbf{S}_{k+1} &= \mathbb{E}_{x_k}[\mathbb{V}_{\Delta x_{k+1}}[\mathbf{x}_{k+1}]] + \mathbb{V}_{x_k}[\mathbb{E}_{\Delta x_{k+1}}[\mathbf{x}_{k+1}]] \\ &= \mathbb{E}_{x_k}[\Delta\boldsymbol{\Sigma}] + \mathbb{V}_{x_k}[\mathbf{f}_{phy}] + \mathbb{V}_{x_k}[\Delta\boldsymbol{\mu}] + \mathbb{C}_{x_k}[\mathbf{f}_{phy}, \Delta\boldsymbol{\mu}] + \mathbb{C}_{x_k}[\mathbf{f}_{phy}, \Delta\boldsymbol{\mu}]^T, \end{aligned} \quad (4-22)$$

verwendet werden. Sofern keine analytische Berechnung für die einzelnen Terme möglich ist, wird im Rahmen dieser Arbeit der Einsatz der UT zur Approximation verwendet. Die Erwartung und Varianz bezüglich f_{phy} werden über die UT-Zusammenhänge aus (2-26), (2-27) und (2-28) berechnet, woraus $\mathbf{m}_{f_{phy}} \approx \mathbb{E}_{x_k}[\mathbf{f}_{phy}]$ und $\mathbf{S}_{f_{phy}} \approx \mathbb{V}_{x_k}[\mathbf{f}_{phy}]$ folgen. Eine ähnliche Approximation wird in [VG15] betrachtet, wobei die Herleitung jedoch aus den kontinuierlichen DGLn des physikalischen Modells aufgestellt und kein zusätzliches datengetriebenes Modell berücksichtigt wird. Für die verbleibenden, vom GP abhängigen Größen, wird erneut die UT angewendet:

$$\begin{aligned} [\Delta\boldsymbol{\mu}_s^{(i)}, \Delta\boldsymbol{\Sigma}_s^{(i)}] &= \mathbf{f}_{gp}(\mathbf{x}_{s,k}^{(i)}, \mathbf{u}_k), \quad i = 1, \dots, 2n_x + 1, \\ \mathbb{E}_{x_k}[\Delta\boldsymbol{\mu}] &\approx \sum_{i=1}^{2n_x+1} W^{(i)} \Delta\boldsymbol{\mu}_s^{(i)} =: \mathbf{m}_{\Delta\mu}, \\ \mathbb{E}_{x_k}[\Delta\boldsymbol{\Sigma}] &\approx \sum_{i=1}^{2n_x+1} W^{(i)} \Delta\boldsymbol{\Sigma}_s^{(i)}, \\ \mathbb{V}_{x_k}[\Delta\boldsymbol{\mu}] &\approx \sum_{i=1}^{2n_x+1} W^{(i)} (\Delta\boldsymbol{\mu}_s^{(i)} - \mathbf{m}_{\Delta\mu}) (\Delta\boldsymbol{\mu}_s^{(i)} - \mathbf{m}_{\Delta\mu})^T, \\ \mathbb{C}_{x_k}[\mathbf{f}_{phy}, \Delta\boldsymbol{\mu}] &\approx \sum_{i=1}^{2n_x+1} W^{(i)} (\mathbf{x}_{s,k+1}^{(i)} - \mathbf{m}_{f_{phy}}) (\Delta\boldsymbol{\mu}_s^{(i)} - \mathbf{m}_{\Delta\mu})^T, \end{aligned} \quad (4-23)$$

wobei dieselben Sigma-Punkte und Gewichtungsfaktoren wie in (2-26) und (2-28) verwendet werden. Ein ähnliches Berechnungsverfahren wurde in [KKFH07] verwendet, um einen rein datengetriebenen GP-UKF-Beobachter zu entwerfen, wobei jedoch dieselbe Varianz $\Delta\boldsymbol{\Sigma}_s^{(i)}$ $\forall i$ angenommen wurde und daher die Varianz der Approximation unterschätzt wird. Zu beachten ist dabei, dass trotz der diagonalen Struktur der GP-

Kovarianzmatrix (4-18) bzw. der Unabhängigkeitsannahme, die Dimensionen des Nachfolgezustands für eine unsichere Eingabeverteilung kovariieren können. Die gleiche Beobachtung wird auch für den analytischen MM-Ansatz in [DFR15] gemacht, sodass der hier gewählte numerische Ansatz diesbezüglich konsistent ist. Die Gleichungen aus (4-21), (4-22) und (4-23) werden zusammen als eine mathematische Funktion definiert. Diese Funktion wird mit

$$[\mathbf{m}_{k+1}, \mathbf{S}_{k+1}] = \mathbf{f}_{hyb}(\mathbf{m}_k, \mathbf{S}_k, \mathbf{u}_k), \quad (4-24)$$

beschrieben, wobei die aktuelle Zustandsverteilung zusammen mit der Stellgröße auf die nächste Zustandsverteilung abgebildet werden. Im Gegensatz zum rein physikalischen Modell gleicht das hybride Modell Fehler über den GP-Teil aus, abhängig von der Menge der aufgezeichneten Daten. Darüber hinaus berücksichtigt die probabilistische Formulierung die Unsicherheit über Fehler bei Langzeitvorhersagen und verhindert so die sogenannte Modellverzerrung, die für ein schlechteres Explorationsverhalten verantwortlich ist.

Eine typische Zielgröße, die in der Regelungstechnik und im Zusammenhang mit einer Optimalsteuerung oft verwendet wird, ist wie folgt definiert

$$J(\boldsymbol{\theta}) = \underbrace{\Delta t \sum_{k=0}^H w_t(k)(\mathbf{x}_k - \mathbf{x}_G)^T \mathbf{W}_x (\mathbf{x}_k - \mathbf{x}_G)}_{=: J_x(\mathbf{x}_k)} + \underbrace{\Delta t \sum_{k=0}^H \mathbf{u}_k^T \mathbf{W}_u \mathbf{u}_k}_{=: J_u(\mathbf{u}_k)}, \quad (4-25)$$

die einen Term für die Distanz zum Zielzustand $(\mathbf{x}_k - \mathbf{x}_G)$ und einen Term für die verbrauchte Energie enthält. $\mathbf{W}_x \succ 0$ und $\mathbf{W}_u \succeq 0$ sind Gewichtungsmatrizen. $w_t(k) \in [0, 1]$ steigt monoton an und gibt somit ein höheres Gewicht für Abweichungen, die in einem späteren Verlauf der Trajektorie auftreten. Da der Zustand normalverteilt ist, wird der Erwartungswert in dieser Arbeit als eigentliche Zielgröße verwendet. Für die Approximation des Erwartungswerts kann erneut die UT genutzt werden. Dies würde wiederum die Einbindung beliebiger Zielgrößen $J(\boldsymbol{\theta})$ ermöglichen. Zur Vereinfachung wird in diesem Kontext die Verwendung von (4-25) fortgesetzt und die analytische Lösung aus [DFO20]

$$\mathbb{E}_{x_k}[J(\boldsymbol{\theta})] = \Delta t \sum_{k=0}^H w_t(k)(\text{Spur}(\mathbf{S}_k \mathbf{W}_x) + (\mathbf{m}_k - \mathbf{x}_G)^T \mathbf{W}_x (\mathbf{m}_k - \mathbf{x}_G)) + \mathbf{u}_k^T \mathbf{W}_u \mathbf{u}_k, \quad (4-26)$$

genutzt. Im Folgenden werden die Einschränkungen diskutiert. Steuerungsbeschränkungen aufgrund von technischen Rahmenbedingungen können auf herkömmliche Weise durch

$\mathbf{u}_{min} \leq \mathbf{u}_k \leq \mathbf{u}_{max}$ berücksichtigt werden. Zustandsbeschränkungen können durch den Erwartungswert eingebunden werden, z. B. über $\mathbf{x}_{min} \leq \mathbf{m}_k \leq \mathbf{x}_{max}$, was jedoch aufgrund der Vernachlässigung der Varianz nicht besonders zuverlässig wäre. Daher wird ein wahrscheinlichkeitsbasierter Ansatz weiterverfolgt, der in Hinblick auf die Inbetriebnahme geeigneter ist. Die Wahrscheinlichkeit, die Zustandsbeschränkungen unter der konservativen Annahme, dass die Dimensionen unabhängig sind, zu erfüllen, wird durch

$$\begin{aligned} P(\mathbf{x}_{min} \leq \mathbf{x}_k \leq \mathbf{x}_{max}) &\approx \prod_{i=1}^{n_x} P(x_{min}^{(i)} \leq x_k^{(i)} \leq x_{max}^{(i)}) \\ &= \prod_{i=1}^{n_x} \Phi\left(\frac{x_{max}^{(i)} - m_k^{(i)}}{\sqrt{S_k^{(i,i)}}}\right) - \Phi\left(\frac{x_{min}^{(i)} - m_k^{(i)}}{\sqrt{S_k^{(i,i)}}}\right) \geq P_x, \end{aligned} \quad (4-27)$$

ausgedrückt, wobei $\Phi(\cdot)$ die Verteilungsfunktion der Standardnormalverteilung ist und $P_x \in (0, 1]$ ein vordefinierter Wahrscheinlichkeitsgrenzwert ist. Die Zustandsbeschränkungen sind in der Regel mit sicherheitskritischen Aspekten verbunden. Aus diesem Grund sollte der Wert von P_x relativ groß gewählt werden, damit die Zustandsbeschränkungen konservativ eingehalten werden. Da sie jedoch nur auf dem Modell und seiner Vorhersagegenauigkeit basieren, müssen sie nicht mit der festgelegten Wahrscheinlichkeit für das reale System erfüllt sein. Aus diesem Grund sollte für praktische Anwendungen immer ein zusätzlicher Puffer an den Zustandsgrenzen und Sicherheitsmechanismen vorgesehen werden.

An dieser Stelle wird das vollständige hybride Optimalsteuerungsproblem, mit welchem die Inbetriebnahme durchgeführt wird, definiert

$$\begin{aligned} \min_{\theta} \mathbb{E}_{x_k}[J(\theta)] \quad \text{u.B.v.} \quad \mathbf{h}(\theta) = 0, \quad \mathbf{g}(\theta) \leq 0, \\ \mathbf{h}(\theta) = \begin{cases} [\mathbf{m}_{k+1}, \mathbf{S}_{k+1}] = \mathbf{f}_{hyb}(\mathbf{m}_k, \mathbf{S}_k, \mathbf{u}_k), & k = 0, \dots, H-1, \\ \mathbf{m}_0 - \mathbf{m}_I, \quad \mathbf{S}_0 - \mathbf{S}_I, & \end{cases} \\ \mathbf{g}(\theta) = \begin{cases} P_x - P(\mathbf{x}_{min} \leq \mathbf{x}_k \leq \mathbf{x}_{max}), \\ \mathbf{u}_k - \mathbf{u}_{max}, \quad \mathbf{u}_{min} - \mathbf{u}_k, \quad k = 0, \dots, H, \end{cases} \end{aligned} \quad (4-28)$$

mit einer initialen Zustandsverteilung $\mathbf{x}_I \sim \mathcal{N}(\mathbf{m}_I, \mathbf{S}_I)$. Ähnlich dem Ansatz des Multiple Shooting für deterministische dynamische Modelle (4-2) erzwingen die Gleichheitsbeschränkungen $\mathbf{h}(\theta)$, dass die probabilistischen Dynamikgleichungen des hybriden Modells (4-24) für alle diskreten Zeitschritte erfüllt sind. Dies erfordert, dass die propagierten Momente $\mathbf{f}_{hyb}(\mathbf{m}_k, \mathbf{S}_k)$ mit den nächsten Momenten $(\mathbf{m}_{k+1}, \mathbf{S}_{k+1})$ übereinstimmen,

bei gegebener Steuerung \boldsymbol{u}_k . Der Optimierungsvektor setzt sich damit aus den folgenden Größen zusammen:

$$\boldsymbol{\theta} := [\boldsymbol{m}_0^T, \boldsymbol{L}_{\Delta,0}^T, \boldsymbol{u}_0^T, \dots, \boldsymbol{m}_H^T, \boldsymbol{L}_{\Delta,H}^T, \boldsymbol{u}_H^T]^T \in \mathbb{R}^{\left(\frac{1}{2}n_x^2 + \frac{3}{2}n_x + n_u\right)(H+1)}, \quad (4-29)$$

wobei die Cholesky-Zerlegung mit

$$\boldsymbol{S}_k = \boldsymbol{L}_k \boldsymbol{L}_k^T, \quad \boldsymbol{L}_{\Delta,k} = \text{nonzeros}(\boldsymbol{L}_k),$$

verwendet wird. Die Funktion $\text{nonzeros}(\cdot)$ bildet dabei die Elemente der unteren Dreiecksmatrix von \boldsymbol{L}_k auf einen Spaltenvektor $\boldsymbol{L}_{\Delta,k} \in \mathbb{R}^{\frac{1}{2}(n_x^2 + n_x)}$ ab. Die Formulierung über die Cholesky-Zerlegung erzwingt implizit eine weitere Nebenbedingung, bei der Symmetrie und positive Definitheit der Matrizen der Zustandsvarianzen erzwungen werden.

Im direkten Vergleich zur Single-Shooting-Formulierung, bei der die Dynamikgleichungen durch Vorwärtsintegration implizit exakt erfüllt sind und nur die Steuervariablen die Optimierungsvariablen repräsentieren ($\dim(\boldsymbol{\theta}) = n_u(H + 1)$), enthält der Multiple-Shooting-Ansatz wesentlich mehr Optimierungsvariablen und wirkt daher komplexer bzw. aufwendiger. Der Multiple-Shooting-Ansatz bietet jedoch mehrere Vorteile, welche die Anzahl der Optimierungsvariablen in Relation setzen. Im Allgemeinen wird die SQP-Methode [NW06; GK02] (s. A2 für eine Einführung des Verfahrens) verwendet, um das hybride Optimalsteuerungsproblem numerisch zu lösen. Dies erfordert insbesondere die Ableitungen der Funktionen in (4-28) nach (4-29), wodurch die Dimensionen über

$$\nabla \mathbb{E}_{x_k}[J(\boldsymbol{\theta})] \in \mathbb{R}^{n_\theta}, \quad \nabla \boldsymbol{h}(\boldsymbol{\theta}) \in \mathbb{R}^{\left(\frac{1}{2}n_x^2 + \frac{3}{2}n_x\right)(H+1) \times n_\theta}, \quad \nabla \boldsymbol{g}(\boldsymbol{\theta}) \in \mathbb{R}^{(2n_u + H + 1)(H + 1) \times n_\theta},$$

mit der Substitution $n_\theta := \dim(\boldsymbol{\theta})$, angegeben werden können. Aufgrund der Definition (4-29) sind diese Ableitungen einfach und schnell zu berechnen, weil die Jakobimatrizen eine dünnbesetzte Struktur aufweisen, bei der nur Elemente in der Nähe der Diagonalen ausgefüllt sind. Dies trifft nicht auf das Single-Shooting zu, was insbesondere für (4-26) und (4-27) offensichtlich ist. Aufgrund der mehrfachen aufeinanderfolgenden Ausführung der Zustandspropagation werden die Einträge der Ableitungen mit zunehmenden Zeitschritten immer komplexer und aufwendiger zu berechnen. Ein weiterer Vorteil des Multiple- gegenüber dem Single-Shooting lässt sich im Zusammenhang mit dem SQP-Verfahren erklären. Die zugrundeliegenden Lagrange-Multiplikatoren, die für die probabilistischen Zustandsbeschränkungen in (4-28) verwendet werden, gewähren dem Optimierungssolver eine gewisse Flexibilität, damit er von den genauen Trajektorien abweichen und bessere Lösungen für das hybride Optimalsteuerungsproblem finden kann. Dies könnte eine Erklärung für die, im Allgemeinen bekannten besseren Konvergenzeigenschaften des Multiple- im Vergleich zum Single-Shooting sein, welche auch bei den

Algorithmus 4 Hybride Optimalsteuerung

- 1: **Initialisierung:** Berechne Optimalsteuerung $\{\mathbf{u}_k^*\}_{k=0}^H$ nach (4-2) auf Basis des physikalischen Modells \mathbf{f}_{phy} und teste sie am realen System, wobei initiale Messdaten gesammelt werden.
 - 2: **Wiederhole** bis ein Konvergenzkriterium oder ein Iterationsbudget erreicht ist:
 - 3: Lerne Modellfehler über GPs (4-19) auf Basis der vorhandenen Daten (4-16).
 - 4: **Wiederhole** die nachfolgenden Schritte im Rahmen einer SQP-Optimierung (A2)
 - 5: Berechne den Gradienten $\nabla \mathbb{E}[J(\boldsymbol{\theta})]$ und die Jacobimatrizen der Nebenbedingungen $\nabla \mathbf{h}(\boldsymbol{\theta}), \nabla \mathbf{g}(\boldsymbol{\theta})$ ▷ Automatische Differenzierung (CasADi)
 - 6: Aktualisiere den Optimierungsvektor $\boldsymbol{\theta}$ und die zugehörigen Lagrange-Multiplikatoren $(\boldsymbol{\lambda}, \boldsymbol{\mu})$.
 - 7: Erprobe die erhaltene Optimalsteuerung $\{\mathbf{u}_k^*\}_{k=0}^H$ am realen System und nehme weitere Messdaten auf, die den bestehenden Daten hinzugefügt werden.
-

Experimenten dieser Arbeit zu beobachten war. Obwohl die Ableitungen $\nabla \mathbb{E}_{x_k}[J(\boldsymbol{\theta})]$, $\nabla \mathbf{h}(\boldsymbol{\theta})$ und $\nabla \mathbf{g}(\boldsymbol{\theta})$ mit entsprechendem Aufwand analytisch aufgestellt werden könnten, werden sie im Rahmen dieser Arbeit mit dem automatischen Differenzierungswerkzeug CasADi [AGH⁺19] berechnet. Außerdem wird als Programmierumgebung MATLAB verwendet und dessen integrierte Implementierung der SQP-Methode (Algorithmus 6) innerhalb des *fmincon*-Solvers [Mat24] genutzt.

Der Gesamタルgorithmus der hybriden Optimalsteuerung ist in Algorithmus 4 zusammengefasst. Aus einer übergeordneten Perspektive verwendet die Methodik einen konventionellen Trial-and-Error Ansatz. Der Prozess beginnt mit der Lösung des Optimalsteuerungsproblems, das ausschließlich auf dem etablierten physikalischen Modell basiert. Anschließend wird die aktuelle optimale Steuersequenz auf das reale System angewendet und dabei Zustandsdaten aufgezeichnet und gesammelt (Zeile 1). Nachfolgend wird eine Zwei-Schritt-Schleife ausgeführt (Zeilen 2-7 und 4-6). Die äußere Schleife beginnt mit dem Erlernen der Modellfehler im Zusammenhang mit dem physikalischen Modellteil unter Verwendung der Gauß-Prozess-Regression und aller verfügbaren Daten (Zeile 3). Danach wird die innere Schleife ausgelöst und das hybride Optimalsteuerungsproblem basierend auf dem hybriden Modell bis zur Konvergenz gelöst (Zeilen 4-6). Dies beinhaltet die wiederholte Berechnung der Ableitungen sowie die Aktualisierung der Optimierungsvariablen und vorhandener Lagrange-Multiplikatoren. Die neu berechnete optimale Steuersequenz wird dann auf das reale System angewendet, wobei mit jeder Iteration die Datenmenge zunimmt. Der sich abwechselnde Lern-, Verbesserungs- und Interaktionszyklus wird wiederholt, bis die Optimalsteuerungsaufgabe erfolgreich abgeschlossen ist.

Anwendung auf ein simuliertes voll-aktuiertes Doppelpendel

Im Folgenden wird die Anwendung von Algorithmus 4 zur Trajektorienplanung eines simulierten vollständig aktuierten Doppelpendels illustriert. Die Ergebnisse³ sind ebenfalls in der Vorveröffentlichung [HTT23] enthalten. Das System besteht aus $n_x = 4$ Zuständen, $\mathbf{x} = [\varphi_1, \varphi_2, \dot{\varphi}_1, \dot{\varphi}_2]^T$ und $n_u = 2$ Steuerungseingängen $\mathbf{u} = [u_1, u_2]^T$, welche die Drehmomente in den Gelenken repräsentieren. Das Szenario beinhaltet eine zeitliche Diskretisierung von $\Delta t = 0.05$ s mit $H = 40$ Zeitschritten, sodass die Gesamtzeit $T := \Delta t H = 2$ s beträgt. Die Aufgabe besteht darin, die optimale Steuerungsabfolge $\{\mathbf{u}_k^*\}_{k=0}^H$ zu finden, sodass der obere instabile Zielzustand $\mathbf{x}_G = \mathbf{x}(t = T) = \mathbf{0}$, ausgehend von dem unteren stabilen Gleichgewichtszustand $\mathbf{x}(t = 0) = \mathbf{m}_I = [\pi, \pi, 0, 0]^T$, am Ende der Trajektorie erreicht wird. Die Steuerungs- und Zustandsbeschränkungen lauten

$$\begin{aligned}\mathbf{u}_{min} &= [-10 \text{ N m}, -10 \text{ N m}]^T \leq \mathbf{u} \leq [10 \text{ N m}, 10 \text{ N m}]^T = \mathbf{u}_{max}, \\ \mathbf{x}_{min} &= [-1 \text{ rad}, -1 \text{ rad}, -4 \text{ rad/s}, -4 \text{ rad/s}]^T \leq \mathbf{x} \leq \mathbf{x}_{max}, \\ \mathbf{x}_{max} &= [4 \text{ rad}, 4 \text{ rad}, 2 \text{ rad/s}, 2 \text{ rad/s}]^T\end{aligned}\quad (4-30)$$

und müssen während des Übergangs erfüllt sein. Die anfängliche Zustandsvarianz $\mathbf{S}_I = 10^{-3} \mathbf{I}_4$ wird als gering angenommen. Die Gewichtungen $\mathbf{W}_x = \text{diag}([25, 25, 1, 1])$, $\mathbf{W}_u = 10^{-2} \mathbf{I}_2$, $w_t(k) = \frac{e^{3\Delta t k} - 1}{e^{3\Delta t H} - 1}$ und der Wahrscheinlichkeitsgrenzwert $P_x = 95\%$ definieren das hybride Optimalsteuerungsproblem (4-28) vollständig.

Die Bewegungsgleichungen für das mechanische System werden in allgemeiner Form über

$$\mathbf{M}(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{C}(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}} + \mathbf{G}(\mathbf{q}) = \mathbf{F}(\dot{\mathbf{q}}, \mathbf{u}), \quad (4-31)$$

angegeben, mit dem Positionsvektor $\mathbf{q} = [\varphi_1, \varphi_2]^T \in \mathbb{R}^{n_q}$, $n_q = 2$, der Massenmatrix $\mathbf{M} \in \mathbb{R}^{n_q \times n_q}$, dem Kraftvektor $\mathbf{F} \in \mathbb{R}^{n_q}$ (der eine lineare Dämpfung für die Reibung und die steuerbaren Drehmomente enthält), der Zentrifugalmatrix $\mathbf{C} \in \mathbb{R}^{n_q \times n_q}$ und dem Gravitationsvektor $\mathbf{G} \in \mathbb{R}^{n_q}$. Die Gleichungen können unter Verwendung des Lagrange-Formalismus ähnlich zu (4-5) aufgestellt werden. Eine Prinzipskizze des Systems ist in [HTT23] enthalten. Die mechanischen Parameter sind identisch zu denen in [TKOT11], jedoch mit einem um den Faktor 10 höheren Reibungskonstanten. Eine Simulation von (4-31) wird als Repräsentation für das reale System verwendet. Im Folgenden werden zwei Anwendungsfälle unterschieden. Im ersten Fall erfolgt das Lernen mit Vorwissen, bei dem nur bestimmte Terme von (4-31) a-priori bekannt sind (partielle Systemkenntnis)

³Der zugehörige Code ist unter <https://github.com/mh510/Hybrid-Optimal-Control> zur Verfügung gestellt.

und im zweiten Fall erfolgt das Lernen ohne dieses spezielle Expertenwissen, bei dem die gesamte Dynamik datengetrieben erlernt werden muss. Das Ziel ist es, die Auswirkung des Expertenwissens auf den Lernfortschritt und das Endresultat auszuwerten, sowie Erkenntnisse für die Inbetriebnahme abzuleiten.

Als Nächstes wird das Vorwissen des ersten Falls genauer erläutert. Es wird angenommen, dass das physikalische Dynamikmodell mit

$$\mathbf{f}_{phy}(\mathbf{x}_k, \mathbf{u}_k) = \mathbf{x}_k + \Delta t \begin{bmatrix} \dot{\mathbf{q}}_k \\ \mathbf{M}^{-1}(\mathbf{q}_k)(\mathbf{F}(\dot{\mathbf{q}}_k, \mathbf{u}_k) - \mathbf{G}(\mathbf{q}_k)) \end{bmatrix}, \quad (4-32)$$

und $\mathbf{x}_k = [\mathbf{q}_k^T, \dot{\mathbf{q}}_k^T]^T$, aufgestellt wurde. Im Vergleich zu (4-31) wird deutlich, dass die Zentrifugalkräfte fehlen, die im Rahmen dieser Erprobung als Modellfehler definiert werden. Der fehlende Term ist nicht trivial, da er die Inversion der Massenmatrix (ähnlich zu (4-6)) beinhaltet. Aufgrund des Modellfehlers kann der gewünschte Zielzustand nicht allein durch das Lösen eines Optimalsteuerungsproblems auf Grundlage von (4-32) erreicht werden bzw. die Anwendung von Zeile 1 in Algorithmus 4 ist nicht ausreichend zur Erfüllung der Steuerungsaufgabe. Aus diesem Grund werden die weiteren Schritte der hybriden Optimalsteuerung ausgeführt und zunächst der GP-Modellteil gelernt, wobei weiteres Systemwissen ausgenutzt werden kann. Die ersten n_q Gleichungen in (4-32) spiegeln das Integratorverhalten des Systems mit $\mathbf{q}_{k+1} = \mathbf{q}_k + \Delta t \dot{\mathbf{q}}_k$ wider und erfordern keine Korrektur durch den GP-Modellteil. Integratorverhalten ist vielen technischen (insbesondere mechanischen) Systemen inhärent und kann daher als bekannte Struktur ausgenutzt werden. Die zugehörige Reduzierung des Lernproblems wird durch die Einführung einer Kopplungsmatrix B_{gp} umgesetzt

$$\mathbf{x}_{k+1} = \mathbf{f}_{phy}(\mathbf{x}_k, \mathbf{u}_k) + \mathbf{B}_{gp} \Delta \mathbf{x}_{k+1}(\mathbf{x}_k), \quad \text{mit} \quad \mathbf{B}_{gp} = \left[\mathbf{0}_{n_q}, \mathbf{I}_{n_q} \right]^T,$$

wobei zudem berücksichtigt wird, dass die Zentrifugalkräfte nicht vom Steuereingang abhängen. Die GP-Abbildung (4-19) wird somit von $\mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \mapsto \mathbb{R}^{2n_q} \times \mathbb{R}^{2n_q \times 2n_q}$ auf $\mathbb{R}^{n_x} \mapsto \mathbb{R}^{n_q} \times \mathbb{R}^{n_q \times n_q}$ reduziert, wodurch der Berechnungsaufwand gesenkt und die Effektivität des Verfahrens erhöht wird. In Hinblick auf (4-21) und (4-22) müssen der Mittelwert und die Varianz des nächsten Zeitschritts gemäß der linearen Transformation von Normalverteilungen folgendermaßen angepasst werden

$$\begin{aligned} \mathbf{m}_{k+1} &= \mathbb{E}[\mathbf{f}_{phy}] + \mathbf{B}_{gp} \mathbb{E}[\Delta \boldsymbol{\mu}], \\ \mathbf{S}_{k+1} &= \mathbf{B}_{gp} \mathbb{E}[\Delta \Sigma] \mathbf{B}_{gp}^T + \mathbb{V}[\mathbf{f}_{phy}] + \mathbf{B}_{gp} \mathbb{V}[\Delta \boldsymbol{\mu}] \mathbf{B}_{gp}^T + \\ &\quad \mathbb{C}[\mathbf{f}_{phy}, \Delta \boldsymbol{\mu}] \mathbf{B}_{gp}^T + \mathbf{B}_{gp} \mathbb{C}[\mathbf{f}_{phy}, \Delta \boldsymbol{\mu}]^T, \end{aligned}$$

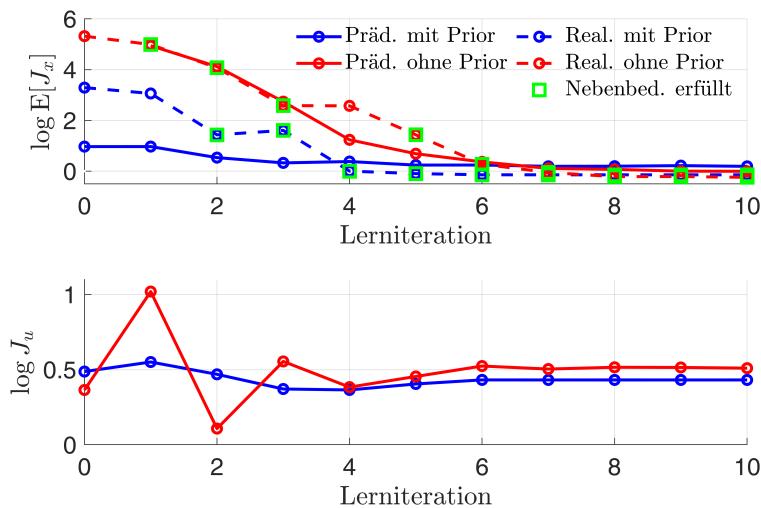


Bild 4-4: Fortschritt in der Zustands- und Energiezielgröße (logarithmische Skala) für den vorgestellten hybriden Optimalsteuerungsentwurf mit (blau) und ohne (rot) Einbeziehung von physikalischen Vorwissen (Prior).

um die Dimension des physikalischen Modellteils zu erhalten. Für einen zulässigen Vergleich wurde bei den nachfolgenden Untersuchungen zumindest das angesprochene Integratorverhalten, für den Fall mit fehlendem physikalischen Vorwissen spezifiziert.

Als Nächstes werden die Ergebnisse vorgestellt. Bild 4-4 zeigt hierzu zunächst den Fortschritt bzgl. der Teilziele J_x und J_u (s. (4-25)) für 10 Lerniterationen. Im oberen Diagramm wird die Vorhersage (Prädiktion) des hybriden Modells vor dem Experiment mit der tatsächlich gemessenen Trajektorie vom realen System (4-31) (Realität) verglichen. Die Prädiktion der Gütfunktion ist für den Fall mit Vorwissen (Prior) in den ersten Iterationen zu optimistisch, aufgrund der vergleichsweise niedrigen Werte. Als Begründung kann angeführt werden, dass die Datenlage in den ersten Iterationen noch zu gering ist, um die Unsicherheit über den Modellfehler korrekt beschreiben zu können und die grundsätzlich zu optimistischen Schätzungen auf das bereits bekannte Dynamikmodell (4-32) zurückzuführen sind. Wie erwartet, führt jedoch die Berücksichtigung von Vorwissen dazu, dass in weniger Lerniterationen eine optimale Steuerung für das System gefunden wird. Trotz der anfangs vorhandenen Modellfehler ist das Konvergenzverhalten stabil und die Steuerungsaufgabe wird bereits in der 4. Lerniteration gelöst. Die Anwendung ohne Prior benötigt hierzu 3 weitere Experimente und der endgültige Energieverbrauch ist zudem höher, was auf eine Konvergenz zu einer geringfügig suboptimalen Steuerung hindeutet.

Dies zeigt sich auch in Bild 4-5, in dem die tatsächlichen Trajektorien der 4. Iteration dargestellt sind. Die vordefinierten Zustandsbeschränkungen aus (4-30) sind in der linken Spalte ausgeschöpft, werden jedoch vollständig eingehalten, während sie auf der rechten Seite unvorhergesehen am Ende der Trajektorie immer noch überschritten werden. Beide

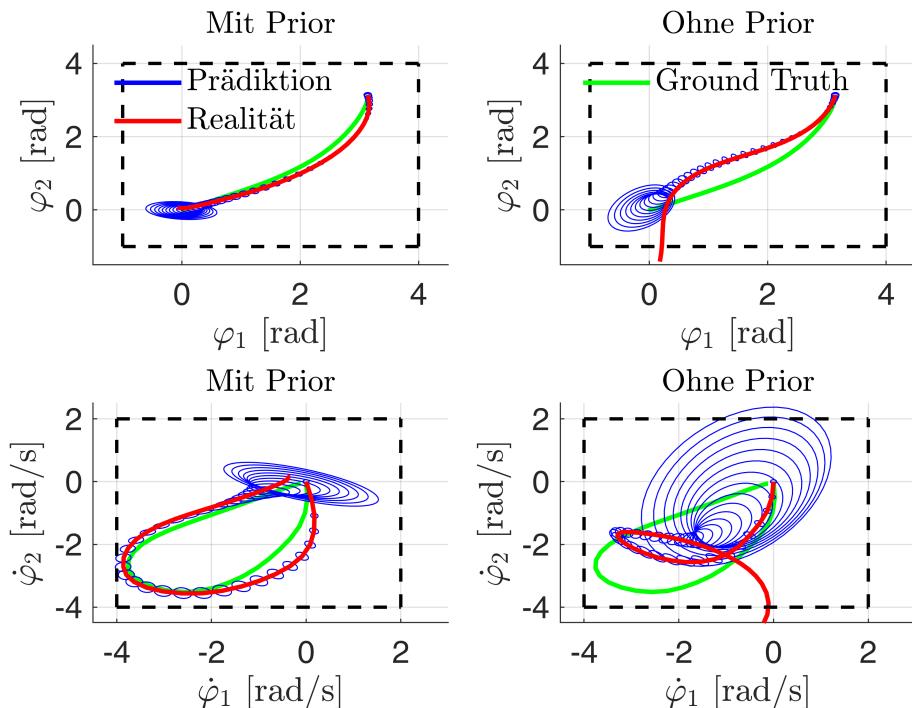


Bild 4-5: Gegenüberstellung von prädizierter (blau) und (fiktiv) gemessener (rot) Zustandstrajektorie aus der 4. Lerniteration der hybriden Optimalsteuerung. Die linke Spalte zeigt den Fall mit und die rechte ohne physikalisches Vorwissen (Prior). Die ideale (Ground Truth) Lösung ist in grün dargestellt. Die blauen Ellipsen deuten die multivariaten Normalverteilungen des Zustands in jedem Zeitschritt an, wobei jeweils 86% der Gesamtwahrscheinlichkeitsdichte abdeckt wird. Die schwarzen gestrichelten Linien repräsentieren die Zustandsbeschränkungen.

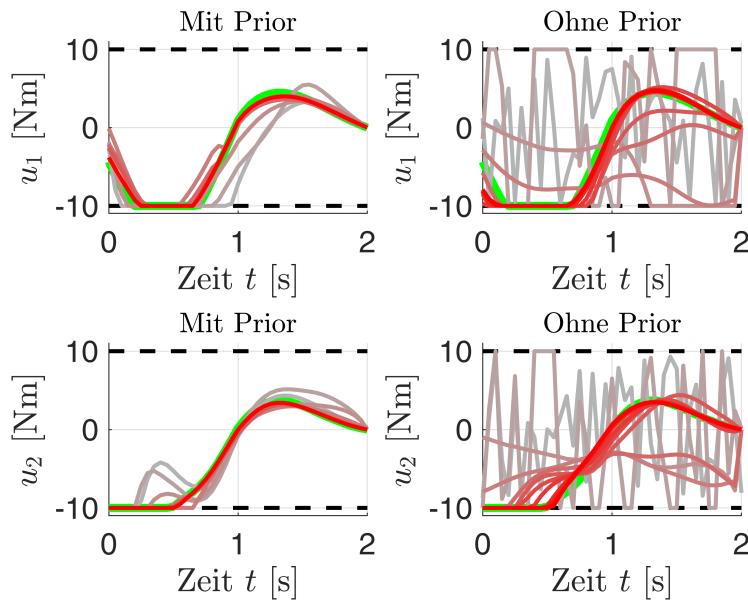


Bild 4-6: Entwicklung der gefundenen optimalen Steuerungen. Das Farbspektrum ändert sich allmählich von Grau für die ersten Iterationen zu Rot für die letzten Iterationen. Die hellroten Linien zeigen entsprechend die Steuerung in der letzten bzw. 10. Lerniteration. Gestrichelte schwarze Linien repräsentieren die Stellgrößenbeschränkung. Die dick gedruckten grünen Linien zeigen den Ground Truth.

Trajektorien haben gemeinsam, dass ihre Zustandsvarianz im Laufe der Zeit und besonders in der Nähe der oberen instabilen Ruhelage zunimmt, was nachvollziehbar ist, da mehrere Zustände, die durch die Zustandsverteilung repräsentiert werden, nicht mit nur einer bestimmten Stellgröße stabilisiert werden können. Obwohl die Varianz am Ende der Trajektorie relativ groß ist, folgt die tatsächliche Trajektorie genau der wahrscheinlichsten vorhergesagten Trajektorie, die durch den Mittelwert ausgedrückt wird (Zentrum der Ellipsen). Auf der rechten Seite (ohne Prior) zeigt sich ein anderes Bild, denn die reale Trajektorie verlässt für die hohen Winkelgeschwindigkeiten ihren Korridor der Prädiktion, was zeigt, dass die Zentrifugalkräfte für die optimale Steuerung des Systems entscheidend sind. Wie bereits erwähnt, zeigt dies außerdem, dass die gegebene Wahrscheinlichkeitsbegrenzung (4-27) des Zustands nur so zuverlässig ist, wie die Qualität des verwendeten Modells. Bild 4-6 zeigt die Variationen der berechneten Steuersignale in jeder Lerniteration. Die Veränderungen sind von Iteration zu Iteration für den Fall mit Vorwissen signifikant kleiner und somit hinsichtlich der Vorhersagbarkeit des Systemverhaltens und der praktischen Anwendbarkeit günstiger. Die ursprüngliche Steuerung, welche nur auf der Basis des physikalischen Modells berechnet wurde, wird nur geringfügig an einigen Stellen des Verlaufs partiell korrigiert, bspw. bei $t = 0.35$ s für u_2 oder $t = 1.25$ s für u_1 . Der Regelungstechniker kann während des Inbetriebnahmeprozesses somit besser abschätzen, wie sich das reale System verhalten wird. In dem Fall ohne Vorwissen ist dies deutlich

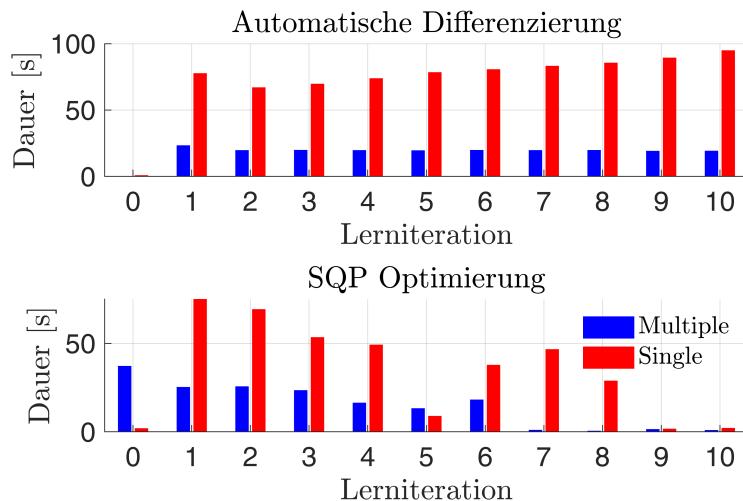


Bild 4-7: Vergleich von Multiple und Single Shooting hinsichtlich der benötigten Berechnungszeit für die automatische Differenzierung von $\nabla \mathbb{E}_{x_k}[J(\theta)]$, $\nabla h(\theta)$, $\nabla g(\theta)$ und die Optimierung mit dem SQP-Verfahren.

schwieriger, da die Steuerungen beider Aktoren sich in den ersten Iterationen zum Teil deutlich voneinander unterscheiden.

Abschließend wurde die benötigte Berechnungszeit für den Multiple und Single Shooting Ansatz gemessen und untersucht, siehe Bild 4-7. Beide Methoden führen zu beinahe identischen Trajektorien und Zielfunktionswerten innerhalb jeder Iteration, allerdings benötigt der Multiple Shooting Ansatz trotz seiner höheren Komplexität insgesamt weniger Rechenzeit, sowohl für die automatische Differenzierung mittels Casadi, als auch die SQP-Optimierung. Dabei ist vorteilhaft, dass die Lösung der vorherigen Lerniteration als Initialisierung für die nächste Optimierung verwendet wurde. Eine Ausnahme stellt die Initialisierung dar, in welcher das Single Shooting vergleichsweise schnell zu einer Lösung führt. Ein Grund hierfür ist, dass dem Multiple Shooting Ansatz eine einfache Initialisierung mit Einheitsmatrizen für die Zustandsvarianzmatrizen vorgeben wurde und es daher lange dauert, bis die vorgegebene Genauigkeit der Nebenbedingungen erreicht wird. Ein interessantes Verbesserungspotential besteht darin, das Multiple Shooting mit der Lösung des Single Shootings zu initialisieren, um die gesamte Berechnungszeit in Hinblick auf Algorithmus 4 weiter zu senken.

4.5 Erweiterung um Zwei-Freiheitsgrade-Regelungsstruktur und Validierung am Doppelpendel auf einem Wagen

Um einen gesamtheitlichen Entwurf zu ermöglichen, wird Algorithmus 4 zur hybriden Optimalsteuerung in diesem Abschnitt um eine Zwei-Freiheitsgrade-Regelungsstruktur (s. Bild 2-1) erweitert. Zur Lösung der Steuerungsaufgabe beim voll-aktuierten Doppelpendel war im vorherigen Abschnitt eine reine Steuerung ausreichend. Abgesehen von

dem Modellfehler bzgl. der Zentrifugalkräfte, wurden keine weiteren Störungen innerhalb der Simulationsumgebung betrachtet, sodass keine zusätzliche Regelung erforderlich war. In der Praxis treten jedoch unvorhergesehene äußere Störungen, u. a. in Form von System- und Messrauschen, auf, welche das System von seiner geforderten Soll-Trajektorie abbringen. Zudem unterliegt das gelernte hybride Modell gewissen Annahmen, bspw. durch die angesetzte Kovarianzfunktion des GPs, sodass es kein ideales Abbild des betrachteten Systems darstellt. Der Einsatz der Zwei-Freiheitsgrade-Regelungsstruktur sorgt in diesem Zusammenhang für eine Stabilisierung der gewünschten Soll-Trajektorie und wirkt den angesprochenen Störungen adäquat entgegen. Der nachfolgende Abschnitt hat das Ziel einen erweiterten hybriden Optimalsteuerungsentwurf allgemein zu formulieren und am realen Doppelpendelprüfstand (s. Bild 1-4 (links)) auf der Grundlage eines nur partiell bekannten Modells zu realisieren. Der anvisierte Fokus liegt damit auf der Einbindung von Vorwissen zur Lösung der anspruchsvollen Aufgabe des optimalen Aufschwungs und stabilen Trajektorienfolgeregelung. Auf diese Weise lässt sich das vorhergehende Ergebnis von PILCO (vgl. Abschnitt 4.3) mit denen des in dieser Arbeit entwickelten Verfahrens vergleichen und ein Rückschluss bzgl. der Effizienz und Praktikabilität für die Inbetriebnahme ziehen. Die grundsätzliche Ausgangssituation ist dabei identisch zu jener des voll-aktuierten Doppelpendels. Es soll von einem physikalischen Modell ausgegangen werden, welches die fundamental wichtigen Zentrifugalkräfte nicht enthält (vgl. (4-31)), sodass sie durch das ML-Modell ausgeglichen werden müssen. Die Entwicklung und Validierung des Verfahrens sind Teil der Vorveröffentlichung [HSTT24].

Eine erste notwendige Anpassung ist in Bezug auf das Optimierungsproblem (4-28) zu nennen. Hierbei wurde bisher nicht der Zielzustand x_G als explizite Nebenbedingungen berücksichtigt. In den durchgeföhrten Experimenten am Doppelpendelprüfstand hat sich beispielsweise gezeigt, dass eine solche Nebenbedingung aufgrund der hohen Sensitivität des Systems für eine erfolgreiche Realisierung des Aufschwungs notwendig ist. Da es sich bei (4-28) um eine probabilistische Formulierung handelt, liegt der Endzustand x_H nicht eindeutig vor, sondern in Form einer Normalverteilung $\mathcal{N}(m_H, S_H)$. Ähnlich zum deterministischen Fall des Multiple Shootings erscheint es im ersten Moment sinnvoll, eine Nebenbedingung bzgl. der Zustandsverteilung zu formulieren. Hierbei ist jedoch unklar, wie eine solche Ziel-Verteilung in Hinblick auf die Varianz gewählt werden sollte. Eine exakte Einhaltung ist zudem äußerst unwahrscheinlich hinsichtlich der probabilistischen Langzeitprädiktion. Eine andere Möglichkeit besteht darin, die Endzustandsverteilung nicht über die Gleichheits- $h(\cdot)$ sondern über die Ungleichheitsnebenbedingungen $g(\cdot)$ zu berücksichtigen, wobei die Endverteilung mit einer gewünschten Verteilung über die Berechnung der Kullback-Leibler-Distanz [Bis06] unter einem bestimmten Grenzwert gehalten wird. Somit würde die strikte Einhaltung entschärft und der Optimierung mehr Flexibilität eingeräumt werden, sodass eine zielführende und plausible Lösung er-

reicht wird. Im Rahmen dieser Arbeit konnten allerdings keine zufriedenstellenden Ergebnisse mit den angeführten Ansätzen erzielt werden. Aus diesem Grund wurde als praktikabler Kompromiss nur die Einhaltung des Mittelwertes der Endzustandsverteilung $\mathbf{m}_H - \mathbf{x}_G \stackrel{!}{=} \mathbf{0}$ über eine weitere Gleichheitsnebenbedingung gefordert und $\mathbf{h}(\cdot)$ in (4-28) dementsprechend erweitert.

Als Nächstes wird die Regelung der Soll-Trajektorie erläutert. In jeder Iteration des Verfahrens liegt vor dem Experiment am realen System eine Lösung der Form $\{\mathbf{m}_k^*, \mathbf{S}_k^*, \mathbf{u}_k^*\}$ mit $k = 0, \dots, H$ vor. Da die Zwei-Freiheitsgrade-Regelungsstruktur eingesetzt werden soll, wird im Folgenden ein Unterschied zwischen dem zugehörigen Steuerungs- und Regelungsanteil vorgenommen. Das Stellgrößensignal aus der Optimierung lautet somit $\mathbf{u}_k^* \rightarrow \mathbf{u}_{S,k}^*$. Zunächst ist die Frage zu klären, auf welchen Sollwert eine Regelung erfolgen soll, denn im Gegensatz zum deterministischen Fall ist nur eine Zustandsverteilung bekannt und bei der Regelung am realen System muss zu jedem Zeitpunkt eine exakte Wahl des Sollwerts vorliegen. Am naheliegendsten ist die Regelung auf der Trajektorie, welche durch das hybride Modell am wahrscheinlichsten erscheint. Diese wird durch den Verlauf des Zustandsmittelwerts \mathbf{m}_k^* repräsentiert. Der Anteil der Regelung wird durch einen zeitvarianten Riccati-Regler [Föl94] (in Anlehnung an das modellbasierte Regelungskonzept aus [TKOT11] für das betrachtete Anwendungsbeispiel) mit

$$\mathbf{u}_{R,k}(\mathbf{x}_k) = \mathbf{K}_k(\mathbf{m}_k^* - \mathbf{x}_k), \quad (4-33)$$

realisiert, sodass sich die gesamte Stellgröße zu $\mathbf{u}_k = \mathbf{u}_{S,k}^* + \mathbf{u}_{R,k}$ ergibt. Die Verstärkungsmatrix $\mathbf{K}_k \in \mathbb{R}^{n_x \times n_u}$ stellt hierbei einen linearen Zusammenhang zwischen der Regeldifferenz und der Stellgröße her und wird nach Riccati mit

$$\mathbf{K}_k = (\mathbf{W}_u + \mathbf{B}_k^T \mathbf{P}_k \mathbf{B}_k)^{-1} \mathbf{B}_k^T \mathbf{P}_k \mathbf{A}_k, \quad (4-34)$$

berechnet. Hierin stellt $\mathbf{P}_k \in \mathbb{R}^{n_x \times n_x}$ die Riccati-Matrix zum Zeitschritt k dar, welche mittels zeitlicher Rückwärtsintegration der riccatischen Differentialgleichung bestimmt wird [Föl94]. In einer bereits diskretisierten Form nach [RT10] lautet der Zusammenhang

$$\mathbf{P}_k = \mathbf{W}_x + \mathbf{A}_k^T \mathbf{P}_{k+1} \mathbf{A}_k - \mathbf{A}_k^T \mathbf{P}_{k+1} (\mathbf{B}_k (\mathbf{W}_u + \mathbf{B}_k^T \mathbf{P}_{k+1} \mathbf{B}_k)^{-1} \mathbf{B}_k^T \mathbf{P}_{k+1}) \mathbf{A}_k, \quad (4-35)$$

$k = H - 1, \dots, 0$, wobei sich die Anfangs- bzw. Endbedingung \mathbf{P}_H aus der Lösung der statischen Riccatigleichung ergibt und $\mathbf{W}_x \geq 0 \in \mathbb{R}^{n_x \times n_x}$ und $\mathbf{W}_u > 0 \in \mathbb{R}^{n_u \times n_u}$ Gewichtungsmatrizen bzgl. des Zustands und der Stellgröße darstellen. Des Weiteren sind $\mathbf{A}_k \in \mathbb{R}^{n_x \times n_x}$ und $\mathbf{B}_k \in \mathbb{R}^{n_x \times n_u}$ die Dynamik- und Eingangsmatrix, welche sich aus einer Linearisierung des hybriden Modells ergeben. Aufgrund der gewählten additiven

Zusammensetzung des hybriden Modells (4-20) lassen sich die entsprechenden Ableitungen des physikalischen und datengetriebenen Modellteils separat voneinander bilden über

$$\mathbf{A}_k = \frac{d\mathbf{f}_{phy}}{d\mathbf{x}_k} \Big|_{\substack{\mathbf{x}_k=\mathbf{m}_k^*, \\ \mathbf{u}_k=\mathbf{u}_k^*}} + \frac{d\Delta\boldsymbol{\mu}}{d\mathbf{x}_k} \Big|_{\substack{\mathbf{x}_k=\mathbf{m}_k^*, \\ \mathbf{u}_k=\mathbf{u}_k^*}}, \quad \mathbf{B}_k = \frac{d\mathbf{f}_{phy}}{d\mathbf{u}_k} \Big|_{\substack{\mathbf{x}_k=\mathbf{m}_k^*, \\ \mathbf{u}_k=\mathbf{u}_k^*}} + \frac{d\Delta\boldsymbol{\mu}}{d\mathbf{u}_k} \Big|_{\substack{\mathbf{x}_k=\mathbf{m}_k^*, \\ \mathbf{u}_k=\mathbf{u}_k^*}}.$$

Bei den in dieser Arbeit eingesetzten Kernelfunktionen lässt sich der Posterior des GPs analytisch ableiten. Eine detaillierte Darstellung ist im Anhang A1 unter dem Abschnitt - Ableitung des GP-Posterior-Erwartungsvektors nach dem erweiterten Zustand - zu finden. In Bezug auf Algorithmus 4 ist die Berechnung des Reglers nach Zeile 6 und vor 7 einzuordnen. Der Regelungstechniker steht dementsprechend vor dem Experiment am Prüfstand der Frage gegenüber, ob nur die reine Steuerung $\mathbf{u}_k = \mathbf{u}_{S,k}^*$ oder die Zwei-Freiheitsgrade-Regelungsstruktur $\mathbf{u}_k = \mathbf{u}_{S,k} + \mathbf{u}_{R,k}$ mit einem möglicherweise instabilen und unsicheren Verhalten verwendet werden soll. Hierfür liegt in Verbindung mit der Steuerung \mathbf{u}_k^* die probabilistische Langzeitprädiktion $\{t_k, \mathbf{m}_k^*, \mathbf{S}_k^*\}$ aus der Lösung des hybriden Optimalsteuerungsproblems vor, anhand derer eine Bewertung vorgenommen werden kann. Um zusätzlich den Einsatz der Regelung bewerten zu können, ist eine ähnliche Prädiktion für das geregelte System wünschenswert. Aus diesem Grund wird vor dem Experiment eine weitere Simulation berechnet, wobei über die UT der geschlossene Regelkreis am hybriden Modell getestet wird. Für das Verständnis der nachfolgenden Punkte ist es daher notwendig, eine Unterscheidung zwischen der Prädiktion des offenen $(\mathbf{m}_k^*, \mathbf{S}_k^*) \rightarrow (\mathbf{m}_{o,x_k}^*, \mathbf{S}_{o,x_k}^*)$ und dem geschlossenen Regelkreis $(\mathbf{m}_{g,x_k}, \mathbf{S}_{g,x_k})$ einzuführen und die Zuordnung zum Zustand kenntlich zu machen.

Eine Besonderheit bei der Berechnung des geschlossenen Regelkreises ist, dass die Stellgröße $\mathbf{u}_k(\mathbf{x}_k)$ über das Regelgesetz (4-33) nun vom Zustand abhängt und diese damit als Zufallsvariable zu betrachten ist. Entsprechend induziert die Zustandsverteilung eine Unsicherheit in die Stellgröße, wobei durch die lineare Abhängigkeit und über (2-10), (2-11) geschlussfolgert wird, dass

$$\begin{aligned} \mathbf{u}_k &\sim \mathcal{N}(\mathbf{m}_{u_k}, \mathbf{S}_{u_k}), \\ \mathbf{m}_{u_k} &= \mathbf{u}_{S,k}^* + \mathbf{K}_k(\mathbf{m}_{o,x_k}^* - \mathbf{m}_{g,x_k}), \\ \mathbf{S}_{u_k} &= \mathbf{K}_k \mathbf{S}_{g,x_k} \mathbf{K}_k^T, \end{aligned} \tag{4-36}$$

gilt. Die Varianz der Stellgröße \mathbf{S}_{u_k} ist ein hilfreiches Maß, um abschätzen zu können, wie sehr sich die Unsicherheit über die Dynamik des Systems im Regelkreis bemerkbar macht. Eine geringe Varianz weist darauf hin, dass die Modellfehler durch das GP-Modell gut abgebildet werden und ausreichend Daten vorhanden sind. Des Weiteren lässt sich an dieser Stelle der klassische regelungstechnische Stabilitätsbegriff (bspw. nach Ljapunov (2-5)) um eine wahrscheinlichkeitsbasierte Sichtweise erweitern. Grundsätzlich wird

sich der Regler bei der Zustandspropagation so bemerkbar machen, dass er die Varianz des Zustands deutlich reduziert. Das liegt daran, dass er die Zustände, die abseits der gewünschten Trajektorie liegen, auf ihren Pfad zurückführt und sich damit insbesondere für das sensitive Doppelpendel das Bewegungsverhalten insgesamt besser vorhersagen lässt. Für den Fall der reinen Steuerung ist die vergleichsweise hohe Zustandsvarianz darin begründet, dass sich mehrere unterschiedliche Zustände, welche durch die Zustandsverteilung repräsentiert werden, in der Nähe der oberen instabilen Ruhelage nicht durch eine einzige Stellgröße stabilisieren lassen. Der Regler stellt in diesem Zusammenhang zu jedem Zustand, der sich abseits der Soll-Trajektorie befindet, eine passende Stellgröße her. Per Definition bzw. Konstruktion des Verfahrens sollte der berechnete Riccati-Regler bzgl. des hybriden Modells immer stabil sein [Fö194]. Für die Inbetriebnahme ist die Stabilität bzgl. des Modells allerdings nicht entscheidend, sondern nur die Stabilität bzgl. des realen Systems (vgl. Bild 4-2). Die Zustandspropagation des geschlossenen Regelkreises ermöglicht es, die Auswirkung der Unsicherheit über die Systemdynamik zu berücksichtigen. Die Fortpflanzung von Modellfehlern, welche sich in einer hohen Varianz des GP-Modellteils äußert, sorgt über (4-36) zu einer anwachsenden Zustandsvarianz, welche sich gegebenenfalls immer weiter verstärken kann. Ähnlich zum deterministischen Fall, ergibt sich somit unter Umständen ein instabiles Regelungsverhalten, welches auf die Tatsache zurückzuführen ist, dass der eingeführte Riccati-Regler nicht die Unsicherheit des Modells berücksichtigt. Eine Möglichkeit, diese Information beim Entwurf nutzbar zu machen, wäre es indem die Verstärkungsmatrix K_k eine Abhängigkeit von der Zustandsvarianz S_{x_k} aufweisen würde. Bspw. könnte die Einflussnahme des Reglers bei einer zu hohen Zustandsvarianz reduziert werden, wodurch sich ein instabiles Verhalten verhindern ließe. Im Rahmen dieser Arbeit wurde dieser Ansatz nicht weiter erforscht, wird jedoch als sehr vielversprechend eingestuft, sodass er ein Potential für weiterführende Arbeiten bietet. Eine Orientierung innerhalb der Literatur könnte hierzu [EM23] bereitstellen, worin eine probabilistische Betrachtungsweise des Riccati-Reglers behandelt wird.

Ausgehend von der ursprünglichen Fragestellung, welche den Einsatz der Regelung am Prüfstand behandelt, bietet die Zustandspropagation des geschlossenen Regelkreises durch die Zustandsvarianz einen Anhaltspunkt darüber, wie sich das reale System verhalten wird. Eine rasch anwachsende Zustandsvarianz deutet dabei ein instabiles Verhalten am Prüfstand an, sodass dem Anwender über diese Information ein objektives Bewertungskriterium zur Verfügung steht. Eine Automatisierung lässt sich durch die Einführung einer Grenzwertmatrix $S_{max} \in \mathbb{R}^{n_x \times n_x}$ herstellen, welche die Zustandsvarianz der einzelnen Dimensionen überprüft und unter Umständen mit den geforderten technischen Rahmenbedingungen abgleicht. Algorithmus 5 fasst die Schritte der erweiterten hybriden Opti-

Algorithmus 5 Erweiterte Hybride Optimalsteuerung

- 1: **Initialisierung:** Berechne Optimalsteuerung $\{\boldsymbol{u}_k^*\}_{k=0}^H$ nach (4-2) auf Basis des physikalischen Modells \boldsymbol{f}_{phy} und teste sie am realen System, wobei initiale Messdaten gesammelt werden.
 - 2: **Wiederhole** bis ein Konvergenzkriterium oder ein Iterationsbudget erreicht ist:
 - 3: Lerne Modellfehler über GPs (4-19) auf Basis der vorhandenen Daten (4-16).
 - 4: **Wiederhole** die nachfolgenden Schritte im Rahmen einer SQP-Optimierung (A2)
 - 5: Berechne den Gradienten $\nabla \mathbb{E}[J(\boldsymbol{\theta})]$ und die Jacobimatrizen der Nebenbedingungen $\nabla \boldsymbol{h}(\boldsymbol{\theta}), \nabla \boldsymbol{g}(\boldsymbol{\theta})$ \triangleright Automatische Differenzierung (CasADi)
 - 6: Aktualisiere den Optimierungsvektor $\boldsymbol{\theta}$ und die zugehörigen Lagrange-Multiplikatoren $(\boldsymbol{\lambda}, \boldsymbol{\mu})$.
 - 7: Erhalte Optimalsteuerung und Prädiktion (offener Rk) $\{\boldsymbol{m}_{o,x_k}^*, \boldsymbol{S}_{o,x_k}^*, \boldsymbol{u}_{S,k}^*\}_{k=0}^H$.
 - 8: Berechne zeitvarianten Riccati-Regler $\{\boldsymbol{K}_k\}_{k=0}^H$ nach (4-34), (4-35) und Prädiktion für den geschlossenen Regelkreis $\{\boldsymbol{m}_{g,x_k}, \boldsymbol{S}_{g,x_k}, \boldsymbol{m}_{u_k}, \boldsymbol{S}_{u_k}\}_{k=0}^H$.
 - 9: **Falls** $S_{g,x_k}^{(i,i)} \leq S_{max}^{(i,i)}, k = 0, \dots, H, i = 1, \dots, n_x$ dann \triangleright Überprüfe Stabilität
setze $\boldsymbol{u}_k \leftarrow \boldsymbol{u}_{S,k}^* + \boldsymbol{K}_k(\boldsymbol{m}_{o,x_k}^* - \boldsymbol{x}_k)$,
sonst
setze $\boldsymbol{u}_k \leftarrow \boldsymbol{u}_{S,k}^*$.
 - 10: Erprobe \boldsymbol{u}_k am realen System und nehme weitere Messdaten auf, die den bestehenden Daten hinzugefügt werden.
-

malsteuerung zusammen und ergänzt Algorithmus 4 um einen zusätzlichen Regelungsentwurf (Zeile 8), sowie einer automatisierten Überprüfung der Stabilität (Zeile 9).

Anwendung auf ein reales Doppelpendel auf einem Wagen

In diesem Abschnitt werden die Ergebnisse der hybriden Optimalsteuerung mit dem erweiterten Regelungskonzept für das Doppelpendel auf einem Wagen vorgestellt. In Unterkapitel 4.2 wurde hierzu die physikalische Modellbildung mittels Lagrange-Formalismus für das System erläutert. Ähnlich zu den Experimenten am simulierten voll-aktuierten Doppelpendel, soll von einem unvollständigen Modell als Vorwissen ausgegangen werden. Vor dem Hintergrund des Prüfstands stellt jedes Modell nur eine Annäherung dar, sodass selbst das hier vorgestellte Gesamtmodell als unvollständig zu bezeichnen ist. Um die Wirkungsweise der hybriden Optimalsteuerung besser erläutern zu können und ein anspruchsvoller Szenario zu schaffen, wurden die hoch relevanten Zentrifugalkräfte als unbekannt angenommen. Ohne diese Kräfte im Modell ist eine erfolgreiche Inbetriebnahme am Prüfstand nach dem klassischen Vorgehen nicht möglich, sodass ein Szenario geschaffen wurde, in welchem die hybride Optimalsteuerung für den Regelungstechniker ein geeignetes Werkzeug darstellt. Des Weiteren ist durch die Vernachlässigung der Zentrifugalkräfte sichergestellt, dass sich in den aufgenommenen Messdaten ein systematischer und damit für das GP-Teilmodell lernbarer Fehler befindet. Der prognostizierte

Modellfehler bzgl. der partiell zustandslinearisierten Form des Modells (4-10) lautet somit

$$\Delta\dot{\varphi}_{k+1} = -\Delta t \mathcal{M}_{1,1}^{-1}(\varphi_k) \mathcal{C}_{1,1}(\varphi_k, \dot{\varphi}_k),$$

wobei er auf der Ebene der Winkelgeschwindigkeiten $\dot{\varphi} = [\dot{\varphi}_1, \dot{\varphi}_2]^T$ definiert ist und die enthaltenen Teilmatrizen durch

$$\begin{aligned}\mathcal{M}_{1,1}(\varphi) &= \begin{bmatrix} J_1 + a_1^2 m_1 + l_1^2 m_2 & a_2 l_1 m_2 \cos(\varphi_1 - \varphi_2) \\ a_2 l_1 m_2 \cos(\varphi_1 - \varphi_2) & J_2 + a_2^2 m_2 \end{bmatrix}, \\ \mathcal{C}_{1,1}(\varphi, \dot{\varphi}) &= \begin{bmatrix} -a_2 l_1 m_2 \sin(\varphi_1 - \varphi_2) \dot{\varphi}_2^2 \\ -a_2 l_1 m_2 \sin(\varphi_1 - \varphi_2) \dot{\varphi}_1^2 \end{bmatrix},\end{aligned}\tag{4-37}$$

beschrieben werden. Anhand dieser Gleichungen bzw. der Inversion der Teilmassenmatrix $\mathcal{M}_{1,1}$ und der Querverbindung beider separaten Zentrifugalkräfte untereinander wird ersichtlich, dass es sich bei dem angenommenen Modellfehler um keinen trivialen Term handelt, sondern dieser einen großen Einfluss auf die Bewegung des Systems nimmt. Dies macht sich insbesondere für hohe Winkelgeschwindigkeiten bemerkbar, welche für den Aufschwung des Doppelpendels unabdingbar sind. Da von der partiell zustandslinearisierten Form ausgegangen wird, findet keine Korrektur der zugehörigen Gleichungen der Wagengeschwindigkeit statt, da diese bei einer ausreichend schnellen Geschwindigkeitsregelung als einfacher Integrator modelliert wird. Die zugrundeliegende Geschwindigkeitsregelung wird hier als eine weitere Art des Vorwissens betrachtet, da sich diese unabhängig vom Doppelpendel auf dem klassischen Entwurfsweg auslegen lässt. Am Prüfstand sorgt die Geschwindigkeitsregelung dafür, dass sich die Kräfte, welche durch die Bewegung der Pendelarme erzeugt werden, nur geringfügig auf das Verhalten des Wagens auswirken. Diese Kopplung findet in der partiell zustandslinearisierten Form keine Berücksichtigung, sodass das Integratorverhalten des Wagens keine Abhängigkeit von den Zuständen der Winkel hat. Für das hybride Modell und die Zustandspropagation bedeutet dies, dass sich die Zustandsvarianz nur auf die Zustände der Winkel auswirkt. Die Dynamik des Wagens bleibt bei der Verwendung einer eindeutig vorgegebenen Steuerung deterministisch. Bei der Verwendung der Regelung ist dies nicht mehr der Fall, da über das Regelgesetz (4-33) und die damit verbundene probabilistische Stellgröße (4-36) eine Abhängigkeit besteht. Aus diesem Grund pflanzt sich die Unsicherheit der Pendelarme über die Stellgröße in die Wagenbewegung fort.

Das betrachtete Szenario beinhaltet eine zeitliche Diskretisierung von $\Delta t = 0.01$ s mit einer fest vorgegebenen Aufschwungzeit von $T = 2$ s. Das System startet aus dem Anfangszustand $\mathbf{m}_I = [-\pi, -\pi, \mathbf{0}]^T$, $\mathbf{S}_I = 10^{-3} \mathbf{I}_6$ und soll zum Zielzustand $\mathbf{x}_G = \mathbf{0}$ überführt

werden. Um die technischen Rahmenbedingungen mit einem ausreichend großen Sicherheitsabstand zu adressieren, werden die Nebenbedingungen mit

$$|y| \leq 0.5 \text{ m}, \quad |\dot{y}| \leq 3 \frac{\text{m}}{\text{s}}, \quad |u| = |\ddot{y}| \leq 50 \frac{\text{m}}{\text{s}^2}, \quad \text{und} \quad P_x = 95\% \quad (4-38)$$

berücksichtigt. Die Gewichtungen des Optimalsteuerungsproblems (4-28) haben die Werte

$$\mathbf{W}_x = \text{diag}(100, 100, 250, \mathbf{0}), \quad W_u = 1, \quad w_t(k) = \frac{e^{25\Delta t k} - 1}{e^{25Hk} - 1}.$$

Bevor der Algorithmus am realen System erprobt worden ist, wurde eine unabhängige Analyse in einer Simulationsumgebung durchgeführt. Hierbei wurde das Gesamtmodell des Doppelpendels auf einem Wagen, d. h. inkl. Zentrifugalkräfte, stellvertretend für das reale System verwendet und an fiktiven Messdaten gelernt. In der Simulation zeigte sich, dass mit Hilfe des unvollständigen (physikalischen) Modells kein Aufschwung des Gesamtmodells realisiert werden konnte. Auch eine zusätzliche Regelung war nicht in der Lage, das System entlang der vorgeschlagenen Aufschwungtrajektorie zu stabilisieren und führte in allen Untersuchungen zur Instabilität. Dieser Ausgangspunkt entspricht der Initialisierung der hybriden Optimalsteuerung bzw. Zeile 1 in Algorithmus 5. Um den Modellfehler zu korrigieren und den Aufschwung zu realisieren, waren insgesamt lediglich zwei Lerniterationen bzw. weitere fiktive Experimente am Gesamtmodell erforderlich.

Weiterführend werden die Ergebnisse am Doppelpendelprüfstand erläutert. Im Vergleich zum Lernen am Simulationsmodell, ist am Prüfstand ein schwierigerer Inbetriebnahmeprozess zu erwarten, denn es muss von anderweitigen Modellierungsfehlern und Störungen ausgegangen werden. In der Simulation fand bspw. keine Berücksichtigung von Rauschprozessen durch die Sensorik statt. Diese ist nicht nur in Hinblick auf die Messdaten negativ zu bewerten, sondern auch in Bezug auf die Regelung, denn diese setzt einen vollständig bekannten Zustandsvektor voraus. Am Prüfstand werden allerdings nur die Positionen durch verschiedene Sensoren erfasst, sodass die Geschwindigkeiten durch Finite-Differenzen-Berechnungen approximiert werden müssen. Das vorhandene Messrauschen wird dabei verstärkt. Eine bessere Abschätzung stellt die Verwendung eines Beobachters dar, welcher dem Rauschprozess entgegenwirkt und ihn unterdrückt. Allerdings ergibt sich durch den Einsatz eines Beobachters eine weitere Fehlerquelle, welche im Zusammenhang mit der Regelung und dem hybriden Modell gegebenenfalls schwierig zu analysieren ist. Um diese Fehlerquelle auszuschließen, wurde die Approximation über die Finite-Differenzen-Methode beibehalten. Ein weiterer Grund ist, dass die Gauß-Prozess-Regression bereits durch ihre Herleitung eine natürliche Handhabung von Messrauschen in Form des zugehörigen Hyperparameters σ_n^2 vornimmt. Das Messrauschen ist, unter der

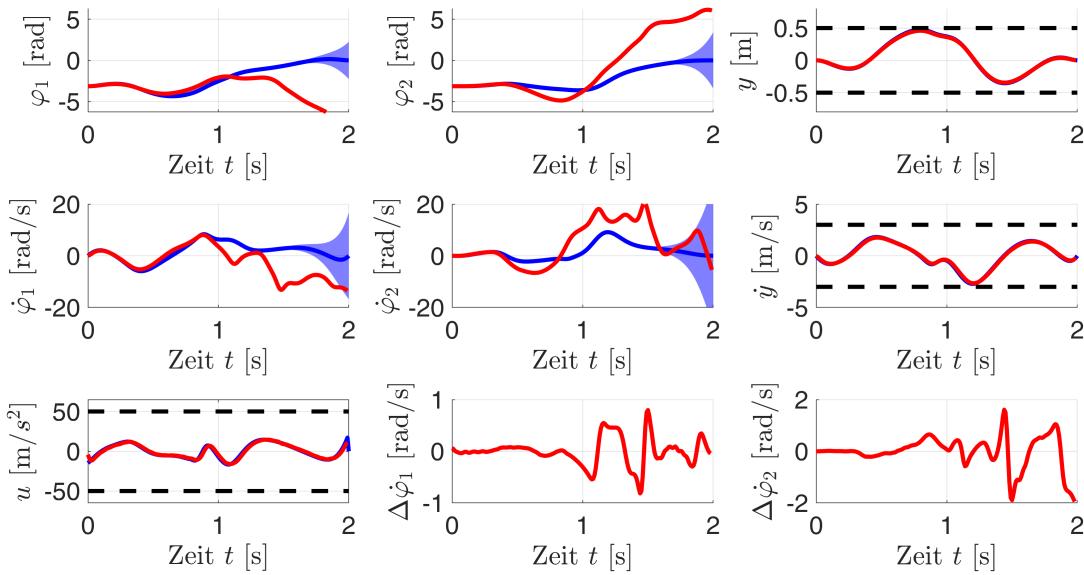


Bild 4-8: Ergebnisse der Initialisierung der erweiterten hybriden Optimalsteuerung am Doppelpendel auf einem Wagen. Die probabilistische Prädiktion ist in Blau und die Prüfstandsmessung in Rot dargestellt.

Voraussetzung einer ausreichend großen Datenmenge, damit als unproblematisch zu bewerten. Des Weiteren soll der hier entwickelte Ansatz mit dem PILCO-Verfahren und dessen Ergebnissen am Prüfstand verglichen werden (s. Abschnitt 4.3). Dabei wurde ebenfalls auf den Einsatz eines Beobachters verzichtet, sodass ein fairer Vergleich unter den gleichen Voraussetzungen stattfinden kann.

Bild 4-8 zeigt die zeitlichen Verläufe der initialen Erprobung. Die blauen Verläufe repräsentieren die Zustandsprädiktion, welche alleine auf der Basis des physikalischen Modells berechnet wurde. Die mittlere durchgezogene Linie steht für den Erwartungswert, wohingegen die helleren Flächen die zweifache Standardabweichung repräsentieren. Die roten Verläufe zeigen in diesem Zusammenhang die Messung am Prüfstand, wobei die gleiche Steuerung der Prädiktion eingesetzt wird. Neben den sechs Zustandsgrößen, ist der Stellgrößenverlauf (unten, links) und der Modellfehler bzgl. der beiden Winkelgeschwindigkeiten (unten, mittig und rechts) zu sehen. Die schwarz-gestrichelten Linien stehen für die technischen Beschränkungen (4-38) und müssen während des Aufschwungmanövers unbedingt eingehalten werden. Da bei der initialen Berechnung noch keine Messdaten vorliegen und damit auch kein GP, ist die einzige Unsicherheit die des Anfangszustands ($\mathbf{m}_I, \mathbf{S}_I$). Aus diesem Grund nimmt die Zustandsvarianz durch die Zustandspropagation mit der Zeit zu. Insbesondere zum Ende der Trajektorie, nimmt die Unsicherheit stark zu, denn die instabile obere Ruhelage sorgt für das Auseinanderdriften der möglichen Systemzustände. Da das physikalische Modell noch keinen probabilistischen Anteil über die beinhaltenden Fehler enthält, ist die Prädiktion insgesamt zu optimistisch. Dies drückt sich in der relativ niedrigen Zustandsvarianz bis etwa 1.5 s

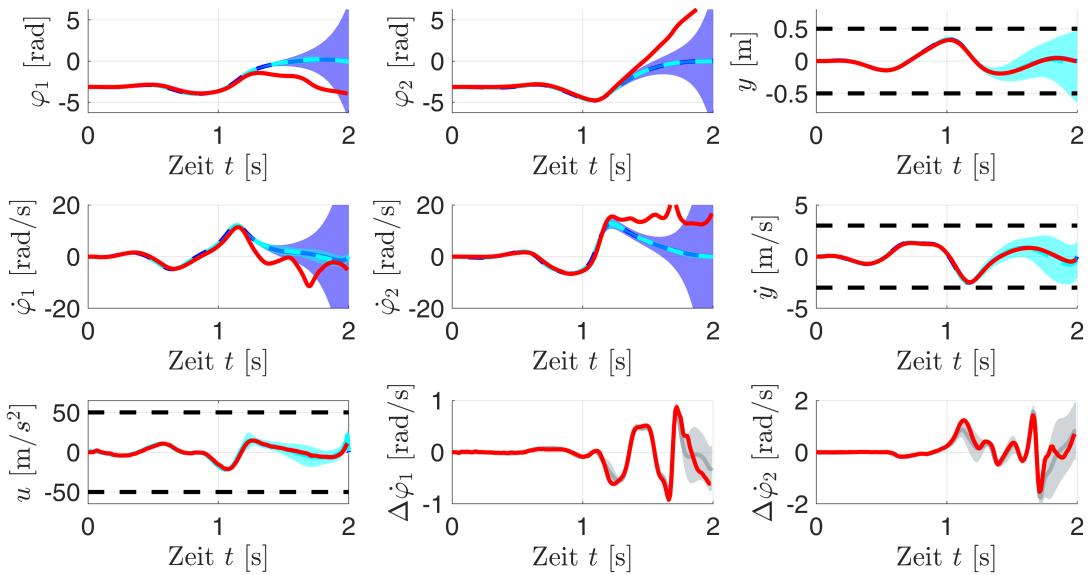


Bild 4-9: Ergebnisse der 1. Lerniteration der erweiterten hybriden Optimalsteuerung am Doppelpendel auf einem Wagen. Die Prädiktion des offenen Regelkreises ist in Blau, die des geschlossenen in Türkis und die Prüfstandsmessung in Rot dargestellt. Zudem ist die Vorhersage des Modellfehlers in Grau abgebildet.

aus. Die zeitlichen Verläufe, welche dem Wagen zugeordnet sind, weisen in diesem Zusammenhang, wie erläutert, einen eindeutig definierten Ablauf auf. Ihre deterministische Prädiktion stimmt nahezu ideal mit der Messung überein und kann daher vor dem Experiment gut durch den Ingenieur eingeschätzt werden. Bei den Winkeln zeichnet sich hingegen ein anderes Bild ab, wobei die Prädiktion bereits nach wenigen Zeitschritten ihre Gültigkeit verliert und das reale Systemverhalten relativ unvorhergesehen erscheint. Anhand der zeitlichen Verläufe der gemessenen Modellfehler wird ersichtlich, dass die ersten 0.5 s nur durch geringe Abweichungen geprägt sind. Innerhalb dieser Zeitspanne werden die Pendelarme aus der ruhenden Position durch den Wagen in Bewegung versetzt. Aufgrund der quadratischen Abhängigkeit der Zentrifugalkräfte von den Winkelgeschwindigkeiten nehmen die Fehler nach etwa 1 s rasch zu und wechseln gleichermaßen schnell ihre Richtung angesichts der trigonometrischen Terme in (4-37). Basierend auf diesen ersten Prüfstandsmessungen wird die 1. Iteration von Algorithmus 5 ausgeführt und der datengetriebene GP-Modellteil zur Fehlerkompensation aufgestellt.

Das Ergebnis der 1. Lerniteration ist in Bild 4-9 zu sehen. Neben den schon besprochenen Verläufen, treten nach der Initialisierung zwei weitere Verläufe hinzu. Zum einen wird in der 1. Iteration die zeitvariante Riccati-Regelung berechnet, woraus sich über eine nachträgliche Simulation des geschlossenen Regelkreises eine weitere Zustandsprädiktion (türkisfarbene Gebiete) ergibt. Anhand der Winkelverläufe wird hierbei besonders deutlich, dass der Eingriff der Regelung im Vergleich zur reinen Steuerung eine Reduzierung der Zustandsvarianz bewirkt. Insgesamt betrachtet, sinkt zwar die Varianz in Bezug auf

die Winkel, jedoch wird durch die Regelung eine zusätzliche Unsicherheit in die Stellgröße und die Wagendynamik induziert. Des Weiteren sind in den Diagrammen bzgl. der Modellfehler zwei neue graue Gebiete hinzugekommen, welche die Vorhersage des gelehnten GP-Modells über den Fehler widerspiegeln. Die Vorhersage wird im Anschluss an das Experiment auf der Basis der neu aufgenommenen Daten berechnet. Im Allgemeinen sind die neuen Daten nicht Teil der bestehenden Datenlage und damit für den momentan bekannten GP als unbekannt einzustufen. Die Auswertung des GPs an den Stellen der neuen Daten liefert somit eine Prädiktion über die noch ungesehenen Daten und entspricht damit in den überwiegenden Fällen einer Extrapolation. Diese Prädiktion wird mit den neuen Messdaten verglichen, um einen Eindruck über die Qualität des datengetriebenen Modellteils bzgl. der Modellfehler zu erhalten. Obwohl die Prädiktion des geschlossenen Regelkreises als stabil zu werten ist, wurde sich zunächst dafür entschieden, die Regelung am Prüfstand nicht zu verwenden und weiterhin nur die reine Steuerung auszuprobieren. Eine Begründung hierfür ist, dass die Menge an Daten noch relativ gering ist und dementsprechend dem datengetriebenen Teil noch nicht zu viel Vertrauen entgegengesetzt werden sollte. Das Verfahren befindet sich noch in einer explorativen Phase, in welcher ein frühzeitig instabiles Verhalten am Prüfstand hinderlich wäre. Ein Indiz hierfür liefert zudem die Prädiktion in Bezug auf die Wagenposition y . Am Ende der Trajektorie ist die zugehörige Varianz relativ groß, sodass nicht sichergestellt ist, dass die Wagenbeschränkungen eingehalten wird. Ein verfälschter Eindruck kann zudem auf die Approximation mittels UT zurückgeführt werden. Um ein genaueres Bild zu bekommen, wurde nachträglich eine komplexere MC-Simulation des geschlossenen Regelkreises nach dem Vorgehen in 2.3.3 durchgeführt. Die Ergebnisse wurden dem Anhang unter A3 hinzugefügt und zeigen im Gegensatz zur UT ein instabiles Systemverhalten.

Bei der Erprobung der Steuerung am Prüfstand wurden die roten Verläufe aufgezeichnet. Im direkten Vergleich zu Bild 4-8 bzw. dem initialen Experiment zeigt sich, dass die Prädiktion deutlich länger mit dem realen Bewegungsverhalten übereinstimmt. Der Pfad der Prädiktion wird erst beim Erreichen eines kritischen Punktes innerhalb der Trajektorie verlassen. Dieser Punkt befindet sich bei etwa 1.25 s, wobei die Winkelgeschwindigkeiten ihr kurzzeitiges Maximum anstreben. Erwartungsgemäß reagiert das chaotische System besonders stark an dieser Stelle auf kleinste Modellfehler, die beim modellbasierten Entwurf nicht berücksichtigt wurden, sodass das weitere Verhalten in Anbetracht der probabilistischen Prädiktion als relativ unwahrscheinlich erscheint. Abseits des anspruchsvollen Systems wurden im Rahmen der Gauß-Prozess-Regression und der Zustandspropagation verschiedenartige Annahmen getroffen, die bei der Inbetriebnahme gegeneinander abgewägt werden müssen. Zum einen spielt die Bestimmung der GP-Hyperparameter hierbei eine wichtige Rolle, welche aufgrund der geringen Datenlage als nicht perfekt anzusehen ist. Eine entsprechende korrekte Einstufung der Unsicherheit über die Systemdyna-

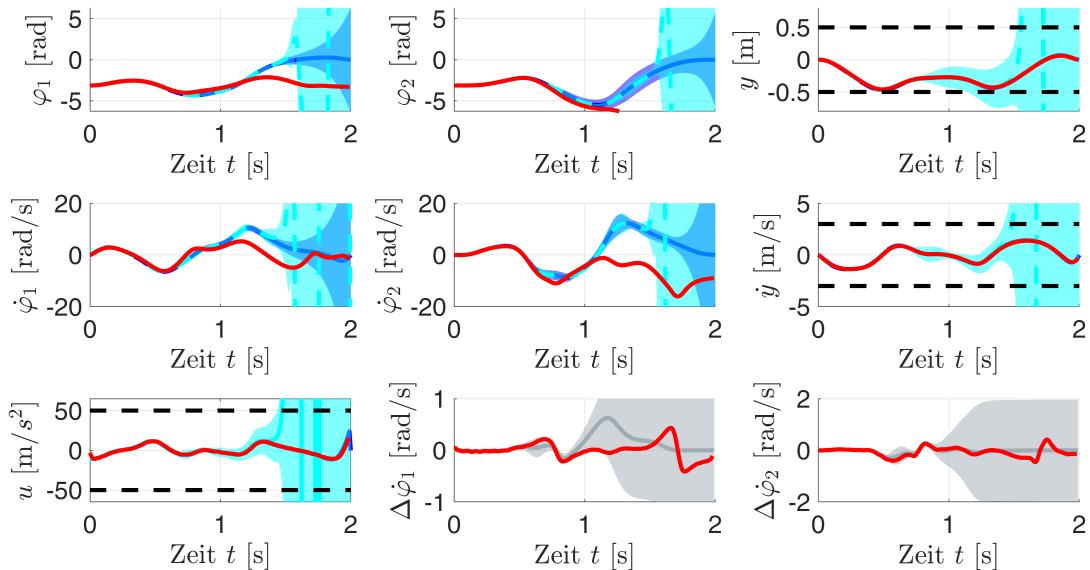


Bild 4-10: Ergebnisse der 3. Lerniteration der erweiterten hybriden Optimalsteuerung am Doppelpendel auf einem Wagen. Die Prädiktion des offenen Regelkreises ist in Blau, die des geschlossenen in Türkis und die Prüfstandsmessung in Rot dargestellt. Zudem ist die Vorhersage des Modellfehlers in Grau abgebildet.

mik ist damit innerhalb der ersten Iteration des Verfahrens in Frage gestellt. Zum anderen basiert die Zustandspropagation auf der UT und stellt damit lediglich eine Approximation an die wahre zugrundeliegende Zustandsverteilung dar. Die Approximationsfehler eines jeden Zeitschritts summieren sich entsprechend auf und verfälschen somit das Ergebnis zum Ende der Trajektorie. Dies sind die Gründe dafür, dass sich das reale System nicht auf den wahrscheinlichsten Trajektorien, welche durch das hybride Modell berechnet wurden, aufgehalten hat. Vor diesem Hintergrund bieten jedoch die Diagramme bzgl. der Modellfehler tiefere Einblicke. Nur anhand der Daten aus dem initialen Experiment ist der GP bereits sehr gut in der Lage den Modellfehler aus dem nächsten Experiment wiederzugeben. Bis auf wenige Ausnahmen stimmt der Mittelwert mit den Messungen überein. Die hohe Varianz nach etwa 1.6 s deutet jedoch an, dass gewisse Bereiche innerhalb des Zustandsraums noch unbekannt sind. Weiterführend werden die neuen Messdaten dafür genutzt, um die Prädiktionsgenauigkeit des GP weiter zu steigern, woraufhin die nächste Lerniteration angestoßen wird.

Die Ergebnisse der 2. sind ähnlich zu denen der 1. Iteration, sodass sie keine neuen Erkenntnisse beinhalten und übersprungen werden. Eine merkliche Veränderung ist erst in der 3. Iteration zu beobachten, welche in Bild 4-10 dargestellt ist. Eindeutig zu erkennen ist, dass die Zustandspropagation des geschlossenen Regelkreises bei etwa 1.6 s instabil wird, in dem die Zustandsvarianz sprunghaft stark zunimmt. Zu diesem Zeitpunkt befindet sich das System wiederum an der kritischen Stelle des Aufschwungmanövers mit den hohen Winkelgeschwindigkeiten. Aus dieser Perspektive betrachtet, ist die Instabilität da-

her nachvollziehbar und plausibel. Aufgrund der höheren Datenmenge ist das GP-Modell in der Lage, die Unsicherheit über die fehlenden Zentrifugalkräfte qualitativ besser auszudrücken, woraus die instabile Vorhersage resultiert. Hier sei nochmals darauf hingewiesen, dass der Riccati-Regler aufgrund seiner Herleitung bzgl. des hybriden Modells entlang der Trajektorie theoretisch stabil sein sollte. Bei dem herkömmlichen deterministischen Entwurf würde der Regelungstechniker also fälschlicherweise von einem stabilen Systemverhalten ausgehen. Die probabilistisch berechnete Instabilität kann daher nur das Resultat eines nicht vollständig bekannten realen Systems sein und ist für den Entwurfsingenieur ein wichtiges Hilfsmittel, um das Verhalten am Prüfstand vor dem Experiment besser einschätzen zu können. Dabei gibt die Simulation nicht nur Aufschluss darüber, ob das reale System instabil wird, sondern auch, zu welchem Zeitpunkt höchstwahrscheinlich mit einer Instabilität gerechnet werden kann. Tatsächlich hat sich bei verschiedenen separaten Experimenten gezeigt, dass die erwähnte kritische Stelle der Trajektorie dafür verantwortlich ist, dass es zu einem Fehlversuch am Prüfstand kommt. Auf der Basis der Prädiktionen bzgl. des offenen und geschlossenen Regelkreises hat der Ingenieur nun die Möglichkeit weiterhin nur die Optimalsteuerung zu testen oder zusätzlich die Regelung bis zum Zeitpunkt $t = 1.6$ s dazuzuschalten und das Aufschwungmanöver dort vorzeitig zu beenden und ggf. ein anschließendes sicheres Abschwungmanöver in die untere stabile Ruhelage einzuleiten. Die Stabilität des Abschwungmanövers müsste entsprechend separat überprüft werden und kann unter Umständen zu weiterführenden Problemstellungen führen. Des Weiteren herrscht bereits eine gewisse Sicherheit über die Trajektorie vor $t = 1.6$ s, sodass der Informationsgewinn aus den zugehörigen Daten eher gering wäre. Zwar könnten nach diesem Vorgehen informationsreiche Daten im Bereich der kritischen Stelle gesammelt werden, jedoch erscheint es für die Praxis sinnvoll, einen gewissen Sicherheitszeitraum, z. B. von 0.1 s, einzuführen, sodass genügend Abstand zur Instabilität besteht. Das Dilemma ist in diesem Zusammenhang, dass die wertvollsten Daten gerade innerhalb des Sicherheitszeitraums aufgenommen werden und das System somit nicht in die Nähe des informationsreichen Gebiets kommt. Schlussendlich ist hier das Expertenwissen des Regelungstechnikers entscheidend, sodass eine Abwägung immer vor dem Hintergrund des betrachteten Systems getroffen werden sollte. Aufgrund der hohen Empfindlichkeit des Doppelpendels und der noch vergleichsweise geringen Datenlage wurde in dieser Iteration entschieden, das System weiterhin nur mit einer reinen Steuerung zu betreiben. Ein weiterer Grund hierfür war, dass bei der Verwendung der Steuerung die vollen 2 s für die Zustandsmessung zur Verfügung stehen und damit innerhalb der Lerniteration mehr Daten aufgenommen werden. Die Auswertung der gemessenen Verläufe zeigt, dass das hybride Modell weiterhin keine vollständige Korrektur der Zentrifugalkräfte enthält. Sowohl der Mittelwert, als auch die hohe Standardabweichung der beiden Modellfehlerverläufe ist hierfür ein deutliches Indiz. Im direkten Vergleich zur 1. Iterati-

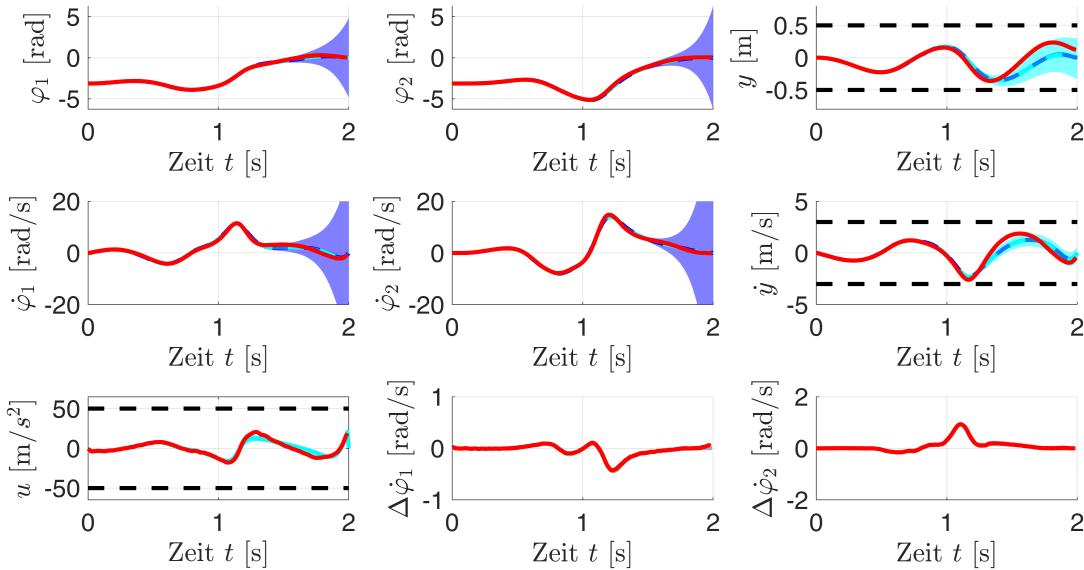


Bild 4-11: Ergebnisse der 5. bzw. letzten Lerniteration der erweiterten hybriden Optimalsteuerung am Doppelpendel auf einem Wagen. Die Prädiktion des offenen Regelkreises ist in Blau, die des geschlossenen in Türkis und die Prüfstandsmessung in Rot dargestellt. Zudem ist die Vorhersage des Modellfehlers in Grau abgebildet.

on sind die neu aufgenommenen Daten damit jedoch als wertvoller anzusehen und tragen dazu bei, die Vorhersagequalität in Hinblick auf die nächste Iteration zu erhöhen.

Die Ergebnisse in der 4. ähneln jenen aus der 1. Iteration. Ein deutlicher Unterschied ist bezüglich der kritischen Stelle zu nennen. Am Prüfstand wird diese Stelle das erste Mal erfolgreich überwunden und stimmt mit der Vorhersage überein, obwohl das System nur gesteuert wurde. Im näheren Vergleich zur Aufschwungtrajektorie aus der 3. Iteration fällt auf, dass sich die Bewegungen voneinander unterscheiden, wenn auch nur geringfügig. Die 3. Iteration kann somit der Exploration des Zustandsraums zugeordnet werden, wobei sich das Verfahren im Sinne eines lernfähigen Algorithmus eigenständig dafür entschieden hat, das ursprüngliche Aufschwungmanöver aus der 1. Iteration weiter zu fokussieren. Bild 4-11 zeigt die Ergebnisse der 5. und damit letzten Iteration des Verfahrens. In dieser Iteration wurde die Regelungsaufgabe erfüllt, in dem das Aufschwungmanöver in die obere Ruhelage am Prüfstand realisiert worden ist und der Zielzustand durch die Regelung über den Zeitpunkt von 2 s hinaus stabil eingehalten wurde. Die Zustandsprädiktion des geschlossenen Regelkreises ist durch ein schmales Varianzband um die geforderte Soll-Trajektorie gekennzeichnet. Dies zeigt auch eine komplexere MC-Simulation, welche dem Anhang A3 hinzugefügt wurde. Dementsprechend wurde das Risiko für eine Instabilität am Prüfstand als gering eingestuft. Um ein automatisiertes Vorgehen umzusetzen, wurde a-priori ein gewisser Grenzwert für die Varianz aller Zustände eingeführt, welcher bei einer Unterschreitung auf die mögliche Verwendung der Rege-

lung am Prüfstand hinweist. Bei den anderen Iterationen wurde dieser Grenzwert noch nicht unterschritten. Ausschlaggebend ist hier vor allem die Varianz der Wagenposition, welche nur innerhalb der 5. Iteration gering ist und einen gewissen Sicherheitsabstand zu den Enden der Strecke aufweist. Anhand des Stellgrößenverlaufs ist ersichtlich, dass ein Großteil des Aufschwungs auf die Steuerung zurückzuführen ist und die Regelung nur eine geringfügige aber wirksame Korrektur vornimmt. Infolgedessen kann geschlussfolgert werden, dass das hybride Modell im Rahmen des Optimalsteuerungsproblems einen qualitativ hochwertigen Ersatz für den Prüfstand darstellt. Dies zeigt sich zudem in den Verläufen der Modellfehler, welche nach dem Experiment durch das GP-Modell nahezu ideal wiedergegeben werden und eine vernachlässigbar kleine Standardabweichung beinhalten. Unter anderem die Zentrifugalkräfte sind damit durch die Messdaten vom GP zufriedenstellend abgebildet worden. Abschließend ist die Inbetriebnahme am Doppelpendel auf einem Wagen damit als erfolgreich abgeschlossen zu betrachten.

Fazit zu der Validierung am Doppelpendelprüfstand und kritische Analyse

Die hybride Optimalsteuerung mit dem erweiterten Regelungsentwurf stellt ein hilfreiches Werkzeug für die Inbetriebnahme von anspruchsvollen Systemen dar, welche eine exakte mathematische Beschreibung der Dynamik benötigen. Das Doppelpendel auf einem Wagen ist in diesem Zusammenhang gesondert hervorzuheben, denn es zeichnet sich durch ein empfindliches chaotisches Bewegungsverhalten aus, wodurch eine realistische Prädiktion nur auf der Basis eines genauen Modells berechnet werden kann. Kleinste Fehler innerhalb der zugrundeliegenden Gleichungen haben dementsprechend einen großen Effekt auf den Prozess der Inbetriebnahme und sind daher entscheidend für den Erfolg und Misserfolg des regelungstechnischen Entwurfs. Die entwickelte hybride Optimalsteuerung setzt bei dieser Problemstellung an und zeichnet sich durch eine hohe Relevanz für die Praxis aus. Die Validierung am Doppelpendel hat gezeigt, dass die Einführung einer wahrscheinlichkeitsbasierten Sichtweise die Inbetriebnahme für den Ingenieur besser einschätzbar macht. Hierzu ist insbesondere der neu eingeführte Begriff der probabilistischen Stabilität zu nennen, der den herkömmlichen Stabilitätsbegriff in Bezug auf ein funktionierendes Regelungskonzept erweitert und einen direkten Bezug zum realen System herstellt. Über die Zustandsvarianz lässt sich das Verhalten am Prüfstand a-priori deutlich besser abschätzen, sodass entsprechende Sicherheitsmaßnahmen vor dem Experiment vorgenommen werden können. Ein weiterer Vorteil ist die Einbeziehung von Vorwissen in Form einer partiell bekannten Dynamik und einer nachvollziehbaren bzw. interpretierbaren Regelungsstruktur. Bei dem rein datengetriebenen PILCO-Verfahren wurden hierfür keine Schnittstellen vorgesehen, wohingegen die erweiterte hybride Optimalsteuerung solche Schnittstellen besitzt. Im Endergebnis wurden mit PILCO dementsprechend weitaus mehr Iterationen für den erfolgreichen Aufschwung benötigt. Insgesamt ergibt sich durch die Einbringung des Vorwissens damit eine Einsparung von knapp über 80%

bzgl. der erforderlichen Experimente am Prüfstand. Dies verdeutlicht bereits die enorme Effizienzsteigerung, welche durch die hybride Optimalsteuerung erzielt werden kann. Ein besonderes Potenzial wird hier vor allem bei Systemen gesehen, für welche die Auswertung eines Experiments aufwendiger und kostenintensiver ausfällt. Abseits der Zeit für ein Experiment, ist auch die Zeit für die Berechnung des nächsten Experiments zu beurteilen. Da hierzu keine Echtzeitanforderungen eingehalten werden müssen bzw. die Berechnungen offline durchgeführt werden, ist die benötigte Zeit nur ein sekundäres Kriterium. Nichtsdestotrotz wurde bei PILCO für eine Iteration in etwa eine Berechnungszeit von 2 Stunden benötigt, welche ein Resultat des angesetzten hoch parametrisierten GP-Reglers ist und auf die Tatsache zurückzuführen ist, dass keine Kenntnis über eine Regelungsstruktur vorgegeben wurde. Eine vergleichbare Iteration der erweiterten hybriden Optimalsteuerung dauerte lediglich 10 Minuten, wobei diese Zeitangabe nicht nur die Lösung des Optimierungsproblems umfasst, sondern auch die Zustandspropagation des geschlossenen Regelkreises. In dieser Metrik liegt der Zeitgewinn bei etwa 98% und ist damit besonders bzgl. der Einsparung von Ressourcen hervorzuheben.

Ein Kritikpunkt des entwickelten Verfahrens ist in Bezug auf die Skalierbarkeit auf größere Datenmengen, welche unter Umständen für Systeme mit vielen Zustands- und Stellgrößendimensionen erforderlich sind, zu nennen. Das Training eines GPs, also die Bestimmung der Hyperparameter, benötigt die Invertierung der Datenmatrix $\mathbf{K}(\eta) \in \mathbb{R}^{n_d \times n_d}$ (vgl. (3-5)), wobei n_d die Anzahl der Daten ist, in jedem Optimierungsschritt. Die Berechnungskomplexität des Trainings beträgt damit $\mathcal{O}(n_d^3)$. Des Weiteren skaliert eine anschließende Auswertung der GP-Posterior-Gleichungen (4-17) mit $\mathcal{O}(n_d^2)$. Eine praktische Anwendbarkeit für große Datenmengen ist damit in Frage gestellt. Eine anwendungsorientierte Lösungsmöglichkeit besteht darin, die Datenmenge durch *Clustering*-Methoden geeignet zu reduzieren und damit einen Genauigkeitsverlust zu akzeptieren, wobei jedoch das grundsätzliche Problem der schlechten Skalierbarkeit bestehen bleibt. Vor diesem Hintergrund ist ein besserer Lösungsweg ein probabilistisches ML-Verfahren mit besseren Skalierungseigenschaften zu verwenden. Im Ausblick dieser Arbeit (Kap. 6) werden hierzu diverse Kandidaten genannt. Unter der Voraussetzung einer normalverteilten Prädiktion eines anderen ML-Verfahrens für den datengetriebenen Modellteil behalten die Gleichungen der hybriden Optimalsteuerung ihre Gültigkeit. Zudem bleibt das grundsätzliche Vorgehen in Bezug auf Algorithmus 5 unverändert und die bei der Inbetriebnahme besprochenen Zusammenhänge bleiben bestehen. Ein weiterer Kritikpunkt betrifft Systeme, die abrupte Änderungen ihrer Dynamik aufweisen, wie sie beispielsweise bei Systemen mit Haftreibung oder hartem Kontakt beobachtet werden können. Ohne eine entsprechende Anpassung der Gauß-Prozess-Regression kann es bei diesen Systemen zu einem Versagen des Verfahrens kommen. Eine beispielhafte Darstellung dieser Problematik wurde im Grundlagenkapitel unter Abschnitt 2.5.1 erläutert, wobei die hohen

Glattheitsanforderungen der SE-Kovarianzfunktion zu einer fehlerhaften Abbildung der gesuchten Sprungfunktion geführt haben. Dementsprechend ist in diesem Fall eine Anpassung der Kernelfunktion (z. B. stückweise) aufgrund des Expertenwissens des Regelungstechnikers notwendig. Im Gegensatz zu vergleichbaren Veröffentlichungen (s. Stand der Forschung) sind durch die Verwendung der UT für die Zustandspropagation beliebige Kernelfunktion einsetzbar, womit die hybride Optimalsteuerung auch für die angesprochene Systemklasse anwendbar bleibt. Eine Voraussetzung ist jedoch, dass anstelle des SQP-Verfahrens, welches auf kontinuierlichen Funktionen basiert, ein Solver aus dem Bereich der nicht glatten Optimierung [GK02] zur Lösung eingesetzt wird.

5 Hybride Zustandslinearisierung für mechanische Mehrkörpersysteme

Der Regelungsentwurf nach dem Verfahren der Zustandslinearisierung ist ein gängiger Regelungstechnischer Ansatz, welcher nur unter besonderen Voraussetzungen angewendet werden kann [Ada14]. Die Methode zielt darauf ab, ein nichtlineares System durch eine geeignete Wahl der Stellgröße, als ein lineares System zu behandeln. Hierbei werden die vorhandenen Nichtlinearitäten im Idealfall vollständig kompensiert - eine Linearisierung durch eine Taylorreihenentwicklung findet dabei nicht statt, sodass das System im gesamten Zustandsraum als linear betrachtet werden darf. Dies hat den großen Vorteil, dass die Methoden der linearen Regelungstechnik zum Entwurf und Analyse angewandt werden dürfen und entsprechend ein linearer Zustandsregler zur Stabilisierung und Sollwertfolge ausreichend ist. Damit die Methodik anwendbar ist, muss die Voraussetzung gelten, dass es sich um ein eingangsaffines System handelt, d. h. dass die Stellgröße linear in die Systemgleichungen eingeht. Diese Voraussetzung ist für viele technische Systeme erfüllt, da physikalische Gesetzmäßigkeiten häufig auf einem solchen Zusammenhang beruhen. Um die Kompensation der Nichtlinearitäten zu erreichen, ist die passende Wahl der Stellgröße am realen System erforderlich. Die Stellgröße hängt hierbei von bestimmten Modelltermen ab, welche die Nichtlinearitäten des realen Systems genau genug wiedergeben müssen. Entsprechend der Grundidee, werden die Terme über die Stellgröße dem realen System so aufgeschaltet, sodass sie invers zur realen Systemdynamik wirken und sich bei einer gesamtheitlichen Betrachtung möglichst ideal kompensieren. Weist das Modell dahingehend Ungenauigkeiten auf, so wirken sich diese über das aufgestellte Regelgesetz direkt auf die Regelgüte aus. Unter Umständen werden durch das ungenaue Modell zusätzliche unerwünschte Nichtlinearitäten eingebracht, welche negative Effekte, wie Grenzzyklen oder Instabilitäten, nach sich ziehen können. Das Ziel dieses Kapitels ist, eine hybride Zustandslinearisierung zu entwickeln, welche die angesprochenen Ungenauigkeiten auf der Basis von Messdaten und der Gauß-Prozess-Regression korrigiert. In diesem Szenario ist ein teilweise unvollständiges physikalisches Modell zulässig, auf dessen Grundlage eine ideale Kompensation nicht funktioniert. Der Fokus dieses Kapitels liegt auf mechanischen Systemen, da diese die Forderung nach einer linear wirkenden Stellgröße erfüllen und eine breite Klasse innerhalb der Mechatronik darstellen. Des Weiteren wird die hybride Zustandslinearisierung anhand eines Hexapoden beispielhaft erprobt. Zur besseren Einordnung und Übersicht wird die klassische Zustandslinearisierung im nächsten Abschnitt zunächst mathematisch näher formalisiert.

5.1 Klassischer Regelungsentwurf durch Zustandslinearisierung

Um das Grundprinzip nachvollziehen zu können, ist die Betrachtung eines eindimensionalen Systems mit dem Zustand $x(t) \in \mathbb{R}$ und der Stellgröße $u(t) \in \mathbb{R}$ zunächst ausreichend. In den nachfolgenden Abschnitten erfolgt die Verallgemeinerung auf mehrdimensionale Systeme. Das betrachtete eingangsaffine System in zeitkontinuierlicher Form lautet

$$\dot{x} = \alpha(x) + \beta(x) u, \quad (5-1)$$

worin $\alpha(x)$ und $\beta(x)$ die erwähnten separierten Nichtlinearitäten darstellen, welche lediglich von der Zustandsgröße abhängen. Der Grundidee folgend, wird das Regelungsgesetz

$$u_R(x, w) = \frac{1}{\hat{\beta}(x)}(-\hat{\alpha}(x) + K(w - x)),$$

so angesetzt, dass sich die Nichtlinearitäten des realen Systems $(\alpha(x), \beta(x))$ und der Stellgröße $(\hat{\alpha}(x), \hat{\beta}(x))$ bzw. des Modells gegenseitig aufheben. Bei einer Gleichsetzung von $u = u_R(w, x)$ ergibt sich somit für das geregelte Gesamtsystem eine Dynamik von

$$\dot{x} = \alpha(x) - \frac{\beta(x)}{\hat{\beta}(x)}\hat{\alpha}(x) + \frac{\beta(x)}{\hat{\beta}(x)}K(w - x), \quad (5-2)$$

wobei unter der Voraussetzung eines genauen Modells $\hat{\alpha}(x) \rightarrow \alpha(x)$ und $\beta(x)/\hat{\beta}(x) = 1$ gelten würde, was dem Idealfall einer vollständigen Kompensation entspricht. Für die nachgelagerte Regelung wurde beispielhaft ein linearer Regler mit der Verstärkung K und dem Sollwert w angesetzt. Dieser bleibt nach der Konstruktion des Verfahrens innerhalb der Differentialgleichung bestehen und sorgt für die Stabilisierung bzw. Sollwerteinhaltung. Anhand von (5-2) wird zudem ersichtlich, dass selbst kleinste Ungenauigkeiten des Modells dazu führen können, dass eine Kompensation nicht vollständig stattfindet und unter Umständen unerwünschte nichtlineare Effekte in das Gesamtsystem eingebracht werden. Im Nachfolgenden soll davon ausgegangen werden, dass eben solche Ungenauigkeiten vorhanden sind, sodass der Einsatz der hybriden Zustandslinearisierung motiviert wird. Zunächst wird die hybride Zustandslinearisierung ähnlich zur hybriden Optimalsteuerung formal entwickelt und anschließend an dem ausgewählten Anwendungsbeispiel des Hexapoden simulativ erprobt.

5.2 Herleitung der hybriden Zustandslinearisierung

Den Ausgangspunkt der Herleitung stellt ein klassisches mechanisches System mit dem Positionsvektor $\mathbf{q}(t) \in \mathbb{R}^{n_q}$ dar, welches in seiner allgemeinen Darstellung über

$$\begin{aligned} \mathbf{M}(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{C}(\dot{\mathbf{q}}, \mathbf{q})\dot{\mathbf{q}} + \mathbf{G}(\mathbf{q}) &= \mathbf{J}(\mathbf{q})^{-T}\mathbf{u} \\ \Leftrightarrow \ddot{\mathbf{q}} &= \underbrace{\mathbf{M}(\mathbf{q})^{-1}\mathbf{J}(\mathbf{q})^{-T}\mathbf{u}}_{=: \beta(\mathbf{q})} - \underbrace{\mathbf{M}(\mathbf{q})^{-1}(\mathbf{C}(\dot{\mathbf{q}}, \mathbf{q})\dot{\mathbf{q}} + \mathbf{G}(\mathbf{q}))}_{=: \alpha(\mathbf{q}, \dot{\mathbf{q}})} \end{aligned} \quad (5-3)$$

beschrieben wird. Hierin sind $\mathbf{M}(\mathbf{q}) \in \mathbb{R}^{n_q \times n_q}$ die Massenmatrix, $\mathbf{C}(\mathbf{q}, \dot{\mathbf{q}}) \in \mathbb{R}^{n_q \times n_q}$ die Matrix der Zentrifugal- und Corioliskräfte, $\mathbf{G}(\mathbf{q}) \in \mathbb{R}^{n_q}$ der Vektor der Gewichtskräfte und $\mathbf{J}(\mathbf{q}) \in \mathbb{R}^{n_q \times n_u}$ die geometrische Jacobimatrix, welche die kinematische Wirkung der Stellgrößen $\mathbf{u} \in \mathbb{R}^{n_u}$ auf die Beschleunigungsgrößen $\ddot{\mathbf{q}}$ definiert. Anhand der Äquivalenzumformung in (5-3) wird ersichtlich, dass das mechanische System in die gewünschte Form für eine Zustandslinearisierung gebracht werden kann. Hierbei handelt es sich im Vergleich zu (5-1) jedoch um den mehrdimensionalen Fall mit den Nichtlinearitäten $\alpha(\mathbf{q}, \dot{\mathbf{q}}) : \mathbb{R}^{n_q} \times \mathbb{R}^{n_q} \mapsto \mathbb{R}^{n_q}$ und $\beta(\mathbf{q}) : \mathbb{R}^{n_q} \mapsto \mathbb{R}^{n_q \times n_u}$. Für den Übergang zur Zustandsraumdarstellung wird der Zustandsvektor über $\mathbf{x} = [\mathbf{q}^T, \dot{\mathbf{q}}^T]^T \in \mathbb{R}^{2n_q}$ eingeführt, sodass sich insgesamt

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}) = \begin{bmatrix} \dot{\mathbf{q}} \\ \alpha(\mathbf{x}) + \beta(\mathbf{x})\mathbf{u} \end{bmatrix}, \quad (5-4)$$

ergibt, wobei zur Verallgemeinerung eine zusätzliche Abhängigkeit zwischen β und $\dot{\mathbf{q}}$ angenommen wurde (vgl. (5-3)). Die ersten n_q Gleichungen repräsentieren ein Integratorverhalten bezüglich \mathbf{q} und $\dot{\mathbf{q}}$. Das Regelungsgesetz wird nun entsprechend zu (5-2) über

$$\mathbf{u}_R(\mathbf{x}, \mathbf{w}) = \beta(\mathbf{x})^{-1}(-\alpha(\mathbf{x}) + \mathbf{K}(\mathbf{w} - \mathbf{x})),$$

beschrieben, wobei für den Fall von $n_q \neq n_u$ bzw. eines nicht voll-aktuierten Systems die Pseudoinverse von $\beta(\mathbf{x})$ benötigt wird. Eine separate Aufstellung der Differentialgleichungen aus (5-3)

$$\begin{aligned} \ddot{q}_1 &= \beta_{1,1}(\mathbf{q})u_1 + \dots + \beta_{1,n_u}(\mathbf{q})u_{n_u} + \alpha_1(\mathbf{q}, \dot{\mathbf{q}}), \\ &\vdots \\ \ddot{q}_{n_q} &= \beta_{n_q,1}(\mathbf{q})u_1 + \dots + \beta_{n_q,n_u}(\mathbf{q})u_{n_u} + \alpha_{n_q}(\mathbf{q}, \dot{\mathbf{q}}), \end{aligned} \quad (5-5)$$

zeigt, dass es sich im Kontext des maschinellen Lernens um n_q unabhängige Regressionsprobleme bzgl. der Beschleunigungen $\ddot{q}_i, i = 1, \dots, n_q$ handelt. Dabei gilt es, die einzelnen Funktionen $\alpha_i, \beta_{i,j}, i = 1, \dots, n_q, j = 1, \dots, n_u$ bzw. deren Abweichung zu den bereits bekannten Funktionen des bestehenden physikalischen Modells zu identifizieren. Eine Herausforderung bei der Lösung der Regressionsprobleme (5-5) wird bei der Be- trachtung der zugehörigen Messdaten $(\dot{\mathbf{x}}, \mathbf{x}, \mathbf{u})$, welche am realen System aufgezeichnet werden, ersichtlich. Basierend auf den Messdaten sollen die einzelnen nichtlinearen Funktionen $\alpha_i(\mathbf{x}), \beta_{i,j}(\mathbf{x})$ identifiziert werden, wobei jedoch unklar ist wie sich die Beschleu- nigungen \ddot{q}_i geeignet zerlegen lassen. Die gesuchten Funktionen $\alpha_i(\mathbf{x}), \beta_{i,j}(\mathbf{x})$ können nicht ohne Weiteres direkt gemessen werden, sondern sind additiv im Messsignal von $\ddot{\mathbf{q}}$ enthalten und müssen entsprechend (5-5) separiert werden.

Die Autoren aus [UBKH17; UH20] stellen hierzu ein Konzept auf der Basis der Gauß- Prozess-Regression vor. Hierfür nutzen sie zwei verschiedene Eigenschaften von GPs aus. Zum einen führt die Summe von mehreren GPs wiederum auf einen GP und zum anderen führt das Produkt eines GPs mit einer bekannten Funktion gleichermaßen auf einen GP. In diesem Zusammenhang werden für den eindimensionalen Fall (vgl. (5-1)) für α und β zwei separate GPs angesetzt, welche über die erwähnten Eigenschaften miteinander ver- bunden werden können und über die eine Aufspaltung der Messungen zugänglich wird. Die Grundidee aus [UBKH17; UH20] wird im Rahmen dieser Arbeit aufgegriffen und im Kontext der Themenstellung erweitert. Im nachfolgenden Abschnitt wird dazu das Kon- zept auf den mehrdimensionalen Fall übertragen, sodass es auf das betrachtete Anwen- dungsbeispiel angewendet werden kann. Im Gegensatz zu der rein theoretischen Arbeit in [UBKH17; UH20] wird damit die praktische Anwendbarkeit der hybriden Zustandsli- nearisierung hervorgehoben. Zudem werden die Mittelwertfunktionen der einzelnen GPs nicht zu null angenommen, sondern bilden die Schnittstellen zu dem vorhanden dynami- schen Vorwissen, welche sich aus den Überlegungen zu (5-3) berücksichtigen lässt. Damit wird sowohl die Effizienz in Hinblick auf ein iteratives Vorgehen ähnlich zu PILCO und der hybriden Optimalsteuerung erhöht, als auch die Extrapolationseigenschaften im Ver- gleich zu einem rein datengetriebenen Ansatz verbessert. Des Weiteren wird in dieser Arbeit die Auswirkung der probabilistischen GPs auf das Regelgesetz und die Stellgröße ausformuliert. Erwartungsgemäß lassen sich so während des Betriebs des Systems mit dem entwickelten Regelungskonzept Aussagen darüber treffen, wie zuverlässig der Reg- ler in Bezug zu seiner Stellgröße und vor dem Hintergrund von möglicherweise vorhan- denen Modellfehlern ist. Weiterführend lassen sich somit Handlungsanweisungen für die Aufnahme zusätzlicher Daten ableiten und mögliche Instabilitäten, die auf die Regelung zurückzuführen sind, frühzeitig erkennen (ähnlich zum Fall in Bild 4-10).

Da für alle n_q Differentialgleichungen in (5-5) die gleiche Gauß-Prozess-Regression angesetzt wird, wird im Folgenden lediglich die i -te Gleichung mit $i \geq n_q + 1$ aus der Zustandsraumdarstellung (5-4)

$$\dot{x} = \beta_1(\mathbf{x})u_1 + \dots + \beta_{n_u}(\mathbf{x})u_{n_u} + \alpha(\mathbf{x}),$$

betrachtet. Zur besseren Übersicht wird der Index i weggelassen. Den Ausgangspunkt des Verfahrens bilden die n_d aufgenommenen bzw. berechneten Datenpaare $(\dot{\mathbf{x}}, \mathbf{x}, \mathbf{u})$, welche mit dem standardmäßigen Ansatz für das Messrauschen $\epsilon_n \sim \mathcal{N}(0, \sigma_n^2)$ additiv verfälscht sind und in den nachfolgenden Matrizen und dem Ausgangsdatenvektor zusammengefasst werden:

$$\begin{aligned}\mathbf{X} &= [\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n_d)}] \in \mathbb{R}^{n_x \times n_d}, \\ \mathbf{U}_j &= \text{diag}(u_j^{(1)}, \dots, u_j^{(n_d)}) \in \mathbb{R}^{n_d \times n_d}, \quad j = 1, \dots, n_u, \\ \mathbf{y} &= [\dot{x}^{(1)} + \epsilon_n^{(1)}, \dots, \dot{x}^{(n_d)} + \epsilon_n^{(n_d)}]^T \in \mathbb{R}^{n_d}.\end{aligned}$$

Der Grundidee folgend, werden die gesuchten Funktionen für die Zustandslinearisierung als Zufallsvariablen, die von einem jeweiligen GP stammen, über

$$\begin{aligned}\alpha(\mathbf{x}) &\sim \mathcal{GP}(m_\alpha(\mathbf{x}), k_\alpha(\mathbf{x}, \mathbf{x}')), \\ \beta_j(\mathbf{x}) &\sim \mathcal{GP}(m_{\beta_j}(\mathbf{x}), k_{\beta_j}(\mathbf{x}, \mathbf{x}')), \quad j = 1, \dots, n_u,\end{aligned}\tag{5-6}$$

definiert. Ein möglicherweise vorhandenes Vorwissen kann hierbei über die Mittelfunktions $m_\alpha(\mathbf{x})$ und $m_{\beta_j}(\mathbf{x})$ berücksichtigt werden. Für die Kovarianzfunktionen wird standardmäßig der SE-Kernel angesetzt, sofern über die Gestalt der Dynamikfunktionen keine anderweitigen Informationen vorliegen. Es wird nun ein Zustand \mathbf{x}^* eingeführt, an welchem die Funktionen ausgewertet werden sollen. Zusammen mit den Ausgangsdaten, lässt sich durch die Definition als GPs für die Prior-Verteilung folgern, dass

$$\begin{bmatrix} \alpha(\mathbf{x}^*) \\ \beta_1(\mathbf{x}^*) \\ \vdots \\ \beta_{n_u}(\mathbf{x}^*) \\ \mathbf{y} \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} m_\alpha(\mathbf{x}^*) \\ m_{\beta_1}(\mathbf{x}^*) \\ \vdots \\ m_{\beta_{n_u}}(\mathbf{x}^*) \\ \mathbf{m}_y \end{bmatrix}, \begin{bmatrix} k_\alpha^* & 0 & 0 & \dots & \mathbf{k}_\alpha^T \\ 0 & k_{\beta_1}^* & 0 & \dots & \mathbf{k}_{\beta_1}^T \mathbf{U}_1^T \\ \vdots & \vdots & \ddots & \dots & \vdots \\ 0 & 0 & \dots & k_{\beta_{n_u}}^* & \mathbf{k}_{\beta_{n_u}}^T \mathbf{U}_{n_u}^T \\ \mathbf{k}_\alpha & \mathbf{U}_1 \mathbf{k}_{\beta_1} & \dots & \mathbf{U}_{n_u} \mathbf{k}_{\beta_{n_u}} & \mathbf{K}_y \end{bmatrix} \right)$$

gilt, mit $k_{(\cdot)}^*(\mathbf{x}^*) = k_{(\cdot)}(\mathbf{x}^*, \mathbf{x}^*)$, $\mathbf{k}_{(\cdot)}(\mathbf{x}^*) = [k_{(\cdot)}(\mathbf{x}^*, \mathbf{x}^{(1)}), \dots, k_{(\cdot)}(\mathbf{x}^*, \mathbf{x}^{(n_d)})]^T \in \mathbb{R}^{n_d}$ und wobei die erwähnten Regeln zur Addition und Produktbildung von GPs angewandt

wurden. Dadurch ergeben sich insbesondere für die Ausgangsdaten der Erwartungsvektor und die Kovarianzmatrix zu

$$\begin{aligned}\mathbf{m}_y &= \mathbf{m}_\alpha(\mathbf{X}) + \mathbf{U}_1 \mathbf{m}_{\beta_1}(\mathbf{X}) + \dots + \mathbf{U}_{n_u} \mathbf{m}_{\beta_{n_u}}(\mathbf{X}), \quad \text{und} \\ \mathbf{K}_y &= \mathbf{K}_\alpha + \mathbf{U}_1^T \mathbf{K}_{\beta_1} \mathbf{U}_1 + \dots + \mathbf{U}_{n_u}^T \mathbf{K}_{\beta_{n_u}} \mathbf{U}_{n_u} + \sigma_n^2 \mathbf{I}_{n_d},\end{aligned}$$

mit $\mathbf{m}_{(\cdot)}(\mathbf{X}) = [m_{(\cdot)}(\mathbf{x}^{(1)}), \dots, m_{(\cdot)}(\mathbf{x}^{(n_d)})]^T \in \mathbb{R}^{n_d}$ und $\mathbf{K}_{(\cdot)} \in \mathbb{R}^{n_d \times n_d}$ mit den Elementen $K_{(\cdot),i,j} = k_{(\cdot)}(\mathbf{x}^{(i)}, \mathbf{x}^{(j)}), i, j = 1, \dots, n_d$. Es handelt sich hierbei nicht um die Verstärkungsmatrix des Reglers. Auf der Grundlage der Prior-Verteilung werden die bedingten bzw. Posterior-Verteilungen nach

$$\begin{aligned}\alpha(\mathbf{x}^*) \mid \mathbf{y} &\sim \mathcal{N}(\mu_\alpha(\mathbf{x}^*), \sigma_\alpha^2(\mathbf{x}^*)), \\ \mu_\alpha(\mathbf{x}^*) &= m_\alpha(\mathbf{x}^*) + \mathbf{k}_\alpha^T(\mathbf{x}^*) \mathbf{K}_y^{-1}(\mathbf{y} - \mathbf{m}_y), \\ \sigma_\alpha^2(\mathbf{x}^*) &= k_\alpha^*(\mathbf{x}^*) - \mathbf{k}_\alpha^T(\mathbf{x}^*) \mathbf{K}_y^{-1} \mathbf{k}_\alpha(\mathbf{x}^*),\end{aligned}$$

$$\begin{aligned}\beta_j(\mathbf{x}^*) \mid \mathbf{y} &\sim \mathcal{N}(\mu_{\beta_j}(\mathbf{x}^*), \sigma_{\beta_j}^2(\mathbf{x}^*)), \quad j = 1, \dots, n_u, \\ \mu_{\beta_j}(\mathbf{x}^*) &= m_{\beta_j}(\mathbf{x}^*) + \mathbf{k}_{\beta_j}^T(\mathbf{x}^*) \mathbf{U}_j^T \mathbf{K}_y^{-1}(\mathbf{y} - \mathbf{m}_y), \\ \sigma_{\beta_j}^2(\mathbf{x}^*) &= k_{\beta_j}^*(\mathbf{x}^*) - \mathbf{k}_{\beta_j}^T(\mathbf{x}^*) \mathbf{U}_j^T \mathbf{K}_y^{-1} \mathbf{U}_j \mathbf{k}_{\beta_j}(\mathbf{x}^*),\end{aligned}$$

bestimmt. Es liegt somit eine probabilistische Schätzung in Form einer eindimensionalen Normalverteilung für jede der gesuchten Funktionen der Zustandslinearisierung vor. Das Training der GPs wird durch die Identifikation der Hyperparameter vorgenommen. Beispielsweise wird für alle GPs der SE-Kernel als Kovarianzfunktion mit

$$\begin{aligned}k_\alpha(\mathbf{x}, \mathbf{x}'; \boldsymbol{\eta}) &= \sigma_\alpha^2 \exp \left(\sum_{i=1}^{n_x} \frac{(x_i - x'_i)^2}{-2l_{\alpha,i}^2} \right), \\ k_{\beta_j}(\mathbf{x}, \mathbf{x}'; \boldsymbol{\eta}) &= \sigma_{\beta_j}^2 \exp \left(\sum_{i=1}^{n_x} \frac{(x_i - x'_i)^2}{-2l_{\beta_j,i}^2} \right), \quad j = 1, \dots, n_u,\end{aligned}$$

angesetzt. Der Hyperparametervektor umfasst dann jeden einzelnen GP aus (5-6) über

$$\begin{aligned}\boldsymbol{\eta} &= [l_{\alpha,1}, l_{\beta_1,1}, \dots, l_{\beta_{n_u},1}, \dots, l_{\alpha,n_x}, l_{\beta_1,n_x}, \dots, l_{\beta_{n_u},n_x}, \\ &\quad \sigma_\alpha^2, \sigma_{\beta_1}^2, \dots, \sigma_{\beta_{n_u}}^2, \sigma_n^2]^T \in \mathbb{R}^{(n_x+1)(n_u+1)+1}\end{aligned}$$

und wird standardmäßig über die Maximierung der logarithmischen Likelihoodfunktion

$$\begin{aligned}\boldsymbol{\eta}^* &= \arg \max_{\boldsymbol{\eta}} \log p(\mathbf{y} \mid \boldsymbol{\eta}) \\ &= \arg \max_{\boldsymbol{\eta}} -\frac{1}{2} (\mathbf{y} - \mathbf{m}_y)^T \mathbf{K}_y^{-1} (\mathbf{y} - \mathbf{m}_y) - \frac{1}{2} \log |\mathbf{K}_y| - \frac{n_d}{2} \log 2\pi\end{aligned}$$

eindeutig bestimmt (vgl. (2-33)).

Die beschriebene Methode wird für alle n_q Differentialgleichungen (5-5) gleichermaßen durchgeführt, sodass eine Zusammenführung aller Posterior-Mittelwerte kurz über

$$\boldsymbol{\mu}_\alpha(\mathbf{x}) = \begin{bmatrix} \mu_{\alpha_1}(\mathbf{x}) \\ \vdots \\ \mu_{\alpha_{n_q}}(\mathbf{x}) \end{bmatrix} \in \mathbb{R}^{n_q} \quad \boldsymbol{\mu}_\beta(\mathbf{x}) = \begin{bmatrix} \mu_{\beta_{1,1}}(\mathbf{x}) & \cdots & \mu_{\beta_{1,n_u}}(\mathbf{x}) \\ \vdots & \ddots & \vdots \\ \mu_{\beta_{n_q,1}}(\mathbf{x}) & \cdots & \mu_{\beta_{n_q,n_u}}(\mathbf{x}) \end{bmatrix} \in \mathbb{R}^{n_q \times n_u},$$

beschrieben wird. Diese Größen stellen, auf der Basis des Vorwissens und der Datenlage, die beste Schätzung für die gesuchten Funktionen dar. Aus diesem Grund wird hinsichtlich der Anwendung am realen System das Regelungsgesetz über diese Schätzungen mit

$$\mathbf{u}_R(\mathbf{x}, \mathbf{w}) = \boldsymbol{\mu}_\beta(\mathbf{x})^{-1}(-\boldsymbol{\mu}_\alpha(\mathbf{x}) + \mathbf{K}(\mathbf{w} - \mathbf{x})) =: \boldsymbol{\mu}_u(\mathbf{x}) \quad (5-7)$$

definiert, sodass zu jedem Zeitpunkt eine eindeutige Stellgröße bekannt ist. Die Bildung der (Pseudo-)Inversen kann unter Umständen zu Problemen bei der Berechnung der Stellgröße führen. Im Zusammenhang mit der Gauß-Prozess-Regression wurde kein Mechanismus dafür vorgesehen, dass die Matrix $\boldsymbol{\mu}_\beta(\mathbf{x})$ für beliebige \mathbf{x} gut gestellt bzw. konditioniert ist. Eine mögliche Singularität kann daher nicht ausgeschlossen werden. Bei den bisherigen Erprobungen des entwickelten Verfahrens konnten jedoch keine Probleme in Bezug auf das Invertieren der Matrix festgestellt werden. Ein möglicher Grund hierfür könnte sein, dass in den überwiegenden Fällen ein bestimmtes Vorwissen über $m_\beta(\mathbf{x})$ berücksichtigt wurde, sodass sich im Vergleich zu einem rein datengetriebenen Ansatz eine gute Konditionierung der Matrix tendenziell eher einstellt. Im Rahmen dieser Arbeit wurde abseits der experimentellen Erprobung die Konditionierung nicht weiter untersucht, sodass die Fragestellung als Ausblick für weiterführende Forschungsarbeiten offen bleibt.

Vor dem Hintergrund von (5-7) ist ersichtlich, dass die Varianzangabe bzgl. der einzelnen Funktionen nicht im Regelgesetz enthalten ist und als weitere Informationsquelle bisher nicht genutzt wird. Da es sich bei $\alpha(\mathbf{x})$ und $\beta(\mathbf{x})$, nach der hier eingeführten Definition, um Zufallsvariablen handelt, induzieren sie eine Unsicherheit in die Stellgröße, wodurch diese ebenfalls als Zufallsvariable zu betrachten ist. Ähnlich zum MM-Ansatz im Bezug auf die Zustandsverteilung, kann eine Approximation der Stellgrößenverteilung mit $p(\mathbf{u}) \approx \mathcal{N}(\boldsymbol{\mu}_u, \sigma_u^2)$ eingeführt werden. Dies ermöglicht tiefere Einsichten in den Regler, sodass sein Verhalten besser abgeschätzt werden kann. Für die Praxis und den Betrieb des realen Systems bedeutet dies, dass weiterhin das Regelgesetz aus (5-7) eingesetzt wird, jedoch zur Laufzeit parallel die Varianz der Stellgröße σ_u^2 ausgewertet wird, ähnlich zu

dem Verhalten eines Zustandsbeobachters. Die Varianz σ_u^2 bildet in diesem Zusammenhang die Unsicherheit des Reglers aufgrund der (noch) vorhandenen Modellfehler ab und kann als Maß dafür benutzt werden, um einzuschätzen ob ein sicheres Systemverhalten zu erwarten ist oder im Bereich des aktuellen Systemzustands \mathbf{x} weitere Messdaten aufgenommen werden sollten.

Im eindimensionalen Fall $\alpha(\mathbf{x}), \beta(\mathbf{x}) \in \mathbb{R}$ lässt sich die Varianz, ähnlich zu (3-13), über

$$\sigma_u^2(\mathbf{x}) = \frac{1}{\mu_\beta^2(\mathbf{x})} \left(\sigma_\alpha^2(\mathbf{x}) + \frac{(-\mu_\alpha(\mathbf{x}) + \mathbf{K}(\mathbf{w} - \mathbf{x}))^2 \sigma_\beta^2(\mathbf{x})}{\mu_\beta^2(\mathbf{x})} \right)$$

abschätzen, wobei wiederum die Bedingung $\sigma_\beta/\mu_\beta \leq 0.1$ zu berücksichtigen ist, damit die Approximation ihre Gültigkeit behält. Für den überwiegenden höherdimensionalen Fall, erscheint ein MC-Ansatz als geeignet. Dabei werden n_s Realisierungen der Stellgröße $\mathbf{u}_s^{(i)}, i = 1, \dots, n_s$ erzeugt, welche dem Zusammenhang

$$\mathbf{u}_s^{(i)} \sim (\boldsymbol{\mu}_\beta(\mathbf{x}) + \boldsymbol{\sigma}_\beta(\mathbf{x}) \circ \boldsymbol{\epsilon}_{s,\beta})^{-1} (-\boldsymbol{\mu}_\alpha(\mathbf{x}) - \boldsymbol{\sigma}_\alpha(\mathbf{x}) \circ \boldsymbol{\epsilon}_{s,\alpha} + \mathbf{K}(\mathbf{w} - \mathbf{x})),$$

folgen, wobei \circ die elementweise Multiplikation darstellt. Des Weiteren sind die Matrizen der Standardabweichungen

$$\boldsymbol{\sigma}_\alpha(\mathbf{x}) = \begin{bmatrix} \sigma_{\alpha_1}(\mathbf{x}) \\ \vdots \\ \sigma_{\alpha_{n_q}}(\mathbf{x}) \end{bmatrix} \in \mathbb{R}^{n_q}, \quad \boldsymbol{\sigma}_\beta(\mathbf{x}) = \begin{bmatrix} \sigma_{\beta_{1,1}}(\mathbf{x}) & \cdots & \sigma_{\beta_{1,n_u}}(\mathbf{x}) \\ \vdots & \ddots & \vdots \\ \sigma_{\beta_{n_q,1}}(\mathbf{x}) & \cdots & \sigma_{\beta_{n_q,n_u}}(\mathbf{x}) \end{bmatrix} \in \mathbb{R}^{n_q \times n_u},$$

enthalten, sowie die Größen $\boldsymbol{\epsilon}_{s,\alpha} \in \mathbb{R}^{n_q}$ und $\boldsymbol{\epsilon}_{s,\beta} \in \mathbb{R}^{n_q \times n_u}$ mit den normalverteilten Elementen

$$\epsilon_{s,\alpha}^{(i)}, \epsilon_{s,\beta}^{(i,j)} \sim \mathcal{N}(0, 1), \quad i = 1, \dots, n_q, j = 1, \dots, n_u.$$

Nach der Generierung der Stellgrößenrealisierungen wird die zugehörige Varianz empirisch mittels

$$\boldsymbol{\sigma}_u^2 \approx \frac{1}{n_s} \sum_{i=1}^{n_s} (\mathbf{u}_s^{(i)} - \boldsymbol{\mu}_u)(\mathbf{u}_s^{(i)} - \boldsymbol{\mu}_u)^T, \tag{5-8}$$

approximiert. Analog zu den Überlegungen der hybriden Optimalsteuerung, liegen vor der Erprobung des Regelgesetzes (5-7) am realen System über die gelernten $\alpha_i(\mathbf{x}), \beta_{i,j}(\mathbf{x})$ mit $i = 1, \dots, n_q, j = 1, \dots, n_u$ nicht nur ein hybrides probabilistisches Dynamikmodell, sondern auch ein probabilistischer Regler vor. Dementsprechend kann bereits vor der Er-

probung eine probabilistische Langzeitprädiktion ($t, \mu_x(t), \Sigma_x(t), \mu_u(t), \sigma_u^2(t)$) anhand eines gewünschten Sollwertverlaufs $w(t)$ berechnet werden, die dem Entwurfsingenieur die Möglichkeit eröffnet, das System tiefergehend zu untersuchen und die Auswirkungen der möglicherweise noch vorhandenen Modellfehler besser abschätzen zu können. Zudem kann die probabilistische Stabilität des Gesamtsystems auf der Basis der Zustandsvarianz $\Sigma_x(t)$ untersucht werden. Da diese Zusammenhänge bereits vor dem Hintergrund der hybriden Optimalsteuerung mit dem erweiterten Riccati-Regelungsentwurf und den Ergebnissen am Doppelpendelprüfstand erläutert wurden, wird die detaillierte Aufstellung einer Zustandspropagation nicht weitergehend vertieft. Im nächsten Abschnitt wird die Anwendung der hybriden Zustandslinearisierung für ein praxisbezogenes System vorgestellt.

5.3 Konzeptionierung einer Regelung für einen Hexapod mittels hybrider Zustandslinearisierung

Ein bedeutender Prüfstand für die Forschung an aktiven Fahrzeugachsen, der am Lehrstuhl für Regelungstechnik und Mechatronik der Universität Paderborn entwickelt wurde, ist in Bild 5-1 zu sehen und kann Straßenbedingungen realitätsnah in allen sechs Freiheitsgraden nachbilden und arbeitet unter Echtzeitbedingungen⁴. Um dies zu erreichen, wird unter anderem ein hochdynamisch geregelter Hexapod als Anregungseinheit eingesetzt. Um ein vollständiges Fahrzeugverhalten in einer *Hardware-in-the-Loop*(HiL)-Simulation nachzubilden, werden neben der Fahrzeugdynamik auch die Straße und die Räder auf einem Echtzeitsystem synchron mitberechnet. Die notwendige Sensorik umfasst unter anderem sechs Weg- und Druckaufnehmer sowie einen Winkelsensor. Zusätzlich werden zwölf Ventilschieberwegsensoren und ein Kraftmessrad verwendet, um die Kräfte und Momente zwischen Hexapod und aktiver Fahrzeugachse zu erfassen. Des Weiteren werden sechs Steuerspannungen für die Servoventile ausgegeben. Der Hexapod selbst gehört zur Gruppe der Parallelkinematiken [Koh20]. Anders als bei seriellen Kinematiken, die man beispielsweise bei Industrierobotern findet, sind hier alle Antriebe direkt mit dem Endeffektor verbunden. Dadurch entsteht eine komplexe, koordinierte Bewegung, bei der die sechs Hydraulikzylinder gleichzeitig agieren, um den Endeffektor in allen sechs räumlichen Freiheitsgraden zu bewegen. Der Hexapod wird durch insgesamt sechs Hydraulikzylinder aktuiert. Jeder Zylinder ist am unteren Ende über ein Kardangelenk, das mit einer Ölzuflussleitung ausgestattet ist, mit der Bodenplatte verbunden, während das obere Ende über ein Kugelgelenk mit der Arbeitsplattform des Hexapoden gekoppelt ist. Diese Struktur ermöglicht dem Prüfstand unter der Hinzunahme einer Positionsregelung, vielfältige Fahrbahnbedingungen präzise nachzubilden. Das Gesamtsystem wird

⁴Unter https://youtu.be/cN7_Epprgeg?si=4tkj1w-TLf0OApqx ist der Betrieb gezeigt.

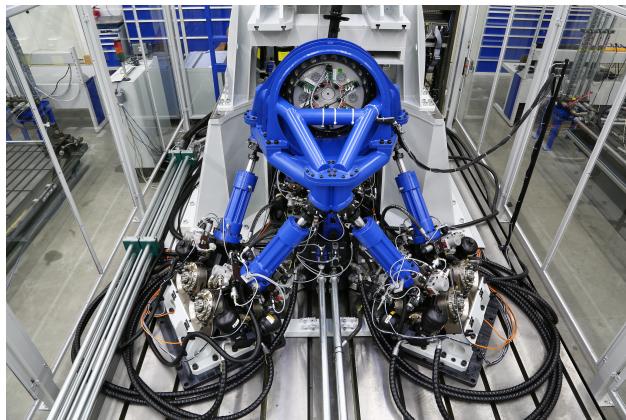


Bild 5-1: Fahrzeugachsprüfstand mit hydraulischem Hexapod (© Heinz Nixdorf Institut)

dazu eingesetzt, um aktive Fahrzeugachsen zu testen und hinsichtlich ihrer Regelgüte bzw. Komforts für den Fahrzeuginsassen zu bewerten. Des Weiteren kann die Betriebsfestigkeit der Achse überprüft werden. Das Regelungskonzept des Achsprüfstands kann auf mehreren unterschiedlich hierarchischen Ebenen betrachtet werden. Bspw. wird jeder Hydraulikzylinder durch seine eigene Kraftregelung angesteuert. Unterlagert werden hierzu die Steuerspannungen für die Servoventile der hydraulischen Pumpen vorgegeben. Die Sollwerte für die Kräfte der Hydraulikzylinder wird durch eine hierarchisch höhere Positionsregelung festgelegt. Auf dieser Positionsregelung liegt der anvisierte Fokus dieses Unterkapitels. Sie sorgt dafür, dass das vorgegebene Straßenprofil bestmöglich nachgebildet wird und hat damit als Ziel, die Position der Endeffektor-Plattform einzuregeln. Die Position ist hierbei durch die räumlichen Koordinaten und die Orientierung im Raum definiert. Da es sich bei dem Hexapod um ein klassisches mechanisches System handelt, wird für die Positionsregelung der beschriebene Ansatz der hybriden Zustandslinearisierung (vgl. (5-3)) eingesetzt, wobei insbesondere für hochdynamische Sollwertverläufe bzw. Fahrbahnprofile mit einem hochfrequenten Anteil ein exaktes Modell des Systems notwendig ist, um eine hohe Regelgüte erreichen zu können. Da es sich bei dem Achsprüfstand um ein hoch komplexes System mit vielen Komponenten aus verschiedenen Disziplinen handelt und damit eine Modellierung als umständlich und schwierig einzustufen ist, bietet sich, auf der Grundlage des umfassenden Sensorkonzepts, der Einsatz der hybriden Zustandslinearisierung besonders an.

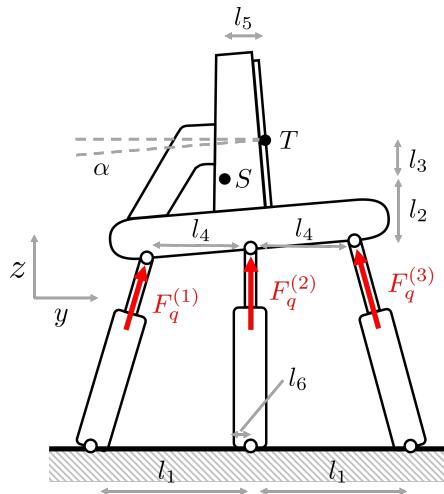


Bild 5-2: Prinzipskizze des Hexapod-Ersatzmodells nach [Koh20]. Der Schwerpunkt wird durch S und der Tool Center Point mit T repräsentiert.

Der Beitrag der vorliegenden Arbeit ist, in diesem Zusammenhang eine erste Konzeptionierung für die Anwendung der hybriden Zustandslinearisierung vorzunehmen. Hierzu wird der dreidimensionale Hexapod in ein äquivalentes zweidimensionales System simplifiziert und eine entsprechende Simulationsumgebung implementiert. Das zweidimensionale System ist dabei eine Abstraktion, welche die gleichen Eigenschaften wie der dreidimensionale Hexapod aufweist, jedoch weniger Zustandsgrößen beinhaltet. Im Rahmen dieser Simulationsumgebung erfolgt dann die Erprobung der entwickelten Positionsregelung auf der Basis der hybriden Zustandslinearisierung. Diesbezüglich wird ein Modell als Ground Truth und Ersatz für das reale System definiert und ein weiteres verändertes Modell zur Repräsentation des vorhandenen Vorwissens eingeführt. Bild 5-2 zeigt den schematischen Aufbau des betrachteten Ersatzsystems. Die Herleitung der Bewegungs-gleichungen ist dem Anhang der Arbeit [Koh20] zu entnehmen und wurde nach dem Prinzip von Jourdain durchgeführt. Der Positionsvektor wird über $\mathbf{q}(t) = [r_y(t), r_z(t), \alpha(t)]^T$ beschrieben und enthält zwei translatorische Koordinaten, sowie die Verkipfung der oberen Plattform. Die Stellgröße $\mathbf{u}(t) = [F_q^{(1)}(t), F_q^{(2)}(t), F_q^{(3)}(t)]^T$ setzt sich aus den drei Kräften, die an den Kolben der Hydraulikzylinder wirken, zusammen. Entsprechend der Herleitung ergibt sich das System dann in nichtlinearer Zustandsraumdarstellung über

$$\dot{\mathbf{x}} = \mathbf{f}_{hex}(\mathbf{x}, \mathbf{u}; \mathbf{p}) = \begin{bmatrix} \dot{\mathbf{q}} \\ \mathbf{M}(\mathbf{q})^{-1}(\mathbf{J}(\mathbf{q})^{-T}\mathbf{u} - \mathbf{F}_d(\dot{\mathbf{q}}) - \mathbf{C}(\dot{\mathbf{q}}, \mathbf{q})\dot{\mathbf{q}}), \end{bmatrix}$$

mit dem Zustandsvektor $\mathbf{x} = [\mathbf{q}^T, \dot{\mathbf{q}}^T]^T \in \mathbb{R}^6$ und dem linearen Dämpfungskraftvektor $\mathbf{F}_d = [d_y \dot{r}_y, d_z \dot{r}_z, d_\alpha \dot{\alpha}]^T$. Die Gewichtskraft wurde dabei vernachlässigt, damit keine konstante Kompensation durch die Stellgröße vorgenommen werden muss, um das System auf einer gewissen Höhe halten zu können. Insgesamt wird das System durch elf mecha-

nische Parameter beschrieben, welche in dem Vektor $\mathbf{p} = [l_1, \dots, l_6, m, I, d_y, d_z, d_\alpha]^T$ zusammengefasst werden. Darin sind $l_{(.)}$ geometrische Längen, m die Masse der Plattform, I das Trägheitsmoment um den Schwerpunkt S und $d_{(.)}$ die Dämpfungskonstanten zur vereinfachten Nachbildung der Reibungseffekte. Des Weiteren wird eine realitätsnahe Stellgrößenbeschränkung von $u_{max} = 250$ N m berücksichtigt. Im Gegensatz zu den bisherigen Beispielen, wird in diesem Szenario kein Modellfehler durch vernachlässigte Terme angenommen, sondern in Form einer fehlerbehafteten Parametrierung \mathbf{p} dargestellt. Auf diese Weise sind alle Terme in einem gewissen Umfang mit einem Fehler versehen, der durch den hybriden Ansatz ausgeglichen werden muss. Ausgehend von zwei unterschiedlichen Parametrierungen \mathbf{p}_1 und \mathbf{p}_2 lautet der zu lernende Modellfehler somit

$$\Delta \dot{\mathbf{x}} = \mathbf{f}_{hex}(\mathbf{x}, \mathbf{u}; \mathbf{p}_1) - \mathbf{f}_{hex}(\mathbf{x}, \mathbf{u}; \mathbf{p}_2),$$

wobei \mathbf{p}_1 dem unbekannten Ground Truth Modell zuzuordnen ist. In Bezug auf die l_1 -Norm beträgt die Abweichung $\|\mathbf{p}_1 - \mathbf{p}_2\|_1 / \|\mathbf{p}_1\|_1$ zwischen den Parametrisierungen ungefähr 15% und ist damit als relativ gering einzustufen.

Simulationsergebnisse

Nachfolgend werden die exemplarischen Simulationsergebnisse besprochen, welche die Anwendung der hybriden Zustandslinearisierung für das zwei-dimensionale Hexapod-Ersatzsystem darstellen. Der Hexapod befindet sich zum Start der Simulation in der Anfangsposition $\mathbf{q}(t=0) = [0.05 \text{ m}, 0.44 \text{ m}, 0]^T$. Die Simulationszeit beträgt $T = 5$ s und die zeitliche Schrittweite $\Delta t = 0.02$ s. Der zeitabhängige Sollwert für den Zustandsvektor lautet $\mathbf{w}(t) = [\mathbf{q}_{Soll}(t), 0, 0, 0]^T$, wobei für die Soll-Positionen sprungförmige Verläufe angenommen werden. Für jedes der nachfolgenden Szenarien wird die gleiche Verstärkungsmatrix $\mathbf{K} = [100 \ \mathbf{I}_3, 17 \ \mathbf{I}_3] \in \mathbb{R}^{3 \times 6}$ für die zugrundeliegende Zustandsregelung verwendet. Des Weiteren wird zur Einhaltung der Stellgrößenbeschränkung der ausgegebene Wert der Regelung nach oben bzw. unten hin begrenzt.

Bild 5-3 zeigt die Zustands- und Stellgrößenverläufe des Hexapod-Ersatzmodells für verschiedene Szenarien. Die grünen Linien stellen den geforderten sprungförmigen Sollwertverlauf der Positionen dar. Eine Änderung erfolgt dabei schrittweise und nur für eine Zustandsgröße. Aufgrund der Kopplung der Bewegungsgleichungen untereinander, wirken sich Änderungen innerhalb eines Zustands jedoch auch auf die anderen Zustände aus und können somit zu negativen Abweichungen zum Sollwert dieser Zustände führen. Der orangefarbene gepunktete Verlauf stellt den gewünschten idealen Fall dar, in welchem die Zustandslinearisierung anhand des Ground Truth Modells vorgenommen wird. Dieses Szenario dient als Vergleich zu den anderen Szenarien und repräsentiert die bestmögliche Regelgüte, die erreicht werden kann. Da die Kompensation der Nichtlinearitäten exakt ist, werden die Sollwerte rasch eingeregelt und lediglich durch das Integratorverhalten leicht

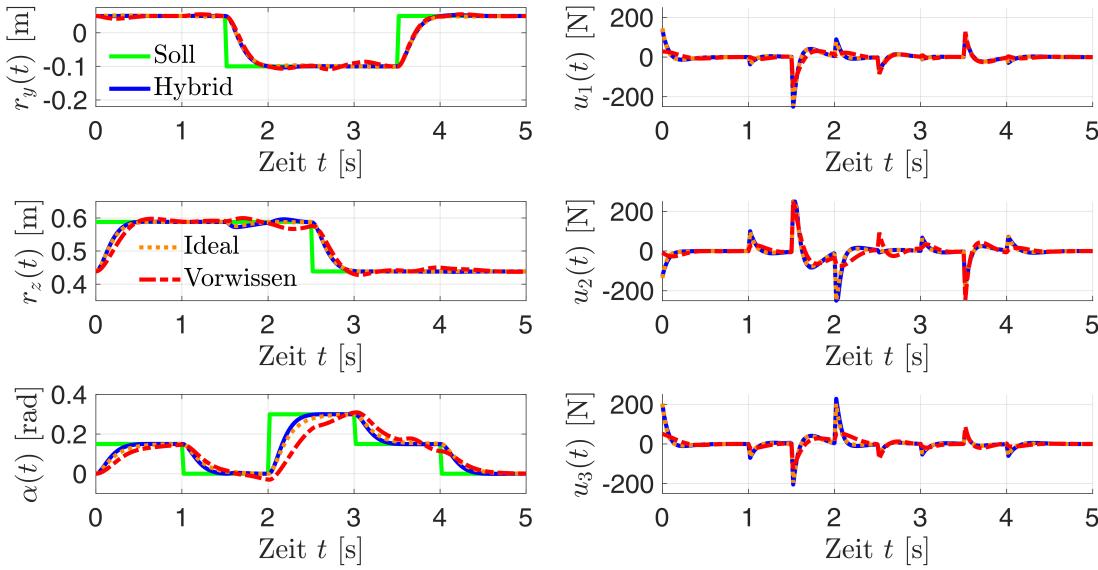


Bild 5-3: Dynamisches Verhalten des Hexapod-Ersatzmodells unter Verwendung verschiedener Regelungskonzepte.

verzögert. Des Weiteren sorgen die präzisen Stellgrößen dieses Szenarios dafür, dass die Kopplung der Zustände aufgehoben wird und sich eine Änderung in einem Zustand nicht auf die anderen Zustände auswirkt. Die roten gestrichelten Linien gehören zu einer Zustandslinearisierung, die alleine auf der Basis des verfälschten Modells (Vorwissen), welches der Parametrierung p_2 zuzuordnen ist, durchgeführt wird. Im Hinblick auf die Praxis stellt dies den rein physikalisch basierten Modellteil dar, welche unter Umständen signifikante Modellfehler enthalten kann. Die Sollwerte stellen sich für dieses Szenario zwar mit der Zeit ein, jedoch ist die Regelgüte insgesamt als schlecht zu bewerten, da die Kompen-sation aufgrund der abweichenden Parametrierung deutliche Mängel aufweist. Auffällig ist insbesondere die verbleibende Wechselwirkung zwischen den Zuständen, als auch die schlechte Sollwertfolge des Winkels α innerhalb von $t \in [2 \text{ s}, 3 \text{ s}]$. Trotz der Verwendung der gleichen Verstärkungsmatrix \mathbf{K} strebt das System hier nur sehr langsam auf seinen Sollwert zu und weist einen zu stark gedämpften Verlauf auf. Der Grund hierfür ist, dass für eine rasche Änderung nicht genug Energie durch die Stellgrößen bzw. die Regelung bereitgestellt wird. Vor dem Hintergrund des eindimensionalen Falls (vgl. (5-1)), wird der nominale Wert der Nichtlinearität β in diesem Zustandsbereich durch das Modell als zu groß angenommen. Bezuglich (5-2) kann damit argumentiert werden, dass das Verhältnis $\beta(x)/\hat{\beta}(x)$ (idealerweise 1) kleiner als 1 ist und der Reglereingriff somit insgesamt abgeschwächt wird. Ausgehend von dem roten Szenario, ist die Zielsetzung nun die Regelgüte zu verbessern. In der Praxis würde nun oftmals eine Verbesserung der Zustandslinearisierung durch eine genauere physikalische Modellierung oder ein manuelles Einstellen der Verstärkungsfaktoren $K_{i,j}, i = 1, \dots, n_u, j = 1, \dots, n_x$ angestrebt werden. Beide

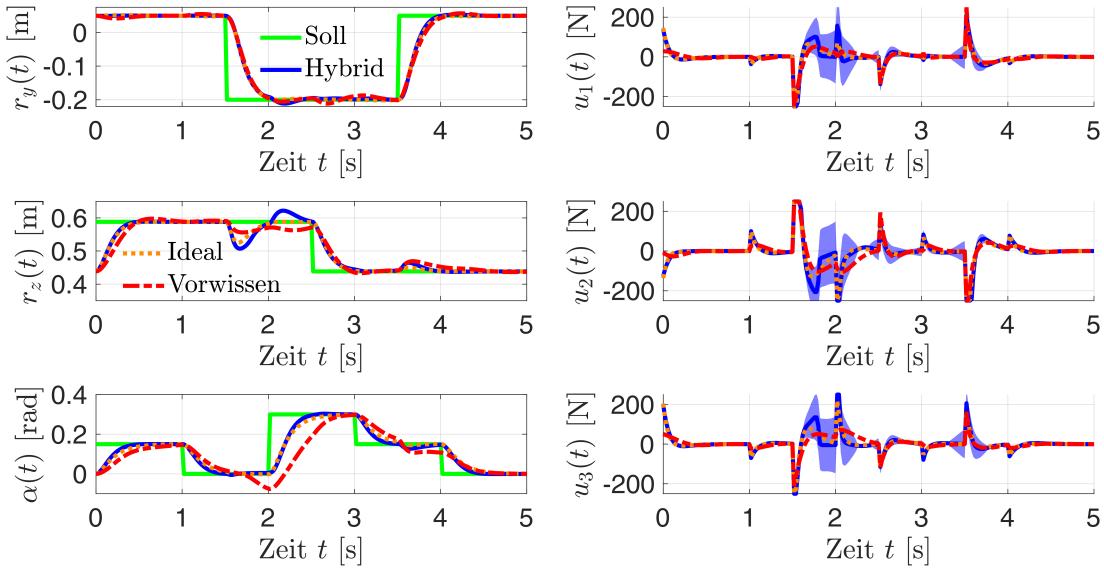


Bild 5-4: Dynamisches Verhalten des Hexapod-Ersatzmodells für einen veränderten Sollwertverlauf.

Lösungsansätze sind in der Regel mit einem hohen Zeitaufwand verbunden. Zudem löst das manuelle Einstellen der Reglerparameter nicht das grundlegende Problem einer fehlerbehafteten Kompensation der Nichtlinearitäten. Eine bessere Lösung stellt die entwickelte hybride Zustandslinearisierung dar. Die fiktiven Messdaten des roten Verlaufs und das bereits bekannte Dynamikmodell werden dabei im Rahmen des GP-Trainings ausgenutzt. Die Ergebnisse des hybriden Ansatzes sind in blau dargestellt. Die Regelgüte kommt sehr nah an die der idealen Referenz heran, sodass die Zustandsverläufe fast vollständig übereinanderliegen. Geringe Abweichungen sind nur für die Höhenposition $r_z(t)$ im Bereich um $t = 2$ s festzustellen. Hier macht sich die relativ große Veränderung des Winkels weiterhin geringfügig bemerkbar. Für einen objektiven Vergleich der erreichten Regelgüten, wird ein Gütekriterium über $J = \int_{t=0}^T \|q_{Soll}(t) - q(t)\|_2^2 dt$ ausgewertet, welches die zeitlich gemittelte quadratische Regelabweichung darstellt. Die Regelung des Vorwissens kommt hierbei auf einen Wert von 0.055, wohingegen die hybride Regelung 0.031 erzielt, was einer Verbesserung von 44% entspricht. In Bezug auf die einzelnen Stellgrößen $u_i, i = 1, 2, 3$ wurde für den hybriden Ansatz zudem die Auswertung der Varianz nach (5-8) vorgenommen. Um den Mittelwert der Stellgrößen wurde das 95%-Konfidenzintervall abgebildet, jedoch ist dieses so gering, dass es kaum erkennbar ist. Die niedrige Stellgrößenvarianz ist eine Folge dessen, dass das GP-Training anhand der Daten durchgeführt wurde, welche im Bereich um den Sollwertverlauf liegen. Für die Praxis bedeutet dies, dass von einer hohen Sicherheit der gelernten Regelung ausgegangen werden kann und sich der gewünschte Sollwertverlauf trotz der fehlerbehafteten Parametrierung des physikalischen Modellteils gut einregeln lässt.

Im Allgemeinen sollte davon ausgegangen werden, dass sich das Regelverhalten beim Verlassen des Bereichs, in welchem die Daten vorliegen, wieder verschlechtert. In diesem Zusammenhang stellt sich die Frage, wie es um die Extrapolationseigenschaften des hybriden Reglers bestellt ist. Zur Beantwortung dieser Frage wurde eine weitere Simulation durchgeführt. Im Vergleich zu Bild 5-3 wurde der Sollwertverlauf für die horizontale Position $r_y(t)$ geändert, in dem der Sprung im mittleren Bereich anstelle von -0.1 m auf -0.2 m erfolgt. Der Hexapod befindet sich somit in einem Gebiet im Zustandsraum, in welchem noch keine Daten vorliegen. Die erzielten Ergebnisse der besprochenen Regelungen sind in Bild 5-4 zu sehen. Selbst die ideale Regelung kann nun für die neue Höhe $r_z(t)$ um $t = 1.8 \text{ s}$ keine exakte Entkopplung mehr herstellen und führt hier zu einer größeren Abweichung zum Sollwert. Der Grund hierfür ist das Erreichen der Stellgrößenbeschränkung. Die roten Verläufe und die damit verbundene inkorrekte Zustandslinearisierung ist weiterhin als vergleichsweise schlecht zu bewerten. Demgegenüber ist die Performance des hybriden Ansatzes deutlich besser, hat sich aber im direkten Vergleich zu Bild 5-3 leicht verschlechtert. Dies entspricht der ursprünglichen Erwartung und zeigt sich ebenfalls durch das objektive Gütekriterium. Die Regelung des Vorwissens erreicht einen Wert von 0.078, während die hybride Regelung 0.047 realisiert, was dem gleichen Wert der idealen Regelung entspricht. Die Verbesserung beträgt damit 40%. Eine signifikante Veränderung ist in Bezug zu den Stellgrößen festzustellen. Hier ist die Varianz der Stellgrößen im mittleren Zeitbereich deutlich erhöht und weist darauf hin, dass sich der Regler für die jeweiligen Zustände äußerst unsicher über seine Stellgröße ist. Ein sicheres und stabiles Regelverhalten kann daher nicht vollständig gewährleistet werden. Bei einem herkömmlichen deterministischen Entwurf stünde diese Information nicht zur Verfügung, sodass eine vergleichbare Bewertung der Regelung entsprechend nicht stattfinden könnte. Aus den Zustandsverläufen, welche in diesem Fall lediglich zur Verfügung stehen würden, könnte aufgrund der hohen Übereinstimmung mit dem Sollwert nicht geschlussfolgert werden, dass eine hohe Unsicherheit bzgl. der Stellgrößen vorherrscht. Durch die Kenntnis der Stellgrößenvarianz können bereits zur Laufzeit (online) geeignete Maßnahmen ergriffen werden, um ein unerwartetes Systemverhalten zu verhindern. Bspw. kann bei der Überschreitung eines Grenzwertes die Reglerverstärkung reduziert werden, damit sich insgesamt eine höhere Stabilitätsreserve einstellt. Weiterführend liegt es nahe, die Varianzangabe dafür zu nutzen, um einen weiteren Lernvorgang anzustoßen. Hierbei wird ähnlich zu den bisher vorgestellten Lernverfahren die Datenbasis durch die neuen Daten erweitert, um so einen größeren Bereich des Zustandsraums abzudecken. In diesem Zusammenhang muss das Training nicht zwangsläufig offline und in bestimmten Iterationen durchgeführt werden, sondern kann zur Laufzeit, während des Betriebs und parallel zur Regelung, automatisiert vorgenommen werden.

Die durchgeführten Simulationen sind ein erster Funktionsnachweis der hybriden Zustandslinearisierung für den Hexapod. Eine Adaption für das reale dreidimensionale System ist durch die allgemeingültige Formulierung des Ansatzes mit einem geringen Aufwand möglich. Die Ergebnisse am zweidimensionalen Hexapod-Ersatzmodell zeigen, dass bereits geringfügige Modellfehler einen großen Einfluss auf die Regelgüte haben können. Für das reale System, welches aufgrund der Vielzahl an Komponenten eine deutlich komplexere (nicht nur mechanische) Modellierung erfordert, ist die Wahrscheinlichkeit für Fehler besonders erhöht. In Verbindung mit dem ausgiebigen Sensorkonzept des Prüfstands bietet sich der Einsatz der hybriden Zustandslinearisierung für die Inbetriebnahme speziell an. Der hier erbrachte prinzipielle Funktionsnachweis stellt eine entsprechend solide Grundlage für weiterführende Arbeiten dar.

6 Zusammenfassung und Ausblick

Durch die Inbetriebnahme von Steuerungen und Regelungen wird sicherstellt, dass ein mechatronisches System ordnungsgemäß funktioniert und den Anforderungen, sowie Erwartungen gerecht wird. In diesem Zusammenhang wurde in Kapitel 1 die Anwendung des V-Modells und des modellbasierten Entwurfs als bewährte Vorgehensweise in der Regelungstechnik vorgestellt. Der Entwurf basiert dabei auf der Herleitung eines ausreichend genauen Simulationsmodells. Insbesondere vor dem Hintergrund von komplexen mechatronischen Systemen oder anspruchsvollen Regelungsaufgaben ist die klassische Vorgehensweise des modellbasierten Entwurfs allerdings nur begrenzt einsetzbar, da eine analytische Modellierung in vielen Fällen zu kompliziert, zeitaufwendig und unter gewissen Umständen nicht praktikabel umsetzbar ist. Ein Beispiel hierfür ist der besprochene Ultraschalldrahtbondprozess, dessen komplizierte Modellierung in Abschnitt 3.4 behandelt wurde. Im Rahmen der Regelungstechnik und des modellbasierten Entwurfs gibt es kaum Handlungsanweisungen für den Fall, dass eine Modellbildung für das betrachtete System nicht oder nur eingeschränkt durchführbar ist. Dieser Sachverhalt wurde in dieser Arbeit als vorhandene Forschungslücke identifiziert, welche auf die Abweichungen zwischen theoretischen Überlegungen und der praxisnahen Inbetriebnahme zurückzuführen ist. Das Ziel der Arbeit bestand darin, die Forschungslücke zu untersuchen und geeignete Verfahren (im Kontext der Arbeit auch als regelungstechnische Werkzeuge bezeichnet) zu entwickeln, welche eine effiziente und sichere Inbetriebnahme einer Steuerung oder Regelung zum Ziel haben. Hierfür wurden im einführenden Kapitel 1 diverse Anforderungen formuliert, wie bspw. eine rasche Realisierung, um den Verschleiß und die Beanspruchung des betrachteten Systems zu reduzieren. Um das Ziel zu erreichen, wurde im Rahmen der Arbeit die Grundidee verfolgt, eine Kombination von Regelungstechnik und ML bzw. RL herzustellen, welche das gesamte Vorwissen über die Regelungsaufgabe ausnutzt und Korrekturen auf der Basis von Messdaten und der probabilistischen GP-Regression vornimmt. Das Vorwissen kann dabei als partiell bekanntes physikalisches Modell oder in Form einer geeigneten Steuerungsfunktion bzw. Regelungsstruktur vorliegen. Die Einhaltung von technischen Rahmenbedingungen und stabilen Systemverhalten bei der Inbetriebnahme wird dabei über probabilistische Prädiktionen berücksichtigt. Als konkrete Anwendungsbeispiele dienten der Ultraschalldrahtbondprozess, verschiedene Pendelsysteme und ein Hexapod zur Fahrzeugachsprüfung.

Nachdem in Kapitel 2 die mathematischen Grundlagen erläutert worden sind, wurde in Kapitel 3 zunächst der Steuerungsentwurf mittels BO behandelt. Die BO ist ein iteratives Lernverfahren aus dem Bereich des modellbasierten RL, welches der globalen Op-

timierung zugeordnet wird. Im Kontext des Steuerungsentwurfs hat sie zum Ziel, eine optimale Parametrisierung zu identifizieren. Hierfür werden nach und nach Experimente am realen System durchgeführt, um Messdaten zu sammeln. Die Messdaten werden wiederum dafür genutzt, um ein probabilistisches GP-Modell zu aktualisieren, welches zur Abbildung der Gütfunktion dient. In jeder Iteration wird das nächste Experiment durch eine Auswertung des Erwartungswertes und der Varianz bzgl. des Gütfunktionswert bestimmt. Vor dem Hintergrund des Ultraschalldrahtbondprozesses wurde als Gütfunktion der Prozessfähigkeitsindex betrachtet. Eine klassische modellbasierte Optimalsteuerung nach dem Vorgehen aus Kapitel 4 ist für dieses System praktisch nicht umsetzbar, da sich die physikalische Modellierung zum einen als schwierig herausgestellt hat (s. Kapitel 3.4) und zum anderen die Anforderungen zur Anwendung der Optimalsteuerungsmethoden nicht erfüllt sind. Bspw. ist eine hohe zeitliche Diskretisierung erforderlich, da sich die Dynamik des Systems im Kilohertzbereich bewegt. Des Weiteren sind Messungen des Zustands aufgrund fehlender Sensorik nicht möglich. In diesem Anwendungsfall stehen somit nur Daten der Gütfunktion zur Verfügung, womit die BO ein geeignetes Werkzeug zur Inbetriebnahme darstellt. Das Vorwissen wurde bei diesem Beispiel über die Definition des Steuerungsverlaufs berücksichtigt, welches sich aus stückweise definierten linearen Funktionen für die Spannung und Normalkraft zusammensetzt. Die Ansatzfunktion der Steuerung basiert auf Expertenwissen und hat sich als besonders nützlich für das Ultraschalldrahtboden herausgestellt. In Kapitel 3.5 wurde das Standardverfahren der BO aufgegriffen und etwaige Anpassungen für die Inbetriebnahme des Bondprozesses eingeführt. Unter anderem galt es, Werkzeugaufsetzer, welche die beiden Verbindungspartner schädigen, zu vermeiden. Hierfür wurde ein weiterer GP eingeführt, der das Auftreten solcher Aufsetzer durch eine Wahrscheinlichkeitsangabe wiedergibt. Des Weiteren wurde das Verfahren um Batchelemente erweitert, sodass in einer Iteration mehrere Experimente parallel durchgeführt werden und so der Automatisierungsgrad des Prozesses vollständig ausgenutzt wird. Im Rahmen der Validierung am realen System wurde der entwickelte BO-Algorithmus mit der herkömmlichen manuellen Identifikation der Parametrisierung verglichen. Die BO konnte bei einem gleichen Iterationsbudget doppelt so hohe Prozessfähigkeitswerte erzielen und damit eine bessere Parametrisierung bestimmen. Nach einer kurzen Explorationsphase fokussierte sich das Verfahren dabei auf Gebiete im Suchraum, welche die höchsten Gütfunktionswerte erzielten. Die manuelle Suche des Einrichters stagnierte hingegen in einem vermeintlichen lokalen Optimum, was auf die Dimension des Suchraums zurückzuführen ist. Diese betrug 7 und stellt für den Menschen daher eine Herausforderung aufgrund der zahlreichen Kombinationsmöglichkeiten und Wechselwirkungen mit dem komplexen Prozess dar. Als Fazit zu den Untersuchungen wurde festgehalten, dass die BO für das betrachtete Anwendungsbeispiel einen Mehrwert darstellt, da es vor den Untersuchungen kein geeignetes regelungstechnisches Verfahren

gab, mit dessen Hilfe der Steuerungsentwurf durchgeführt werden konnte. Die Besonderheit dieses Ansatzes ist, dass er auf der gedanklichen Ebene der Gütfunktion angesiedelt ist und als Voraussetzung für das Verfahren nur die Auswertung der Gütfunktion erforderlich ist. Üblicherweise wird in der Regelungstechnik auf der Ebene der Zuständen gearbeitet, die jedoch im Falle des Ultraschalldrahtbondens nicht zugänglich ist.

In den nachfolgenden Kapiteln 4 und 5 wurde der Fall untersucht, bei dem die Zustände des Systems vollständig definiert sind und gemessen bzw. berechnet werden können. In diesem Szenario werden die zugehörigen Messdaten für eine Korrektur auf der Ebene der Dynamikgleichungen verwendet. Ausgehend von einem partiell bekannten System bzw. einem unvollständigen Simulationsmodell werden die beinhaltenden Modellfehler durch ein ML-Verfahren nachgebildet. Als ML-Verfahren wurde hierfür wiederum die GP-Regression eingesetzt, welche eine Quantifizierung des Fehlers im Sinne der Wahrscheinlichkeitstheorie vornimmt. Das hybride Gesamtmodell setzt sich additiv aus einem vorab identifizierten physikalischen und einem rein datengetriebenen Modellteil zusammen. Um die Anforderungen bei der Inbetriebnahme zu adressieren, wurde das Modell probabilistisch formuliert. In der gängigen deterministischen Betrachtungsweise, mit welcher üblicherweise in der Regelungstechnik gearbeitet wird, werden keine Unsicherheiten in Bezug auf die Zustandsübergänge des Systems berücksichtigt. Vor dem Hintergrund eines unvollständigen Modells und der Inbetriebnahme ist es allerdings hilfreich, gewisse Unsicherheiten mit einzubeziehen. Zum einen können auf diese Weise sicherheitskritische Überschreitungen deutlich besser abgeschätzt werden und zum anderen kann ein exploratives Verhalten erzeugt werden. In diesem Zusammenhang wird der Prozess der Inbetriebnahme als Lernvorgang im Sinne des RL aufgefasst, wobei bei jedem neuen Experiment eine explorative oder exploitative Entscheidung getroffen werden muss. Die probabilistische Betrachtungsweise des Modells ist hierfür zielführend. Zur Realisierung einer probabilistischen Langzeitprädiktion wurde das Verfahren der UT genutzt, wobei die Zustände des Systems als normalverteilt angenommen werden. Für eine gegebene Steuerung lassen sich so die Auswirkungen der Modellfehler über einen gewissen Zeithorizont untersuchen und es kann abgeschätzt werden, wo sich das System mit welcher Wahrscheinlichkeit befindet. In Kapitel 4 wurde die Verwendung des hybriden Modellansatzes im Rahmen eines Optimalsteuerungsentwurf behandelt. Die konventionellen Gleichungen des zugrundeliegenden Optimalsteuerungsproblems wurden für das probabilistische Modell mit Hilfe des Multiple Shooting Ansatzes formuliert und ein lernfähiger Algorithmus entwickelt. Dieser arbeitet ähnlich zur BO und nähert sich iterativ durch Interaktion mit dem realen System dem Erfüllen der Steuerungsaufgabe an. Eine erste Untersuchung wurde an einem voll-aktuierten Doppelpendel durchgeführt. Dabei wurde der Einfluss des Vorwissens verdeutlicht. Das Vorwissen bewirkt eine Effizienzsteigerung, sodass die optimale Steuerung in kürzerer Zeit bzw. in weniger Experimenten bestimmt wird.

Des Weiteren war ein deutlicher Unterschied bei der Entwicklung der Steuerung über die Iterationen hinweg zu beobachten. Durch die Hinzunahme von Vorwissen wird der Steuerungsverlauf von Iteration zu Iteration nur geringfügig angepasst und es werden nur leichte Korrekturen vorgenommen, wohingegen bei dem Fall ohne Vorwissen ein breites Spektrum von Steuerungen ausprobiert wird. Für die praktische Inbetriebnahme ist dieses Verhalten problematisch, da von dem Entwurfsingenieur nur schwer abgeschätzt werden kann, was am realen System tatsächlich passiert. Dementsprechend sorgt das Einbringen von Vorwissen nicht nur für eine Effizienzsteigerung, sondern auch für eine Erhöhung der Sicherheit während der Inbetriebnahme. Weiterführend wurde die hybride Optimalsteuerung um einen zusätzlichen Regelungsentwurf erweitert. Hierfür wurde im speziellen die Zwei-Freiheitsgrade Regelung mit einem zeitvarianten Riccati-Regler betrachtet, welcher auf der Basis des hybriden Modells abgeleitet wird. Die Verwendung dieser Struktur entspricht einer weiteren Möglichkeit, Vorwissen über die Aufgabe einzubringen, da durch sie eine geeignete Zusammensetzung des Stellsignals vorgegeben wird, welche durch einen rein datengetriebenen Ansatz erst aufwendig gelernt werden müsste. Als Beispiel wurde hierfür das modellbasierte RL Verfahren PILCO angeführt, dessen resultierendes Steuersignal für den Menschen nicht im Sinne der Zwei-Freiheitsgrade Regelung, also einem Steuerungs- und Regelungsanteil, interpretiert werden kann. In diesem Zusammenhang fand eine weitere Validierung am Doppelpendel auf einem Wagen statt, welche sich durch ein chaotisches und anspruchsvolles Systemverhalten auszeichnet und daher für viele RL Methoden schwer zu lösen ist. Dennoch konnte das PILCO-Verfahren die Regelungsaufgabe nach 27 Iterationen erfolgreich abschließen, wofür insgesamt jedoch eine gesamte Berechnungsdauer von 2 Tagen erforderlich war. Gleichermaßen wurde das entwickelte hybride Verfahren am Prüfstand getestet, wobei ein Teil der Dynamik als bekannt vorausgesetzt wurde und die wesentlichen Zentrifugalkräfte gelernt werden sollten. Das Vorwissen über die Dynamik und die Regelungsstruktur führte zu einer deutlichen Verbesserung des Lernvorgangs. So wurden insgesamt 80% weniger Experimente und zugleich 98% weniger Berechnungszeit benötigt. Zudem konnte das Systemverhalten in Bezug auf die Stabilität und damit die Sicherheit am Prüfstand qualitativ gut vorhergesagt werden. Ein Grund hierfür ist, dass die Unsicherheit über die vorhandenen Modellfehler explizit im Entwurfsverfahren berücksichtigt wurden. Die gezielte Einbringungen von Vorwissen über die Regelungsstruktur wurde in Kapitel 5 fortgesetzt. Für die Klasse der eingangsaffinen Systeme bietet sich der Einsatz einer Zustandslinearisierung an, wobei die nichtlinearen Funktionen der Dynamik über ein Modell invers durch die Stellgröße aufgeschaltet werden, sodass eine Kompensation stattfindet und das Gesamtsystem als linear betrachtet werden darf. Modellfehler können die Regelgüte jedoch stark mindern, sodass eine hybride Zustandslinearisierung hergeleitet wurde, welche in der Lage ist, Modellungleinheiten auszugleichen. Die GP-Regression lässt hierzu eine gezielte Identifikation

der einzelnen nichtlinearen Terme aus einem Datensatz zu, indem das Vorwissen über die Systemstruktur innerhalb der angenommenen Kernelfunktion vorgegeben wird. Als anschauliches Anwendungsbeispiel wurde die hybride Zustandslinearisierung an einem Hexapod bzw. einem abstrakten zweidimensionalen Simulationsmodell getestet. Anhand des Beispiels wurde gezeigt, dass ein ungenaues Modell problematisch für die Regelung ist und der hybride Ansatz eine deutliche Verbesserung herbeiführt. Aufgrund des probabilistischen Charakters der GP-Regression kann während des Betriebs eine Varianzangabe über die Stellgröße berechnet werden. Darauf lassen sich weiterführend Gebiete im Zustandsraum angeben, für welche eine sichere Funktionsweise nicht gewährleistet werden kann bzw. weitere Messdaten aufgenommen werden sollten, um die Unsicherheit über die verbliebenen Modellfehler zu beseitigen. Zudem wurde durch die Erprobung ein Funktionsnachweis erbracht, um eine weiterführende Entwicklung für den realen dreidimensionalen Hexapod vorzunehmen. Die BO, eine hybride Optimalsteuerung mit erweitertem Regelungsentwurf und die hybride Zustandslinearisierung haben gezeigt, wie sich das maschinelle Lernen gewinnbringend innerhalb der Regelungstechnik bzw. während der Inbetriebnahme einsetzen lässt. Der Fokus wurde dabei auf die Einbringung von Vorwissen über die Regelungsaufgabe und die Quantifizierung von Modellfehlern durch probabilistische Angaben gelegt. Für die Berücksichtigung des Vorwissens wurden verschiedene Schnittstellen betrachtet, so unter anderem die Mittelwert- und Kovarianzfunktion des GP seitens des maschinellen Lernverfahrens, als auch eine partiell bekannte Systemdynamik oder geeignete Regelungsstruktur bzw. Steuerungsansatzfunktion. Insgesamt haben die Anwendungsbeispiele gezeigt, dass die Einbringung von Vorwissen einen positiven Effekt auf die Inbetriebnahme nimmt und die Effizienz, Robustheit und die Interpretierbarkeit steigert. Aufgrund des probabilistischen Charakters der Verfahren, welche durch das Themengebiet des RL motiviert ist, kann zudem die Sicherheit während der Inbetriebnahme trotz möglicherweise vorhandener Modellfehler besser eingeschätzt werden. Des Weiteren ergibt sich ein inhärentes exploratives Verhalten der Lernalgorithmen, welches darauf ausgelegt ist, eine erfolgreiche Realisierung zu ermöglichen.

Ausblick

Der Ausblick dieser Arbeit weist auf Potenziale für weiterführende Forschungsthemen hin. Ein primäres Themenfeld betrifft die Weiterentwicklung der vorgestellten Verfahren und deren Kombination auf verschiedenen hierarchischen Ebenen, wobei sie entweder sequenziell oder parallel während der Inbetriebnahme zum Einsatz kommen können. Zur Veranschaulichung wird das Beispiel eines Doppelpendels auf einem Wagen und die Realisierung des Aufschwungs herangezogen. Im Rahmen der Validierung wurde ein unterlagerter Geschwindigkeitsregler für den Wagen verwendet, der auf einem Entwurf des Herstellers basiert. Eine alternative Herangehensweise wäre die eigenständige Auslegung des Geschwindigkeitsreglers gewesen, wofür die hybride Zustandslinearisierung

geeignet ist. Dieser Ansatz bildet die Basis, auf der die nachfolgenden Schritte aufbauen. Nach erfolgreicher Implementierung einer funktionsfähigen Geschwindigkeitsregelung des Wagens nach der hybriden Zustandslinearisierung wird die hybride Optimalsteuerung angewendet, wobei geeignete Gewichtungsmatrizen in Bezug auf (4-28) ausgewählt werden müssten. Die Flexibilität der BO ermöglicht nicht nur die Identifikation von Steuerungs- oder Regelungsparametrisierungen, sondern könnte auch für die Anpassung der Gewichtungsmatrizen in diesem Szenario genutzt werden. Ein ähnlicher Ansatz wird in [MHB⁺16] verfolgt, wobei die Gewichtungen eines Riccati-Reglers als Optimierungsvariablen betrachtet und optimal an das reale System angepasst werden. Demnach besteht der angestrebte Ansatz darin, die Gewichtungsmatrizen der hybriden Optimalsteuerung durch die BO automatisiert anzupassen, um eine realisierbare Trajektorie für das reale System zu finden (vgl. Anmerkungen zu Bild 4-2). Insgesamt betrachtet, befinden sich die drei verschiedenen Verfahren also auf unterschiedlichen gedanklichen Ebenen, wobei die unterlagerte Geschwindigkeitsregelung auf der hybriden Zustandslinearisierung, die Steuerung des Aufschwungs auf der hybriden Optimalsteuerung und die Gewichtungsmatrizen des zugehörigen Optimierungsproblems auf der BO basieren würden. In diesem Zusammenhang wäre es von Interesse, zu untersuchen, wie sich die Verfahren gegenseitig beeinflussen und welche Synergien, beispielsweise bezüglich Datenmengen und Formen des Vorwissens, genutzt werden könnten.

Ein weiterer Aspekt des Ausblicks betrifft das hybride Modell. Hierbei gibt es sowohl im physikalischen, als auch im datengetriebenen Teil verschiedene Möglichkeiten, um weitere Fälle zu betrachten. Im Großteil dieser Arbeit wurde der Zustand als normalverteilte Zustandsvariable definiert. Wie jedoch in Abschnitt 2.3.1 gezeigt wurde, lässt sich das physikalische Modell auch durch weitere Unsicherheiten beschreiben. So könnte beispielsweise die Parametrierung des Modells als zusätzliche Zufallsvariable definiert und dies entsprechend beim Entwurf berücksichtigt werden. Für geometrische Längen ist diese Unsicherheit in den meisten Fällen als sehr gering abzuschätzen, während insbesondere Parameter, welche die Reibung betreffen, oft nicht eindeutiger Werte zugeordnet werden kann. Neben dem Zustand und der Parametrierung kann zudem auch die Stellgröße als Zufallsvariable aufgefasst werden, um darüber eine nicht ideale Aktorik zu beschreiben. Diese Definitionen sind allerdings je nach Anwendungsfall abzuwägen, sodass es sinnvoll sein kann, einen Mix aus probabilistischen und deterministischen Variablen einzuführen. Des Weiteren bietet die Wahrscheinlichkeitstheorie ein breites Spektrum von Wahrscheinlichkeitsverteilungen, die im Rahmen dieser Arbeit noch nicht betrachtet wurden. Über andere Verteilungen könnten, je nach Anwendung, unterschiedliche Effekte abgebildet werden. So könnten beispielsweise die harten Anschläge des Wagens vom Doppelpendelprüfstand als begrenzte Zufallsvariablen definiert werden, wobei sich der Einsatz einer beidseitig beschnittenen Normalverteilung [Bis06] anbietet. Die komplexere Wahrschein-

lichkeitsverteilung sorgt allerdings gleichzeitig dafür, dass die nachfolgenden Berechnungen schwieriger werden und mehr Zeit in Anspruch nehmen. Es ist daher ein geeigneter Kompromiss zwischen Genauigkeit und Praktikabilität herzustellen, wie es für die Anwendungsbeispiele dieser Arbeit gezeigt wurde. Außerdem wurde bisher angenommen, dass sich die Stärke des Messrauschen nicht verändert und sich stationär verhält. In manchen Fällen kann das Messrauschen allerdings auch nicht-stationär sein und sich über den Zustandsraum (oder auch der Zeit) unterschiedlich stark ausprägen. Dementsprechend müsste eine Anpassung der angenommenen Wahrscheinlichkeitsverteilung im Zusammenhang mit der Bayesschen Regel (2-13) vorgenommen werden. Des Weiteren wurde im Rahmen dieser Arbeit ein Fokus auf die GP-Regression gelegt, wobei sich auch der Einsatz anderer maschineller Lernverfahren anbieten würde, um komplexere Phänomene bei der Inbetriebnahme abbilden zu können. Eine Erweiterung des herkömmlichen GP-Ansatzes wird als *Sparse GP* [SG05] bezeichnet und erhält seinen Namen dadurch, dass die ursprüngliche Datenmenge komprimiert wird und dabei die Prädiktionsgenauigkeit bestmöglich erhalten bleibt. Der Vorteil dieses Ansatzes ist, dass er bessere Skalierungseigenschaften besitzt und eine Auswertung deutlich schneller berechnet werden kann. Für Systeme mit vielen Zuständen und einer großen Datenmenge ist dieser Ansatz daher besonders gut geeignet. Eine weitere Variante nennt sich *Deep GP* [DL13] und leitet sich von den gleichnamigen tiefen neuronalen Netzen ab. Die Grundidee ist es, mehrere GPs hintereinander zu verknüpfen und damit die Beschreibungsfähigkeit der Prädiktion zu steigern. Auf diese Weise ist es beispielsweise möglich, das erwähnte nicht-stationäre Rauschen oder andere Verteilungen als die Normalverteilung abzubilden. Eine ähnliche Fähigkeit besitzen die Bayesschen neuronalen Netze [Gal16], welche sich im Gegensatz zu einem herkömmlichen neuronalen Netz dadurch auszeichnen, dass ihre Gewichtungswerte als Zufallsvariablen definiert werden und somit eine probabilistische Prädiktion erlauben. Die Parameter der zugrundeliegenden Wahrscheinlichkeitsverteilung der Gewichte werden über die Bayessche Regel trainiert, womit das Verfahren der Bayesschen nichtlinearen Regression (vgl. Abschnitt 2.3.1) zugeordnet wird. Zuletzt sei auch der *Koopman-Operator* [BBKK22] erwähnt, zu dem es in der Regelungstechnik in den letzten Jahren viele Forschungsarbeiten gab. Der Koopman-Operator wird innerhalb der Regelungstechnik verwendet, um die Dynamik eines nichtlinearen Systems zu beschreiben, indem er es in einen höherdimensionalen Raum abbildet, in dem die Dynamik linear ist. Durch die Transformation in diesen Raum können nichtlineare Systeme mit linearen Methoden analysiert und geregelt werden. Im Zusammenhang mit Gleichung (2-11) und der Zustandspropagation wird ersichtlich, dass eine lineare Dynamik bzw. die des Koopman-Operators dafür besonders gut geeignet ist. Unter der Voraussetzung eines normalverteilten Zustands wird die Klasse der Normalverteilungen auf diese Weise inhärent nicht verlassen, sodass es sich vor dem Hintergrund des MM nicht mehr um eine Approxima-

tion, sondern eine genaue Berechnung handelt. Die Zustandspropagation lässt sich somit besonders schnell für ein nichtlineares System berechnen. Weiterführend ließe sich so das komplexe hybride Optimalsteuerungsproblem (4-28) vereinfachen und eine Realisierung als modellprädiktive Regelung [Ada14], auch für Systeme mit einer schnellen Dynamik, realisieren.

Literaturverzeichnis

- [Ada14] ADAMY, J.: *Nichtlineare Systeme und Regelungen*. Berlin: Springer, 2014.
- [AAA⁺23] ALLEYNE, A.; ALLGÖWER, F.; AMES, A. u. a.: *Control for Societal-scale Challenges: Road Map 2030*. IEEE Control Systems Society, 2023.
- [ADA⁺22] AMADIO, F.; DALLA LIBERA, A.; ANTONELLO, R.; NIKOVSKI, D.; CARLI, R.; ROMERES, D.: Model-Based Policy Search using Monte Carlo Gradient Estimation with Real Systems Application. *IEEE Transactions on Robotics* 38 (2022), Nr. 6, S. 3879–3898.
- [AGH⁺19] ANDERSSON, J. A. E.; GILLIS, J.; HORN, G.; RAWLINGS, J. B.; DIEHL, M.: CasADi: A Software Framework for Nonlinear Optimization and Optimal Control. *Mathematical Programming Computation* 11 (2019), S. 1–36.
- [BG20] BERGMANN, D.; GRAICHEN, K.: Safe Bayesian Optimization under Unknown Constraints. *59th IEEE Conference on Decision and Control (CDC)*. 2020, S. 3592–3597.
- [BKS23] BERKENKAMP, F.; KRAUSE, A.; SCHOELLIG, A. P.: Bayesian Optimization with Safety Constraints: Safe and Automatic Parameter Tuning in Robotics. *Machine learning* 112 (2023), Nr. 10, S. 3713–3747.
- [Bet17] BETANCOURT, M.: *A Conceptual Introduction to Hamiltonian Monte Carlo*. 2017.
- [Bet10] BETTS, J. T.: *Practical Methods for Optimal Control and Estimation using Nonlinear Programming*. 2. Aufl. Advances in Design and Control. Society for Industrial and Applied Mathematics, 2010.
- [Bet98] BETTS, J. T.: Survey of Numerical Methods for Trajectory Optimization. *Journal of Guidance, Control, and Dynamics* 21 (1998), Nr. 2, S. 193–207.
- [Bij18] BIJL, H. J.: *Gaussian Process Regression Techniques with Applications to Wind Turbines*. Dissertation. Technische Universität Delft, 2018.
- [BNv⁺14] BISCHOFF, B.; NGUYEN-TUONG, D.; VAN HOOF, H.; McHUTCHON, A.; RASMUSSEN, C. E.; KNOLL, A.; PETERS, J.; DEISENROTH, M. P.: Policy Search for Learning Robot Control using Sparse Data. *IEEE International Conference on Robotics and Automation (ICRA)*. 2014, S. 3882–3887.
- [Bis06] BISHOP, C. M.: *Pattern Recognition and Machine Learning*. Computer Science. New York: Springer, 2006.

- [BBKK22] BRUNTON, S. L.; BUDIŠIĆ, M.; KAISER, E.; KUTZ, J. N.: Modern Koopman Theory for Dynamical Systems. *SIAM Review* 64 (2022), Nr. 2, S. 229–340.
- [CSPD16] CALANDRA, R.; SEYFARTH, A.; PETERS, J.; DEISENROTH, M. P.: Bayesian Optimization for Learning Gaits under Uncertainty. *Annals of Mathematics and Artificial Intelligence* 76 (2016), Nr. 2, S. 5–23.
- [CVS⁺20] CHATZILYGEROUDIS, K.; VASSILIADES, V.; STULP, F.; CALINON, S.; MOURET, J.-B.: A Survey on Policy Search Algorithms for Learning Robot Controllers in a Handful of Trials. *IEEE Transactions on Robotics* 36 (2020), Nr. 2, S. 328–347.
- [CHJA18] CHEN, Z.; HECKMAN, C.; JULIER, S. J.; AHMED, N. R.: Weak in the NEES?: Auto-Tuning Kalman Filters with Bayesian Optimization. *21st International Conference on Information Fusion (FUSION)* (2018), S. 1072–1079.
- [CH15] CUTLER, M.; HOW, J. P.: Efficient Reinforcement Learning for Robots using Informative Simulated Priors. *IEEE International Conference on Robotics and Automation (ICRA)*. 2015, S. 2605–2612.
- [DL13] DAMIANOU, A.; LAWRENCE, N. D.: Deep Gaussian Processes. *Proceedings of the Sixteenth International Conference on Artificial Intelligence and Statistics*. Bd. 31. Proceedings of Machine Learning Research. 2013, S. 207–215.
- [DFE20] DE ATH, G.; FIELDSEND, J. E.; EVERSON, R. M.: What do you mean? The Role of the Mean Function in Bayesian Optimisation. *Proceedings of the 2020 Genetic and Evolutionary Computation Conference Companion (GECCO)*. 2020, S. 1623–1631.
- [DFR15] DEISENROTH, M.; FOX, D.; RASMUSSEN, C.: Gaussian Processes for Data-Efficient Learning in Robotics and Control. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 37 (2015), S. 408–423.
- [Dei10] DEISENROTH, M. P.: *Efficient Reinforcement Learning using Gaussian Processes*. Dissertation. Karlsruher Institut für Technologie, 2010.
- [DEPF14] DEISENROTH, M. P.; ENGLERT, P.; PETERS, J.; FOX, D.: Multi-Task Policy Search for Robotics. *IEEE International Conference on Robotics and Automation (ICRA)*. 2014, S. 3876–3881.
- [DFO20] DEISENROTH, M. P.; FAISAL, A. A.; ONG, C. S.: *Mathematics for Machine Learning*. Cambridge University Press, 2020.

- [DHH09] DEISENROTH, M. P.; HUBER, M. F.; HANEBECK, U. D.: Analytic Moment-Based Gaussian Process Filtering. *Proceedings of the 26th Annual International Conference on Machine Learning*. New York, 2009, S. 225–232.
- [DRF12] DEISENROTH, M. P.; RASMUSSEN, C. E.; FOX, D.: Learning to Control a Low-Cost Manipulator using Data-Efficient Reinforcement Learning. *Robotics*. The MIT Press, 2012, S. 57–64.
- [DR13] DÍAZ-FRANCÉS, E.; RUBIO, F. J.: On the Existence of a Normal Approximation to the Distribution of the Ratio of two Independent Normal Random Variables. *Statistical Papers* 54 (2013), Nr. 2, S. 309–323.
- [DNM⁺17] DOERR, A.; NGUYEN-TUONG, D.; MARCO, A.; SCHAAAL, S.; TRIMPE, S.: Model-Based Policy Search for Automatic Tuning of Multivariate PID Controllers. *IEEE International Conference on Robotics and Automation (ICRA)* (2017), S. 5295–5301.
- [DVS17] DVS - DEUTSCHER VERBAND FÜR SCHWEISSEN UND VERWANDTE VERFAHREN: *Merkblatt DVS 2811 - Prüfverfahren für Drahtbondverbindungen*. 2017.
- [ED18] EMMERICH, M. T. M.; DEUTZ, A. H.: A Tutorial on Multiobjective Optimization: Fundamentals and Evolutionary Methods. *Natural computing* 17 (2018), Nr. 3, S. 585–609.
- [EPDP13] ENGLERT, P.; PARASCHOS, A.; DEISENROTH, M. P.; PETERS, J.: Probabilistic Model-Based Imitation Learning. *Adaptive Behavior* 21 (2013), Nr. 5, S. 388–403.
- [EM23] ESMZAD, R.; MODARES, H.: Maximum-Entropy Satisficing Linear Quadratic Regulator. *IEEE Control Systems Letters* 7 (2023), S. 3241–3246.
- [Föl94] FÖLLINGER, O.: *Optimale Regelung und Steuerung*. De Gruyter, 1994.
- [FKL⁺22] FÖLLINGER, O.; KONIGORSKI, U.; LOHMANN, B.; ROPPENECKER, G.; TRÄCHTLER, A.: *Regelungstechnik: Einführung in die Methoden und ihre Anwendung*. 13. Aufl. VDE Verlag, 2022.
- [FKDZ19] FRÖHLICH, L. P.; KLENSKE, E. D.; DANIEL, C.; ZEILINGER, M. N.: Bayesian Optimization for Policy Search in High-Dimensional Systems via Automatic Domain Selection. *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (2019), S. 757–764.
- [Gal16] GAL, Y.: *Uncertainty in Deep Learning*. Dissertation. University of Cambridge, 2016.

- [GMR16] GAL, Y.; MCALLISTER, R.; RASMUSSEN, C. E.: Improving PILCO with Bayesian Neural Network Dynamics Models. *Data-Efficient Machine Learning workshop, International Conference on Machine Learning*. 2016.
- [Gau09] GAUL, H.: *Berechnung der Verbindungsqualität beim Ultraschall-Wedge/Wedge-Bonden*. Dissertation. Technische Universität Berlin, 2009.
- [GM02] GAUSEMEIER, J.; MOEHRINGER, S.: VDI 2206- A New Guideline for the Design of Mechatronic Systems. *IFAC Proceedings Volumes* 35 (2002), Nr. 2, S. 785–790.
- [GK02] GEIGER, C.; KANZOW, C.: *Theorie und Numerik restringierter Optimierungsaufgaben*. Berlin: Springer, 2002.
- [Gei09] GEISSLER, U.: *Verbindungsbildung und Gefügeentwicklung beim Ultraschall-Wedge-Wedge-Bonden von AlSi1-Draht*. Dissertation. Technische Universität Berlin, 2009.
- [GDHL15] GONZÁLEZ, J. I.; DAI, Z.; HENNIG, P.; LAWRENCE, N. D.: Batch Bayesian Optimization via Local Penalization. *International Conference on Artificial Intelligence and Statistics*. 2015.
- [GTZ07] GRAICHEN, K.; TREUER, M.; ZEITZ, M.: Swing-up of the Double Pendulum on a Cart by Feedforward and Feedback Control with Experimental Validation. *Automatica* 43 (2007), Nr. 1, S. 63–71.
- [Han09] HANKE-BOURGEOIS, M.: *Grundlagen der Numerischen Mathematik und des Wissenschaftlichen Rechnens*. Wiesbaden: Vieweg+Teubner, 2009.
- [Har10] HARMAN, G. G.: *Wire Bonding in Microelectronics*. 3rd ed. McGraw-Hill's AccessEngineering. New York: McGraw-Hill, 2010.
- [HS12] HENNIG, P.; SCHULER, C.: Entropy Search for Information-Efficient Global Optimization. *Journal of Machine Learning Research* 13 (2012), S. 1809–1837.
- [Her20] HERBERMANN, J.; HESSE, M. (Betreuer/in); TRÄCHTLER, A. (Betreuer/in): *Modellierung des Ultraschalldrahtbondprozesses und Evaluierung der Bayes'schen Optimierung*. Studienarbeit. Universität Paderborn, 2020.
- [HHTT22] HESSE, M.; HUNSTIG, M.; TIMMERMANN, J.; TRÄCHTLER, A.: Batch Constrained Bayesian Optimization for Ultrasonic Wire Bonding Feed-Forward Control Design. *Proceedings of the 11th International Conference on Pattern Recognition, Applications and Methods (ICPRAM)*. 2022, S. 383–394.

- [HSTT24] HESSE, M.; SCHWARZER, L.; TIMMERMANN, J.; TRÄCHTLER, A.: Robust and Efficient Hybrid Optimal Control via Gaussian Process Regression and Multiple Shooting with Experimental Validation on a Double Pendulum on a Cart. *Proceedings of the 94th Annual Meeting of the International Association of Applied Mathematics and Mechanics (GAMM)*. 2024.
- [HTHT18] HESSE, M.; TIMMERMANN, J.; HÜLLERMEIER, E.; TRÄCHTLER, A.: A Reinforcement Learning Strategy for the Swing-up of the Double Pendulum on a Cart. *Procedia Manufacturing* 24 (2018). *4th International Conference on System-Integrated Intelligence: Intelligent, Flexible and Connected Systems in Products and Production (SysInt)*, S. 15–20.
- [HTT23] HESSE, M.; TIMMERMANN, J.; TRÄCHTLER, A.: Hybrid Optimal Control for Dynamical Systems using Gaussian Process Regression and Unscented Transform. *European Control Conference (ECC)*. 2023, S. 1–8.
- [HKZ17] HEWING, L.; KABZAN, J.; ZEILINGER, M. N.: Cautious Model Predictive Control using Gaussian Process Regression. *IEEE Transactions on Control Systems Technology* 28 (2017), S. 2736–2743.
- [HLZ18] HEWING, L.; LINIGER, A.; ZEILINGER, M. N.: Cautious NMPC with Gaussian Process Dynamics for Autonomous Miniature Race Cars. *European Control Conference (ECC)*. 2018, S. 1341–1348.
- [HW21] HÜLLERMEIER, E.; WAEGEMAN, W.: Aleatoric and Epistemic Uncertainty in Machine Learning: An Introduction to Concepts and Methods. *Machine learning* 110 (2021), Nr. 3, S. 457–506.
- [HSB⁺20] HUNSTIG, M.; SCHAERMANN, W.; BRÖKELMANN, M.; HOLTKAEMPER, S.; SIEPE, D.; HESSE, H. J.: Smart Ultrasonic Welding in Power Electronics Packaging. *11th International Conference on Integrated Power Electronics Systems (CIPS)*. 2020.
- [Ism18] ISMAIL, A.; HESSE, M. (Betreuer/in); TIMMERMANN, J. (Betreuer/in); TRÄCHTLER, A. (Betreuer/in): *Aufschwung und Stabilisierung eines Dreifachpendels auf einem Wagen mit experimenteller Validierung*. Masterarbeit. Universität Paderborn, 2018.
- [KD18] KAMTHE, S.; DEISENROTH, M.: Data-Efficient Reinforcement Learning with Probabilistic Model Predictive Control. *Proceedings of the Twenty-First International Conference on Artificial Intelligence and Statistics*. Bd. 84. 2018, S. 1701–1710.
- [Kel17] KELLY, M.: An Introduction to Trajectory Optimization: How to Do Your Own Direct Collocation. *SIAM Review* 59 (2017), Nr. 4, S. 849–904.

- [KKFH07] KO, J.; KLEIN, D. J.; FOX, D.; HAEHNEL, D.: GP-UKF: Unscented Kalman Filters with Gaussian Process Prediction and Observation Models. *IEEE/RSJ International Conference on Intelligent Robots and Systems*. 2007, S. 1901–1907.
- [Koh20] KOHLSTEDT, A.: *Modellbasierte Synthese einer hybriden Kraft-/Positionsregelung für einen Fahrzeugachsprüfstand mit hydraulischem Hexapod*. Dissertation. Universität Paderborn, 2020.
- [Kus64] KUSHNER, H. J.: A New Method of Locating the Maximum Point of an Arbitrary Multipeak Curve in the Presence of Noise. *Journal of Basic Engineering* 86 (1964), Nr. 1, S. 97–106.
- [Kuß06] KUSS, M.: *Gaussian Process Models for Robust Regression, Classification, and Reinforcement Learning*. Dissertation. Technische Universität Darmstadt, 2006.
- [LTW17] LONG, Y.; TWIEFEL, J.; WALLASCHEK, J.: A Review on the Mechanisms of Ultrasonic Wedge-Wedge Bonding. *Journal of Materials Processing Technology* 245 (2017), S. 241–258.
- [MBK⁺21] MARCO, A.; BAUMANN, D.; KHADIV, M.; HENNIG, P.; RIGHETTI, L.; TRIMPE, S.: Robot Learning With Crash Constraints. *IEEE Robotics and Automation Letters* 6 (2021), Nr. 2, S. 1439–1446.
- [MBH⁺17] MARCO, A.; BERKENKAMP, F.; HENNIG, P.; SCHOELLIG, A. P.; KRAUSE, A.; SCHAAL, S.; TRIMPE, S.: Virtual vs. Real: Trading off Simulations and Physical Experiments in Reinforcement Learning with Bayesian Optimization. *IEEE International Conference on Robotics and Automation (ICRA)*. 2017, S. 1557–1563.
- [MHB⁺16] MARCO, A.; HENNIG, P.; BOHG, J.; SCHAAL, S.; TRIMPE, S.: Automatic LQR Tuning based on Gaussian Process Global Optimization. *IEEE International Conference on Robotics and Automation (ICRA)*. 2016, S. 270–277.
- [Mat24] MATHWORKS INC.: *MATLAB*. 2024. <https://www.mathworks.com>.
- [MR17] MCALLISTER, R.; RASMUSSEN, C. E.: Data-Efficient Reinforcement Learning in Continuous State-Action Gaussian-POMDPs. *Advances in Neural Information Processing Systems*. Bd. 30. 2017.
- [Moc74] MOCKUS, J.: On Bayesian Methods for Seeking the Extremum. *Optimization Techniques*. 1974.

- [Mur13] MURPHY, K. P.: *Machine Learning: A Probabilistic Perspective*. 4. Aufl. Adaptive Computation and Machine Learning Series. Cambridge: MIT Press, 2013.
- [NMST20] NEUMANN-BROSIG, M.; MARCO, A.; SCHWARZMANN, D.; TRIMPE, S.: Data-Efficient Autotuning With Bayesian Optimization: An Industrial Control Study. *IEEE Transactions on Control Systems Technology* 28 (2020), Nr. 3, S. 730–740.
- [NW06] NOCEDAL, J.; WRIGHT, S. J.: *Numerical Optimization*. New York: Springer, 2006.
- [NH18] NOÈ, U.; HUSMEIER, D.: On a New Improvement-Based Acquisition Function for Bayesian Optimization. *ArXiv* (2018).
- [PLB12] PAPAGEORGIOU, M.; LEIBOLD, M.; BUSS, M.: *Optimierung*. Berlin: Springer, 2012.
- [RAMA19] RAI, A.; ANTONOVA, R.; MEIER, F.; ATKESON, C. G.: Using Simulation to Improve Sample-Efficiency of Bayesian Optimization for Bipedal Robots. *Journal of machine learning research* 20 (2019), Nr. 49, S. 1–24.
- [Ras06] RASMUSSEN, C. E.: *Gaussian Processes for Machine Learning*. Adaptive Computation and Machine Learning Series. Cambridge: MIT Press, 2006.
- [Rei23] REILING, F. P.; KOPPERT, S. (Betreuer/in); BAUSE, M. (Betreuer/in); TRÄCHTLER, A. (Betreuer/in): *Entwicklung und Umsetzung einer Mehrzieloptimierung am Ultraschalldrahtbondprozess unter Verwendung der Bayes'schen Optimierung*. Masterarbeit. Universität Paderborn, 2023.
- [RT10] REIST, P.; TEDRAKE, R.: Simulation-Based LQR-trees with Input and State Constraints. *IEEE International Conference on Robotics and Automation*. 2010, S. 5504–5510.
- [RWL19] ROWOLD, M.; WISCHNEWSKI, A.; LOHMANN, B.: Constrained Bayesian Optimization of a Linear Feed-Forward Controller. *IFAC-PapersOnLine* 52 (2019), Nr. 29, S. 1–6.
- [SAS⁺18] SCHEMMEL, R.; ALTHOFF, S.; SEXTRO, W.; UNGER, A.; HUNSTIG, M.; BRÖKELMANN, M.: Effects of Different Working Frequencies on the Joint Formation in Copper Wire Bonding. *10th International Conference on Integrated Power Electronics Systems (CIPS)*. 2018.
- [SKHS21] SCHEMMEL, R.; KRIEGER, V.; HEMSEL, T.; SEXTRO, W.: Co-Simulation of MATLAB and ANSYS for Ultrasonic Wire Bonding Process Optimization. *Microelectronics Reliability* 119 (2021), S. 114077.

- [SSH⁺20] SCHEMMEL, R.; SCHEIDEMANN, C.; HEMSEL, T.; KIRSCH, O.; SEXTRO, W.: Experimental Analysis and Modelling of Bond Formation in Ultrasonic Heavy Wire Bonding. *11th International Conference on Integrated Power Electronics Systems (CIPS)*. 2020.
- [Sch19] SCHÖN, O.; HESSE, M. (Betreuer/in); TIMMERMANN, J. (Betreuer/in); TRÄCHTLER, A. (Betreuer/in): *Effiziente Bayessche Optimierung durch Berücksichtigung A-priori-Wissens in Form eines physikalischen Dynamikmodells*. Studienarbeit. Universität Paderborn, 2019.
- [SKOT13] SCHWEERS, C.; KRUSE, D.; OESTERWINTER, T.; TRACHTLER, A.: Automated Design of an Unscented Kalman Filter for State- and Parameter Estimation on Unknown Models. *International Conference on Control, Automation, Robotics and Embedded Systems (CARE)*. 2013, S. 1–6.
- [SSW⁺16] SHAHRIARI, B.; SWERSKY, K.; WANG, Z.; ADAMS, R. P.; FREITAS, N. de: Taking the Human Out of the Loop: A Review of Bayesian Optimization. *Proceedings of the IEEE* 104 (2016), Nr. 1, S. 148–175.
- [Sim06] SIMON, D.: *Optimal State Estimation: Kalman, H [infinity], and Nonlinear Approaches*. Hoboken: Wiley-Interscience, 2006.
- [Sko18] SKOLAUT, W.: *Maschinenbau: Ein Lehrbuch für das ganze Bachelor-Studium*. Springer Vieweg, 2018.
- [SG05] SNELSON, E.; GHAHRAMANI, Z.: Sparse Gaussian Processes using Pseudo-inputs. *Advances in Neural Information Processing Systems*. Bd. 18. 2005.
- [SLA12] SNOEK, J.; LAROCHELLE, H.; ADAMS, R. P.: Practical Bayesian Optimization of Machine Learning Algorithms. *Advances in Neural Information Processing Systems*. Bd. 25. 2012.
- [SKKS12] SRINIVAS, N.; KRAUSE, A.; KAKADE, S. M.; SEEGER, M. W.: Information-Theoretic Regret Bounds for Gaussian Process Optimization in the Bandit Setting. *IEEE Transactions on Information Theory* 58 (2012), Nr. 5, S. 3250–3265.
- [SAA20] STENGER, D.; AY, M.; ABEL, D.: Robust Parametrization of a Model Predictive Controller for a CNC Machining Center using Bayesian Optimization. *IFAC-PapersOnLine* 53 (2020), Nr. 2, S. 10388–10394.
- [SB98] SUTTON, R. S.; BARTO, A.: *Reinforcement Learning: An Introduction*. 1. Aufl. Adaptive Computation and Machine Learning. Cambridge: The MIT Press, 1998.

- [SB18] SUTTON, R. S.; BARTO, A.: *Reinforcement Learning: An Introduction*. 2. Aufl. Adaptive Computation and Machine Learning. Cambridge: The MIT Press, 2018.
- [TKOT11] TIMMERMANN, J.; KHATAB, S.; OBER-BLÖBAUM, S.; TRÄCHTLER, A.: Discrete Mechanics and Optimal Control and its Application to a Double Pendulum on a Cart. *IFAC Proceedings Volumes* 44 (2011), S. 10199–10206.
- [Tim13] TIMMERMANN, J.: *Optimale Steuerung und Mehrzieloptimierung von dynamischen Systemen untersucht am Beispiel des Mehrfachpendels*. Dissertation. Universität Paderborn, 2013.
- [UBKH17] UMLAUFT, J.; BECKERS, T.; KIMMEL, M.; HIRCHE, S.: Feedback Linearization using Gaussian Processes. *IEEE 56th Annual Conference on Decision and Control (CDC)*. 2017, S. 5249–5255.
- [UH20] UMLAUFT, J.; HIRCHE, S.: Feedback Linearization Based on Gaussian Processes with Event-Triggered Online Learning. *IEEE Transactions on Automatic Control* 65 (2020), Nr. 10, S. 4154–4169.
- [Ung17] UNGER, A.: *Modellbasierte Mehrzieloptimierung zur Herstellung von Ultraschall-Drahtbondverbindungen in Leistungshalbleitermodulen*. Dissertation. Universität Paderborn, 2017.
- [USA⁺14] UNGER, A.; SEXTRO, W.; ALTHOFF, S.; MEYER, T.; BRÖKELMANN, M.; NEUMANN, K.; REIMANN, R. F.; GUTH, K.; BOLOWSKI, D.: Data-Driven Modeling of the Ultrasonic Softening Effect for Robust Copper Wire Bonding. *8th International Conference on Integrated Power Electronics Systems (CIPS)*. 2014.
- [vBS⁺20] VAN GOGH, B.; BENNER, T.; SEPPAENEN, H.; TSZENG, C.; SEPEHRBAND, P.: An Oxide Wear Model of Ultrasonic Bonding. *International Symposium on Microelectronics* (2020), S. 222–229.
- [Viv98] VIVARELLI, F.: *Studies on the Generalisation of Gaussian Processes and Bayesian Neural Networks*. Dissertation. Aston University, Birmingham, 1998.
- [VG15] VOLZ, A.; GRAICHEN, K.: Stochastic Model Predictive Control of Non-linear Continuous-Time Systems using the Unscented Transformation. *European Control Conference (ECC)*. 2015, S. 3365–3370.
- [WJ17] WANG, Z.; JEGETKA, S.: Max-Value Entropy Search for Efficient Bayesian Optimization. *Proceedings of the 34th International Conference on Machine Learning*. 2017, S. 3627–3635.

- [WHB⁺21] WEBER, D.; HEID, S.; BODE, H.; LANGE, J. H.; HÜLLERMEIER, E.; WALLSCHEID, O.: Safe Bayesian Optimization for Data-Driven Power Electronics Control Design in Microgrids: From Simulations to Real-World Experiments. *IEEE Access* 9 (2021), S. 35654–35669.
- [YEDB19] YANG, K.; EMMERICH, M.; DEUTZ, A.; BÄCK, T.: Multi-Objective Bayesian Global Optimization using Expected Hypervolume Improvement Gradient. *Swarm and Evolutionary Computation* 44 (2019), S. 945–956.
- [ZBP21] ZHU, M.; BEMPORAD, A.; PIGA, D.: Preference-Based MPC Calibration. *European Control Conference (ECC)*. 2021, S. 638–645.

Anhang

Inhaltsverzeichnis

A1 Mathematischer Anhang	188
A2 Sequentielle Quadratische Programmierung	194
A3 Details zu Experimenten	196

A1 Mathematischer Anhang

Beispielhafte Anwendung des Transformationssatzes für Pendeldynamik

Dieser Abschnitt dient dazu eine beispielhafte Anwendung des Transformationssatzes (2-12) vorzustellen. Den Ausgangspunkt stellt eine einfache Pendeldynamik in der nichtlinearen Zustandsraumdarstellung mit

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) = \begin{bmatrix} x_2 \\ -g_l \sin(x_1) - d_J x_2 \end{bmatrix},$$

dar. Hierin ist $x_1 = \varphi$ der Winkel, $x_2 = \dot{\varphi}$ die Winkelgeschwindigkeit und g_l, d_J sind systemspezifische Parameter. Für die weitere Betrachtung wird ein normalverteilter Zustandsvektor $\mathbf{x} \sim \mathcal{N}(\mathbf{m}, \mathbf{S}) = p_x(\mathbf{x})$ angenommen. Die Aufgabe besteht darin, die Verteilung der zeitlichen Ableitung $p(\dot{\mathbf{x}})$ zu bestimmen. Um diese Verteilung exakt zu bestimmen, wird der Transformationssatz angewandt, welcher in diesem Fall mit

$$p_{\dot{\mathbf{x}}}(\dot{\mathbf{x}}) = p_x(\mathbf{x} = \mathbf{f}^{-1}(\dot{\mathbf{x}})) \left| \frac{d\mathbf{f}^{-1}(\dot{\mathbf{x}})}{d\dot{\mathbf{x}}} \right| \quad (\text{A1-1})$$

angegeben wird. Für die Berechnung wird dementsprechend die Umkehrfunktion \mathbf{f}^{-1} der nichtlinearen Dynamikfunktion benötigt, welche aufgrund der trigonometrischen Funktion mehrere Lösungen

$$\mathbf{x} = \mathbf{f}^{-1}(\dot{\mathbf{x}}; i, k) = \begin{cases} \left[\pi + \arcsin\left(\frac{d_J \dot{x}_1 + \dot{x}_2}{g_l}\right) + 2\pi k, \dot{x}_1 \right]^T, & i = 1, \\ \left[2\pi k - \arcsin\left(\frac{d_J \dot{x}_1 + \dot{x}_2}{g_l}\right), \dot{x}_1 \right]^T, & i = 2, \end{cases}$$

$k \in \mathbb{Z}$, $-1 \leq \frac{d_J \dot{x}_1 + \dot{x}_2}{g_l} \leq 1$, beinhaltet und daher mit einer Fallunterscheidung $i = 1, 2$ und einem Index k angegeben wird. Hieraus kann die Determinante der Jacobimatrix mit

$$\left| \frac{d\mathbf{f}^{-1}(\dot{\mathbf{x}})}{d\dot{\mathbf{x}}} \right| = \frac{1}{\sqrt{|(d_J \dot{x}_1 + \dot{x}_2)^2 - g_l^2|}}, \quad (d_J \dot{x}_1 + \dot{x}_2)^2 \neq g_l^2,$$

angegeben werden. Das Einsetzen der berechneten Größen in (A1-1) liefert die exakte Verteilung

$$p_{\dot{\mathbf{x}}}(\dot{\mathbf{x}}) = \sum_{i=1}^2 \sum_{k=-\infty}^{\infty} \frac{p_x(\mathbf{f}^{-1}(\dot{\mathbf{x}}); i, k)}{\sqrt{|(d_J \dot{x}_1 + \dot{x}_2)^2 - g_l^2|}},$$

der zeitlichen Ableitung des Zustandsvektors.

Produkt und Integration zweier Normalverteilungen

Die nachfolgenden Zusammenhänge [Kuß06] gelten für die Größen $\mathbf{x}, \mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ und die Kovarianzmatrizen $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$. Für das Produkt zweier Normalverteilung kann dann geschlussfolgert werden, dass es proportional zu einer weiteren Normalverteilung nach

$$\mathcal{N}_x(\mathbf{a}, \mathbf{A}) \mathcal{N}_x(\mathbf{b}, \mathbf{B}) = z \mathcal{N}_x(\mathbf{c}, \mathbf{C}), \quad (\text{A1-2})$$

ist, wobei der Erwartungsvektor $\mathbf{c} \in \mathbb{R}^n$ und die Kovarianzmatrix $\mathbf{C} \in \mathbb{R}^{n \times n}$ mit

$$\mathbf{C} = (\mathbf{A}^{-1} + \mathbf{B}^{-1})^{-1} \quad \text{und} \quad \mathbf{c} = \mathbf{C}(\mathbf{A}^{-1}\mathbf{a} + \mathbf{B}^{-1}\mathbf{b})$$

angegeben werden. Die Normalisierungskonstante $z \in \mathbb{R}_+$ beinhaltet außerdem die Form einer Normalverteilung und lautet

$$z = (2\pi)^{-\frac{n}{2}} |\mathbf{ABC}^{-1}|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(\mathbf{a}^T \mathbf{A}^{-1} \mathbf{a} + \mathbf{b}^T \mathbf{B}^{-1} \mathbf{b} - \mathbf{c}^T \mathbf{C}^{-1} \mathbf{c})\right).$$

Für die Integration über zwei Normalverteilungen gilt ein ähnlicher Zusammenhang

$$\int \mathcal{N}_x(\mathbf{a}, \mathbf{A}) \mathcal{N}_a(\mathbf{b}, \mathbf{B}) d\mathbf{a} = \mathcal{N}_x(\mathbf{b}, \mathbf{A} + \mathbf{B}), \quad (\text{A1-3})$$

wobei das Ergebnis wiederum einer Normalverteilung entspricht. Zu beachten ist dabei, dass die zweite angegebene Normalverteilung nicht für die Zufallsvariable \mathbf{x} , sondern für \mathbf{a} definiert ist.

Berechnungen zur Zustandsvarianzmatrix für das Pendelbeispiel

In Abschnitt 2.3.2 wurde der Moment Matching Ansatz für ein Pendel vorgestellt, wobei die ersten beiden Momente exakt berechnet werden können. Der Erwartungsvektor $\mathbf{m}_{x_{k+1}}$ wurde im Hauptteil der Arbeit übersichtlich angegeben, wohingegen die Berechnung der Zustandsvarianzmatrix $\mathbf{S}_{x_{k+1}}$ mehr Platz benötigt und daher dem Anhang hinzugefügt wurde. Die Grundlage zur vollständigen Berechnung stellt die aufgestellte Gleichung für die Zustandsvarianzmatrix des Zeitschritts $k + 1$ mit

$$\mathbf{S}_{x_{k+1}}^{MM} = \mathbf{S}_{x_k} + \Delta t^2 \mathbb{V}[\mathbf{f}(\mathbf{x}_k, \mathbf{u}_k)] + \Delta t \left(\mathbb{C}[\mathbf{x}_k, \mathbf{f}(\mathbf{x}_k, \mathbf{u}_k)] + \mathbb{C}[\mathbf{x}_k, \mathbf{f}(\mathbf{x}_k, \mathbf{u}_k)]^T \right),$$

dar. Nachfolgend werden die einzelnen Bestandteile dieser Gleichung weiter aufgeschlüsselt. Für die Varianz der Dynamikfunktion folgt:

$$\mathbb{V}[\mathbf{f}(\mathbf{x}_k, \mathbf{u}_k)] = \begin{bmatrix} \mathbb{V}[\mathbf{f}(\mathbf{x}_k, \mathbf{u}_k)]^{(1,1)} & \mathbb{V}[\mathbf{f}(\mathbf{x}_k, \mathbf{u}_k)]^{(1,2)} \\ * & \mathbb{V}[\mathbf{f}(\mathbf{x}_k, \mathbf{u}_k)]^{(2,2)} \end{bmatrix}, \quad (\text{A1-4})$$

wobei das \star -Element aus der Symmetrie folgt und sich die anderen Elemente der Matrix zu

$$\begin{aligned}\mathbb{V}[\mathbf{f}(\mathbf{x}_k, \mathbf{u}_k)]^{(1,1)} &= S_{x_k}^{(2,2)}, \\ \mathbb{V}[\mathbf{f}(\mathbf{x}_k, \mathbf{u}_k)]^{(2,2)} &= g_l^2 (\mathbb{E}[\sin(x_k^{(1)})^2] - \mathbb{E}[\sin(x_k^{(1)})]^2) + \\ &\quad d_J^2 S_{x_k}^{(2,2)} + 2g_l d_J (\mathbb{E}[\sin(x_k^{(1)}) x_k^{(2)}] - \mathbb{E}[\sin(x_k^{(1)})] \mathbb{E}[x_k^{(2)}]) \\ &= g_l^2 \left(\frac{1}{2}(1 - \exp(-2S_{x_k}^{(1,1)}) \cos(2m_{x_k}^{(1)})) - (\exp(-\frac{1}{2}S_{x_k}^{(1,1)}) \sin(m_{x_k}^{(1)}))^2 \right) + \\ &\quad d_J^2 S_{x_k}^{(2,2)} + 2g_l d_J (\exp(-\frac{1}{2}S_{x_k}^{(1,1)}) (S_{x_k}^{(2,1)} \cos(m_{x_k}^{(1)}) + m_{x_k}^{(2)} \sin(m_{x_k}^{(1)})) - \\ &\quad \exp(-\frac{1}{2}S_{x_k}^{(1,1)}) \sin(m_{x_k}^{(1)}) m_{x_k}^{(2)}), \\ \mathbb{V}[\mathbf{f}(\mathbf{x}_k, \mathbf{u}_k)]^{(1,2)} &= -g_l \exp(-\frac{1}{2}S_{x_k}^{(1,1)}) S_{x_k}^{(2,1)} \cos(m_{x_k}^{(1)}) - d_J S_{x_k}^{(2,2)},\end{aligned}$$

ergeben. Des Weiteren wird die Kreuz-Kovarianz zwischen dem aktuellen Zustand und der Dynamikfunktion benötigt:

$$\begin{aligned}\mathbb{C}[\mathbf{x}_k, \mathbf{f}(\mathbf{x}_k, \mathbf{u}_k)] &= \mathbb{E}[\mathbf{x}_k \mathbf{f}(\mathbf{x}_k, \mathbf{u}_k)^T] - \mathbb{E}[\mathbf{x}_k] \mathbb{E}[\mathbf{f}(\mathbf{x}_k, \mathbf{u}_k)]^T \\ &= \begin{bmatrix} \mathbb{E}[x_k^{(1)} x_k^{(2)}], & -g_l \mathbb{E}[x_k^{(1)} \sin(x_k^{(1)})] - d_J \mathbb{E}[x_k^{(1)} x_k^{(2)}] + J_i u_k \mathbb{E}[x_k^{(1)}] \\ \mathbb{E}[x_k^{(2)}]^2, & -g_l \mathbb{E}[x_k^{(2)} \sin(x_k^{(1)})] - d_J \mathbb{E}[x_k^{(2)}]^2 + J_i u_k \mathbb{E}[x_k^{(2)}] \end{bmatrix} \\ &- \begin{bmatrix} m_{x_k}^{(1)} \mathbb{E}[x_k^{(2)}] & m_{x_k}^{(1)} (-g_l \mathbb{E}[\sin(x_k^{(1)})] - d_J \mathbb{E}[x_k^{(2)}] + J_i u_k) \\ m_{x_k}^{(2)} \mathbb{E}[x_k^{(2)}] & m_{x_k}^{(2)} (-g_l \mathbb{E}[\sin(x_k^{(1)})] - d_J \mathbb{E}[x_k^{(2)}] + J_i u_k) \end{bmatrix} \\ &= \begin{bmatrix} S_{x_k}^{(2,1)} & -g_l \exp(-\frac{1}{2}S_{x_k}^{(1,1)}) S_{x_k}^{(1,1)} \cos(m_{x_k}^{(1)}) - d_J S_{x_k}^{(2,1)} \\ S_{x_k}^{(2,2)} & -g_l \exp(-\frac{1}{2}S_{x_k}^{(1,1)}) S_{x_k}^{(2,1)} \cos(m_{x_k}^{(1)}) - d_J S_{x_k}^{(2,2)} \end{bmatrix}\end{aligned}\tag{A1-5}$$

Theorem für die Inverse einer partitionierten Matrix

Die Beschreibung der nachfolgenden Theorems folgt derer in [Mur13]. Betrachtet wird eine partitionierte Matrix

$$\mathbf{P} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix},\tag{A1-6}$$

deren Inverse \mathbf{P}^{-1} gesucht wird. Zur Bestimmung wird als Grundidee der Ansatz verfolgt, nach welchem P zunächst in Diagonalform gebracht und anschließend invertiert wird. Als Erstes wird dafür die Matrixmultiplikation

$$\underbrace{\begin{bmatrix} \mathbf{I} & -\mathbf{B}\mathbf{D}^{-1} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}}_{:=\mathbf{X}} \underbrace{\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}}_{=\mathbf{P}} = \begin{bmatrix} \mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C} & \mathbf{0} \\ \mathbf{C} & \mathbf{D} \end{bmatrix},$$

eingeführt. Im zweiten Schritt erfolgt eine weitere Matrixmultiplikation von der rechten Seite über den Zusammenhang

$$\underbrace{\begin{bmatrix} A - BD^{-1}C & 0 \\ C & D \end{bmatrix}}_{:=XP} \underbrace{\begin{bmatrix} I & 0 \\ -D^{-1}C & I \end{bmatrix}}_{:=Z} = \underbrace{\begin{bmatrix} A - BD^{-1}C & 0 \\ 0 & D \end{bmatrix}}_{:=W}.$$

Aufgrund ihrer Diagonalform ist die Inverse der Matrix $\mathbf{W} = \mathbf{XPZ}$ leicht zu bestimbar. Es folgt weiterführend

$$\mathbf{W}^{-1} = (\mathbf{XPZ})^{-1} = \mathbf{Z}^{-1}\mathbf{P}^{-1}\mathbf{X}^{-1} \Leftrightarrow \mathbf{P}^{-1} = \mathbf{ZW}^{-1}\mathbf{X},$$

womit sich die gesuchte Inverse zu

$$\begin{aligned} \mathbf{P}^{-1} &= \begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} I & 0 \\ -D^{-1}C & I \end{bmatrix} \begin{bmatrix} A - BD^{-1}C & 0 \\ 0 & D \end{bmatrix}^{-1} \begin{bmatrix} I & -BD^{-1} \\ 0 & I \end{bmatrix} \\ &= \begin{bmatrix} M^{-1} & -M^{-1}BD^{-1} \\ -D^{-1}CM^{-1} & D^{-1} + D^{-1}CM^{-1}BD^{-1} \end{bmatrix} \end{aligned}$$

ergibt und die Definition des sogenannten Schur-Komplements [Mur13] von \mathbf{P} mit Bezug auf \mathbf{D} über $M := A - BD^{-1}C$ genutzt wurde.

Ableitung der Kovarianzmatrix nach den Hyperparametern des SE-Kernels

Für die Optimierung der Hyperparameter wird im Rahmen der Gauß-Prozess-Regression die logarithmische Likelihoodfunktion (2-33) maximiert. Eine Möglichkeit dieses Optimierungsproblems effizient zu lösen, bietet die Klasse der gradientenbasierten Optimierer. Hierzu ist allerdings die Kenntnis der partiellen Ableitungen (2-34) notwendig. Für den SE-Kernel (2-32) soll die analytische Berechnung in diesem Abschnitt beispielhaft durchgeführt werden. Im Folgenden wird von der SE-Kovarianzfunktion in der Darstellung

$$k(\mathbf{x}, \mathbf{x}'; \boldsymbol{\theta}) = \sigma_f^2 \exp(-\frac{1}{2}(\mathbf{x} - \mathbf{x}')^T \mathbf{W}^{-1}(\mathbf{x} - \mathbf{x}')) + \sigma_n^2 \delta(\mathbf{x}, \mathbf{x}'),$$

ausgegangen, wobei der Rauschterm an dieser Stelle vereinfachend zum SE-Kernel hinzugezählt und die Diracfunktion δ benutzt wird. Es gilt $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^{n_x}$, $\mathbf{W} = \text{diag}(l_1^2, \dots, l_{n_x}^2)$ und $\boldsymbol{\theta} := [l_1, \dots, l_{n_x}, \sigma_f, \sigma_n]^T \in \mathbb{R}^{n_x+2}$. Des Weiteren liegen die Eingangsdaten über die Matrix $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_{n_d}] \in \mathbb{R}^{n_x \times n_d}$ vor. Die eingangsdatenbasierte Kovarianzmatrix lautet somit

$$\mathbf{K}(\boldsymbol{\theta}) = \begin{bmatrix} k(\mathbf{x}_1, \mathbf{x}_1; \boldsymbol{\theta}) & \dots & k(\mathbf{x}_1, \mathbf{x}_{n_d}; \boldsymbol{\theta}) \\ \vdots & \ddots & \vdots \\ k(\mathbf{x}_{n_d}, \mathbf{x}_1; \boldsymbol{\theta}) & \dots & k(\mathbf{x}_{n_d}, \mathbf{x}_{n_d}; \boldsymbol{\theta}) \end{bmatrix} \in \mathbb{R}^{n_d \times n_d},$$

sodass jedes Element von der Hyperparametrierung abhängt. In Hinblick auf (2-34) wird die Ableitung dieser Matrix nach den Hyperparametern benötigt. Die grundsätzliche Gestalt ist somit ein Tensor mit

$$\frac{d\mathbf{K}}{d\boldsymbol{\theta}} = \left[\begin{array}{ccc} \frac{\partial k_{1,1}}{\partial \theta_i} & \cdots & \frac{\partial k_{1,n_d}}{\partial \theta_i} \\ \vdots & \ddots & \vdots \\ \frac{\partial k_{n_d,1}}{\partial \theta_i} & \cdots & \frac{\partial k_{n_d,n_d}}{\partial \theta_i} \end{array} \right]_i \in \mathbb{R}^{n_d \times n_d \times (n_x+2)}, \quad (\text{A1-7})$$

wonach die einzelnen Teilmatrizen $i = 1, \dots, n_x + 2$ die partiellen Ableitungen der Kovarianzfunktion nach dem spezifischen Hyperparameter beinhalten. Zur besseren Übersicht wurde $k_{i,j} := k(\mathbf{x}_i, \mathbf{x}_j; \boldsymbol{\theta})$ mit $i, j = 1, \dots, n_d$ eingeführt.

Weiterführend wird stellvertretend für alle relevanten Elemente dieses Tensors der (i, j) -te Eintrag betrachtet und zunächst die Ableitung nach dem k -ten Lengthscale-Parameter ausgewertet. Es gilt

$$\begin{aligned} \frac{\partial k_{i,j}}{\partial l_k} &= \frac{\partial}{\partial l_k} (\sigma_f^2 \exp(-\frac{1}{2}(\mathbf{x}_i - \mathbf{x}_j)^T \mathbf{W}^{-1}(\mathbf{x}_i - \mathbf{x}_j)) + \sigma_n^2 \delta(\mathbf{x}_i, \mathbf{x}_j)) \\ &= \sigma_f^2 \frac{\partial}{\partial l_k} \exp\left(-\frac{1}{2} \sum_{d=1}^{n_x} (x_{i,d} - x_{j,d})^2 W_{d,d}^{-1}\right) \\ &= \sigma_f^2 \frac{\partial}{\partial l_k} \exp\left(-\frac{(x_{i,k} - x_{j,k})^2}{2l_k^2} - \sum_{d=1, d \neq k}^{n_x} \frac{(x_{i,d} - x_{j,d})^2}{2l_d^2}\right) \\ &= \sigma_f^2 \exp\left(-\sum_{d=1, d \neq k}^{n_x} \frac{(x_{i,d} - x_{j,d})^2}{2l_d^2}\right) \frac{(x_{i,k} - x_{j,k})^2}{l_k^3} \exp\left(-\frac{(x_{i,k} - x_{j,k})^2}{2l_k^2}\right) \\ &= \frac{(x_{i,k} - x_{j,k})^2}{l_k^3} \sigma_f^2 \exp(-\frac{1}{2}(\mathbf{x}_i - \mathbf{x}_j)^T \mathbf{W}^{-1}(\mathbf{x}_i - \mathbf{x}_j)) \\ &= \frac{(x_{i,k} - x_{j,k})^2}{l_k^3} (k(\mathbf{x}_i, \mathbf{x}_j) - \sigma_n^2 \delta(\mathbf{x}_i, \mathbf{x}_j)), \end{aligned}$$

für $i, j = 1, \dots, n_d$ und $k = 1, \dots, n_x$. Insgesamt folgt somit im Zusammenhang mit der Kovarianzmatrix und in Hinblick auf eine effiziente Implementierung:

$$\frac{\partial \mathbf{K}}{\partial l_k} = \frac{1}{l_k^3} \left[\begin{array}{cccc} 0 & (x_{1,k} - x_{2,k})^2 & \dots & (x_{1,k} - x_{n_d,k})^2 \\ (x_{2,k} - x_{1,k})^2 & 0 & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ (x_{n_d,k} - x_{1,k})^2 & (x_{n_d,k} - x_{2,k})^2 & \dots & 0 \end{array} \right] \odot (\mathbf{K} - \sigma_n^2 \mathbf{I}_{n_d}).$$

Die restlichen Ableitungen bzgl. der Hyperparameter σ_f und σ_n sind vergleichsweise einfach zu berechnen und lassen sich kompakt angeben mit

$$\frac{\partial \mathbf{K}}{\partial \sigma_f} = \frac{2}{\sigma_f} (\mathbf{K} - \sigma_n^2 \mathbf{I}_{n_d}) \quad \text{und} \quad \frac{\partial \mathbf{K}}{\partial \sigma_n} = 2\sigma_n \mathbf{I}_{n_d}.$$

Ableitung des GP-Posterior-Erwartungsvektors nach dem erweiterten Zustand

In (4-18) wird der GP-Posterior-Erwartungsvektor $\Delta\boldsymbol{\mu}(\mathbf{x}^u) \in \mathbb{R}^{n_x}$ zur Abbildung der vorhandenen Modellfehler aufgestellt. In diesem Abschnitt wird die Ableitung nach dem erweiterten Zustand $\mathbf{x}^u = [\mathbf{x}^T, \mathbf{u}^T]^T \in \mathbb{R}^{n_x+n_u}$ vorgestellt. Ziel ist es, die Jacobimatrix

$$\frac{d}{d\mathbf{x}^u} \Delta\boldsymbol{\mu}(\mathbf{x}^u) = \frac{d}{d\mathbf{x}^u} [\Delta\mu_1(\mathbf{x}^u), \dots, \Delta\mu_{n_x}(\mathbf{x}^u)]^T = [\mathbf{A}_\Delta, \mathbf{B}_\Delta] \in \mathbb{R}^{n_x \times (n_x+n_u)}$$

und damit die Teilmatrizen $\mathbf{A}_\Delta \in \mathbb{R}^{n_x \times n_x}$ und $\mathbf{B}_\Delta \in \mathbb{R}^{n_x \times n_u}$, welche für eine Linearisierung notwendig sind, aufzustellen. Zunächst wird der Gradient der i -ten Zeile über

$$\frac{d}{d\mathbf{x}^u} \Delta\mu_i(\mathbf{x}^u) = \frac{d}{d\mathbf{x}^u} \mathbf{k}_{X_k^U}(\mathbf{x}^u)^T \underbrace{\mathbf{K}_{\eta_i^*}^{-1} \Delta \mathbf{X}_{k+1}^{(i,:)} T}_{=: \boldsymbol{\alpha}} = \sum_{j=1}^{n_d} \frac{d}{d\mathbf{x}^u} k(\mathbf{x}^u, \mathbf{X}_k^{U(j)}) \alpha_j,$$

berechnet. Dementsprechend muss die verwendete Kovarianzfunktion nach dem erweiterten Zustand abgeleitet werden. Im Folgenden wird diese Ableitung beispielhaft für den SE-Kernel (2-32) berechnet:

$$\begin{aligned} \frac{d}{d\mathbf{x}^u} k_{SE}(\mathbf{x}^u, \mathbf{X}_k^{U(j)}) &= \frac{d}{d\mathbf{x}^u} \sigma_f^2 \exp(-(\mathbf{x}^u - \mathbf{X}_k^{U(j)})^T \mathbf{W}^{-1}(\mathbf{x}^u - \mathbf{X}_k^{U(j)})) \\ &= \frac{d}{d\mathbf{x}^u} (-(\mathbf{x}^u - \mathbf{X}_k^{U(j)})^T \mathbf{W}^{-1}(\mathbf{x}^u - \mathbf{X}_k^{U(j)})) k_{SE}(\mathbf{x}^u, \mathbf{X}_k^{U(j)}) \\ &= 2\mathbf{W}^{-1}(\mathbf{x}^u - \mathbf{X}_k^{U(j)}) k_{SE}(\mathbf{x}^u, \mathbf{X}_k^{U(j)}). \end{aligned}$$

Eine effiziente Berechnung des Gradienten wird über

$$\begin{aligned} \frac{d}{d\mathbf{x}^u} \Delta\mu_i(\mathbf{x}^u) &= \sum_{j=1}^{n_d} 2\mathbf{W}^{-1}(\mathbf{x}^u - \mathbf{X}_k^{U(j)}) k_{SE}(\mathbf{x}^u, \mathbf{X}_k^{U(j)}) \alpha_j \\ &= 2\mathbf{W}^{-1} \underbrace{\begin{bmatrix} \mathbf{x}^u - \mathbf{X}_k^{U(1)} & \dots & \mathbf{x}^u - \mathbf{X}_k^{U(n_d)} \end{bmatrix}}_{=: \tilde{\mathbf{X}}_k^U} \begin{bmatrix} k_{SE}(\mathbf{x}^u, \mathbf{X}_k^{U(1)}) \alpha_1 \\ \vdots \\ k_{SE}(\mathbf{x}^u, \mathbf{X}_k^{U(n_d)}) \alpha_{n_d} \end{bmatrix} \\ &= 2\mathbf{W}^{-1} \tilde{\mathbf{X}}_k^U (\mathbf{k}_{SE, X_k^U}(\mathbf{x}^u; \tilde{\mathbf{X}}_k^U) \odot \boldsymbol{\alpha}), \end{aligned}$$

erreicht. Anschließend setzt sich die Jacobimatrix aus den einzelnen Gradienten mit $i = 1, \dots, n_x$ zusammen.

A2 Sequentielle Quadratische Programmierung

Für die Lösung des hybriden Optimalsteuerungsproblems (4-28) wird im Rahmen dieser Arbeit das SQP-Verfahren [GK02; PLB12; NW06] verwendet, welches in diesem Abschnitt auszugsweise vorgestellt werden soll. Zur Erläuterung wird von einem allgemeinen nichtlinearen Optimierungsproblem der Form

$$\min_x f(\mathbf{x}), \quad \text{u.B.v.} \quad \mathbf{g}(\mathbf{x}) \leq \mathbf{0}, \quad \mathbf{h}(\mathbf{x}) = \mathbf{0},$$

mit der skalaren Zielfunktion $f(\mathbf{x}) : \mathbb{R}^{n_x} \mapsto \mathbb{R}$ und Ungleichheitsnebenbedingungen $\mathbf{g}(\mathbf{x}) : \mathbb{R}^{n_x} \mapsto \mathbb{R}^{n_g}$, sowie Gleichheitsnebenbedingungen $\mathbf{h}(\mathbf{x}) : \mathbb{R}^{n_x} \mapsto \mathbb{R}^{n_h}$, welche von dem Optimierungsvektor $\mathbf{x} \in \mathbb{R}^{n_x}$ abhängen, ausgegangen. Eine Lösung des Problems muss die sogenannten *Karush-Kuhn-Tucker*-(KKT)-Bedingungen, welche formal mit

$$\begin{aligned} \nabla_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) &= \mathbf{0}, \\ \mathbf{h}(\mathbf{x}) &= \mathbf{0}, \\ g_i(\mathbf{x}) &\leq 0, \quad \lambda_i \geq 0, \quad \lambda_i g_i(\mathbf{x}) = 0, \quad i = 1, \dots, n_g, \end{aligned} \tag{A2-1}$$

angegeben werden, erfüllen. Hierin ist L die Lagrange-funktion, welche über die zugehörigen Lagrangenmultiplikatoren $\boldsymbol{\lambda} \in \mathbb{R}^{n_g}$ und $\boldsymbol{\mu} \in \mathbb{R}^{n_h}$ berechnet wird, und über

$$L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) := f(\mathbf{x}) + \sum_{i=1}^{n_g} \lambda_i g_i(\mathbf{x}) + \sum_{j=1}^{n_h} \mu_j h_j(\mathbf{x}),$$

definiert ist.

Algorithmus 6 SQP-Verfahren für restriktierte Optimierungsprobleme nach [GK02]

- 1: **Eingabe:** Initiale Schätzung $\mathbf{x}_0, \boldsymbol{\lambda}_0, \boldsymbol{\mu}_0$ und \mathbf{H}_0 , Iterationsindex $k = 0$.
 - 2: **Wiederhole** bis $(\mathbf{x}_k, \boldsymbol{\lambda}_k, \boldsymbol{\mu}_k)$ ein KKT-Punkt nach (A2-1) ist:
 - 3: Berechne eine Lösung $\Delta \mathbf{x}_k$ des quadratischen Teilproblems (A2-2) mit zugehörigen Lagrangenmultiplikatoren $(\boldsymbol{\lambda}_{k+1}, \boldsymbol{\mu}_{k+1})$.
 - 4: Setze $\mathbf{x}_{k+1} = \mathbf{x}_k + \Delta \mathbf{x}_k$, wähle \mathbf{H}_{k+1} (z. B. BFGS) und inkrementiere $k \leftarrow k + 1$.
 - 5: **Ausgabe:** Optimal Lösung \mathbf{x}^* und zugehörige Lagrangenmultiplikatoren $(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$.
-

Algorithmus 6 fasst die wesentlichen Schritte des SQP-Verfahrens zusammen. Die Grundidee ist es, sich iterativ der optimalen Lösung anzunähern, indem in jeder Iteration das exakte Optimierungsproblem durch ein quadratisches Teilproblem

$$\begin{aligned} \min_{\Delta x} & \nabla_x f(\mathbf{x}_k)^T \Delta \mathbf{x} + \frac{1}{2} \Delta \mathbf{x}^T \mathbf{H}_k \Delta \mathbf{x}, \quad \text{u.B.v.} \\ & g_i(\mathbf{x}_k) + \nabla_x g_i(\mathbf{x}_k)^T \Delta \mathbf{x} \leq 0, \quad i = 1, \dots, n_g, \\ & h_j(\mathbf{x}_k) + \nabla_x h_j(\mathbf{x}_k)^T \Delta \mathbf{x} = 0, \quad j = 1, \dots, n_h, \end{aligned} \quad (\text{A2-2})$$

approximiert wird. Dabei wird die Zielfunktion durch eine quadratische Annäherung und die Nebenbedingungen durch lineare Approximationen beschrieben. Zur Lösung des quadratischen Teilproblems werden weitere Verfahren eingebunden, bspw. die Strategie der aktiven Menge für quadratische Programme, welche zur Bestimmung einer initialen zulässigen Lösung wiederum auf der Lösung eines linearen Programms aufgebaut ist [GK02]. Des Weiteren wird aufgrund der meist aufwendigen Berechnung nicht die exakte Hessematrix $\mathbf{H} = \nabla_{xx}^2 f(\mathbf{x})$ für die quadratische Approximation der Zielfunktion verwendet, sondern eine initiale Schätzung \mathbf{H}_0 (oftmals Einheitsmatrix \mathbf{I}_{n_x}) in Kombination mit einer geeigneten Aufdatierungsvorschrift, bspw. nach *Broyden–Fletcher–Goldfarb–Shanno* (BFGS).

A3 Details zu Experimenten

Parametrisierte Steuerung

Die Stellgrößen des Ultraschalldrahtbondprozesses sind die Normalkraft $F_N(t)$ und die elektrische Spannung $U_S(t)$. Die Spannung wird durch eine hochfrequente Schwingung über $U_S(t) = \hat{U}_S(t) \sin(\omega t)$ beschrieben. Im Rahmen der Experimente dieser Arbeit wurde eine parametrisierte Ansatzfunktion für die Größen $(F_N(t; \boldsymbol{\theta}), \hat{U}_S(t; \boldsymbol{\theta}))$, welche auf Expertenwissen basiert, mit

$$F_N(t; \boldsymbol{\theta}) = \begin{cases} F_0, & t < 0, \\ F_1, & 0 \leq t < T_1, \\ F_1 + \frac{F_2 - F_1}{T_{2,R}}(t - T_1), & T_1 \leq t < T_1 + T_{2,R}, \\ F_2, & T_1 + T_{2,R} \leq t < T_1 + T_2, \\ 0, & \text{sonst,} \end{cases} \quad (\text{A3-1})$$

$$\hat{U}_S(t; \boldsymbol{\theta}) = \begin{cases} \frac{\hat{U}_1}{T_{1,R}}t, & 0 \leq t < T_{1,R}, \\ \hat{U}_1, & T_{1,R} \leq t < T_1, \\ \hat{U}_1 + \frac{\hat{U}_2 - \hat{U}_1}{T_{2,R}}(t - T_1), & T_1 \leq t < T_1 + T_{2,R}, \\ \hat{U}_2, & T_1 + T_{2,R} \leq t < T_1 + T_2, \\ 0, & \text{sonst,} \end{cases}$$

definiert. Diese stückweise definierte Funktion orientiert sich an den Verbindungsphasen beim Bondprozess (vgl. Bild 3-3).

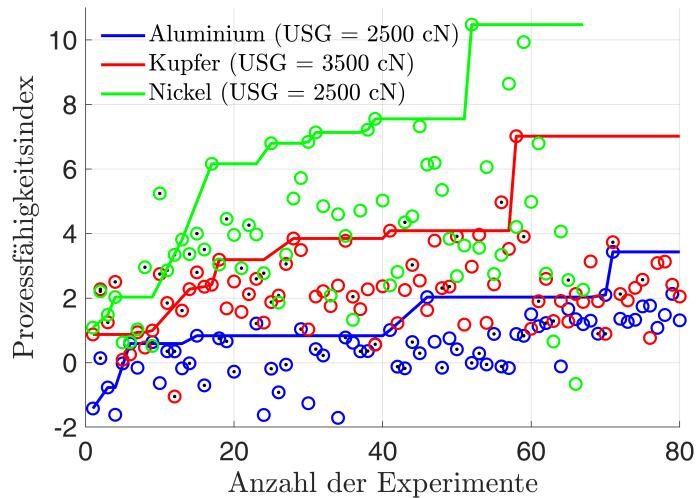


Bild A3-1: Anwendung der Bayesschen Optimierung zur Identifikation einer parametrisierten Steuerung beim Ultraschalldrahtboden. Erweiterte Auswertung für die Materialien Kupfer und Nickel.

Ergebnisse zu weiteren Materialien

Bild A3-1 zeigt eine erweiterte Auswertung der Bayesschen Optimierung (Algorithmus 2). Hierbei wurden die Materialien Kupfer und Nickel als Verbindungspartner verwendet. Als Referenz dient der blaue Verlauf der Aluminium-Experimente des Hauptkapitels (s. Bild 3-7). Die durchgezogenen Linien repräsentieren die Experimente, in denen eine Verbesserung stattgefunden hat. Hierfür muss sich der Prozessfähigkeitsindex erhöhen und es darf zu keinen Werkzeugaufsetzern gekommen sein. Diesbezüglich stellen die Punkte jedes einzelnen Experimente dar, wobei ein schwarzer Punkt innerhalb des Kreises dafür steht, dass es zu einer Werkzeugkollision gekommen ist.

MC-Simulationen zur hybriden Optimalsteuerung am Doppelpendelprüfstand

Um die Stabilität des geschlossenen Regelkreises besser bewerten zu können, wurden ergänzende Zustandspropagationen berechnet. Dabei wurde anstelle der UT das genauere MC-Verfahren mit 100 Partikeln eingesetzt. Die Bilder A3-2 und A3-3 zeigen die verschiedenen MC-Trajektorien in grau und zum Vergleich in türkis die UT-Propagation. In Bezug zu der 1. Iteration ist erkennbar, dass einige Trajektorien nach ungefähr 1.5 s instabil werden und ein sicherer Betrieb der Regelung daher nicht gewährleistet werden kann.

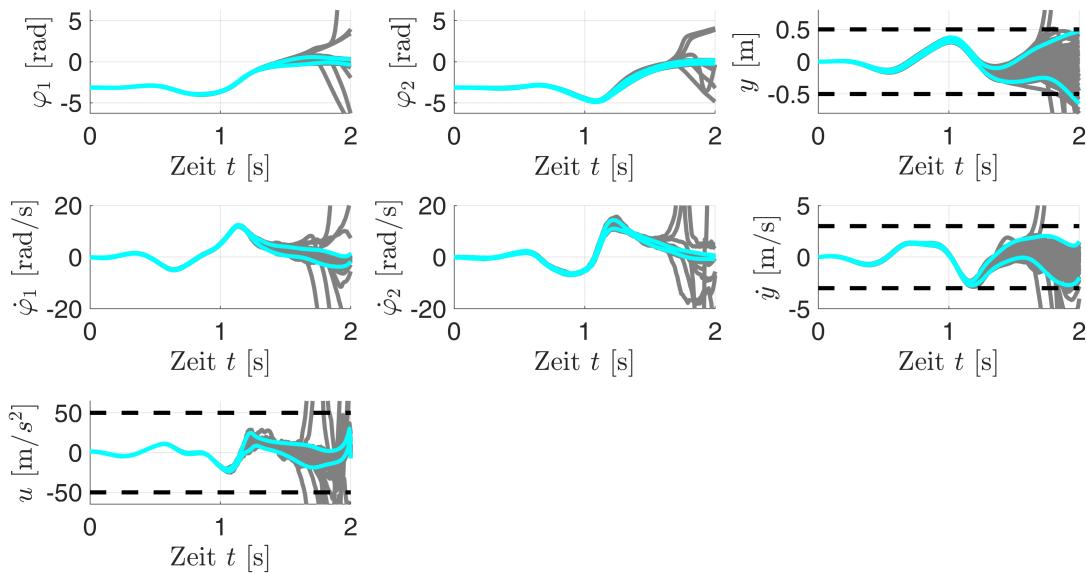


Bild A3-2: Ergänzende Monte-Carlo-Simulation des geschlossenen Regelkreises der 1. Lerniteration.

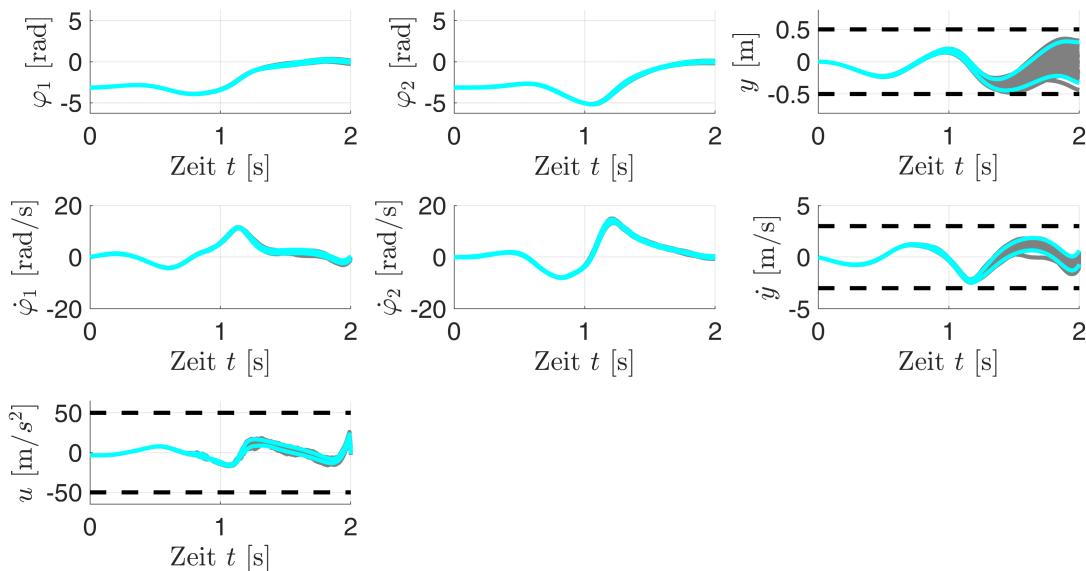


Bild A3-3: Ergänzende Monte-Carlo-Simulation des geschlossenen Regelkreises der 5. Lerniteration.

Erklärung zur Zitation von Inhalten aus studentischen Arbeiten

In Ergänzung zu meinem Antrag auf Zulassung zur Promotion in der Fakultät für Maschinenbau der Universität Paderborn erkläre ich gemäß §11 der Promotionsordnung und unter Beachtung der Regelung zur Zitation studentischer Arbeiten:

Die von mir vorgelegte Dissertation habe ich selbstständig verfasst, **und ich habe keine anderen** als die dort angegebenen Quellen und Hilfsmittel benutzt. Es sind **Inhalte** studentischen Ursprungs (studentische Arbeiten) in dieser Dissertation enthalten.

Ich habe die verwendeten Arbeiten entsprechend der Regelung „Zitation aus studentischen Arbeiten in Dissertationen“ zitiert.

Paderborn, Juni 2024 *Michael Hesse*

Das Heinz Nixdorf Institut – Interdisziplinäres Forschungszentrum für Informatik und Technik

Das Heinz Nixdorf Institut ist ein Forschungszentrum der Universität Paderborn. Es entstand 1987 aus der Initiative und mit Förderung von Heinz Nixdorf. Damit wollte er Ingenieurwissenschaften und Informatik zusammenführen, um wesentliche Impulse für neue Produkte und Dienstleistungen zu erzeugen. Dies schließt auch die Wechselwirkungen mit dem gesellschaftlichen Umfeld ein.

Die Forschungsarbeit orientiert sich an dem Programm „Dynamik, Vernetzung, Autonomie: Neue Methoden und Technologien für die intelligenten technischen Systeme von morgen“. In der Lehre engagiert sich das Heinz Nixdorf Institut in Studiengängen der Informatik, der Ingenieurwissenschaften und der Wirtschaftswissenschaften.

Heute wirken am Heinz Nixdorf Institut acht Professoren/in mit insgesamt 120 Mitarbeiterinnen und Mitarbeitern. Pro Jahr promovieren hier etwa 15 Nachwuchswissenschaftlerinnen und Nachwuchswissenschaftler.

Heinz Nixdorf Institute – Interdisciplinary Research Centre for Computer Science and Technology

The Heinz Nixdorf Institute is a research centre within the University of Paderborn. It was founded in 1987 initiated and supported by Heinz Nixdorf. By doing so he wanted to create a symbiosis of computer science and engineering in order to provide critical impetus for new products and services. This includes interactions with the social environment.

Our research is aligned with the program “Dynamics, Networking, Autonomy: New methods and technologies for intelligent technical systems of tomorrow”. In training and education the Heinz Nixdorf Institute is involved in many programs of study at the University of Paderborn. The superior goal in education and training is to communicate competencies that are critical in tomorrow's economy.

Today eight Professors and 120 researchers work at the Heinz Nixdorf Institute. Per year approximately 15 young researchers receive a doctorate.

Zuletzt erschienene Bände der Verlagsschriftenreihe des Heinz Nixdorf Instituts

- Bd. 401 BRETZ, L.: Rahmenwerk zur Planung und Einführung von Systems Engineering und Model-Based Systems Engineering. Dissertation, Fakultät für Elektrotechnik, Informatik und Mathematik, Universität Paderborn, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 401, Paderborn, 2021 – ISBN 978-3-947647-20-0
- Bd. 402 WU, L.: Ultrabreitbandige Sampler in SiGe-BiCMOS-Technologie für Analog-Digital-Wandler mit zeitversetzter Abtastung. Dissertation, Fakultät für Elektrotechnik, Informatik und Mathematik, Universität Paderborn, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 402, Paderborn, 2021 – ISBN 978-3-947647-21-7
- Bd. 403 HILLEBRAND, M.: Entwicklungssystematik zur Integration von Eigenschaften der Selbstheilung in Intelligente Technische Systeme. Dissertation, Fakultät für Elektrotechnik, Informatik und Mathematik, Universität Paderborn, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 403, Paderborn, 2021 – ISBN 978-3-947647-22-4
- Bd. 404 OLMA, S.: Systemtheorie von Hardware-in-the-Loop-Simulationen mit Anwendung auf einem Fahrzeugachsprüfstand mit parallelkinematischem Lastsimulator. Dissertation, Fakultät für Maschinenbau, Universität Paderborn, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 404, Paderborn, 2022 – ISBN 978-3-947647-23-1
- Bd. 405 FECHTELPETER, C.: Rahmenwerk zur Gestaltung des Technologietransfers in mittelständisch geprägten Innovationsclustern. Dissertation, Fakultät für Elektrotechnik, Informatik und Mathematik, Universität Paderborn, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 405, Paderborn, 2022 – ISBN 978-3-947647-24-8
- Bd. 406 OLEFF, C.: Proaktives Management von Anforderungsänderungen in der Entwicklung komplexer technischer Systeme. Dissertation, Fakultät für Maschinenbau, Universität Paderborn, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 406, Paderborn, 2022 – ISBN 978-3-947647-25-5
- Bd. 407 JAVED, A. R.: Mixed-Signal Baseband Circuit Design for High Data Rate Wireless Communication in Bulk CMOS and SiGe BiCMOS Technologies. Dissertation, Fakultät für Elektrotechnik, Informatik und Mathematik, Universität Paderborn, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 407, Paderborn, 2022 – ISBN 978-3-947647-26-2
- Bd. 408 DUMITRESCU, R., KOLDEWEY, C.: Daten gestützte Projektplanung. Fachbuch. Fakultät für Elektrotechnik, Informatik und Mathematik, Universität Paderborn, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 408, Paderborn, 2022 – ISBN 978-3-947647-27-9
- Bd. 409 PÖHLER, A.: Automatisierte dezentrale Produktionssteuerung für cyber physische Produktionssysteme mit digitaler Repräsentation der Beschäftigten. Dissertation, Fakultät für Maschinenbau, Universität Paderborn, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 409, Paderborn, 2022 – ISBN 978-3-947647-28-6
- Bd. 410 RÜDDENKLAU, N.: Hardware-in-the-Loop Simulation von HD-Scheinwerfer Steuergeräten zur Entwicklung von Lichtfunktionen in virtuellen Nachtfahrten. Dissertation, Fakultät für Maschinenbau, Universität Paderborn, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 410, Paderborn, 2023 – ISBN 978-3-947647-29-3
- Bd. 411 BIEMELT, P.: Entwurf und Analyse modell prädiktiver Regelungsansätze zur Steigerung des Immersionsempfindens in interaktiven Fahrsimulationen. Dissertation, Fakultät für Maschinenbau, Universität Paderborn, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 411, Paderborn, 2023 – ISBN 978-3-947647-30-9
- Bd. 412 HAAKE, C.-J., MEYER AUF DER HEIDE, F., PLATZNER, M., WACHSMUTH, H., WEHRHEIM, H. (Eds.): On-The-Fly Computing - Individualized IT-Services in dynamic markets, Collaborative Research Centre 901 (2011 - 2023), Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 412, Paderborn, 2023 – ISBN 978-3-947647-31-6

Bezugsadresse:
Heinz Nixdorf Institut
Universität Paderborn
Fürstenallee 11
33102 Paderborn

Zuletzt erschienene Bände der Verlagsschriftenreihe des Heinz Nixdorf Instituts

- Bd. 413 DUMITRESCU, R.; HÖLZLE, K. (Hrsg.): Vorausschau und Technologieplanung. 17. Symposium für Vorausschau und Technologieplanung, Heinz Nixdorf Institut, 14./15. September 2023, Berlin-Brandenburgische Akademie der Wissenschaften, Berlin, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 413, Paderborn, 2023 – ISBN 978-3-947647-32-3
- Bd. 414 ABUGHANNAM, S.: Low-power Direct-detection Wake-up Receiver at 2.44 GHz for Wireless Sensor Networks. Dissertation, Fakultät für Elektrotechnik, Informatik und Mathematik, Universität Paderborn, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 414, Paderborn, 2024 – ISBN 978-3-947647-33-0
- Bd. 415 REINHOLD, J.: Systematik zur musterbasierten Transformation von Wertschöpfungssystemen für Smart Services. Dissertation, Fakultät für Maschinenbau, Universität Paderborn, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 415, Paderborn, 2024 – ISBN 978-3-947647-34-7
- Bd. 416 YANG, X.: Eine Methode zur Unterstützung von Entscheidungen bei der Entwicklung modularer Leichtbauprodukte. Dissertation, Fakultät für Maschinenbau, Universität Paderborn, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 416, Paderborn, 2024 – ISBN 978-3-947647-35-4
- Bd. 417 GRÄLER, M.: Entwicklung adaptiver Einrichtassistenzsysteme für Produktionsprozesse. Dissertation, Fakultät für Maschinenbau, Universität Paderborn, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 417, Paderborn, 2024 – ISBN 978-3-947647-36-1
- Bd. 418 RÖSMANN, D.: Menschenzentrierte Montageplanung und -steuerung durch fähigkeitsorientierte Aufgabenzuordnung. Dissertation, Fakultät für Maschinenbau, Universität Paderborn, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 418, Paderborn, 2024 – ISBN 978-3-947647-37-8
- Bd. 419 BAHMANIAN, M.: Optoelectronic Phase-Locked Loop, Theory and Implementation. Dissertation, Fakultät für Elektrotechnik, Informatik und Mathematik, Universität Paderborn, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 419, Paderborn, 2024 – ISBN 978-3-947647-38-5
- Bd. 420 HEIHOFF-SCHWEDE, J.: Spezifikationstechnik zur Analyse, Gestaltung und Bewertung von Engineering-IT-Architekturen. Dissertation, Fakultät für Elektrotechnik, Informatik und Mathematik, Universität Paderborn, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 420, Paderborn, 2024 – ISBN 978-3-947647-39-2
- Bd. 421 MEYER, M.: Systematik zur Planung und Verwertung von Betriebsdaten-Analysen in der strategischen Produktplanung. Dissertation, Fakultät für Elektrotechnik, Informatik und Mathematik, Universität Paderborn, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 421, Paderborn, 2024 – ISBN 978-3-947647-40-8
- Bd. 422 MALENA, K.: Konzipierung, Analyse und Realumsetzung eines mehrstufigen modellprädiktiven Lichtsignalanlagenregelungssystems. Dissertation, Fakultät für Maschinenbau, Universität Paderborn, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 422, Paderborn, 2024 – ISBN 978-3-947647-41-5
- Bd. 423 GÖTTE, R.-S.: Online-Schätzung von Modellgenauigkeiten zur automatischen Modelladaption unter Beibehaltung einer physikalisch-technischen Interpretierbarkeit. Dissertation, Fakultät für Maschinenbau, Universität Paderborn, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 423, Paderborn, 2024 – ISBN 978-3-947647-42-2
- Bd. 424 PIERENKEMPER, C.: Systematik zur Entwicklung Leistungsstufen-basierter Industrie 4.0-Strategien. Dissertation, Fakultät für Elektrotechnik, Informatik und Mathematik, Universität Paderborn, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 424, Paderborn, 2024 – ISBN 978-3-947647-43-9
- Bd. 425 DUMITRESCU, R.; HÖLZLE, K. (Hrsg.): Vorausschau und Technologieplanung. 18. Symposium für Vorausschau und Technologieplanung, Heinz Nixdorf Institut, 5./6. Dezember 2024, Berlin-Brandenburgische Akademie der Wissenschaften, Berlin, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 425, Paderborn, 2024 – ISBN 978-3-947647-44-6

Bezugsadresse:
Heinz Nixdorf Institut
Universität Paderborn
Fürstenallee 11
33102 Paderborn



Die Inbetriebnahme von Steuerungen und Regelungen stellt sicher, dass ein mechatronisches System ordnungsgemäß funktioniert und den Anforderungen gerecht wird. Der modellbasierte Entwurf basiert auf einem genauen Simulationsmodell. Allerdings ist dieser klassische Weg bei komplexen Systemen oft nicht praktikabel, da die analytische Modellierung zu kompliziert und zeitaufwendig ist. Diese Forschungslücke wird durch Verfahren adressiert, die eine effiziente und sichere Inbetriebnahme ermöglichen. Diese Verfahren kombinieren Regelungstechnik und Reinforcement Learning und nutzen vorhandenes Wissen über die Regelungsaufgabe, um Korrekturen basierend auf Messdaten und der probabilistischen Gauß-Prozess-Regression vorzunehmen. Das Vorwissen kann als teilweise bekanntes physikalisches Modell oder als Steuerungsfunktion vorliegen.

Anwendungsbeispiele sind der Ultraschalldrahtbondprozess, verschiedene Pendelsysteme und ein Hexapod. Eine angepasste Bayessche Optimierung wird zur Identifikation einer Steuerparametrisierung für das Ultraschallboden eingesetzt. Außerdem wird eine hybride Optimalsteuerung für das Doppelpendel auf einem Wagen entwickelt und erfolgreich validiert. Für einen Hexapod zur Fahrzeugachsprüfung wird eine hybride Zustandslinearisierung formuliert und ein Funktionsnachweis im Rahmen einer Simulation erbracht. Die Einhaltung technischer Rahmenbedingungen und ein stabiles Systemverhalten werden durch probabilistische Prädiktionen gewährleistet. In allen Anwendungsfällen wird eine Steigerung der Effizienz und Güte erzielt.