

Christoph Schweers

Adaptive Sigma-Punkte-Filter-Auslegung zur Zustands- und Parameterschätzung an Black-Box-Modellen

Adaptive Design of Sigma-point Filters for State- and Parameter Estimation with Black-box Models

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 367 der Verlagsschriftenreihe des Heinz Nixdorf Instituts

© Heinz Nixdorf Institut, Universität Paderborn – Paderborn – 2017

ISSN (Print): 2195-5239

ISSN (Online): 2365-4422

ISBN: 978-3-942647-86-1

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: Christoph Schweers

Hersteller: readbox unipress in der readbox publishing GmbH
Münster

Printed in Germany

Adaptive Sigma-Punkte-Filter-Auslegung zur Zustands- und Parameterschätzung an Black-Box-Modellen

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

genehmigte
DISSERTATION

von
Dipl.-Ing. Christoph Schweers
aus Walchum

Tag des Kolloquiums: 19. Dezember 2016

Referent: Prof. Dr.-Ing. habil. Ansgar Trächtler

Korreferent: Prof. Dr.-Ing. Sören Hohmann

Zusammenfassung

In dieser Arbeit wird eine neuartige Methodik zur Zustands- und Parameterschätzung an nicht-analytischen, als Black-Box vorliegenden Multi-Domänen-Modellen von technischen Systemen vorgestellt sowie alle notwendigen Werkzeuge zur Modelleinbindung, Filtertest und Bewertung der Schätzungsgüte entwickelt. Herausragendes Merkmal der entwickelten Methodik stellt die vollkommene Unabhängigkeit von Expertenwissen über das zugrundeliegende Filtermodell und die Filtertechnologie dar. Dies macht die Filterauslegung an Multi-Domänen-Modellen möglich, die mittels moderner Entwicklungswerkzeuge auf Basis von Bibliotheken oder sogar teil automatisiert erstellt wurden und über die somit keine Informationen über Zustände, Struktur und Nichtlinearitäten vorliegen.

In dieser Arbeit wurden echtzeitfähige Varianten von Sigma-Punkte-Kalman-Filtern erweitert, damit als Black-Box vorliegende Filtermodelle verwendbar werden. Es wurde ein Interface entwickelt, das in der Lage ist, Modelle aus einer Vielzahl an Modellierungswerkzeugen zu nutzen. Diese Arbeit liefert einen wesentlichen Neuerungswert, um einen Filter mit hoher Schätzungsgüte auch bei stark nichtlinearen, als Black-Box vorliegenden Systemen ohne Expertenwissen im zugrundeliegenden Filtermodell oder der Filtertechnologie auszulegen, da der Filterentwurf vollständig automatisiert auf Basis von Szenarien erfolgt. Alle erarbeiteten Algorithmen und Methoden wurden in einer unter MATLAB zur Verfügung stehenden Toolbox zusammengefasst, um so ein Werkzeug für die Zustands- und Parameterschätzung an unbekannten, stark nichtlinearen Modellen zur Verfügung zu stellen.

Abstract

The thesis presents a new methodology for state- and parameter estimation with non-analytical multi-domain models of highly nonlinear technical systems that exist only in the shape of black-box models and details the development of all necessary tools. An outstanding characteristic of the methodology is its complete independence of expert knowledge of the underlying filter model and filter technology. It enables filter design with multi-domain models that were constructed either by means of modern development tools on the basis of libraries or even in a partly automated way so there is no information available on their states, structure, and nonlinearities.

Variations of Sigma-point Kalman filters to be used in real time were extended in such a way as to enable use of filter models that exist only in the shape of black-box models. An interface was developed that is able to use models made up by any one of many different modelling tools. All algorithms and procedures thus developed were gathered in a toolbox available in MATLAB so as to provide a tool for state- and parameter estimation on unknown and highly nonlinear models.

Danksagung

Die vorliegende Dissertation entstand während meiner Tätigkeit am Lehrstuhl für Regelungstechnik und Mechatronik (RtM) und in der Fraunhofer Projektgruppe Entwurfstechnik für Mechatronik. Bedanken möchte ich mich bei meinem Doktorvater, Herrn Prof. Dr.-Ing. habil. Ansgar Trächtler, der mir die Möglichkeit gegeben hat, diese Arbeit an seinem Lehrstuhl durchführen zu können. Herrn Prof. Dr.-Ing. Sören Hohmann danke ich für die Übernahme des Korreferats. Ich bedanke mich auch bei Herrn Prof. Dr.-Ing. Jürgen Gausemeier und Herrn Prof. Dr.-Ing. Eugeny Kenig für ihre Mitwirkung in der Promotionskommission.

Allen Mitarbeitern des Lehrstuhls danke ich für die konstruktive und angenehme Zusammenarbeit. Insbesondere den Kollegen Daniel Kruse, Alexander Löffler und Martin Krüger möchte ich hiermit ein herzliches Dankeschön aussprechen für die fachlichen Diskussionen und zahlreichen Ratschläge. Ein großer Dank gebührt Frau Annette Bökamp-Gros für die sorgfältige Korrektur dieser Arbeit. Weiterhin danke ich den studentischen Hilfskräften, Studien-, Diplom-, Bachelor- und Masterarbeitern für ihre Unterstützung und Beiträge.

Besonders bedanken möchte ich mich bei meinen Eltern Margret und Wilhelm, meinen Brüdern Johannes und Bernd sowie meinen Freunden, auf deren Unterstützung ich mich während meines Studiums und in meiner Zeit als wissenschaftlicher Mitarbeiter immer verlassen konnte.

Mein größter Dank gilt meiner Frau Christine, die mir stets den Rücken freigehalten hat und durch ihr Verständnis und ihre positive Art maßgeblich zum Gelingen dieser Arbeit beigetragen hat.

Burgdorf, im Januar 2017

Christoph Schweers

Vorveröffentlichungen

- [1] LÖFFLER, A.; KOERT, D.; SCHWEERS, C.; TRÄCHTLER, A.: *Einführung in die modellbasierte Entwicklung im Bereich der Haugerätetechnik*. In: VDI Mechatronik-Tagung, Dresden, 2011.
- [2] LÖFFLER, A.; SCHWEERS, C.; FAST, V.; KRUSE, D.; TRÄCHTLER, A.: *Multidomänen-Modell eines Waschvollautomaten für einen Hardware-in-the-Loop-Prüfstand*. In: 8. Paderborner Workshop: „Entwurf mechatronischer Systeme“, Paderborn, 2011.
- [3] KRUSE, D.; SCHWEERS, C.; FAST, V.; TRÄCHTLER, A.: *Einheitliche Testumgebung für MiL und RCP mittels NI VeriStand am Beispiel eines Waschautomaten*. In: Virtuelle Instrumente in der Praxis 2012, München, 2012.
- [4] SCHWEERS, C.; KRUSE, D.; TRÄCHTLER, A.: *Entwurf eines Unscented-Kalman-Filters zur Zustands- und Parameterschätzung an Dymola-Modellen*. In: VDI Mechatronik-Tagung, Aachen, 2013.
- [5] SCHWEERS, C.; KRUSE, D.; FAST, V.; TRÄCHTLER, A.: *Online-Zustands- und Parameterschätzung an Dymola-Modellen auf NI-Echtzeithardware*. In: Virtuelle Instrumente in der Praxis 2013, München, 2013.
- [6] FAST, V.; SCHWEERS, C.; KRUSE, D.; TRÄCHTLER, A.: *Methoden zur Einbindung von Fehlersimulationen in die XiL-Techniken*. In: Virtuelle Instrumente in der Praxis 2013, München, 2013.
- [7] SCHWEERS, C.; KRUSE, D.; TRÄCHTLER, A.: *Automated Design of an Unscented Kalman Filter for State- and Parameter Estimation on unknown Models*. In: IEEE Conference on Control, Automation, Robotics & Embedded Systems (CARE), Jabalpur, Indien, 2013.
- [8] SCHWEERS, C.; BRUNSTEIN, F.; TRÄCHTLER, A.: *An Automated Approach To Filter Design For Online State- And Parameter Estimation On Unknown, Non-Analytic Models*. In: 2nd International Conference on System-integrated Intelligence (SysInt2014), Bremen, 2014.
- [9] KRUSE, D.; SCHWEERS, C.; TRÄCHTLER, A.: *Methodology for a partly automated parameter identification to validate Multi-Domain Models*. In: ASME International Mechanical Engineering Congress and Exposition, Montreal, Canada, 2014.

Inhaltsverzeichnis

1	Motivation und Zielsetzung	1
1.1	Entwicklungssystematiken mechatronischer Systeme	2
1.2	Fachdisziplinspezifische Auslegung in der Regelungstechnik . .	3
1.2.1	Steuerungs- und Reglerentwurf	6
1.2.2	Zustandsschätzung	7
1.3	Problemstellung und Anforderungen	9
1.4	Stand der Wissenschaft und der Technik	12
1.5	Handlungsbedarf	14
1.6	Aufbau der Dissertation	15
2	Zustands- und Parameterschätzung	19
2.1	Statistische Grundlagen	20
2.1.1	Wahrscheinlichkeit	20
2.1.2	Zufallsvariable und Dichtefunktion	21
2.1.3	Statistische Momente	22
2.2	Das Kalman-Filter	24
2.2.1	Das Extended-Kalman-Filter	27
2.2.2	Least-Square- und Weighted-Least-Square-Methode . .	29
2.2.3	Bayessche Zustandsschätzung	32
2.3	Particle Filter	33
2.4	Sigma-Punkte-Filter	35
2.5	Machine Learning	36
2.6	Zusammenfassung der Schätzverfahren	36
3	Sigma-Punkte-Filter	39
3.1	Unscented-Kalman-Filter	40
3.2	Unscented Transform	46
3.2.1	Second-Order Transformation	48
3.2.2	Scaled-Unscented Transformation	54
3.2.3	Spherical-Simplex Transformation	60
3.2.4	Zusammenfassung der Sigma-Punkte-Verfahren	64
3.3	Erweiterung des UKF-Algorithmus	65
3.4	Das Central-Difference-Kalman-Filter	68
3.5	Parameterschätzung mittels Sigma-Punkte-Filter	71
3.5.1	Das Joint-Unscented-Kalman-Filter	72

3.5.2	Wahl der Kovarianz einer Parameterschätzung	74
3.6	Sigma-Punkte-Filter unter Nebenbedingungen	75
3.7	Zusammenfassung der Sigma-Punkte-Filter	80
4	Wahl der Filter- und Entwurfparameter	83
4.1	Optimaler Parametersatz zur Filterauslegung	84
4.1.1	Szenariodefinition	84
4.2	Ermittlung von Prozess- und Messrauschkovarianzen	85
4.3	Optimierung der Entwurfparameter	88
4.3.1	Gaußprozess	90
4.3.2	Kernel-Funktionen und Hyperparameter	95
4.3.3	Optimale Vorgabe von Hyperparametern des Kernels	97
4.3.4	Gaußprozessoptimierung	98
4.4	Anwendung am Unscented-Kalman-Filter	100
4.4.1	Trainingsdaten	102
4.4.2	Optimierung zur Maximafindung und Abbruchbedingung	104
5	Auslegungsmethodik	109
5.1	Das Functional-Mockup-Interface	110
5.1.1	MATLAB-Interface zur FMU-Ansteuerung	111
5.2	Modelleinbindung	114
5.3	Filterentwurf	115
5.4	Systemtest	116
5.5	Zusammenfassung Auslegungsmethodik	118
6	Anwendungsbeispiele	121
6.1	Doppelpendel	121
6.1.1	Modellbildung Doppelpendel	122
6.1.2	Zustands- und Parameterschätzung unter idealen Bedingungen	123
6.1.3	Optimierter Satz von Entwurfparametern	123
6.1.4	Wahl von Prozess- und Messrauschkovarianz	127
6.1.5	Einbeziehung von Nebenbedingungen	129
6.1.6	Vergleich unterschiedlicher Filtervarianten	133
6.2	Schätzung zeitvarianter Unwucht in Waschautomaten	137
6.2.1	Modellbildung Waschautomat	138
6.2.2	Unwuchtschätzung im Schleuderprozess	140
6.2.3	Robustheit gegenüber Modellungenauigkeiten	141
6.3	Zustandsschätzung Dingo 2	147
6.3.1	Modellbildung Dingo	147
6.3.2	Zustandsschätzung bei Hindernisfahrt	148
6.4	Zusammenfassung Anwendungsbeispiele	151
7	Resümee und Perspektiven	153

8	Literaturverzeichnis	157
A	Anhang	167
A.1	Square-Root Unscented Kalman Filter	167
A.2	Central-Difference-Filter	168
A.3	MATLAB-Toolbox	169
A.4	Doppelpendel	174
A.5	Schätzung zeitvarianter Unwucht im Waschautomaten	178
A.6	Zustandsschätzung Dingo 2	182

Symbolverzeichnis

A	Dynamikmatrix
A_k^{lin}	linearisierte Dynamikmatrix
α	Skalierungsparameter der SUT-Sigma-Punkte-Methode
α_{RM}	Verteilungsparameter Robbins-Monro Methode
B	Eingangsmatrix
β	Gewichtungsparameter der SUT-Sigma-Punkte-Methode
C_k	Matrixwurzel
$D_{\delta_{x_{k-1}}}$	totales Differential von F um x_{k-1}
$\Delta\mathcal{X}$	maximaler Abstand Sigma-Punkte zu Nebenbedingungen
ϵ	Fehlerschranke zur Anwendung der modellbasierten Optimierung des \mathcal{GP}
$F()$	Verteilungsfunktion
γ	Zusammenfassender Skalierungsparameter der Sigma-Punkte als
γ_W	Exzess einer Verteilung
γ'	neuer Skalierungsparameter der Sigma-Punkte nach Korrektur von ungültigen Sigma-Punkten
\mathcal{GP}	Gaußprozess
H	Ausgangsmatrix
H_k^{lin}	linearisierte Messmatrix
J	Gütefunktion der Gaußprozessoptimierung, bestehend aus $J_F(\theta)$ und $\text{var}[J_F(\theta)]$
J_F	Funktionswert der Schätzgüte
K	Kovarianzmatrix \mathcal{GP} bekannter Stützstellen
K_*	Kreuzkovarianzmatrix \mathcal{GP} bekannter und unbekannter Stützstellen
K_{**}	Kovarianzmatrix \mathcal{GP} unbekannter Stützstellen
$k()$	Kovarianzfunktion \mathcal{GP}
k	Zeitindex
k_w	Gewichtungsparameter der Gaußprozesskostenfunktion zur explorativen oder schnellen Optimierung
\mathcal{K}_k	Kalman-Korrektur
κ	Entwurfsparameter der SOT-Sigma-Punkte-Methode
l	Anzahl Sigma-Punkte
l_f	Längenskalierungsparameter des \mathcal{GP}
λ	Vereinfachungsparameter der SUT-Sigma-Punkte-Methode
λ_{RLS}	Fensterfunktion Prozessrauschen Schätzparameter

M	Diagonalmatrix der Längenskalierungsfaktoren von $k()$ bei mehrdimensionalem \mathcal{GP}
m	Element aus der Menge \mathcal{M}
\mathcal{M}	Menge aller in den Filternebenbedingungen zu überprüfenden Vektoren
$NB_{P,u,o}$	untere und obere Nebenbedingung der Sigma-Punkte
n	Systemordnung
n_k	Messrauschen
n_z	Anzahl zu korrigierender Sigma-Punkte
$\mathcal{N}()$	Normalverteilung
ν	Schiefe einer Verteilung
$P_{ww,k}$	Kovarianzmatrix Schätzparameter
$P(\mathcal{A})$	Wahrscheinlichkeit des Ereignisses \mathcal{A}
$P_{xx,k}$	A-posteriori-Schätzfehlerkovarianz
$P_{xx,k}^-$	A-priori-Schätzfehlerkovarianz
$P_{xy,k}$	Kreuzkovarianzmatrix vorhergesagter Messungen und Zustände
$P_{yy,k}$	Kovarianzmatrix vorhergesagter Messungen
$p()$	Wahrscheinlichkeitsdichtefunktion
$\hat{p}()$	approximierte bedingte Dichteverteilung
Ψ	Schiefemaß nach Mardia
R_n	Prozessrauschkovarianzmatrix
R_v	Messrauschkovarianzmatrix
$R_{v,w}$	Prozessrauschen Parameter
s	Abstand der Sigma-Punkte zum Erwartungswert
$s_{x_i,k-1}$	Abstand der Sigma-Punkte des CDKF vom Erwartungswert
\mathcal{S}	Sigma-Punkte-Satz
σ^2	Rauschkovarianz der in \mathcal{GP} eingehenden Daten
σ_f^2	maximale Kovarianz der Kovarianzfunktion $k()$ des \mathcal{GP}
t_{end}	Endzeit Simulation
t_{start}	Startzeit Simulation
θ_{SUT}	Entwurfparameter der SUT-Sigma-Punkte-Berechnungsmethode
θ_{SST}	Entwurfparameter der SST-Sigma-Punkte-Berechnungsmethode
θ_{CDKF}	Entwurfparameter des CDKF
θ_{CKF}	Entwurfparameter des CKF
u_k	Systemeingang
v_k	Prozessrauschen
W	Gewichtungen
$W^{(m)}, W^{(c)}$	Gewichtungen SUT-Sigma-Punkte-Methode
$W_\zeta^{(m)} = W_\zeta^{(c)}$	Gewichtungen der zu korrigierenden Sigma-Punkte

$W_\tau^{(m)} = W_\tau^{(c)}$	Gewichtungen der beizubehaltenden Sigma-Punkte
w_k	zu schätzender Parameter
\hat{w}_k	Schätzparameter
X	Gesamtheit Eingang Stützstellen \mathcal{GP}
\mathcal{X}	Sigma-Punkte
X_*	Gesamtheit der Eingangs-Stützstellen \mathcal{GP} an vorherzusagenden Funktionswerten
\mathcal{X}^*	Transformierte Sigma-Punkte
$\hat{\mathcal{X}}$	Sigma-Punkt mit maximaler Verletzung der Filternebenbedingungen
\mathcal{X}_{k-1}	A-priori-Sigma-Punkte
$\mathcal{X}_{k k-1}^*$	transformierte A-priori-Sigma-Punkte
$\mathcal{X}_{k k-1}$	A-priori-Sigma-Punkte unter Einbeziehung des Prozessrauschens
$\mathcal{X}_{w,k}$	Sigma-Punkte der Schätzparameter
\mathcal{X}_τ	beizubehaltende Sigma-Punkte
\mathcal{X}_ς	zu korrigierende Sigma-Punkte
\bar{x}	Mittelwert der Zufallsvariablen x
x_k	zu schätzender Zustandsvektor
\hat{x}_k	A-posteriori-Zustandsschätzung
\tilde{x}_k	Schätzfehler
\hat{x}_k^-	A-priori-Zustandsschätzung
ξ	Entwurfparameter CDKF
Y	Gesamtheit Ausgang Stützstellen \mathcal{GP}
Y_*	Gesamtheit vorherzusagender Funktionswerte der Eingangs-Stützstellen \mathcal{GP}
y_k	Messungen
\hat{y}_k	geschätzte Messung
$\mathcal{Y}_{k k-1}$	transformierte Messungen auf Sigma-Punkte
z	Standardnormalverteilung
\mathcal{Z}	Hilfsmatrix der SST-Sigma-Punkte-Methode
ζ	Gesamtheit der Hyperparameter der Kovarianzfunktion $k()$

1 Motivation und Zielsetzung

Für eine erfolgreiche Wirtschaft ist die Entwicklung von kundenseitig nachgefragten Produkten genauso wichtig wie die vom Anbieter sicherzustellende Funktionsfähigkeit. Auf Produkte, die im maschinenbaulichen Umfeld oder verwandten Branchen wie der Automobilindustrie oder Medizintechnik entwickelt werden, wirken dabei immer mehr Kommunikations- und Informationstechniken ein, was zu einer immer stärkeren Ausrichtung zu komplexen mechatronischen Produkten führt. Die VDI-Richtlinie 2206 *Entwicklungsmethodik für mechatronische Systeme* definiert:

Mechatronik bezeichnet das synergetische Zusammenwirken der Fachdisziplinen Maschinenbau, Elektrotechnik und Informationstechnik beim Entwurf und der Herstellung industrieller Erzeugnisse sowie bei der Prozessgestaltung

In mechatronischen Systemen wirken Elemente des Maschinenbaus, der Elektrotechnik und Elektronik sowie der Informatik und Informationstechnik zusammen, was zu einer erhöhten technischen Komplexität des Produktes selbst, aber auch in dessen Entwicklungsprozess führt. Beispiele reichen vom Sondermaschinenbau (Werkzeugmaschinen) über Serienprodukte (Automotive) bis hin zu Massenprodukten (Wälzlager).

Diese Komplexität des Produktes und der Entwicklung selbst wird bei neuen Produkten noch durch die Faktoren einer immer kürzen Entwicklungszeit, hohen Kostendrucks und hoher Qualität verschärft, wobei auf Störungen wie Wettbewerb oder geänderte Kundenwünsche flexibel reagiert werden muss [Lin09]. In [SEF04] wird die Entwicklung komplexer mechatronischer Systeme als Kunst und Wissenschaft bezeichnet, ein funktionierendes System zu schaffen, das bestehenden Anforderungen gerecht wird. Die Entwicklung stellt dabei einen ganzheitlichen, unterschiedliche Fachdisziplinen einbeziehenden Prozess dar, wobei die Disziplinen aufeinander abgestimmt werden müssen, um ein kohärentes System zu schaffen, das nicht von einer Fachdisziplin dominiert wird.

Zur Erreichung dieser Ziele gewinnen der Aspekt der Systemdynamik und die simulationsbasierte Analyse dieser Systeme immer mehr an Bedeutung. Neben gängigen 3D-CAD-Modellen kommen dabei vermehrt physikalisch motivierte Verhaltensmodelle zum Einsatz, wobei besonderer Fokus auf einer durchgängig modellbasierten Entwicklung liegt. Gleichzeitig vollziehen Unternehmen immer mehr den Wandel, sich die Entwicklungsarbeit mit spezialisierten Anbietern von Lösungselementen zu teilen, wobei diese Elemente durch realisierte und bewährte Lösungen zur Erfüllung einer Funktion gegeben sind [GTS14].

Eine frühzeitige fachdisziplinübergreifende Beschreibung des Systemverhaltens ist dabei zwingend erforderlich, um notwendige Kooperationen zwischen den beteiligten Fachdisziplinen sicherzustellen und die Grundlage für eine spätere, parallel verlaufende Entwicklung in den Fachbereichen zu legen. Bei der frühzeitigen

Beschreibung des Systemverhaltens wird dabei in immer stärkerem Maße das Internet genutzt, um auf das breite Innovationspotenzial von vielen Anbietern zurückgreifen zu können. Diese Fülle an Informationen mit einer gleichzeitig immer weiter steigenden Komplexität und der fachdisziplinübergreifenden Systembeschreibung stellt im Besonderen neue Anforderungen an die IT-Unterstützung des Entwicklungsprozesses und der fachdisziplinspezifischen Auslegung im Form von Werkzeugen, Vorgehensmodellen und Methoden.

1.1 Entwicklungssystematiken mechatronischer Systeme

Der Entwurf komplexer mechatronischer Produkte unter Einwirkung von vielen Domänen stellt besondere Herausforderungen an das Vorgehen, für das eine Reihe unterschiedlicher Systematiken existieren, die oftmals auf einzelne Fachdisziplinen ausgerichtet sind.

Eine Grundlage für methodisches Konstruieren im Maschinenbau legten Pahl/Beitz in [PBF⁺06], wobei den Kern mechanische Systeme bilden und andere Domänen unbeachtet bleiben. Andere domänenspezifische Methoden sind mit dem V-Modell der Softwareentwicklung und dem Y-Modell der Schaltungstechnik gegeben. Domänenübergreifende Ansätze sind mit der Entwicklungssystematik nach Isermann [Ise02], dem Münchener Vorgehensmodell nach Lindemann [Lin09], der integrierten Produktentwicklung nach Ehrlenspiel [Ehr09] und der *Entwicklungsmethodik für mechatronische Systeme* nach [VDI04] vorhanden. In der zuletzt genannten Richtlinie werden fachdisziplinübergreifende Lösungskonzepte erarbeitet, die in der weiterführenden fachdisziplinspezifischen Ausarbeitung getrennt voneinander zur detaillierten Auslegung genutzt werden. Die getrennt voneinander entwickelten Bausteine der unterschiedlichen Fachdisziplinen werden im Anschluss in der Systemintegration zu einem Gesamtsystem zusammengefügt.

Ein lösungsneutraler und funktionsorientierter Entwurfsprozess ist allerdings oftmals gefährdet, da der Entwickler in einer frühen Entwurfsphase dazu verleitet werden kann, sich auf bekannte Lösungen zu konzentrieren, was dem immensen Informationsaufkommen, unternehmensspezifischen Beschreibungsformen, Detaillierungsgraden etc. geschuldet ist [GTS14]. Eine Weiterentwicklung der bestehenden Entwurfssystematiken ist durch das Verbundprojekt ENTIME gegeben, das einen domänenübergreifenden Ansatz zur Beseitigung der bestehenden Nachteile bietet und ein Vorgehensmodell und Spezifikationstechniken für den durchgängigen und integrierten Entwurf mechatronischer Systeme darstellt [GTS14]. Die Struktur des Instrumentariums ist in Bild 1-1 dargestellt.

Im Rahmen der Konzipierung erfolgt frühzeitig eine fachdisziplinübergreifende Systemdefinition, die auch den Ausgangspunkt für die fachdisziplinspezifische Ausarbeitung darstellt. Auf Basis einer Funktionsbeschreibung erfolgt die Suche nach möglichen Lösungsmustern, die Zulieferer und Domänenexperten einschließt und eine vom Entwickler losgelöste, lösungsneutrale Möglichkeit zur Realisierung

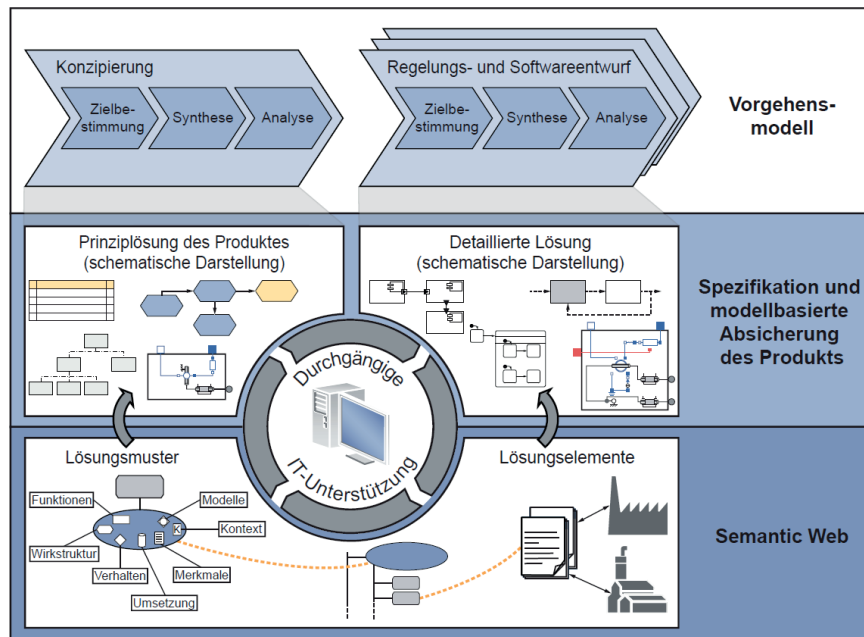


Bild 1-1: Struktur der Entwurfsmethode ENTIME [GTS14]

der Funktionsbeschreibung bietet, für die das Semantic Web¹ genutzt wird. Der Schwerpunkt des Projektes lag auf der frühzeitigen und durchgängigen Modellierung des Systemverhaltens, wofür den auf Basis der Funktionsbeschreibung ausgewählten Lösungsmustern bereits idealisierte Simulationsmodelle zugeordnet sind.

Nach der ganzheitlichen, fachdisziplinübergreifenden Modellierung des Systemverhaltens folgt die parallel verlaufende Entwicklung in den unterschiedlichen Disziplinen, wie der Elektronik, dem Softwareentwurf und der Regelungstechnik. Auf die Fachdisziplin der Regelungstechnik wird im folgenden Abschnitt eingegangen.

1.2 Fachdisziplinspezifische Auslegung in der Regelungstechnik

In der fachdisziplinspezifischen Auslegung der Regelungstechnik besteht die Aufgabe in der Systemanalyse mit dem Ziel, der *Ausgangsgröße eines technischen Systems durch eine Stellgröße ein Sollverhalten, d. h. ein gewünschtes Verhalten, aufzuprägen, und zwar gegen den Einfluss von Störgrößen, die nur unvollständig bekannt sind* [Föl13]. Hierzu muss das System fortlaufend beobachtet und Änderungen des Systems dahingehend genutzt werden, eine Stellgröße zu generieren, die das gewünschte Systemverhalten einstellt. Hierfür werden auch eine Auslegung

¹Bedeutung (Semantik) von Informationen im Netz muss so definiert sein, dass sie auch von Maschinen gelesen und verstanden werden kann.

und eine Optimierung der mechanischen Grundstruktur, Aktorik und Sensorik durchgeführt. Eine grundlegende Reglerstruktur zur Erzielung eines gewünschten Streckenverhaltens ist in Bild 1-2 dargestellt. Die beobachtete Systemgröße y wird zurückgeführt und mit der Führungsgröße w verglichen. Die Differenz von Führungsgröße und Rückführgröße wird über den Regler, der Stabilität, Dämpfung, stationäre Genauigkeit und Schnelligkeit sicherstellt, an die Stelleinrichtung übergeben und wirkt auf das dynamische System.

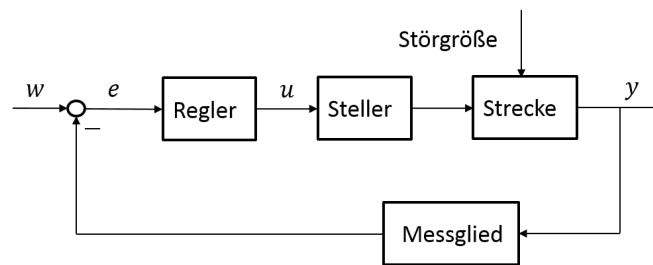


Bild 1-2: Grundlegende Reglerstruktur [Föl13]

Einen wichtigen Erfolgsfaktor für den Entwurf von Regelungen stellt dabei die Beschreibung der Systemdynamik durch mathematische Modelle dar. Auf Basis dieser Modelle kann eine systematische Auslegung von Mechanik, Aktorik, Sensorik und Regelung erfolgen. Im Sinne der VDI-Richtlinie 2206 handelt es sich bei diesen Modellen um Multi-Domänen-Modelle, die wichtige Informationen aus allen Fachdisziplinen beinhalten und deren Entwicklungsprozess in Bild 1-3 dargestellt ist.

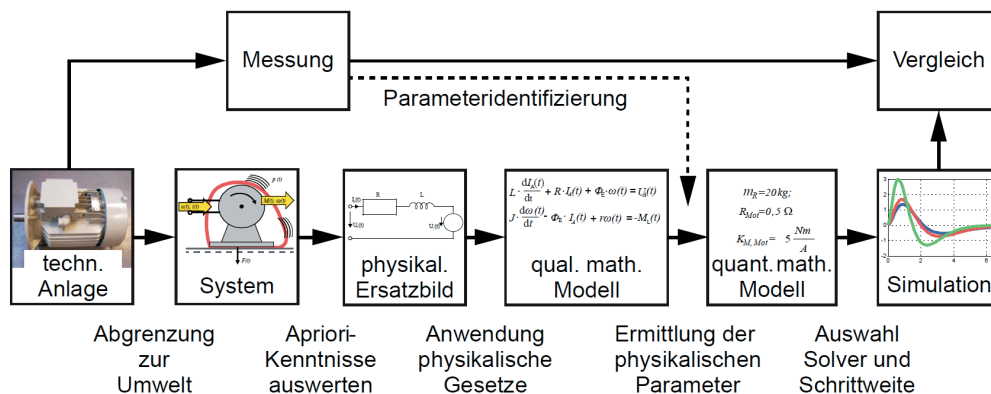


Bild 1-3: Vorgehensmodell zur Ableitung von Dynamikmodellen [Trä12a]

Im ersten Schritt erfolgt die Umweltabgrenzung zur Definition der auf das System wirkenden Ein- und Ausgänge. Im zweiten Schritt wird auf Basis von A-priori-Wissen das technische System in ein physikalisches Ersatzschaltbild überführt. Dies wird oftmals mehr als Kunst denn als Ingenieurwissenschaft bezeichnet, da dieser Schritt den Detaillierungsgrad festlegt und ob alle Anforderungen durch

das Modellverhalten abgedeckt werden. Das physikalische Ersatzsystem kann im Anschluss durch die Anwendung physikalischer Gesetze in ein qualitatives mathematisches Modell überführt und durch Einbeziehung der physikalischen Parameter des realen Systems in ein quantitatives Modell übertragen werden.

Als Beschreibungssprachen können eine Vielzahl von Varianten eingesetzt werden. Für die Regelungstechnik hat sich MATLAB/SIMULINK wegen seiner signalflussorientierten Modellierungsart, die der regelungstechnischen Blockschalt-darstellung stark ähnelt, etabliert. Hierbei werden Blöcke mit Ein- und Ausgängen genutzt, die durch rückwirkungsfreie Leitungen miteinander verbunden werden. Oftmals vorhandene Rückwirkungen müssen durch Rückkopplungen realisiert werden. Die Ableitung der mathematischen Beschreibungsform aus dem physikalischen Ersatzschaltbild und die Implementierung in SIMULINK erfolgen fast ausschließlich manuell, was einen schwierigen und fehleranfälligen Prozess darstellt. Im Gegensatz zum Signalfluss werden in der topologieorientierten Modellierung ungerichtete Verbindungen genutzt. Erste Ansätze hierzu wurden durch Elmqvist [Elm78] mit der Entwicklung des Simulationstools Dymola gezeigt. Das Grundprinzip besteht darin, physikalische Systeme möglichst direkt auf Modellobjekte abzubilden, die wiederum untereinander verknüpft werden können. Diese Verknüpfung ist allerdings akausal², und es werden nur Beziehungen zwischen Objekten hergestellt [Ott93]. Simulationstools können die mathematischen Beschreibungen dieser Modellierung automatisiert erstellen und optimieren. Aufgrund des objektorientierten Charakters und einer mittlerweile weiten Verbreitung kann man auf eine breite Klasse von vorgefertigten Bibliotheken zurückgreifen, die komplexe Einzelelemente oder komplette technische Systeme enthalten und nur noch durch Ein-/ Ausgangsverhalten zu bewerten sind [Ott93]. Auch Isermann stellt in [Ise95] die Sprache Modelica, auf der das Simulationstool Dymola basiert, für einen domänenübergreifenden Entwurf als objektorientierte Software unter Verwendung allgemeiner Modellbildungsgesetze als vorteilhaft dar.

War bei einer fachdisziplinspezifischen Modellbildung der Ingenieur Experte in seinem Modell, so ist dies bei einer topologieorientierter Modellierung nicht mehr gewährleistet. Bei einem Vorgehen nach ENTIME erfolgt die Modellableitung teil- oder vollautomatisiert auf Basis von Lösungselementen, die von Zulieferern oder Entwicklungspartnern erstellt wurden und deren innere Struktur nicht mehr vollständig bekannt ist. Aber auch bei der Nutzung von erforderlichen Multi-domänensprachen wie Modelica greift der Ingenieur auf vorgefertigte Elemente zurück, deren inneren Aufbau er nicht im Detail kennt oder für den er aufgrund der domänenübergreifenden Modellierung kein Experte ist. Bevor in Abschnitt 1.3 auf die daraus resultierenden Probleme bei der regelungstechnischen Auslegung eingegangen wird, werden in den nächsten Abschnitten regelungstechnische

²Keine zuweisungs-basierte Modellierung im Sinne von Ursache - Wirkung, wodurch keine explizite Festlegung von Ein- und Ausgang notwendig ist.

Grundlagen kurz erläutert, die zu der in dieser Arbeit betrachteten Problemstellung führen.

1.2.1 Steuerungs- und Reglerentwurf

Die Realisierung von Reglern zur Erzielung der gewünschten dynamischen Eigenschaften bei Eingrößensystemen bietet eine Fülle möglicher Strukturen wie Proportionalregler (P-Regler), Integratoranteile (PI-Regler) oder auch differenzierende Strukturen (PD- oder PID-Regler) und Auslegungsverfahren hinsichtlich Stabilität, Schnelligkeit und Dämpfung wie beispielsweise Nyquist, Hurwitz oder Kessler. Die Wahl der Strukturen und Auslegungsverfahren ist dabei problem-spezifisch je nach System und Regelungsanforderung. Erweiterungen der oben genannten Regler zu Kaskadenregelungen, Zwei-Freiheitsgradsstrukturen oder auf Mehrgrößenregler, beispielsweise mittels Entkopplung, bieten viele Möglichkeiten, gewünschtes Systemverhalten für komplexe regelungstechnische Aufgabenstellungen einzustellen.

Eine weitere Möglichkeit der Regelung von dynamischen Systemen besteht in der Nutzung von Zustandsreglern. Die Grundlage stellt ein im Folgenden als zeitkontinuierlich angenommenes System in Zustandsraumdarstellung dar:

$$\begin{aligned}\dot{x} &= Ax + Bu + x_0 \\ y &= Cx,\end{aligned}\tag{1-1}$$

mit der auf das System wirkenden Anfangsstörung $x(t_0) = x_0$. Die Durchgriffsmatrix D wird vernachlässigt, wie es bei der überwiegenden Mehrzahl realer Systeme der Fall ist [Föl13]. Zur Beseitigung dieser auf das System wirkenden Anfangsstörung ist eine Rückführung erforderlich, was der regelungstechnischen Grundsituation entspricht [Föl13] und durch eine Zustandsrückführung realisiert werden kann. Bei einem linearen Rückführgesetz der Form

$$u = R(x^* - x) + M_u w\tag{1-2}$$

ergibt sich das in Bild 1-4 abgebildete Blockschaltbild mit den Vorsteuermatrizen M_u und M_x , welche die Erzielung eines gewünschten Sollverhaltens w sicherstellen. Die Auslegung der Rückführmatrix R kann dabei auf unterschiedlichen Wegen erfolgen, allerdings unabhängig von den Vorsteuermatrizen M_u und M_x . Eine Möglichkeit zur Auslegung stellt die Polvorgabe oder Eigenwertvorgabe dar, die auf der Erkenntnis basiert, dass die Lage der Eigenwerte eines dynamischen Systems entscheidend für dessen Verhalten ist. Das Ziel besteht demnach darin, die Eigenwerte in ein Stabilitätsgebiet zu bringen, das gleichzeitig die gewünschte Systemdynamik aufweist und die Stellgrößenbeschränkungen des Systems indirekt berücksichtigt [Föl13]. Eine weitere Möglichkeit der Bestimmung der Rückführmatrix R besteht in der Minimierung eines quadratischen Gütemaßes, was

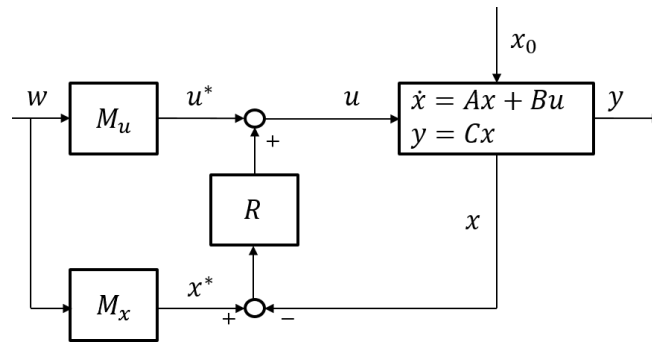


Bild 1-4: Zustandsregler mit Zustandsrückführung und Vorfilter [Föl13]

auf den Riccati-Regler führt, bei dem eine Kostenfunktion J über den zeitlichen Verlauf aller Zustände und Stellgrößen gebildet wird.

Diese und weitere Verfahren zur Realisierung eines Zustandsreglers haben eine Eigenschaft gemein: Der Zustandsvektor muss vollständig bekannt sein! Dies wird allerdings in fast keiner realen Anwendungen möglich sein oder nur mit nicht vertretbarem Aufwand, was auf die Entwicklung von Beobachterstrukturen führte, um den unbekannten Zustand eines dynamischen Systems auf Basis von Messungen und den bekannten Eingangsgrößen zu rekonstruieren.

1.2.2 Zustandsschätzung

Eine Möglichkeit der Realisierung von Zustandsreglern stellt die Nutzung von Beobachterstrukturen wie des Kalman-Filters dar, mit dessen Hilfe sich der Zustand eines dynamischen Systems unter Einwirkung von Anfangsstörungen und fortwährenden stochastischen Störungen rekonstruieren lässt. Hierzu ist ein mathematisches Modell der zu regelnden Strecke notwendig, auf dessen Basis unter Einwirkung der bekannten Systemeingänge und der realen Messungen am System der interne Zustand geschätzt wird, wie in Bild 1-5 abgebildet. Eine direkte Nutzung des vorhandenen Modells zur Ermittlung der Systemzustände ist zum Einen aufgrund der bereits erwähnten Anfangsstörung und der fortwährend auftretenden stochastischen Störungen nicht möglich, zum Zweiten aber auch der Tatsache geschuldet, dass man nie in der Lage ist, ein Modell zu erstellen, das vollkommen dem Verhalten eines realen technischen Systems entspricht [May79].

Neben der Nutzung zur Bereitstellung der Istgröße für den Zustandsregler können und werden Zustandsschätzer in annähernd allen Bereichen des Ingenieurwesens und der Wissenschaft eingesetzt. Eine Google-Suche zum Zeitpunkt der Entstehung dieser Arbeit liefert bei den Schlagworten „state estimation“ und „application“ über 20 Millionen Treffer. Die Eingabe von „Kalman Filter“ und „application“ liefert immer noch über 5 Millionen Ergebnisse. Jede Disziplin, in der mathematische Modelle von Systemen genutzt werden, stellt ein mögliches Anwendungsgebiet der Zustandsschätzung dar: Elektrotechnik und Elektronik, Maschinenbau,

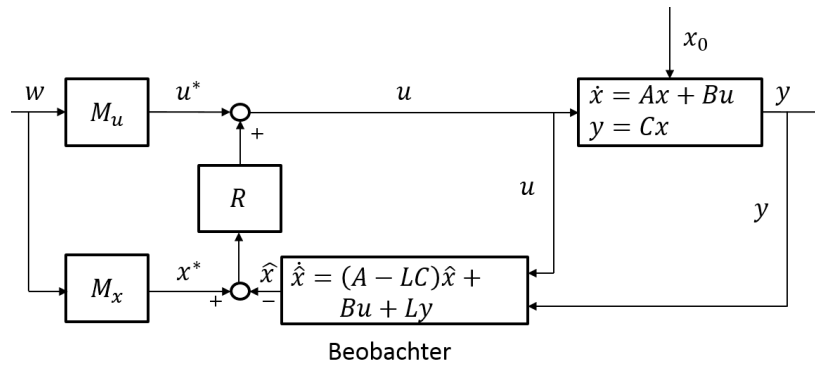


Bild 1-5: In Zustandsregelung eingebundener Beobachter zur Schätzung des Zustandes eines dynamischen Systems [Föl13]

Chemieindustrie, Ökonomie, Biologie, Luft- und Raumfahrt und Robotik, um nur einige zu nennen [Sim06], und führte auf die Entwicklung von unzähligen Verfahren, Varianten und Anwendungen.

In [LMW14] wendet man Beobachterstrukturen auf Zustandsraummodelle von Finanzstrukturen an, um Vorhersagen der Beeinflussung von Kursgewinnen und Inflation zu treffen. Kaakinen et al. nutzen Kalman-Filter-Techniken in [KHP⁺14] zur Beobachtung von Zellbewegungen unter Kontrastmikroskopen. In [SLPP⁺13] wird von Sanchez-Lopez et al. ein Extended-Kalman-Filter zur Lokalisierung und Routenplanung von flugfähigen Roboterschwärmen eingesetzt. Um eine Vielzahl von verrauschten Sensoren zu einer verwertbaren Größe zu verbinden, werden von Kim et al. in [KKS14] Beobachter eingesetzt und eine Möglichkeit geschaffen, mechanische Strukturen auf Schäden zu überwachen. Zur Regelung der Zielführung von Lenkwaffen ist das Kalman-Filter ein oftmals eingesetztes Mittel, wie von Patra et al. in [KKS14] gezeigt. Neben diesen Spezialanwendungen hat das Kalman-Filter Einzug in die Standardelektronik gehalten. So wäre beispielsweise die Nutzung des GPS-Systems zur Fahrzeugnavigation oder die Positionsbestimmung mit Smart-Phones ohne das Kalman-Filter nicht möglich.

Die Optimalität, auf die in späteren Abschnitten noch eingegangen wird, die einfache Implementierbarkeit, die rekursive Funktionsweise und die Möglichkeit, dynamische Systeme einzubinden und jedwede Sensorsignale dieses Systems verarbeiten zu können, führten in den letzten Jahrzehnten zum Siegeszug des von Rudolf Kalman 1960 in [Kal60] vorgestellten Filters. Den Durchbruch brachte schließlich die Weiterentwicklung des Kalman-Filters durch Schmidt zum Kalman-Schmidt-Filter [SM85] oder dem besser bekannten Extended-Kalman-Filter, das für nicht-lineare Systeme geeignet ist und seine Bestimmung in der Navigationsregelung der Apollo-Raumfähre fand, allerdings mit einigen strukturbedingten Nachteilen einhergeht. Hierzu zählen die Notwendigkeit der Berechnung von Jacobimatrizen um aktuelle Arbeitspunkte, die schwierig zu berechnen und zu implementieren sind. Darüber hinaus stellen sie nur eine Approximation der von Kalman be-

wiesenen Optimalität dar, und eine korrekte Funktionsweise des Filters ist nur in engen Grenzen möglich, worauf in den folgenden Abschnitten noch eingegangen wird und was beispielsweise von van der Merwe in [VDM04] gezeigt wurde. Da allerdings fast alle natürlichen Systeme nichtlinearer Natur sind, stellt das Extended-Kalman-Filter die am stärksten genutzte Variante des Kalman-Filters dar.

Alle oben genannten Beispiele basieren auf den von Kalman hergeleiteten Gleichungen und stellen Weiterentwicklungen für spezifische Problemstellungen dar. Eines haben allerdings alle genannten Anwendungen gemeinsam: Sie besitzen als Grundlage für die Nutzung der Beobachterstrukturen ein analytisches Modell in Zustandsraumdarstellung.

Doch wie ist eine Zustands- und Parameterschätzung an Systemen zu realisieren, deren Entwicklung mit modernen Entwicklungstools und -systematiken durchgeführt wurde, wie in den Abschnitten 1.1 und 1.2 beschrieben?

1.3 Problemstellung und Anforderungen

In 1.1 wurde gezeigt, dass die Entwicklung komplexer mechatronischer Systeme einen domänenspezifischen Entwurf nicht mehr zulässt und die modellbasierte Vorgehensweise in allen etablierten Entwicklungssystematiken den Stand der Technik darstellt. Die Verbindung der fachdisziplinspezifischen Auslegung in der Regelungstechnik zur Entwicklung von Zustands- und Parameterschätzern für etwa Reglerrealisierung, Systemidentifikation und Model-Predictive-Control-Anwendungen stellt dabei allerdings neue Herausforderungen an den Entwicklungsprozess an sich und an notwendige Werkzeuge, deren Anforderungen im Folgenden hergeleitet werden.

Die notwendige domänenübergreifende Modellierung von komplexen mechatronischen Systemen stellt immer höhere Anforderungen an zu nutzende Entwicklungsumgebungen. Eine signalflussorientierte Modellierung geht mit der Notwendigkeit der Herleitung von systembeschreibenden Differentialgleichungen und deren manueller Implementierung in Tools wie beispielsweise MATLAB/Simulink[®] einher. Dieses Vorgehen stellt allerdings einen äußerst fehleranfälligen Prozess dar, der nur das Expertenwissen des Modellbildungsingenieurs berücksichtigt. Dies hat in den letzten Jahren zur Entwicklung von topologieorientierten Modellierungswerkzeugen wie beispielsweise Dymola, MATLAB/Simscape und MapleSim geführt, deren Modelle im Weiteren für Echtzeitanwendungen genutzt werden können. Die Modellierung erfolgt dabei mit Hilfe von Bibliothekselementen, die von Domänenexperten, wie Zulieferern, Dienstleistern oder Universitäten, entwickelt und bereitgestellt werden und die in Abschnitt 1.2.1 erläuterten Vorteile mit sich bringen. Der Nachteil in der Nutzung von vorgefertigten Modellen besteht allerdings in der Tatsache, dass der damit arbeitende Ingenieur nicht mehr Experte in seinem Modell ist. Das Verständnis für genutzte physikalische Effekte,

für deren Modellierung und Detailsystemverhalten geht verloren. Beispielhaft zu nennen ist die Nutzung einer Gleichstrommaschine aus der Standardbibliothek des Tools Dymola, wie in Bild 1-6 dargestellt. Die Funktionsweise des genutzten Elementes ist zwar bekannt, aber den wenigsten Anwendern werden die grundlegenden Systemgleichungen und deren Nichtlinearitäten in den untergeordneten Ebenen des Elementes bekannt sein, wie die nichtlinearen Reibungseffekte in der Rotorlagerung einer Gleichstrommaschine.

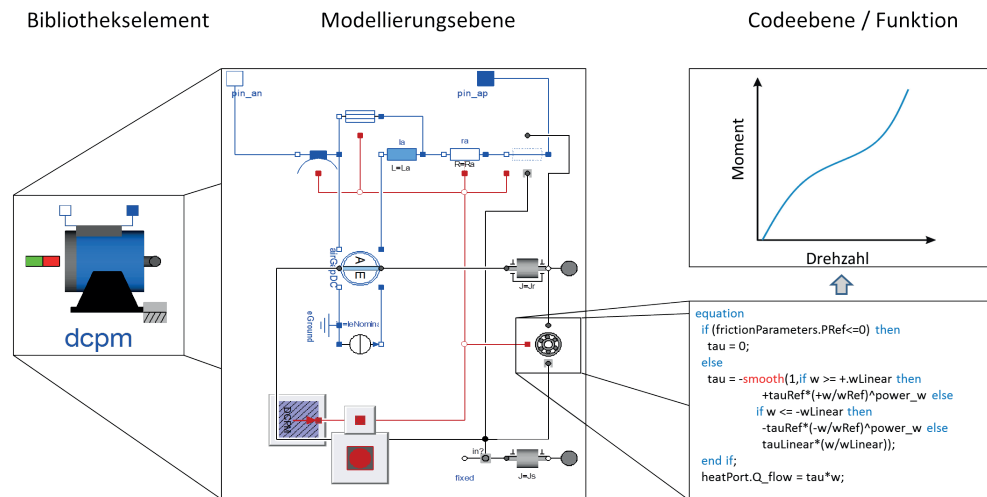


Bild 1-6: Element aus einer Dymola-Bibliothek, bestehend aus einer Gleichstrommaschine (links), der unterlagerten Modellierungsebene (Mitte), der Gleichungsebene eines nichtlinearen Wälzlagers zur Wellenlagerung (rechts unten) und der Funktionsebene zur Momentenberechnung in Abhängigkeit der Drehzahl (rechts oben)

Verschärft wird diese Eigenschaft durch die Nutzung von Entwicklungsmethodiken wie ENTIME, die, wie in Abschnitt 1.1 aufgeführt, zu einer domänenübergreifenden, teil- oder vollautomatisierten Modellerstellung auf Basis von Lösungsmustern und Lösungselementen führt. Im ungünstigsten Fall kennt der Regelungstechniker nicht nur einzelne Elemente nicht im Detail, sondern nur die grundlegende Funktion bzw. das Ein-/ Ausgangsverhalten. Im Allgemeinen sind weder dessen interne Struktur noch enthaltene Nichtlinearitäten bekannt, was eine aufwändige Analyse oder eine detaillierte Einarbeitung in das Modell zur Folge hat und den Vorteil der automatisierten Modellableitung zunichtemacht. Hieraus resultieren die erste und die zweite Anforderung an die Auslegungsmethodik und an Werkzeuge einer Zustands- und Parameterschätzung.

Anforderung I

Methoden und Werkzeuge zur Zustands- und Parameterschätzung für Systeme, deren Entwurf modellbasiert auf Basis von teilautomatisierten Entwurfssystematiken und/oder auf Basis von Komponentenbibliotheken durchgeführt wurde, müs-

sen in der Lage sein, unbekannte Modellstrukturen in Form von Systemordnung, Frequenz- und Zeitverhalten zu nutzen.

Anforderung II

Der Filteralgorithmus muss robust sein gegen unbekannte, starke Nichtlinearitäten des Modells, sodass eine möglichst hohe Genauigkeit und Filterstabilität garantiert werden kann.

Da die Modelle nicht im Detail bekannt sind, können Entwurfparameter von Schätzverfahren nicht auf Basis von Expertenwissen ausgelegt werden, was zur dritten Anforderung des Entwurfsprozesses führt.

Anforderung III

Die Entwurfparameter des Schätzalgorithmus müssen auf Basis des Systemmodells und auf Basis von Anwendungsszenarien mit einem maximalen Automatisierungsgrad ausgelegt werden.

Der Nutzer muss die Möglichkeit haben, während aller Schritte des Entwurfsprozesses Genauigkeit und Funktionsfähigkeit des automatisiert ausgeführten Entwurfsverfahrens zu kontrollieren und das ausgelegte Verfahren modellbasiert zu testen.

Anforderung IV

Frühestmögliche Kontrolle der Ergebnisse und Offline-Verfahrenstest, um Robustheit, Genauigkeit und Schnelligkeit zu bewerten.

Da das Ziel der eingesetzten Verfahren zur Zustands- und Parameterschätzung beispielsweise in der Regelung von mechatronischen Systemen liegt, müssen alle genutzten Verfahren zu Zustandsschätzern führen, die im Betrieb nach der Auslegungsphase Echtzeitfähigkeit aufweisen.

Anforderung V

Alle eingesetzten Werkzeuge zum Betrieb nach der Auslegungsphase müssen Echtzeitfähigkeit aufweisen.

Insbesondere im Bereich der Parameterschätzung darf es nicht zur Identifikation von nichtphysikalischen Werten kommen, etwa zur Bestimmung von negativen Massen. Um dies sicherzustellen, muss das Schätzverfahren in der Lage sein, lineare und nichtlineare Nebenbedingungen von Parametern und Zuständen verarbeiten zu können.

Anforderung VI

Das Verfahren zur Zustands- und Parameterschätzung muss in der Lage sein, Nebenbedingungen in Form linearer oder nichtlinearer Gleichheits- und Ungleichheitsnebenbedingungen verarbeiten zu können.

Auf Basis der gestellten Anforderungen wird in den folgenden Abschnitten der Stand der Technik vorgestellt, der Handlungsbedarf identifiziert und das Ziel der vorliegenden Arbeit definiert.

1.4 Stand der Wissenschaft und der Technik

In diesem Abschnitt wird auf den derzeitigen Stand der Wissenschaft und der Technik bezüglich der in den letzten Abschnitten hergeleiteten Anforderungen eingegangen. Der Fokus liegt dabei auf der Anwendung von Kalman-Filter-Techniken, da diese in ihrer rekursiven Form durch zahlreiche Anwendungen ihre praktische Bedeutung zur numerischen Ermittlung der Momente einer Verteilungsdichtefunktion bewiesen haben [BS94], worauf in Kapitel 2 detailliert eingegangen wird.

Wie bereits in 1.2.2 aufgeführt, zählt das Kalman-Filter mit seiner Erweiterung für nichtlineare Systeme, dem Extended-Kalman-Filter, zu den am weitesten verbreiteten Algorithmen zur Zustands- und Parameterschätzung [JU96]. Der Einsatz des Extended-Kalman-Filters geht allerdings mit einer Vielzahl an strukturbedingten Nachteilen einher (die in 2.2.1 erläutert werden), wodurch es nicht in der Lage ist Modelle mit starken Nichtlinearitäten mit gesicherter Stabilität zu nutzen.

In [LBG⁺12] werden von Li et al. Luenberger-Beobachter, Extended-Kalman-Filter und Sigma-Punkte-Filter für die Zustands- und Parameterschätzung von Lithium-Ionen-Batterien eingesetzt und verglichen. Die Grundlage ist ein analytisches Modell des Systems, mit dem Ergebnis, dass Sigma-Punkte-Filter bei Weitem am Besten mit den starken Systemnichtlinearitäten umgehen können. Xiong et al. erweitern den Ansatz in [XSC⁺14], um online unbekannte Systemparameter eines Black-Box-Modells³ zu identifizieren. Hierfür kombiniert er ein Extended-Kalman-Filter auf Basis des bekannten Systemmodells mit Neuronalen Netzen zur Erfassung des unbekannten Black-Box-Modells. Die Kombination dieser beiden Ansätze kann für stark nichtlineare Systeme eingesetzt werden, bedarf allerdings eines bekannten Systemmodells als Grundlage für das Extended-Kalman-Filter, das bei starken Nichtlinearitäten divergieren kann.

Ein Vergleich von KI-Methoden⁴ in Form von Neuronalen Netzen mit Kalman-Filtern wird von Psychogios und Ungar in [PU92] durchgeführt. Hierbei wird die Strecke zur Nutzung von Neuronalen Netzen aufgeteilt in einen gut bekannten, vorab modellierbaren Teil und ein System, das nicht zu modellieren oder nicht ausreichend bekannt ist und online durch das Neuronale Netz erlernt wird. Der zweite Ansatz besteht in der vollständigen Implementierung aller Effekte in das physikalische Modell und der Anwendung eines Extended-Kalman-Filters zur Schätzung der unbekannten Parameter. Beide Methoden bieten dabei Vor- und Nachteile. Die Anwendung des Neuronalen Netzes zeigt dessen Stärken, wenn das modellierte Teilsystem eine sehr gute Übereinstimmung mit dem realen Teilsystem besitzt, und übertrifft in diesen Fällen die Performance des Kalman-Filters. Treten allerdings kleine Abweichungen in diesem Modell auf, ist das Kalman-Filter besser in

³Modell, von dem nur das äußere Verhalten beobachtbar ist

⁴Künstliche Intelligenz

der Lage, diesen Effekt durch Schätzung der Parameter zu neutralisieren. Gleichzeitig wird auch aufgezeigt, dass bei der Nutzung von Neuronalen Netzen das Problem der Stabilität und des ausreichenden bzw. übertrainierten Netzes bei nicht im Detail bekannten Modellen auftritt.

In [KHW12] und [IKS10] werden von Kittilsen et al. Zustandsbeobachter in Form eines Extended-Kalman-Filters eingesetzt. Das Testsystem stellt eine Unterwasserpipeline mit einer Systemordnung von 48 dar. Zur Modelleinbindung in das Extended-Kalman-Filter ist ein generisches Interface entwickelt worden, und die Schätzung von Systemzuständen zeigt sehr gute Ergebnisse, wobei Filter- und Simulationsmodell durch das gleiche Modell gegeben sind. Bei der Einbindung von starken Nichtlinearitäten kann es beim Einsatz eines Extended-Kalman-Filters zu unvorhersehbaren Ergebnissen kommen, was den Einsatz für automatisch erstellte oder unbekannte Modelle nicht empfehlenswert macht.

Für die Anwendung in einer nichtlinearen modellprädiktiven Regelung (NMPC) werden von Gräber et al. in [GKS⁺12] im Detail bekannte Dymola-Modelle genutzt, die mittels des Functional-Mockup-Interfaces, auf das in Abschnitt 5.1 noch detailliert eingegangen wird, in eine Softwareumgebung zur Auslegung des NMPC eingebunden werden. Für die notwendige Zustandsschätzung werden wiederum Extended-Kalman-Filter eingesetzt, die nicht mit starken Modellnichtlinearitäten genutzt werden können. Die Bestimmung der Entwurfparameter erfolgt manuell, was mit der Nutzung eines bekannten Modells begründet wird.

Ebenfalls zur Schätzung des Ladezustandes von elektrischen Fahrzeugen wird von Brembeck et al. in [BOZ11] ein Dymolamodell in Verbindung mit einem Unscented-Kalman-Filter genutzt. Die Adaption des Modells an den Filteralgorithmus erfolgt, wie von Gräber vorgestellt, mittels des FMI-Standards. Allerdings ist eine schwierige manuelle Anpassung von mehreren Entwurfparametern durchzuführen, die entweder Expertenwissen über das Modell, wie in der Veröffentlichung vorhanden, oder einen aufwändigen Trial-and-error Prozess nach sich zieht.

Zhong et al. geben in [ZZH⁺14] einen guten Überblick über verschiedene Kalman-Filter-Typen und nutzen ein Particle-Filter für die Schätzung des Ladezustandes von elektrischen Fahrzeugen. Die Methode ist zwar in der Lage, ein unbekanntes Systemmodell aufzunehmen, allerdings wird keine automatisierte Auslegung genutzt, und gleichzeitig ist die Genauigkeit des Particle-Filters stark von der Anzahl der genutzten Particle abhängig, was die Nutzung in Echtzeitanwendungen bei großen Systemen und komplexen Dichtverteilungen fragwürdig macht.

Zur Identifikation unbekannter Parameter in nichtlinearen unbekannten Modellen werden von Wigren et al. in [WBT05] Kalman-Filter eingesetzt, die für eine korrekte Initialisierung eines RPEM-Algorithmus⁵ genutzt werden und bereits die Möglichkeit der Nutzung von unbekannten Modellen zeigen. Allerdings ist eine

⁵Recursive Prediction-Error Minimization

manuelle Auslegung des Filters notwendig und keine Online-Anwendung vorgehen.

Einen sehr guten Überblick über die Einbindung von linearen und nichtlinearen Nebenbedingungen für die Zustandsschätzung via Kalman-Filter-Techniken gibt Simon in [Sim10]. Er unterscheidet hierbei zwischen dem Standard-Kalman-Filter, Moving-Horizon-Estimation⁶, dem Unscented-Kalman-Filter und Particle-Filtern. Die Analyse von Simon zeigt, dass bereits eine Reihe von Verfahren für unterschiedliche Problemstellungen existieren, wie die Nutzung von zwei Unscented-Kalman-Filtern, wobei die Realisierung der Beschränkungsfunktionen über das zweite Filter erfolgt [TCT⁺09]. In [JL07] wird von Julier selber eine Möglichkeit vorgestellt, nicht nur die A-posteriori Zustandsschätzung, sondern auch die berechneten Sigma-Punkte nach definierten Nebenbedingungen zu begrenzen. In [XLD⁺13] werden die oben genannten Verfahren für ein dynamisches System angewandt und verglichen, wobei das Filtermodell im Detail bekannt ist.

1.5 Handlungsbedarf

Aus dem in 1.4 definierten Stand von Wissenschaft und Technik wird deutlich, dass bisher keine geschlossene Methodik für die Erfüllung der in 1.3 definierten Anforderungen zur Zustands- und Parameterschätzung an unbekannte Systeme mit unbekannten, starken Nichtlinearitäten existiert.

Der Handlungsbedarf besteht in der Erstellung einer durchgängigen Methode zur Einbindung von Modellen in einen echtzeitfähigen Algorithmus zur Zustands- und Parameterschätzung und der Entwicklung aller notwendigen Werkzeuge zu Modelleinbindung, Filterauslegung und Systemtest. Hierbei muss sowohl mit unbekannten Modellstrukturen als auch mit starken Nichtlinearitäten umgegangen werden, wofür bereits eine Reihe von Algorithmen genutzt werden können. Die Auslegung aller notwendigen Entwurfsparameter muss aufgrund des nicht vorhandenen Expertenwissens über das Modell automatisiert auf Basis des Ein-/Ausgangsverhaltens erfolgen, was bisher nicht methodisch aufgearbeitet worden ist, da eine Filterauslegung immer an bekannten Systemmodellen durchgeführt wurde. Die Einbeziehung von Nutzerinformationen muss allerdings weiterhin möglich sein, um beispielsweise die Identifikation von zu schätzenden Parametern zu ermöglichen.

Die Modelleinbindung, die Filterauslegung und der Systemtest sollen in dem Tool MATLAB erfolgen, da es das am Weitesten verbreitete Tool in den Ingenieurwissenschaften und insbesondere in der Regelungstechnik darstellt und somit die breiteste Basis der Zugänglichkeit bietet. Hierzu ist eine Schnittstelle zu schaffen, welche die Einbindung von Modellen unterschiedlicher Entwicklungstools wie Dymola und Adams ermöglicht. Alle notwendigen Informationen zur automatisierten Filteradaption und Auslegung müssen bereitgestellt werden. Es ist eine

⁶Iterativ rechnender Ansatz zur Zustands- und Parameterschätzung

Methode zur Einbindung von linearen und nichtlinearen Nebenbedingungen zu entwickeln, wobei zu beachten ist, dass diese ohne Expertenwissen über das Modell auszulegen und einzubinden sind, sowie Echtzeitbedingungen erfüllen muss, was bisher noch nicht möglich ist.

Das Ziel ist es, regelungstechnischen Anwendern alle notwendigen Werkzeuge und Methoden zur Verfügung zu stellen, um ohne Expertenwissen im zugrundeliegenden Modell und in den Algorithmen von Filtern zur Zustands- und Parameterschätzung eine Filterauslegung durchzuführen. Die Güte der Auslegung muss zu jeder Zeit bewertbar und die Ergebnisse im weiteren Projektverlauf nutzbar sein. Das entstandene Vorgehensmodell dieser Arbeit und dessen grundlegende Handlungsschritte sind in Bild 1-7 dargestellt. Gleichzeitig müssen bei unzureichender Filtergüte Handlungsempfehlungen vorhanden sein, um eine verbesserte Schätzung zu erreichen.

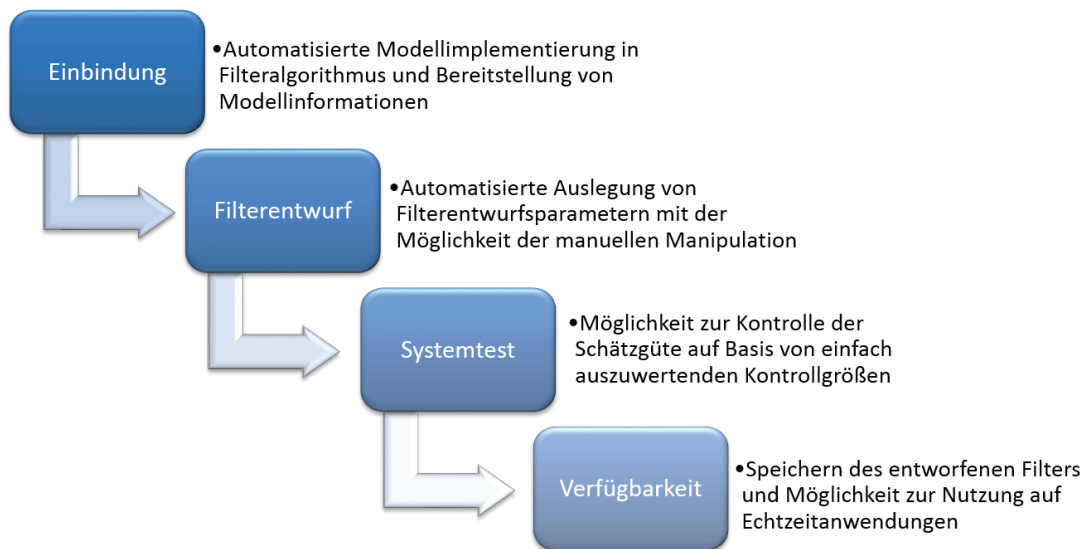


Bild 1-7: Vorgehensmodell zur Filterauslegung an unbekannten Systemmodellen mit einem maximalen Grad an Automatisierung zur Zustands- und Parameterschätzung

1.6 Aufbau der Dissertation

Diese Arbeit gliedert sich in fünf Bereiche. In Kapitel 2 wird zu Beginn eine kurze Einführung in Grundbegriffe der Statistik gegeben, die für das Verständnis dieser Arbeit wichtig sind. Nachfolgend werden einige Techniken zur Zustands- und Parameterschätzung vorgestellt, die von der Methode der kleinsten Quadrate über das Kalman-Filter und neue, darauf basierende Techniken der Sigma-Punkte- und Particle-Filter zu Verfahren des maschinellen Lernens führen. Abschließend werden eine kurze Zusammenfassung gegeben und die Auswahl der in dieser Arbeit genutzten Sigma-Punkte-Filter begründet.

In Kapitel 3 wird die Klasse der Sigma-Punkte-Filter ausführlich vorgestellt, sowohl mit unterschiedlichen Filtervarianten als auch Sigma-Punkte-Berechnungsmethoden, die hinsichtlich Echtzeitfähigkeit und Schätzungsgüte verglichen werden. Für die Nutzung in Parameterschätzungsaufgaben werden die jeweiligen Filtererweiterungen vorgestellt und ein Verfahren zur Einführung von Nebenbedingungen für die Begrenzung der Zustands- und Parameterschätzung und berechneten Sigma-Punkten hergeleitet, bevor eine Zusammenfassung der Sigma-Punkte-Verfahren folgt.

Methodik und Werkzeuge für eine automatisierte Ermittlung von Entwurfsparametern werden in Kapitel 4 vorgestellt. Hierzu wird ein Optimierungsproblem definiert, das mittels Kernel-Techniken des maschinellen Lernens gelöst wird. Die notwendigen Grundbegriffe des maschinellen Lernens werden erläutert, die Anpassung an das Optimierungsproblem vorgestellt und eine Methode mit allen benötigten Algorithmen hergeleitet. Das oftmals auftretende Problem des *over-* und *underfittings* wird durch ein Lernverfahren vermieden, das die Hyperparameter des Kernel-Verfahrens problemspezifisch bestimmt. Abschließend erfolgt die Anwendung zur Bestimmung der Entwurfsparemeter zur Zustands- und Parameterschätzung, ohne dass der Benutzer Experte in der Filtertechnik, dem eingesetzten Modell oder der Optimierungsstrategie sein muss.

Nachdem die genutzten Filteralgorithmen der Sigma-Punkte-Filter vorgestellt und ein modellbasiertes Lernverfahren auf Basis eines Optimierungsproblems hergeleitet wurden, wird in Kapitel 5 auf Basis der Anforderungen der unterschiedlichen Werkzeuge eine entwickelte Schnittstelle präsentiert. Durch diese ist die Einbindung einer breiten Klasse von Modellen unterschiedlicher Entwicklungstools möglich, und alle notwendigen Informationen für eine Filterauslegung ohne Expertenkenntnisse in dem Modell werden bereitgestellt. Im Anschluss werden alle entwickelten Algorithmen in eine Gesamtmethodik eingebunden, die dem Anwender eine Schritt-für-Schritt Anleitung zur Filterauslegung bietet, mit der Möglichkeit des manuellen Eingreifens, wenn Expertenwissen über bestimmte Teilgebiete vorhanden ist. Die Methodik gliedert sich dabei in die Unterpunkte der Modelleinbindung, der Bestimmung der Filterentwurfsparemeter, des Offlinesystemtests und der Codegenerierung sowie Handlungsempfehlungen bei unzureichender Schätzungsgüte. Abschließend wird die entstandene MATLAB-Toolbox vorgestellt, die alle genannten Werkzeuge enthält und eine automatisierte Auslegung unter MATLAB ermöglicht.

Auf Basis der vorgestellten Methodik werden in Kapitel 6 unterschiedliche Anwendungsbeispiele sowie die implementierten Filter und deren Varianten hinsichtlich Schätzungsgüte und Rechenbelastung vorgestellt. Im ersten Fall handelt es sich um ein Laborbeispiel, bestehend aus einem Doppelpendel, an dem eine Zustands- und Parameterschätzung mit unterschiedlichen Implementierungsvarianten vorgestellt wird. Den zweiten und dritten Fall stellen Industrieanwendungen dar. Zum Einen wird die Schätzung einer zeitvarianten Unwucht in einem Waschautomaten durchgeführt, anhand derer die Schleuderstrategie ausgelegt werden kann.

Zum Zweiten wird eine Zustands- und Parameterschätzung an dem leicht gepanzerten Mehrzweckschutzfahrzeug DINGO 2 der Bundeswehr vorgenommen, die auf realen Messungen basiert.

In Kapitel 7 werden die in dieser Dissertation vorgestellten Methoden, Filteralgorithmen und entwickelten Verfahren zusammengefasst und ein Ausblick für weitere Arbeiten in diesem Umfeld gegeben.

2 Zustands- und Parameterschätzung

Da der Kern der Arbeit in der Entwicklung einer Methode und aller notwendigen Werkzeuge bzw. Algorithmen für die automatische Auslegung von Zustands- und Parameterschätzern für unbekannte, nicht-analytische Modelle liegt, wird in diesem Abschnitt eine Einführung in die Zustands- und Parameterschätzung gegeben. Es werden zu Beginn kurz einige statistische Grundbegriffe erläutert, die in dieser Arbeit genutzt werden.

Das am Weitesten verbreitete Werkzeug zur Zustands- und Parameterschätzung ist das Kalman-Filter, das als Grundlage für die in dieser Arbeit angewandten Techniken zur Bestimmung von Zuständen und Parametern eines dynamischen Systems dient. Aus diesem Grund wird in diesem Kapitel eine Einführung in die Funktionsweise der Schätzung mittels Kalman-Filtern gegeben und dessen Entstehung erläutert, um zu zeigen, dass das Kalman-Filter auf statistischem Weg hergeleitet werden kann. Dies soll den Grundstein legen für die Herleitung von Sigma-Punkte-Kalman-Filtern, die das herkömmliche Kalman-Filter erweitern und als Basis zur Realisierung der in Abschnitt 1.5 hergeleiteten Anforderungen dienen. Ihre Herleitung erfolgt in Kapitel 3.

In den darauf folgenden Abschnitten wird ein kurzer Überblick über weitere mögliche Techniken zur Zustands- und Parameterschätzung gegeben und erläutert, warum diese in der vorliegenden Arbeit nicht angewandt wurden.

In Abschnitt 1.1 wurde gezeigt, dass mit Hilfe der Entwurfssystematik ENTIME die Möglichkeit besteht, mit maximalem Automatisierungsgrad dynamische Modelle von technischen Systemen zu erzeugen. Es gibt eine Vielzahl von Gründen warum diese Modelle in der Analyse wie auch in der Synthese nicht ausreichend sind für einen Ingenieur. Als wichtigster Grund ist anzugeben: Kein mathematisches Modell eines Systems ist hinsichtlich Güte, Genauigkeit und Einbeziehung von Störgrößen perfekt modelliert [May79].

Alle entwickelten Modelle beinhalten nur die Charakteristiken der Realität, die auf den Interessen des entwickelnden Ingenieurs beruhen. Beispielsweise sind bei der Modellierung eines Waschautomaten eine Vielzahl an Effekten vorhanden, die auf die Schwingung des Aggregates einwirken. Der Sinn des Modells besteht allerdings darin, die essentiellen oder kritischen Effekte des Waschautomaten abzubilden, sodass viele Effekte wissentlich vernachlässigt werden. Modelle, die im späteren Verlauf für den Einsatz in HiL-Umgebungen oder für eine direkte Implementierung auf dem Steuergerät vorgesehen sind, müssen dabei auf das Notwendigste beschränkt werden, um einen noch in Echtzeit ausführbaren Algorithmus zu liefern. Die notwendigen, abzubildenden Effekte werden dabei allerdings auch nur durch mathematische Modelle approximiert. Beispielfhaft zu nennen sind die Newtonschen Gesetze, die zwar eine sehr gute Struktur bieten, um ein beobachtetes Verhalten zu approximieren, aber etliche Parameter in diesen Strukturen

enthalten, die wiederum schwierig auszulegen sind und Unsicherheiten in das zu nutzende Modell übertragen.

Ein weiterer wichtiger Grund liegt in den auf das System wirkenden Eingängen. Diese bestehen nicht nur aus den dem Ingenieur bekannten Größen, sondern ebenfalls aus Störungen, die der Ingenieur nicht kennt oder gar kontrollieren könnte.

Ein dritter Grund liegt in den Sensordaten. Messungen eines realen Systems unterliegen immer Störungen, die nicht vorherzusehen sind. Zudem können Sensoren oftmals nicht alle notwendigen Größen eines Systems liefern. Beispielhaft zu nennen ist die Regelung eines Reaktors. Prozesse, die im Inneren dieses Reaktors ablaufen, sind mit messtechnischen Mitteln nur mit extremem Aufwand oder gar nicht zu erfassen. In anderen Fällen liegen eine Reihe von Sensoren vor, die redundante Informationen liefern. Wie sind solche Sensorgrößen zu einer Größe zu verarbeiten bzw. welcher Größe sollte vertraut werden?

Das am Weitesten verbreitete Werkzeug zur Bestimmung von nicht messbaren oder nur schwer messbaren Systemzuständen, der Unterdrückung von Rauschprozessen in Messungen, des Ausgleichs von Modellungenauigkeiten im Gegensatz zum realen System und der Schätzung von unbekannten Systemzuständen sowie Systemparametern ist das in Online-Anwendungen nutzbare Kalman-Filter.

Im nächsten Abschnitt wird zunächst ein kurzer Überblick über die in dieser Arbeit genutzten statistischen Begriffe gegeben. Im Anschluss werden das Grundprinzip des Kalman-Filters vorgestellt, seine Funktionsweise auf das Verfahren von Gauss aus dem Jahre 1795 zurückgeführt sowie die Verwandtschaft zum Bayesschen Schätzproblem hergestellt. Nachfolgend werden die Nachteile des Kalman-Filters aufgezeigt und dessen Erweiterung zur Anwendung auf nichtlineare Systeme mittels des Extended-Kalman-Filters erläutert. Abschließend werden kurz alternative Verfahren zur Bestimmung von Zuständen und Parametern von dynamischen Systemen aufgeführt und gezeigt, warum in dieser Arbeit auf dem Kalman-Filter basierende Verfahren eingesetzt werden.

2.1 Statistische Grundlagen

In diesem Abschnitt sollen kurz einige Grundbegriffe der Statistik erläutert werden, die in dieser Arbeit genutzt werden. Schwerpunkte stellen die Begriffe der Wahrscheinlichkeit, der Zufallsvariable, der Verteilungsfunktion und der statistischen Momente dar, die zur Herleitung von Schätzalgorithmen auf Basis von wahrscheinlichkeitstheoretischen Betrachtungen genutzt werden.

2.1.1 Wahrscheinlichkeit

Der Begriff der Wahrscheinlichkeit definiert die Sicherheit einer Vorhersage für das Eintreten eines bestimmten Ereignisses [WB06] und ist definiert als eine nicht-

negative Zahl [Frä05]:

$$P(\mathcal{A}) = \frac{\text{Mögliche Ereignisse bezüglich des Ereignisses } \mathcal{A}}{\text{Gesamtzahl aller möglichen Ereignisse}} \geq 0. \quad (2-1)$$

Geht man von zwei disjunkten Ereignissen aus, so berechnet sich die Wahrscheinlichkeit, dass das Experiment mindestens eines der beiden Ereignisse hervorbringt, als Summe der Einzelwahrscheinlichkeiten:

$$P(\mathcal{A} \cup \mathcal{B}) = P(\mathcal{A}) + P(\mathcal{B}). \quad (2-2)$$

Falls die Wahrscheinlichkeiten zweier Ereignisse voneinander unabhängig sind und die Ereignisse sich nicht gegenseitig beeinflussen können, dann ist die Wahrscheinlichkeit, dass beide Ereignisse eintreten, wie folgt definiert;

$$P(\mathcal{A} \cap \mathcal{B}) = P(\mathcal{A})P(\mathcal{B}). \quad (2-3)$$

Aus diesen Definitionen lässt sich direkt der Begriff der bedingten Wahrscheinlichkeit angeben, der in der stochastischen Beschreibung von Systemprozessen eine wichtige Rolle spielt und in Abschnitt 2.2.3 genutzt wird. Er beschreibt die Wahrscheinlichkeit des Auftretens eines Ereignisses \mathcal{A} , wenn das Auftreten des Ereignisses \mathcal{B} gesichert ist [Frä05]:

$$P(\mathcal{A}|\mathcal{B}) = \frac{P(\mathcal{A} \cap \mathcal{B})}{P(\mathcal{B})}. \quad (2-4)$$

2.1.2 Zufallsvariable und Dichtefunktion

Eine Zufallsvariable ξ beschreibt den Wert einer Größe, deren Realisierung vom Zufall abhängig ist [BN06] und der eine Eintrittswahrscheinlichkeit $P(\mathcal{A})$ zugeordnet ist. Einer Zufallsvariablen ist eine Verteilungsfunktion $G(x)$ zugeordnet, die angibt, mit welcher Wahrscheinlichkeit die Zufallsvariable einen Wert kleiner oder gleich dem Wert x annimmt:

$$G(x) = P\{\xi \leq x\}. \quad (2-5)$$

Als Dichtefunktion, oft als Wahrscheinlichkeitsdichtefunktion, Wahrscheinlichkeitsdichte oder Dichte bezeichnet, wird die Ableitung der Verteilungsfunktion $G(x)$ angegeben [Frä05], die eine wesentliche Zielgröße in der stochastischen Herleitung von Zustandsschätzern ist:

$$p(x) = \frac{dG(x)}{dx}. \quad (2-6)$$

Die Integration der Dichtefunktion einer Zufallsvariablen ξ über ein gegebenes Intervall $[a, b]$ gibt die Wahrscheinlichkeit an, dass sich ξ in dem Wertebereich von $[a, b]$ befindet, d. h. es gilt:

$$\int_a^b p(x)dx = P(a \leq \xi \leq b) \quad (2-7)$$

und damit insbesondere

$$\int_{-\infty}^{\infty} p(x) dx = 1. \quad (2-8)$$

2.1.3 Statistische Momente

Im Allgemeinen ist die Angabe der in den letzten Abschnitten gezeigten Wahrscheinlichkeitsdichtefunktion nicht möglich, was in Abschnitt 2.2.3 gezeigt wird. Daher wird sich oft auf die statistischen Momente dieser Verteilung konzentriert, wie z.B. dem Erwartungswert $E[x]$. Bei einer endlichen Anzahl N an Stichproben ergibt sich der Mittelwert zu

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_N}{N} \quad (2-9)$$

und strebt für $N \rightarrow \infty$ gegen den Erwartungswert $E[x]$. Ist die Dichtefunktion der Verteilung bekannt, so lässt sich der Erwartungswert einer diskreten Zufallsvariablen nach

$$\hat{x} = E[x] = \sum_{i=1}^{\infty} x_i p(x_i) \quad (2-10)$$

bestimmen. Der Erwartungswert für eine kontinuierliche Zufallsvariable ist durch

$$\hat{x} = E[x] = \int_{-\infty}^{\infty} x p(x) dx \quad (2-11)$$

definiert. Zur Beschreibung der Streuung einer Zufallsvariablen wird die Varianz $\text{var}[x]$ bzw. deren Wurzel, die Standardabweichung, genutzt, die für eine diskrete Verteilung als

$$\text{var}[x] = \sum_{i=1}^{\infty} (x_i - E[x])^2 p(x_i) \quad (2-12)$$

definiert ist und nach

$$\text{var}[x] = \int_{-\infty}^{\infty} (x - E[x])^2 p(x) dx \quad (2-13)$$

für eine kontinuierliche Zufallsvariable. Die Streuung ist ein Maß, wie hoch die Wahrscheinlichkeit ist, dass die Zufallsvariable Werte in der Nähe des Erwartungswertes annimmt oder weiter entfernte Werte. Im weiteren Verlauf der vorliegenden Arbeit wird in Ermangelung der Wahrscheinlichkeitsdichte die Kovarianzmatrix einer Zufallsvariablen in der Form

$$P_{xx} = E \left[(x - \hat{x}) (x - \hat{x})^T \right] \quad (2-14)$$

genutzt. Erwartungswert und Varianz stellen Spezialfälle einer allgemeinen Klasse von Parametern zur Charakterisierung von Wahrscheinlichkeitsverteilungen (pdf) dar [EKT08]. Hierbei handelt es sich um die Momente einer Zufallsvariablen, die unterschieden werden in die Momente um Null und die Momente um den Erwartungswert, die sogenannten zentralen Momente.

Das zentrale Moment k -ter Ordnung ist durch

$$E \left[(x - E[x])^k \right] = \sum_{i=1}^{\infty} (x_i - E[x])^k p(x_i) \quad (2-15)$$

für den diskreten Fall und durch

$$E \left[(x - E[x])^k \right] = \int_{-\infty}^{\infty} (x - E[x])^k p(x) dx \quad (2-16)$$

für den kontinuierlichen Fall definiert. Für $k = 3$ erhält man nach Normierung das dritte Moment der Schiefe ν (siehe Bild 2-1) einer Verteilung, die ein Symmetriemaß darstellt und rechts- oder linksschief sein kann, wie in Bild 2-1 dargestellt. Beispielhaft zu nennen ist das Pro-Kopf-Einkommen, bei dem einige wenige Personen sehr hohe Einkommen erzielen und sehr viele Personen eher geringe Einnahmen aufweisen können.

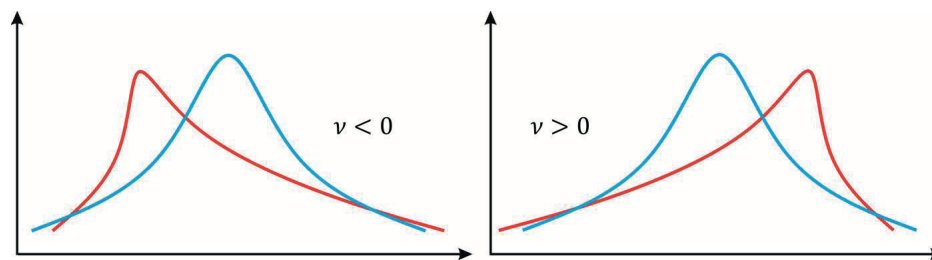


Bild 2-1: Links: Rechtsschiefe Verteilung, in Rot dargestellt gegenüber Normalverteilung in Blau. Rechts: Linksschiefe Verteilung, in Rot dargestellt gegenüber Normalverteilung in Blau

Die Wölbung, auch Kurtosis genannt, ist ein Maß für die relative Steilheit einer Verteilung und bildet das vierte, normierte Moment einer Verteilung. Die Angabe erfolgt häufig mittels des Exzesses γ . Positive Werte γ beschreiben eine leptokurtische Verteilung, was einer spitz zulaufenden Kurve entspricht (vgl. Bild 2-2 links). Negative γ entsprechen einer platykurtischen Verteilung und beschreiben einen eher flachen Verlauf, wie in 2-2 rechts dargestellt.

Die Wölbung einer Normalverteilung kann mit 3 angegeben werden. Um allerdings eine einfache Unterscheidung der Wölbung anderer Verteilungen nach dem Exzess $\gamma_W > 0$ bzw. $\gamma_W < 0$ zu ermöglichen, werden diese mittels

$$\gamma_W = \text{Exzess} = \text{Wölbung} - 3 \quad (2-17)$$

auf die Wölbung der Normalverteilung bezogen. Bei $\gamma_W = 0$ spricht man demnach von einer normalgipfligen Verteilung, auch mesokurtisch genannt [Ros06].

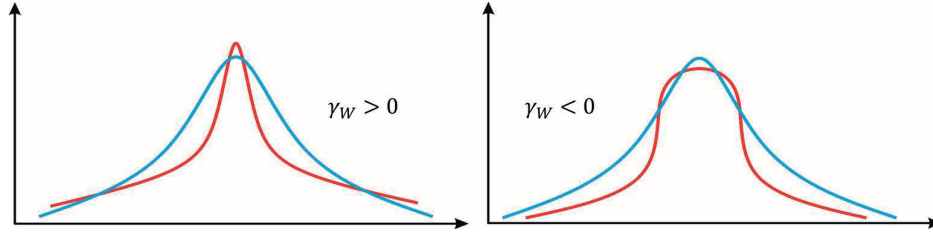


Bild 2-2: Links: Verteilung mit einer positiven Wölbung, einem steilen Verlauf, in Rot gegenüber eine Normalverteilung in Blau. Rechts: Verteilung mit einer negativen Wölbung, einem flachen Verlauf, in Rot gegenüber einer Normalverteilung in Blau

2.2 Das Kalman-Filter

Ziel des Kalman-Filters ist die Schätzung von Zuständen $x_k \in \mathbb{R}^{n \times 1}$ mit n als Systemordnung eines Systems zum diskreten Zeitpunkt¹ k auf Basis von verrauschten Sensorinformationen. Im Vergleich zu anderen Verfahren zur Bestimmung der Zustände eines Systems betrachtete Kalman das Problem von einer stochastischen Sichtweise, wodurch dynamische Systeme eingebunden werden können und Modellunsicherheiten sowie fortwährende stochastische Störungen durch Rauschprozessen repräsentiert werden. Die stellte R. Kalman erstmalig in [Kal60] vor und erlangte damit große Aufmerksamkeit. Das übergeordnete Ziel des Kalman-Filters ist die Schätzung des Zustandsvektors zum Zeitpunkt k als Erwartungswert der unbekannten Zufallsvariablen x_k mit

$$\hat{x}_k = E[x_k] \quad (2-18)$$

unter Minimierung der Schätzfehlerkovarianz, mit dem Schätzfehler $\tilde{x}_k = x_k - \hat{x}_k$

$$P_{xx,k} = E \left[(x_k - \hat{x}_k) (x_k - \hat{x}_k)^T \right] = E \left[\tilde{x}_k \tilde{x}_k^T \right]. \quad (2-19)$$

Die eigentliche Schätzung des Kalman-Filters wird in einem zweistufigen Prozess als Prädiktor- und Korrektor-Schritt durchgeführt, wie in Bild 2-3 dargestellt. Nach der Initialisierung erfolgen die Vorhersage der Zustände \hat{x}_k^- auf Basis eines vorhandenen Modells sowie die Schätzung der Fehlerkovarianz $P_{xx,k}^-$. Die Prädiktion unter gegebenen Startbedingungen oder anhand eines vorherigen Filterschrittes wird als A-priori-Schätzung bezeichnet und mit der Superposition – gekennzeichnet.

Die Korrektur der A-priori-Schätzung \hat{x}_k^- und $P_{xx,k}^-$ nach dem Auftreten von neuen Messungen y_k vom realen System wird als A-posteriori-Schätzung bezeichnet und ist in Bild 2-3 rechts dargestellt.

Der Kern der Vorhersage von Zuständen im Rahmen der A-priori-Schätzung ist durch ein Systemmodell der Form

$$x_k = Ax_{k-1} + Bu_{k-1} + v_{k-1} \quad (2-20)$$

¹Für zeitkontinuierliche Systeme kann das Kalman-Bucy-Filter genutzt werden

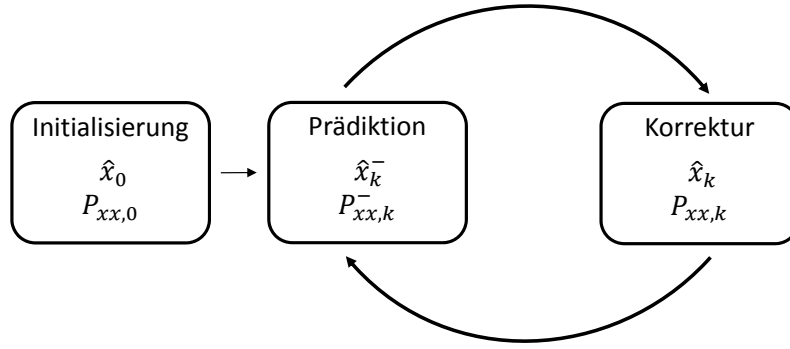


Bild 2-3: Prädiktor-Korrektor-Struktur des Kalman-Filters

gegeben, das hier als lineares Modell im Zustandsraum angegeben ist, mit v_k als Prozessrauschen, für das $p(v) \sim \mathcal{N}(0, R_v)$ gilt. Hierbei umfasst die Matrix A die Dynamik des Systems und wird als zeitinvariant angesehen, was allerdings keinesfalls eine Notwendigkeit des Kalman-Filters darstellt. Die Matrix B modelliert die Wirkung des als bekannt vorausgesetzten System- bzw. Kontrolleingangs u_{k-1} und wird ebenfalls zur Vereinfachung als zeitinvariant angesehen. Das Messmodell des Systems ist durch

$$y_k = Cx_k + n_k \quad (2-21)$$

gegeben. Die Matrix C stellt die Verteilung der Zustände x_k auf die Messgrößen y_k dar mit dem Messrauschen n_k für das $p(n) \sim \mathcal{N}(0, R_n)$ gilt.

Die beste A-priori-Schätzung des Zustandes zum Zeitpunkt k sei durch \hat{x}_k^- gegeben und basiert auf der besten A-posteriori-Schätzung zum letzten Zeitschritt $k-1$ bzw. der Initialisierung und dem bekannten Modell der Strecke:

$$\hat{x}_k^- = A\hat{x}_{k-1} + Bu_{k-1}. \quad (2-22)$$

Der Einfluss des Prozessrauschens v_{k-1} spielt demnach in der Berechnung nach (2-22) des A-priori-Zustandes keine Rolle, was in der Annahme als Gaußscher, mittelwertfreier Prozess mit $E[v_{k-1}] = 0$ begründet ist. Die beste Schätzung der Messungen ist durch das Messmodell aus (2-21) auf Basis der A-priori-Zustandschätzung \hat{x}_k^- gegeben, wobei n_k durch die Mittelwertfreiheit wieder vernachlässigt werden kann:

$$\hat{y}_k = C\hat{x}_k^-. \quad (2-23)$$

Die Schätzung des A-priori-Schätzfehlers $P_{xx,k}^-$ erfolgt bei Vorliegen eines linearen, zeitdiskreten Systems in Zustandsformdarstellung mittels

$$\begin{aligned} P_{xx,k}^- &= E[\tilde{x}_k^-(\tilde{x}_k^-)^T] \\ &= AP_{xx,k-1}A^T + R_v \end{aligned} \quad (2-24)$$

mit $\tilde{x}_k^- = x_k - \hat{x}_k^-$ als Schätzfehler zwischen realem Zustand x_k und der Schätzung \hat{x}_k^- und $R_v = E[v_{k-1}v_{k-1}^T]$. Die A-posteriori-Schätzung \hat{x}_k , also die Korrektur der A-priori-Schätzung \hat{x}_k^- , auf Basis der geschätzten Messgrößen \hat{y}_k beim Auftreten neuer Messungen y_k , realisierte Kalman durch eine lineare, affine Korrektur:

$$\hat{x}_k = \hat{x}_k^- + \mathcal{K}_k(y_k - \hat{y}_k). \quad (2-25)$$

Die Verstärkung \mathcal{K}_k bestimmte Kalman durch das Orthogonalitätsprinzip. Die optimale Lösung \hat{x}_k , welche die Kovarianz des Schätzfehlers $\tilde{x}_k = x_k - \hat{x}_k$ minimiert, ist demnach gegeben, wenn der Vektor des Schätzfehlers $x_k - \hat{x}_k$ statistisch orthogonal auf dem Raum der Messungen y_k aufsetzt. Anschaulich bedeutet dies, dass, wenn kein Messrauschen vorhanden ist, den Messungen demnach voll vertraut werden kann und die Schätzung des A-posteriori-Zustandes \hat{x}_k ausschließlich eine Abbildung zu den aktuellen Messungen ist. Da dies allerdings nie der Fall ist, liegt der optimale Schätzwert (im minimum-variance bzw. MMSE-Sinn) im geringsten Abstand zum Raum der Messungen, was einem orthogonalen Vektor entspricht und durch

$$E[(x_k - \hat{x}_k)y_k^T] = 0 \quad (2-26)$$

gegeben ist. Nach Einsetzen von (2-25) für \hat{x}_k und der Messgleichung (2-21) für y_k und \hat{y}_k in (2-26) gelangt man auf

$$\mathcal{K}_k = P_{xx,k}^- C^T (C P_{xx,k}^- C^T + R_n)^{-1}, \quad (2-27)$$

was in [Gib11] nachvollzogen werden kann. Nach der Berechnung der Kalman-Verstärkung \mathcal{K}_k kann die Korrektur der A-priori-Schätzung mittels

$$\hat{x}_k = \hat{x}_k^- + \mathcal{K}_k(y_k - C\hat{x}_k^-) \quad (2-28)$$

durchgeführt werden. Abschließend kann die Aktualisierung der Schätzfehlerkovarianz durch

$$\begin{aligned} P_{xx,k} &= E[\tilde{x}_k \tilde{x}_k^T] \\ &= (I - \mathcal{K}_k C) P_{xx,k}^- \end{aligned} \quad (2-29)$$

erfolgen, um eine angepasste Bewertung des Schätzfehlers zu ermöglichen. Durch (2-22)-(2-29) hat Kalman einen Prädiktor-Korrektor-Algorithmus entwickelt, der für ein dynamisches System mit zeitvarianten Rauschprozessen eine optimale, rekursive Schätzung im Minimum-Mean-Squared-Error (MMSE)-Sinn bietet. Dabei wird auf die Kenntnisse der Prozess- und Messdynamik sowie auf die statistischen Eigenschaften der Rauschprozesse zurückgegriffen. Es werden Unsicherheiten im Prozessmodell und jegliche Information über die Anfangsbedingungen der Schätzung genutzt [May79].

Die Annahme einer Gaußverteilung für die Rauschprozesse hat nicht nur den praktischen Aspekt der mathematischen Vereinfachung als Ursache. Ein Ingenieur

kennt im Allgemeinen oder im besten Fall “nur“ die Statistik erster und zweiter Ordnung der vorhandenen Rauschprozesse. Wenn keine weiteren Informationen über die höheren statistischen Momente vorliegen, gibt es allerdings auch keine bessere Annahme als die Gaußverteilung, da diese die ersten beiden Momente vollständig erfasst [May79].

Das Kalman-Filter setzte ab seiner Veröffentlichung im Jahr 1960 zu seinem Siegeszug in der Schätztheorie an und ist das meistgenutzte Filterverfahren. Neben dem Kalman-Filter existieren noch eine Vielzahl weiterer Typen, von denen nur einige kurz genannt werden. Bereits 1942 veröffentlichten Wiener [Wie64] und zeitgleich Kolmogorov ein Schätzverfahren im MMSE-Sinn, das allerdings ein zeitinvariantes Schätzproblem löst [Gib11] und als Ausgangspunkt der modernen Filtertheorie angesehen werden kann [Kre80]. 1955 baute Föllin [CF56] auf dieser Idee auf, entwarf für einen speziellen Anwendungsfall ein rekursiv arbeitendes Verfahren und bereitete den Weg für Kalman. Bucy erarbeitete in Zusammenarbeit mit Kalman den Algorithmus zur Nutzung für zeitkontinuierliche Modelle. 1964 wurde von David Luenberger ein Verfahren entwickelt, das für ein lineares System den Zustand unter einer Anfangsstörung rekonstruiert, allerdings keine dauerhaft einwirkenden Störungen berücksichtigt [Föll13]. Eine interessante Herleitung des Kalman-Filters aus dem Luenberger-Beobachter abseits der Stochastik kann in [Föll78] nachvollzogen werden und [May79] bietet ein anschauliches Beispiel der Funktionsweise des Kalman-Filters.

2.2.1 Das Extended-Kalman-Filter

Im letzten Abschnitt wurde gezeigt, dass Kalman sein Schätzverfahren für lineare Systeme angewendet hatte. Trotzdem war die erste nennenswerte Anwendung ein nichtlineares Problem: die Navigation der Apollo-Raumfähre. Als direktes Ergebnis der Zusammenarbeit von Kalman und Schmidt entstand das Kalman-Schmidt-Filter oder, in seiner weiter verbreiteten Form, das Extended-Kalman-Filter (EKF) [Gib11].

Die Grundidee besteht in der Linearisierung der nichtlinearen Systemgleichungen um eine Referenztrajektorie. Wie bei allen Linearisierungen ist dabei die Wahl des Linearisierungspunktes von entscheidender Bedeutung [Frä05]. Ein minimaler Linearisierungsfehler ist anzunehmen, wenn als Entwicklungspunkt der aktuelle Zustand genutzt wird, der aber nicht bekannt ist, da man ihn ja sonst nicht schätzen müsste. Als beste Wahl der Referenztrajektorie kann daher die Schätzung des Kalman-Filters selbst herangezogen werden [Sim06]. Ausgangspunkt stellt die nichtlineare Systemfunktion F dar, mit

$$x_k = F(x_{k-1}, u_{k-1}, v_{k-1}) \quad (2-30)$$

sowie der Messfunktion H , die eine nichtlineare Abbildung der Zustände auf die Messungen durchführt:

$$y_k = H(x_k, n_k). \quad (2-31)$$

Die Terme v_{k-1} und n_k , stellen wie im vorherigen Abschnitt bereits eingeführt, normalverteilte Rauschprozesse mit $p(v) \sim \mathcal{N}(0, R_v)$ bzw. $p(n) \sim \mathcal{N}(0, R_n)$ dar. Die beste A-priori-Schätzung basiert dabei auf der bekannten nichtlinearen Systemfunktion und ergibt sich, aufbauend auf der letzten A-posteriori-Zustandschätzung \hat{x}_{k-1} , zu (2-32):

$$\hat{x}_k^- = F(\hat{x}_{k-1}, u_{k-1}). \quad (2-32)$$

Zur Berechnung der A-priori-Kovarianz $P_{xx,k}^-$ wird angenommen, dass die Systemfunktion F als Taylorreihe mit Abbruch nach dem ersten Glied um den letzten A-posteriori-Schätzwert \hat{x}_{k-1} entwickelt werden kann.

$$F(\hat{x}_{k-1} + \Delta x, u_{k-1}, 0) \approx F(\hat{x}_{k-1}, u_{k-1}) + \left. \frac{\partial F}{\partial \hat{x}} \right|_{\hat{x}=\hat{x}_{k-1}} \cdot \Delta x \quad (2-33)$$

Der Term

$$A_k^{lin} = \frac{\partial F(\hat{x}_{k-1}, u_{k-1}, 0)}{\partial x}. \quad (2-34)$$

bezeichne die Jacobimatrix zur Bestimmung der A-priori-Kovarianz $P_{xx,k}^-$.

$$P_{xx,k}^- = A_k^{lin} P_{xx,k-1}^- (A_k^{lin})^T + R_v. \quad (2-35)$$

Im Anschluss kann die Schätzung der Messungen auf Basis von

$$\hat{y}_k = H(\hat{x}_k^-) \quad (2-36)$$

durchgeführt und die partielle Ableitung der Messfunktion zur Berechnung der Kalman-Verstärkung entsprechend (2-37) durchgeführt werden:

$$H_k^{lin} = \frac{\partial H(\hat{x}_k^-)}{\partial x}. \quad (2-37)$$

Die Berechnung der Kalman-Verstärkung ist danach direkt durch

$$\mathcal{K}_k = P_{xx,k}^{-1} (H_k^{lin})^T (H_k^{lin} P_{xx,k}^- (H_k^{lin})^T + R_n)^{-1} \quad (2-38)$$

möglich, woraufhin die Korrektur der A-priori-Schätzung des Zustandes \hat{x}_k^- und der Schätzfehlerkovarianz $P_{xx,k}^-$ durchführbar ist:

$$\begin{aligned} \hat{x}_k &= \hat{x}_k^- + \mathcal{K}_k (y_k - \hat{y}_k) \\ P_{xx,k} &= P_{xx,k}^- - \mathcal{K}_k H_k^{lin} P_{xx,k}^- \end{aligned} \quad (2-39)$$

Bei dieser sehr einfachen Erweiterung des Kalman-Filter-Algorithmus auf nicht-lineare Systeme muss allerdings eine gewisse Vorsicht beibehalten werden. Im Allgemeinen ist das EKF gegenüber dem Kalman-Filter nicht mehr erwartungstreu [Frä05]. Da die Transformation von Normalverteilungen durch eine nicht-lineare Funktion nicht wieder zu einer Normalverteilung führen muss, kann es in einigen Fällen zur Divergenz des Filters kommen, was anschaulich von Gibbs gezeigt wurde [Gib11]. Nachfolgend werden die Hauptprobleme bei der Nutzung des EKF kurz zusammengefasst.

- Eine wichtige Einschränkung liegt darin, dass die Wahrscheinlichkeitsverteilungen der einzelnen Zustände nicht länger normalverteilt sind, sobald diese durch eine nichtlineare Funktion propagiert werden. Das EKF ist allerdings nur ein einfaches Ad-hoc-Berechnungsverfahren, das allein die Optimalität der Bayesschen Regel über eine Linearisierung approximiert und dabei eine Normalverteilung der Zufallsgrößen voraussetzt [WB06], worauf in den nächsten Abschnitten noch eingegangen wird.
- Weisen Systeme stärkere Nichtlinearitäten auf, so gewinnen die Terme höherer Ordnung in einer Taylorreihe immer mehr an Gewicht. Das EKF bricht die Taylor-Entwicklung aber nach der ersten Ordnung ab, sodass mit stärkerer Nichtlinearität auch ein größerer Fehler in der Schätzung entsteht [WVDM00].
- Die Betrachtung von Diskontinuitäten zeigt eine weitere Schwäche des EKF. Die Linearisierung entspricht der Berechnung von Jacobi-Matrizen. Liegt der Schätzwert, um den linearisiert wird, auf einer Unstetigkeit, so existiert die zugehörige Jacobi-Matrix nicht, und die Kovarianzen können nicht berechnet werden [VDMW⁺04].

Eine Möglichkeit die Genauigkeit zu erhöhen und den genannten Gefahren entgegenzuwirken, besteht in der Nutzung von Termen höherer Ordnung der Taylorreihe, was den Fehler, der durch die Linearisierung entsteht, verringert [Sim06]. Die Berechnung von Hesse-Matrizen ist allerdings ein fehleranfälliger Prozess und ein schwierig zu implementierendes Vorgehen, dessen Nutzen den Aufwand nicht aufwiegt [Frä05]. Ein weiterer Ansatz besteht in der Nutzung des iterierten Kalman-Filters. Bei diesem werden die A-posteriori-Gleichungen des EKF zum jeweiligen Zeitpunkt iteriert. Hierbei wird der Linearisierungspunkt auf Basis der Schätzfehlerkovarianz solange verbessert und die Systemgleichungen linearisiert, bis ein vorgegebenes Abbruchkriterium in Form einer Genauigkeit erfüllt oder die maximale Anzahl an Iterationen erreicht ist [Frä05], [Sim06]. Das iterative Vorgehen besitzt naturbedingt allerdings den Nachteil der fehlenden deterministischen Berechnungszeit bei Vorgabe eines Gütekriteriums bzw. bei Erreichen einer maximalen Iterationsanzahl, weshalb die Verfahren zur Erhöhung der Genauigkeit des EKF bei stark nichtlinearen Systemen in dieser Arbeit nicht angewandt werden.

2.2.2 Least-Square- und Weighted-Least-Square-Methode

Da im weiteren Verlauf dieser Arbeit statistische Methoden als Erweiterung von Kalman-Filtern angewendet werden, soll im Folgenden kurz die Herleitbarkeit des Kalman-Filter-Algorithmus aus rein statistischer Sicht gezeigt werden. Bereits im Jahr 1795 nutzte Karl Friedrich Gauß eine ähnliche Methode zur Beschreibung von Planetenbewegungen, für die es eine Reihe von Parametern zu bestimmen galt. Gauß legte durch seine Analysen hinsichtlich der Least-Square- und Weighted-Least-Square-Methode den Grundstein für heutige Verfahren zur

Bestimmung von Systemparametern, Optimierungsverfahren und Zustandsschätzungen.

Die Anforderungen, die zur Entwicklung der Least-Square-Methode durch Gauß führte, sind in den astronomischen Studien in *Theoria Motus* [Gau57] beschrieben. Gauß führt hier auf, dass bei korrekten astronomischen Beobachtungen nur eine definierte Anzahl notwendig wären, um eine absolut korrekte Vorhersage der Planetenbewegungen zu treffen. Voraussetzung ist, dass die Planetenbewegung exakt den Gesetzen von Kepler folgt. Jede weitere Messung würde demnach nur die Schätzung bestätigen, allerdings nicht korrigieren. Da aber alle Beobachtungen und Messungen nur Approximationen darstellen, stellen auch alle darauf basierenden Erkenntnisse nur solche dar. Das Ziel einer Schätzung liegt demnach darin, sie dem wahren Wert anzunähern, was nur durch die Einbeziehung einer Anzahl an Messungen erfolgen kann, die größer ist als die Anzahl der Unbekannten. Dieses Problem ist nur zufriedenstellend lösbar durch eine approximierte Kenntnis der Unbekannten und eine spätere Korrektur dieser Schätzung, um die Anforderungen aller Messungen in einer zu definierenden Weise zu erfüllen [Sor70].

Als Basis, diese Unbekannten zu erlangen, formulierte Gauß, dass die Summe der quadratischen Abweichungen zwischen Messungen und berechneten Werten, multipliziert mit einem Faktor, der den Grad der Genauigkeit angibt, zu minimieren ist.

Zur Fassung in mathematischer Form seien der zu ermittelnde konstante Vektor $x \in \mathbb{R}^n$ gegeben sowie der diskrete, verrauschte Messvektor $y_i = y(t_i)$, wobei $y \in \mathbb{R}$ eine Linearkombination der Elemente von x ist und mit additivem, mittelwertfreiem Rauschen $v_i \in \mathbb{R}$ überlagert wird:

$$y_i = h_i x + v_i. \quad (2-40)$$

Zur Vereinfachung werden die einzelnen Messungen in einem Vektor gesammelt:

$$\begin{aligned} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} &= \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_m \end{bmatrix} x + \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{bmatrix} \\ y &= H x + v. \end{aligned} \quad (2-41)$$

Die von Gauß formulierte Kostenfunktion zur Erlangung der optimalen Schätzung \hat{x} des unbekannten Vektors x , der ebenso der Zustandsvektor eines technischen Systems sein kann, ist durch

$$J = \sum_{i=1}^m (y_i - h_i \hat{x})^2 \quad (2-42)$$

$$= (y - H\hat{x})^T (y - H\hat{x}) \quad (2-43)$$

gegeben und entspricht der Least-Square-Formulierung. Durch Bildung der partiellen Ableitung nach x und Nullsetzung erlangt man die Lösung der optimalen Schätzung \hat{x} als

$$\hat{x} = (H^T H)^{-1} H^T y. \quad (2-44)$$

Falls $H^T H$ den Rang n besitzt, ist (2-44) explizit lösbar, was bedeutet, dass die Anzahl der Messungen m größer oder gleich der Anzahl der zu ermittelnden Parameter n ist [Sim06].

Was aber, wenn bestimmten Messungen mehr vertraut werden soll als anderen, da die einwirkenden Rauschprozesse v_i zeitvariant sind? Gauss löste dieses Problem durch Einführung von Gewichtungen W der unterschiedlichen Messungen y_i , wodurch sich die Kostenfunktion zu

$$J = (y - H\hat{x})^T W (y - H\hat{x}) \quad (2-45)$$

ändert. Die einfachste Form der Besetzung der Gewichtungsmatrix stellt $W = R_v^{-1}$ dar, mit $R_v = E[vv^T]$. Dies entspricht dem einfachen Fall, dass die Messungen invers zu den erwarteten Fehlern gewichtet werden. Die Lösung der Kostenfunktion (2-45) ergibt sich durch Bildung der Ableitung $\partial J / \partial \hat{x}$ und unter der Annahme $W = R_v^{-1}$ zu [Gib11]:

$$\hat{x} = (H^T R_v^{-1} H)^{-1} H^T R_v^{-1} y. \quad (2-46)$$

Die Fehlerkovarianz, welche die Unsicherheit der Schätzung von x mittels \hat{x} angibt, ergibt sich nach [Gib11] zu:

$$\begin{aligned} P_{xx} &= E[\tilde{x}\tilde{x}^T] \\ &= E[(x - \hat{x})(x - \hat{x})^T] \\ &= (H^T R_v^{-1} H)^{-1}. \end{aligned} \quad (2-47)$$

Der Term P_{xx} wird auch als *information*-Matrix bezeichnet, da über ihn die Unsicherheiten in der Schätzung von \hat{x} angegeben werden können. Die Wurzeln aus den Diagonalelementen $\sqrt{P_{xx,ii}}$ mit $i = 1, \dots, n$, stellen die Varianz der Schätzung des i -ten Elementes des zu schätzenden Vektors x dar. Die Nebendiagonalelemente $P_{xx,ij}$ mit $i, j = 1, \dots, n$ und $i \neq j$ beinhalten Informationen über die Korrelationen der Fehler unter den Zuständen, d. h. es kann angegeben werden, inwieweit Fehler in einem Zustand solche in weiteren Zuständen verursachen. Diese Zusammenhänge verdeutlichen direkt die Wirksamkeit des von Kalman entwickelten Algorithmus aus Abschnitt 2.2, da in diesem die Diagonalelemente von P_{xx} minimiert werden, was einer minimalen Varianz der zu schätzenden Zustände und somit einer hohen Sicherheit in den geschätzten Zuständen entspricht.

2.2.3 Bayessche Zustandsschätzung

Die im letzten Abschnitt vorgestellte Schätzung des konstanten, unbekannten Zustandsvektors x wird in diesem Abschnitt um den Blickwinkel erweitert, dass es sich bei x um eine zeitabhängige Zufallsvariable x_k zum Zeitpunkt k handelt. Diese Betrachtungsweise erlaubt es, dem Zustandsvektors x_k statistische Informationen in Form von Erwartungswert und Kovarianz oder sogar einer vollständigen Wahrscheinlichkeitsdichteverteilung (pdf) zuzuordnen. Die nachfolgende, vereinfachte Beschreibung der Bayesschen Schätzung entstammt den Büchern von Gibbs [Gib11] und Simon [Sim06] und kann in diesen detailliert nachvollzogen werden.

Die Grundidee der Bayesschen Schätzung bzw. der wahrscheinlichkeitstheoretischen Interpretation, besteht in der bereits erwähnten Betrachtung des zu schätzenden Zustandsvektors als Zufallsvariable und der darauf aufbauenden Schätzung der gesamten pdf des Zustandsvektors x zum nächsten Zeitschritt auf Basis aller bereits vorhandenen Messungen $y_{0,\dots,k} = Y_k$ vom Zeitpunkt 0 bis k [Kre80]. Gelingt es, die pdf des Zustandsvektors x zum Zeitpunkt k

$$p(x_k|Y_k) \tag{2-48}$$

zu schätzen, so sind alle statistischen Informationen über den Zustandsvektor x zum Zeitpunkt k verfügbar und dadurch nicht nur Erwartungswert und Kovarianz wie beim Kalman-Filter angebar, sondern auch alle Informationen über höhere statistische Momente und Multimodalitäten². Eine Möglichkeit, die pdf von x_k auf Basis von Y_k zu bestimmen, ist die Nutzung von Bayes Theorem

$$p(x_k|Y_k) = \frac{p(y_k|x_k)p(x_k|Y_{k-1})}{p(y_k|Y_{k-1})}, \tag{2-49}$$

mit der zu bestimmenden bedingten pdf des Zustandsvektors x_k auf Basis aller bisherigen Messungen Y_k und

$$\begin{aligned} p(y_k|Y_{k-1}) &= \int p(x_k|Y_{k-1})p(y_k|x_k)dx_k \\ p(x_k|Y_{k-1}) &= \int p(x_k|x_{k-1})p(x_{k-1}|Y_{k-1})dx_{k-1}. \end{aligned} \tag{2-50}$$

Die pdf $p(x_k|Y_k)$ wird bei dieser Art der Schätzung als A-posteriori-pdf auf Basis der A-priori-Informationen $p(x_k|Y_{k-1})$ bezeichnet. Die noch unbekannten pdfs $p(y_k|x_k)$ bzw. $p(x_k|x_{k-1})$ können aus dem vorhandenen Messmodell (2-31) bzw. dem Dynamikmodell (2-30) berechnet werden. Eine geschlossene Lösung der in (2-50) angegebenen Integrale zur Bestimmung von $p(x_k|Y_k)$ ist allerdings im Allgemeinen unmöglich, bzw. eine äußerst schwierige Aufgabe [Kre80].

²Multimodalität stellt im Gegensatz zur unimodalen Gaußverteilung eine mehrgipflige Dichtefunktion dar.

Falls es sich bei den Dichtefunktionen jedoch um Gaußverteilungen handelt, lassen sich in der Notation des Kalman-Filters der A-posteriori bedingte Erwartungswert \hat{x}_k und die Kovarianz $P_{xx,k}$ vereinfacht durch

$$\begin{aligned}\hat{x}_k &= \hat{x}_k^- + P_{xy,k} P_{yy,k}^{-1} (y_k - \hat{y}_k) \\ P_{xx,k} &= P_{xx,k}^- - P_{xy,k} P_{yy,k}^{-1} P_{yx,k}\end{aligned}\tag{2-51}$$

ausdrücken [Gib11], [Frä05], [WVDM01], wobei \hat{x}_k^- die A-priori-Schätzung des Zustandes x_k mit der A-priori-Kovarianz $P_{xx,k}^-$ und \hat{y}_k die Vorhersage der Messungen darstellen. Im Falle eines linearen Messmodells H läßt sich die A-posteriori-Schätzung des Zustandes x_k als \hat{x}_k nach [Gib11] und einigen längeren Rechnungen durch

$$\hat{x}_k = \hat{x}_k^- + P_{xx,k}^- H^T (H P_{xx,k}^- H^T + R_v)^{-1} (y - \hat{y}_k)\tag{2-52}$$

darstellen, was mit $P_{xx,k}^- = 0$ der bereits von Gauß postulierten Lösung aus (2-46) und der von Kalman ermittelten optimalen Anpassung \mathcal{K}_k aus (2-27) entspricht. Diese Herleitungen zeigen, dass das Kalman-Filter im Rahmen seiner Annahmen von linearen Dynamik- und Messmodellen sowie Gaußverteilten Rauschprozessen die ideale Lösung der Bayesschen Schätzung darstellt und eine effiziente Methode ist, die statistischen Eigenschaften einer Zufallsvariablen durch lineare Systemdynamik zu transformieren [VDM04]. Daraus geht allerdings auch hervor, dass das Extended-Kalman-Filter bei nichtlinearer Systemdynamik diese optimale Lösung mittels (2-38) nur noch annähert und es so zu unvorhersehbaren Ergebnissen führen kann.

Eine wesentlich effizientere Methode zur Beibehaltung der Optimalität des Kalman-Filters bei nichtlinearer Systemdynamik ist das im nächsten Abschnitt vorgestellte Particle-Filter, das auch nicht Gaußverteilte, multimodale Dichtefunktionen von Zufallsvariablen durch nichtlineare Sytemdynamik transformieren kann. Eine weitere Möglichkeit bietet die in dieser Arbeit verwendete Klasse der Sigma-Punkte-Filter, welche die optimale Bayessche Schätzung bei vorliegender nichtlinearer Systemdynamik wesentlich effizienter annähert als das Extended-Kalman-Filter und die in Kapitel 3 detailliert vorgestellt wird.

2.3 Particle Filter

In Abschnitt 2.2.3 wurde gezeigt, dass die A-posteriori-pdf $p(x_k|Y_k)$ unter der Annahme von Gaußverteilungen und linearer Systemdynamik mittels des Kalman-Filters bestimmbar ist. Ist dies nicht der Fall, stellt das Particle-Filter eine Möglichkeit dar, diese pdf zu anzunähern. Das Particle-Filter zählt zur Klasse der Sequentiellen Monte-Carlo-Methoden und versucht die Wahrscheinlichkeitsverteilung durch eine endliche Anzahl an Stützstellen mit zugehörigen Gewichtungen zu beschreiben [RC04]. Allgemein ausgedrückt, wird versucht, durch Simulation

einer Vielzahl von Systemzuständen und deren Auswertung eine Schätzung des Zustandes zu erreichen. Durch die nichtparametrische Beschreibung der Wahrscheinlichkeitsdichte ist es dabei möglich, einen Bayesschen Schätzer zu realisieren, dem nicht die Annahme von Normalverteilungen zu Grunde liegt. Das System wird dabei durch die folgenden Gleichungen beschrieben:

$$\begin{aligned} x_k &= F(x_{k-1}, u_{k-1}, v_{k-1}) \\ y_k &= H(x_k, n_k). \end{aligned} \quad (2-53)$$

Das System- und Messrauschen v_k bzw. n_k muss dabei nicht mehr einer Normalverteilung folgen, wie in den letzten Abschnitten vorausgesetzt. Das Ziel besteht wiederum in der Berechnung der bedingten Dichte $p(x_k|Y_k)$ mit $Y_k := [y_i]_{i=0}^k$, wodurch die Messung über den gesamten betrachteten Zeitraum 0 bis k beschrieben werden. Die Beschreibung der gesuchten und daher unbekannten Dichte erfolgt auf Basis von Stichproben einer ihr äquivalenten Importance- oder Trägerdichte $\hat{p}(x_k|Y_k)$ mittels [Mac98] [Frä05]

$$\hat{p}(x_k|Y_k) \approx \sum_{i=1}^N \omega_k^i \delta(x_k - x_k^i) \quad (2-54)$$

mit

$$\omega_k^i \propto \frac{\hat{p}(x_k^i|Y_k)}{p(x_k^i|Y_k)} \quad \text{und} \quad \sum_{i=1}^N \omega_k^i = 1. \quad (2-55)$$

Hierbei entsprechen $(x_k^i; i = 1, 2, \dots, N)$ den zufälligen Stichproben der Dichtefunktion $\hat{p}(x_k|Y_k)$ und $\delta()$ der Dirac-delta-Funktion [VDMW⁺04]. Zur Bestimmung der Gewichtungen ω_k^i erfolgt die Auswertung der Stichproben anhand der vorliegenden Sytem- und Messfunktion F bzw. H und eine rekursive Berechnung nach

$$\omega_k^i \propto \frac{\hat{p}(y_k|x_k^i) \hat{p}(x_k^i|x_{k-1}^i)}{p(x_k^i|x_{k-1}^i, y_k)} \omega_{k-1}^i \quad \text{und} \quad \sum_{i=1}^N \omega_k^i = 1. \quad (2-56)$$

Aus dieser Vielzahl an Stichproben nebst den ihnen zugeordneten Gewichtungen, können mittels

$$\hat{x}_k = \sum_{i=1}^N \omega_k^i x_k^i \quad \text{und} \quad P_{xx,k} = \sum_{i=1}^N \omega_k^i (x_k^i - \hat{x}_k) (x_k^i - \hat{x}_k)^T \quad (2-57)$$

die gesuchten Größen des Erwartungswertes und der Kovarianzmatrix äquivalent zu den vorherigen Abschnitten berechnet werden [Frä05].

Der große Vorteil der sequentiellen Monte-Carlo-Methode ist, dass sie auch auf nicht-normalverteilte Wahrscheinlichkeiten angewendet werden kann. Darüber

hinaus können die System- und die Messdynamik nichtlinear und die Wahrscheinlichkeitsverteilung multimodal sein. Als größter Nachteil ist zu nennen, dass eine Vielzahl an Stützstellen für die Schätzung notwendig ist, was selbst bei Systemen geringer Ordnung Probleme der Echtzeitfähigkeit mit sich bringt. Daher ist das Particle-Filter im Allgemeinen eher für Offline- als für Onlinezwecke einzusetzen. In der vorliegenden Arbeit ist die Systemordnung aufgrund des Black-Box-Charakters der zu nutzenden Modelle im Allgemeinen unbekannt, weshalb die Nutzung von Particle-Filtern ausgeschlossen ist.

Neben diesem Hauptnachteil gibt es eine Reihe unterschiedlicher Ansätze, wodurch inhärente Probleme der hier sehr knapp vorgestellten Monte-Carlo-Methode umgangen werden, wie z.B. die Sample-Degeneration und Resampling, die Sample-Verarmung und Regularisierung sowie die Wahl einer geeigneten Importance-dichte, was einen entscheidenden Entwurfparameter darstellt [VDM04] [GSS93] [Frä05].

2.4 Sigma-Punkte-Filter

Der Begriff Sigma-Punkte-Filter fasst eine Klasse von ableitungsfreien Filtern zusammen, die auf deterministischen Stützstellen, den Sigma-Punkten, basieren. Dies stellt gleichzeitig eine Verbindung zu dem im vorigen Abschnitt angesprochenen Particle-Filter dar, das auf Monte-Carlo-Verfahren basiert. Im Gegensatz zum Particle-Filter wird in den Sigma-Punkte Filtern allerdings eine geringe, deterministische Anzahl an Stützstellen zur Transformation der statistischen Eigenschaften einer Zufallsvariablen durch eine nichtlineare Funktion genutzt, deren Lage allerdings genau bestimmt wird. Zu den häufigsten Vertretern zählen das Central-Difference-Filter [IX00] und das Unscented-Kalman-Filter [JU97b]. Bei beiden Filtern wird die Wahrscheinlichkeitsdichte des Schätzvektors durch die Sigma-Punkte approximiert, die bestimmte Informationen über die Eigenschaften der Verteilung enthalten. Die Sigma-Punkte werden daraufhin durch die nichtlineare System- und Messfunktion transformiert, auf deren Basis die Erwartungswertschätzung und die Fehlerkovarianzschätzung durchgeführt werden.

Die Stärke beider Filter liegt in der Möglichkeit, nichtlineare System- und Messmodelle einzubinden. Weiter wird eine höhere Genauigkeit als beim EKF erzielt, da die Wahl der Sigma-Punkte einem Abbruch der Taylorreihenentwicklung nach dem zweiten Moment entspricht und somit jede Wahrscheinlichkeitsverteilung bis zur zweiten Ordnung erfasst werden kann. Bei Vorliegen einer Gaußverteilung sogar bis zur dritten Ordnung (in Abhängigkeit des gewählten Sigma-Punkte-Berechnungsverfahrens, was in Abschnitt 3.2 gezeigt wird). Da keine Jacobi-Matrizen der System- und Messfunktion benötigt werden, ist der Einsatz der in dieser Arbeit verwandten Dynamikmodelle als Black-Box geeignet. Im Gegensatz zum Particle-Filter sind beide Typen für den Onlineeinsatz denkbar, da wesentlich weniger Stützstellen notwendig sind. Gleichzeitig lässt dabei die Genauigkeit

gegenüber dem Particle-Filter nach, da keine multimodalen Verteilungen approximiert werden können und im Allgemeinen “nur” die statistischen Momente zweiter Ordnung erfasst werden. Auf die Klasse der Sigma-Punkte-Filter wird in Kapitel 3 noch im Detail eingegangen.

2.5 Machine Learning

Da im Verlauf der letzten Jahre das Gebiet des maschinellen Lernens immer weiter an Bedeutung gewinnt und auch Zustands- und Parameterschätzung mittels Techniken des maschinellen Lernens, wie Support Vector Machines durchgeführt werden, soll im Folgenden kurz auf diese Techniken eingegangen und aufgeführt, warum diese nicht in der vorliegenden Arbeit genutzt werden.

Die Idee des maschinellen Lernens zur Zustands- und Parameterschätzung basiert auf bestimmten Eingangs-/Ausgangsinformationen über ein ansonsten unbekanntes System. Falls ein System diese Trainingsdaten erlernt hat, sollte maschinelles Lernen am Besten in der Lage sein, den Ausgang auf Basis eines neuen Einganges zu approximieren. Wie gut ein System qualitativ mit diesen neuen Daten umgeht, wird als *Generalisierung* bezeichnet [BN06]. Ein großes Problem beim Arbeiten mit maschinellem Lernen stellt das Vorhandensein von Rauschprozessen in den Messungen und dem Prozess dar. Da bis auf die Eingangs-/Ausgangsinformationen keine Systemkenntnisse vorliegen, ist die Unterscheidung zwischen systembedingter und messbedingter Dynamik und Rauschen extrem schwierig.

In [YS06], [HE04] und [FP03] werden rekurrente neuronale Netzwerke zur Zustandsschätzung von dynamischen Systemen eingesetzt. Falls die Trainingsdaten allerdings nicht ausreichend waren und somit im Betrieb ein Arbeitspunkt angefahren wird, der nicht ausreichend erlernt wurde, kann es zu ungewolltem Systemverhalten kommen. Ein Nachlernen dieses neuen Systemverhaltens kann allerdings ebenfalls zu unvorhersehbarem Verhalten führen [BN06]. In [CHZ⁺13] wurden neuronale Netze eingesetzt, um das stark nichtlineare Verhalten von menschlichen Muskeln nachzubilden und eine Schätzung des Bewegungszustandes zu ermöglichen. Aber selbst die hier eingesetzten kleinen Netze zum Erlernen eines bestimmten Bewegungsprofiles erfüllen bereits keine Echtzeitanforderungen mehr und stellen den Gegenstand aktueller Arbeiten dar.

2.6 Zusammenfassung der Schätzverfahren

In diesem Kapitel wurde gezeigt, dass eine Reihe von Methoden zur Bestimmung des Zustandes eines Systems existieren. Als am Weitesten verbreitetes Werkzeug wurde das Kalman-Filter vorgestellt mit seiner Erweiterung zum Extended-Kalman-Filter für nichtlineare Systeme. Im Anschluss wurde kurz auf die Ursprünge der Schätztheorie von Zuständen und Parametern durch das Least-Square-

Verfahren von Gauß eingegangen und die Bayessche Zustandsschätzung vorgestellt. Diese bildet die Grundlage, um die vollständige Wahrscheinlichkeitsdichte eines Zustandsvektors auf Basis von A-priori-Informationen zu bestimmen. Die Lösung dieses Schätzproblems ist allerdings nur für Spezialfälle geschlossen möglich und daher nicht für die in dieser Arbeit verwandten unbekannten Systemmodelle denkbar.

Im Anschluss wurde allerdings gezeigt, dass das Kalman-Filter im Falle von linearer Systemdynamik und Gaußverteilten Rauschprozessen und Zustandsvektoren eine optimale Lösung der Bayesschen Schätzung darstellt und das Extended-Kalman-Filter diese optimale Lösung mittels Linearisierung der Systemdynamik nur noch annähert. In Abschnitt 2.3 wurde das Particle-Filter als eine Variante vorgestellt, mit der auch multimodale Verteilungen bei nichtlinearer Systemdynamik transformiert werden können. Allerdings erfolgt dies mittels einer Vielzahl an Stichproben einer Verteilung, die den Online-Einsatz zur Zustands- und Parameterschätzung schwierig machen.

Einen Mittelweg zwischen Extended-Kalman-Filter und Particle-Filter stellt die in 2.4 vorgestellte Klasse der Sigma-Punkte-Filter dar. Diese erfassen die Zustandsverteilung wie die Particle-Filter auch durch Stützstellen, die sogenannten Sigma-Punkte, allerdings werden diese deterministisch berechnet und liegen in einer sehr viel geringeren Anzahl vor. Durch diese geringe Anzahl sind Sigma-Punkte-Filter jedoch im Gegensatz zu Particle-Filtern nicht in der Lage, Multimodalitäten abzubilden. Die Transformation der Sigma-Punkte erfolgt allerdings durch die vollständige nichtlineare Funktion und benötigt keine Jacobimatrizen wie das Extended-Kalman-Filter.

In Abschnitt 2.5 wurden kurz Verfahren des maschinellen Lernens vorgestellt, die wegen ihrer Unwägbarkeiten in der Vorhersage von Zuständen und Parametern in Abhängigkeit des vorab erlernten Verhaltens in dieser Arbeit nicht genutzt werden.

Auf Basis der in 1.3 definierten Anforderungen dieser Arbeit scheidet das Kalman-Filter durch seine alleinige Anwendung bei linearen Systemen aus. Das Extended-Kalman-Filter ist zwar in der Lage, mit nichtlinearer Systemdynamik umzugehen, allerdings müssten in diesem Fall Jacobimatrizen aus den als Black-Box vorliegenden Systemmodellen erzeugt werden, was nur mit erheblichem Aufwand möglich ist. Als wichtigerer Grund ist allerdings zu nennen, dass das Extended-Kalman-Filter aufgrund seiner Linearisierung nur dann brauchbare Ergebnisse liefert, wenn sich das Modell über den Update-Vorgang des Filters annähernd linear verhält. Gleichzeitig muss die A-posteriori-Schätzung des letzten Zeitschrittes, um den die Linearisierung durchgeführt wird, korrekt sein, da ansonsten um den vollkommen falschen Arbeitspunkt linearisiert wird. Im Hinblick darauf, dass die in dieser Arbeit vorliegenden Modelle stark nichtlinear sein können, wird ein Einsatz des Extended-Kalman-Filters nicht in Betracht gezogen.

Das Particle-Filter ist durch seine hohe Genauigkeit zwar in der Lage, mit den stark nichtlinearen Modellen umzugehen und gleichzeitig diese als Black-Box einzubinden, allerdings kommt es aufgrund der Anforderung des Online-Einsatzes und der Vielzahl zu transformierender Stützstellen nicht in Betracht.

Die Klasse der Sigma-Punkte-Filter erfüllt alle in Abschnitt 1.3 gestellten Anforderungen. Es können Modelle, die als Black-Box vorliegen, in den Filteralgorithmen genutzt werden, und gleichzeitig kann es durch seine höhere Genauigkeit gegenüber dem Extended-Kalman-Filter auch mit starken Nichtlinearitäten im Modell umgehen. Durch die reduzierte Anzahl an Sigma-Punkten gegenüber dem Particle-Filter ist es grundsätzlich bei moderaten Systemordnungen der zu nutzenden Modelle auch echtzeitfähig. Liegen zudem explizite oder implizite Informationen über die statistischen Eigenschaften der Zufallsvariablen vor, so können diese über in den Sigma-Punkte-Filtern enthaltene Entwurfsparameter implementiert werden und die Genauigkeit noch weiter gegenüber den Annahmen einer Gaußverteilung erhöht werden.

Aus diesen Gründen werden in dieser Arbeit zur Zustands- und Parameterschätzung an Black-Box-Modellen von dynamischen Systemen Sigma-Punkte-Filter eingesetzt, die im nächsten Kapitel detailliert erläutert werden.

3 Sigma-Punkte-Filter

In Abschnitt 2.2 wurde das Kalman-Filter als das am Weitesten verbreitete Werkzeug zur Online-Zustandsschätzung eingeführt und seine einfache, auf dem Bayesschen Theorem basierende Struktur vorgestellt. Es wurde gezeigt, dass die Anwendung auf nichtlineare Systeme durch die Erweiterung zum Extended-Kalman-Filter (vgl. 2.2.1) das Standardverfahren zur Zustandsschätzung nichtlinearer Systeme oder zur Parameterschätzung darstellt, allerdings nicht den Anforderungen entspricht, die in Abschnitt 1.3 aufgestellt wurden. Im letzten Kapitel wurden unterschiedliche Werkzeugklassen zur Zustands- und Parameterschätzung vorgestellt. Sowohl die Klasse der Neuronalen Netze als auch die der Particle-Filter und Monte-Carlo-Methoden können nicht als Werkzeuge zur Online Zustands- und Parameterschätzung an unbekannten, topologieorientierten Modellen genutzt werden. Als nutzbare Filteralgorithmen wurde die Klasse der Sigma-Punkte-Filter identifiziert, auf die im Folgenden eingegangen wird. Den Schwerpunkt bildet dabei das Unscented-Kalman Filter. Die Ansätze sind deterministisch und ableitungsfrei, wobei die Schätzgüte die des EKF übertrifft ([WVDM00], [WVDMN99], [VDMW01a], [JU96]) und der Algorithmus zudem echtzeitfähig ist. In diesem Kapitel sollen zu Beginn ausführlich die Motivation und das grundsätzliche Vorgehen des Unscented-Kalman-Filters vorgestellt werden. Der Kern des Unscented-Kalman-Filters, die Unscented Transform, und die beinhalteten Verfahren zur Wahl der notwendigen Filtergrößen werden im Anschluss vorgestellt.

Aufgrund häufiger Fragen zur Namensherkunft wird vorab kurz der Begriff *Unscented* erläutert, da die alleinige Übersetzung *unparfümiert* nicht selbsterklärend ist. Hierzu sind nachfolgend Ausschnitte aus einem Interview, das begleitend zur Veröffentlichung von Uhlmann und Julier in *Proceedings of the IEEE* aus dem Jahr 2004 mit dem Titel *Unscented Filtering and Nonlinear Estimation* geführt wurde [GA10], dargestellt.

Auf die Frage, was es mit dem Begriff *Unscented* auf sich hat, antwortet Uhlmann, dass er auf der Suche nach einem Begriff für das neue Filter war, der einfach ohne Begründung akzeptiert wird. Zu Beginn hätten seine Mitarbeiter das Filter als *Uhlmann-Filter* bezeichnet, was er so nicht akzeptieren konnte. Als er eines Abends allein im Labor war, fiel sein Blick auf das zurückgelassene Deo eines Kollegen mit der Aufschrift *Unscented*, was nichts mit dem eigentlichen Filter zu tun hat und somit für Uhlmann den perfekten Begriff darstellte. Zu Beginn sei seine Idee der Namensgebung als absurd abgetan worden; als aber einige Monate später ein Gastredner einen Vortrag über das *Unscented-Filter* hielt, ohne auch nur einen Gedanken an die Bedeutung des Wortes verschwendet zu haben, war für Uhlmann die Wahl perfekt: Die Annahme des Begriffs *unscented* als technischer Begriff, ohne dass auf dessen Ursprung eingegangen werden müsste.

3.1 Unscented-Kalman-Filter

Das Unscented-Kalman-Filter (UKF) stellt einen rekursiven Minimum-Mean-Squared-Error Schätzer dar, der die Nachteile des Extended-Kalman-Filters umgeht bei einer gleichzeitig höheren Genauigkeit hinsichtlich der Approximation von Nichtlinearitäten des Modells. Das UKF basiert dabei auf folgender Annahme, die von Julier und Uhlmann in [JU97b] aufgestellt wurde:

It is easier to approximate a probability distribution than it is to approximate an arbitrary nonlinear function or transformation

Das Extended-Kalman-Filter versucht gerade, die gesamte nichtlineare Systemfunktion mittels einer Linearisierung erster Ordnung zu approximieren. Dies stellt auch direkt den Unterschied des UKF zum EKF dar. Das UKF versucht nicht, die gesamte nichtlineare Systemfunktion zu approximieren, sondern die Momente der Wahrscheinlichkeitsverteilung einer Zufallsvariablen. Im Sinne der Zustandschätzung ist diese Zufallsvariable durch den Zustandsvektor und eventuell zu schätzende Parameter gegeben, wodurch der Zustandsvektor durch eine Wahrscheinlichkeitsdichtefunktion bzw. statische Momente beschreibbar wird.

Die grundlegende Funktionsweise wird im Folgenden an einem einfachen Beispiel erläutert. Hierzu wird das lineare System eines Einmassenschwingers genutzt, der in Bild 3-1 dargestellt ist und aus seiner Anfangslage von 0,2 Metern und einer Geschwindigkeit von 0 m/s gestartet wird. Die Masse schwingt sich daraufhin auf ihre stabile Ruhelage von 0,09 m und 0 m/s ein.

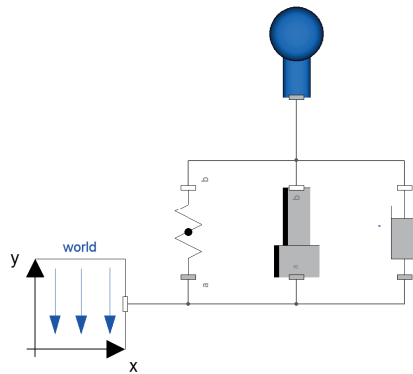


Bild 3-1: Dymola-Modell eines linear gelagerten Einmassenschwingers

Die Zustände des Systems sind Position und Geschwindigkeit der Masse, deren Zustandsraum in Bild 3-2 abgebildet ist.

Jeder Punkt des durchlaufenden Zustandsraumes von Bild 3-2 entspricht dabei dem Zustandsvektor zu einem diskreten Zeitschritt, was, vereinfacht zu definierten Abtastschritten, mit Sternen gekennzeichnet ist. Durch die Annahme von jedem dieser Zustandsvektoren als Zufallsvariablen, was durch die Herleitung in Abschnitt 2.2 gültig ist, lassen sich statistische Informationen des Zustandsvektors zu jedem diskreten Zeitschritt angeben, wie in Bild 3-3 beispielhaft für die

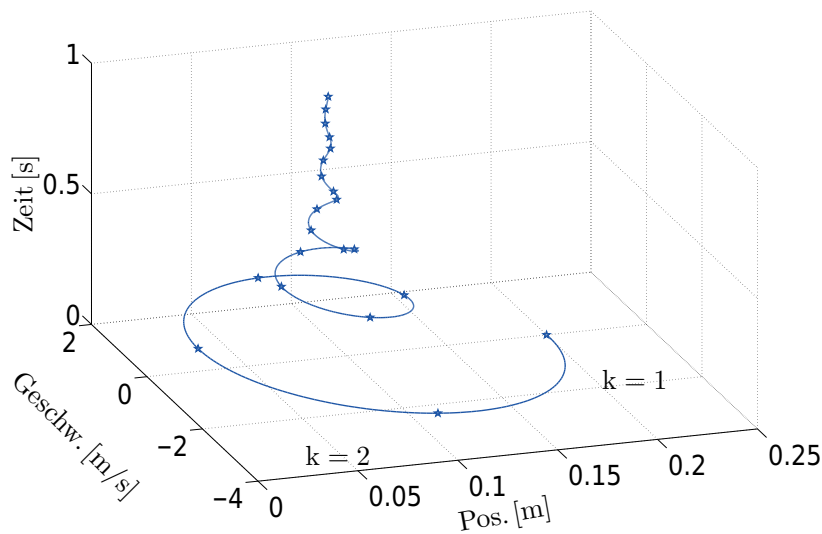


Bild 3-2: Zustandsebene eines Einmassenschwingers über die Zeit

ersten beiden diskreten Zeitpunkte gezeigt, wobei die Verteilungen als Gaußsche angenommen werden.

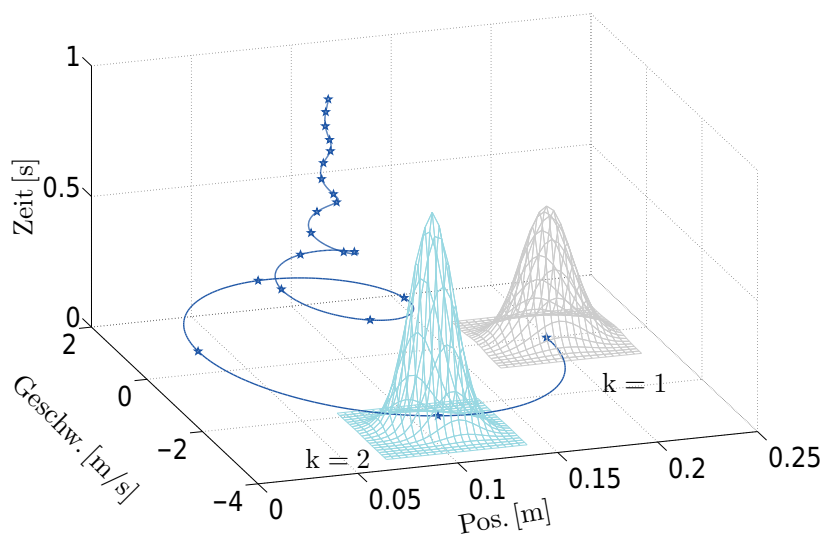


Bild 3-3: Zustandsverlauf des Einmassenschwingers und Gaußsche Dichteverteilungen für die diskreten Zeitpunkte $k = 1$ und $k = 2$

Bei einer Zustands- und Parameterschätzung ist die Aufgabe eines Beobachters, die statistischen Informationen der Verteilung für den Zeitschritt $k = 2$ auf Basis der statistischen Informationen bei $k = 1$ sowie des Streckenmodells zu schätzen. Bei der Nutzung eines Kalman- oder Extended-Kalman-Filters stellt dies die A-priori-Vorhersage des Erwartungswertes vom Zustandsvektor $E[x_k] = \hat{x}_k^-$ sowie der Fehlerkovarianz $E[\tilde{x}_k \tilde{x}_k^T] = P_{xx,k}^-$ zum Zeitschritt $k = 2$ dar, mit $\tilde{x}_k = x_k - \hat{x}_k$. Anders ausgedrückt: Der unbekannte blaue Stern und die unbekannte cyanfarbe-

ne Gaußverteilung für $k = 2$ sind auf Basis des Sterns und der Verteilung für $k = 1$ sowie des Streckenmodells zu bestimmen. Das Extended-Kalman-Filter würde dies durch Vorhersage des A-priori-Zustandes \hat{x}_2^- mittels des bekannten Streckenmodells und einer anschließenden Linearisierung erster Ordnung zur Bestimmung der Fehlerkovarianz $P_{xx,2}^-$ durchführen. Das Extended-Kalman-Filter transformiert demnach nur das erste statistische Moment des Erwartungswertes in Form der A-priori-Zustandsschätzung \hat{x}_k^- korrekt und nähert höhere statistische Informationen nur an.

Das Unscented-Kalman-Filter umgeht diese Linearisierung durch das in den Bildern 3-4 und 3-5 gezeigte Verfahren. Im oberen Bild von 3-4 ist der Zustandsraum dargestellt mit der Initialisierung des Unscented-Kalman-Filters bestehend aus dem Erwartungswert des Zustandsvektors $E[x_1]$ und dessen Kovarianz $P_{xx,1}$. Zur Erfassung dieser ersten beiden statistischen Momente des Zustandsvektors x zum Zeitpunkt $k = 1$ werden im unteren Teil von Bild 3-4 fünf Sigma-Punkte \mathcal{X}_i mit $i = 0, \dots, 4$ berechnet.

Die berechneten Sigma-Punkte \mathcal{X}_i werden im Anschluss mittels des vorhandenen Streckenmodells F zu $\mathcal{X}_i^* = F(\mathcal{X}_i)$ mit $i = 1, \dots, 5$ transformiert, was im oberen Teil von Bild 3-5 dargestellt ist. Auf Basis dieser transformierten Sigma-Punkte \mathcal{X}_i^* können der Erwartungswert zum Zeitschritt $k = 2$ als $E[x_2] = \hat{x}_2^-$ und dessen Kovarianz $P_{xx,2}^-$ berechnet werden, die im unteren Bild als cyanfarbene Gaußverteilung dargestellt ist.

Obwohl dieser Ansatz eine starke Ähnlichkeit zur Monte-Carlo-Methode aufweist, unterscheidet er sich davon doch in einem wichtigen Punkt. Im Gegensatz zur Monte-Carlo-Methode, bei der eine sehr große Anzahl an Punkten durch die Systemfunktion transformiert werden, kommen beim UKF eine geringe, genau definierte Anzahl zur Anwendung. Dies bietet die Möglichkeit, Filterverfahren, die auf einer Unscented-Transformation basieren, auch in Echtzeitanwendungen zu nutzen, was bei der Monte-Carlo-Methode nicht möglich wäre.

Die hier dargestellten Schritte stellen die A-priori-Schätzung des Unscented-Kalman-Filters dar, also die Vorhersage von Erwartungswert und Kovarianz auf Basis des letzten Zeitschrittes und des bekannten Systemmodells. Beim Auftreten neuer Messungen wird diese Vorhersage korrigiert und somit die A-posteriori-Schätzung berechnet.

Der Unterschied zum EKF besteht demnach einzig in der Bestimmung der statistischen Momente des Zustandsvektors, nachdem diese durch die Systemfunktion transformiert wurden. Bild 3-6 zeigt eine Darstellung des Ablaufs des EKF und des UKF bei einer Unterteilung in Vorhersage und Korrektur. Im linken Teil des Bildes ist der Ablauf des EKF mit seinen auf Jacobimatrizen basierenden Berechnungen dargestellt. Der rechte Teil zeigt das UKF mit den gut zu erkennenden Änderungen in der Vorhersage, der A-priori-Schätzung von Erwartungswert und Kovarianz.

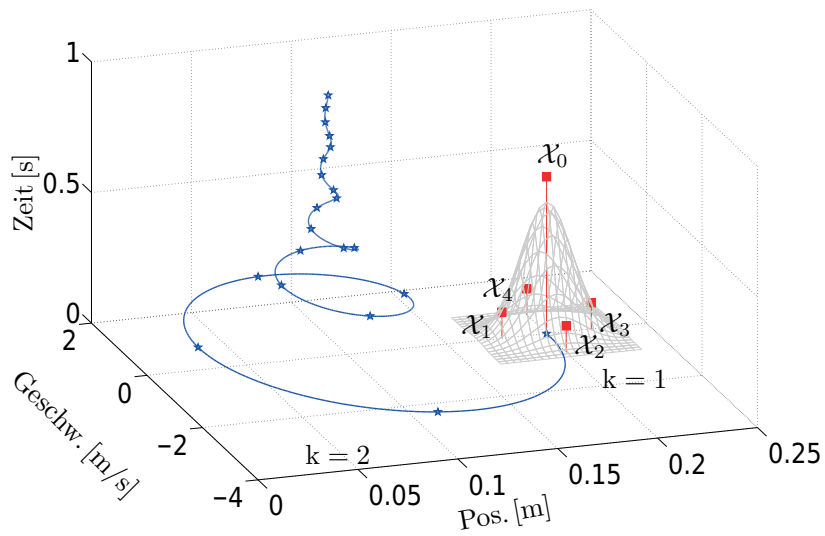
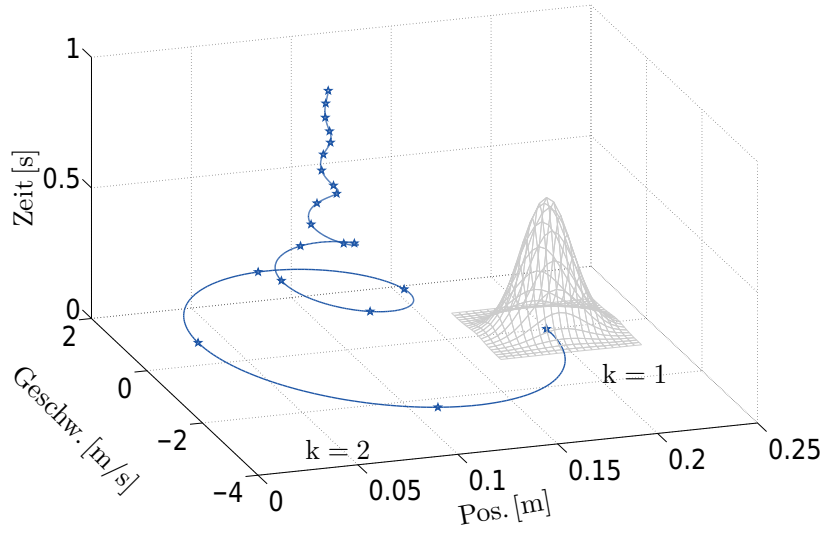


Bild 3-4: A-priori-Schätzung UKF. Oben: Erwartungswert $E[x_1] = \hat{x}_1$ und Fehlerkovarianz $E[\tilde{x}_1\tilde{x}_1^T] = P_{xx,1}$ bei $k = 1$. Unten: Erfassung $E[\hat{x}_1]$ und $P_{xx,1}$ durch fünf Sigma-Punkte.

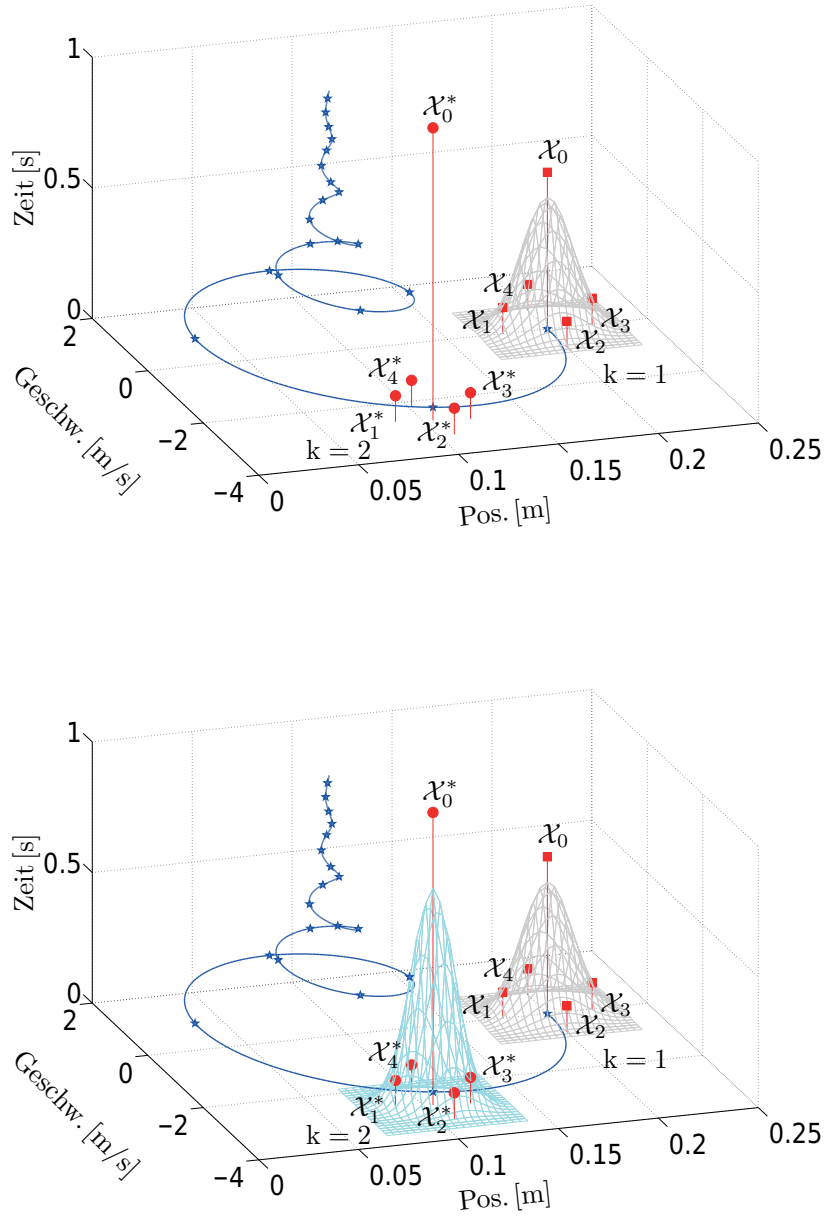


Bild 3-5: A-priori-Schätzung UKF. Oben: Transformation Sigma-Punkte \mathcal{X}_i , $i = 0, \dots, 4$ von $k = 1$ zu $k = 2$. Unten: A-priori-Zustandsschätzung \hat{x}_2^- als transformierter Erwartungswert $E[x_2]$ und Fehlerkovarianz $P_{xx,2}$

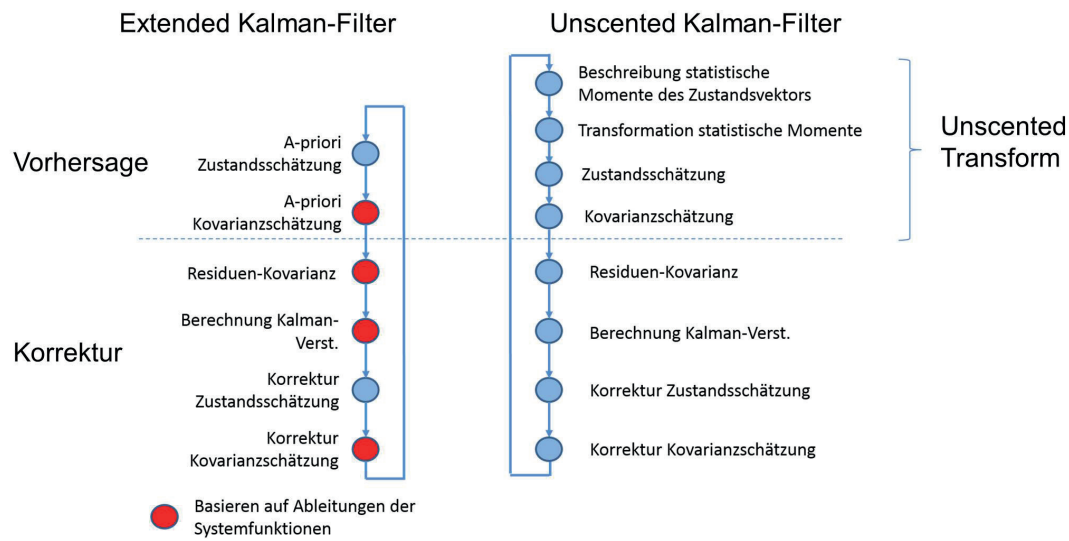


Bild 3-6: Ablaufdiagramm von Extended- und Unscented-Kalman-Filter in Aufteilung zwischen Vorhersage und Korrektur

Ebenfalls gut zu erkennen ist, dass die Struktur der Kalman-Filterung im UKF unverändert ist im Vergleich zu der des EKF, alleinig die Transformation der statistischen Eigenschaften des Systemzustandes erfolgt mittels der Unscented Transform, die in den Algorithmus des Kalman-Filters eingebettet ist und im nächsten Abschnitt vorgestellt wird.

Ein wesentlicher Neuerungsbeitrag der vorliegenden Arbeit besteht in der Erkenntnis, dass für die Transformation der Sigma-Punkte von \mathcal{X}_i zu \mathcal{X}_i^* keine analytischen Modelle notwendig sind. Die berechneten Sigma-Punkte entsprechen deterministisch berechneten Stichproben der Verteilung des Zustandsvektors und stellen somit selbst mögliche Zustandsvektoren im System dar. Die Transformation der Sigma-Punkte wird durch die Einbringung jedes einzelnen Sigma-Punktes als Zustandsvektor in das Streckenmodell und die anschließende Integration des Modells von Zeitschritt $k = 1$ zu $k = 2$ realisiert. Hierdurch sind keine analytischen Modelle in Zustandsform oder die Berechnung von Jacobi- und Hesse-Matrizen notwendig. Für die Transformation können beliebige Modelle genutzt werden, die das dynamische Verhalten beschreiben. Einzig die Möglichkeit der Setzung von Anfangswerten für den Zustandsvektor in das Modell, dessen Simulation und die anschließende Ermittlung des Zustandsvektors sind notwendig, wofür in dieser Arbeit eine Schnittstelle zu dem in MATLAB implementierten Filteralgorithmus entwickelt wurde.

Die in Bild 3-5 dargestellten Sigma-Punkte müssen die Verteilung des Zustandsvektors möglichst exakt erfassen. Hierzu gibt es in der Literatur verschiedene Möglichkeiten, die in Rechenaufwand und Erfassungsgüte variieren, wobei alle Varianten freie Parameter zur Anpassung enthalten. Die unterschiedlichen Vari-

anten und freien Parameter, die im Folgenden als Entwurfsparameter bezeichnet werden, haben dabei einen entscheidenden Einfluss auf die Gesamtfiltergüte, weshalb auf die in dieser Arbeit genutzten Verfahren zur Sigma-Punkte-Berechnung eingegangen und der Einfluss der Entwurfsparameter verdeutlicht wird.

Hierzu wird im nächsten Abschnitt die allgemeine Form zur Berechnung von Sigma-Punkten, die Unscented Transform, vorgestellt und auf deren Basis mögliche Realisierungen, die in dieser Arbeit genutzt werden, hergeleitet.

3.2 Unscented Transform

Ziel der Unscented Transform ist, abstrakt ausgedrückt, die Bestimmung der ersten beiden statistischen Momente einer Zufallsvariablen¹, nachdem diese durch eine nichtlineare Funktion transformiert wurde. Zur Herleitung der Sigma-Punkte-Berechnungen wird im Folgenden die A-priori-Schätzung eines Kalman-Filters genutzt, dessen Ziel es ist, eine Vorhersage des Zustandsvektors \hat{x}_k^- und dessen Kovarianz $P_{xx,k}^-$ zum Zeitpunkt k auf Basis der gegebenen A-posteriori-Schätzung des letzten Zeitschrittes \hat{x}_{k-1} und $P_{xx,k-1}$ abzugeben. Voraussetzung hierfür ist die initiale Vorgabe von \hat{x}_0 und $P_{xx,0}$, das Vorhandensein des als Black Box vorliegenden Systemmodells, das im Folgenden durch die Funktion

$$x_k = F(x_{k-1}, v_k, u_k) \quad (3-1)$$

gegeben ist, sowie die Kenntnis der Systemeingänge u_k . Der zeitlich unkorrelierte Rauschprozess v_k wird als Gaußprozess angenommen, weshalb er nachfolgend keine Berücksichtigung mehr findet. Zu Beginn sind die statistischen Eigenschaften der gegebenen Schätzung \hat{x}_{k-1} und $P_{xx,k-1}$ durch deterministisch bestimmbare Sigma-Punkte zu beschreiben. Die genutzten Sigma-Punkte $\mathcal{X}_{i,k-1}$ seien durch den Satz \mathcal{S}_k , bestehend aus l Vektoren und den zugehörigen Gewichtungen W_i , definiert, mit $i = 0, \dots, l-1$:

$$\mathcal{S}_k = \{i = 0, 1, \dots, l-1 : \mathcal{X}_{i,k-1}, W_i\}. \quad (3-2)$$

Auf die notwendige Anzahl der Sigma-Punkte \mathcal{X}_i wird im weiteren Verlauf eingegangen. Die zeitunabhängigen Gewichtungen W_i sind dabei jeweils den Sigma-Punkten $\mathcal{X}_{i,k-1}$ zugeordnet und entsprechen der Wahrscheinlichkeit des Auftretens von $\mathcal{X}_{i,k-1}$. Daher muss die Summe der Gewichtungen W_i der Normalisierungsfunktion

$$\sum_{i=0}^{l-1} W_i = 1 \quad (3-3)$$

¹Im Falle einer vorliegenden Gaußverteilung wird auch das dritte Moment erfasst, da dieses gleich Null ist.

entsprechen, um insgesamt eine Wahrscheinlichkeit von 1 zu erhalten. Zur Berechnung des A-priori-Zustandsvektors \hat{x}_k^- und dessen Fehlerkovarianz $P_{xx,k}^-$, die im statistischen Sinn dem Erwartungswert des zu schätzenden Zustandsvektors $E[x_k]$ und dessen Kovarianz $E[\tilde{x}_k \tilde{x}_k^T]$ entsprechen, werden die gewählten Sigma-Punkte $\mathcal{X}_{i,k-1}$ einzeln durch die bestehende Systemfunktion F transformiert:

$$\mathcal{X}_{i,k|k-1} = F(\mathcal{X}_{i,k-1}). \quad (3-4)$$

Gemäß den transformierten Sigma-Punkten $\mathcal{X}_{i,k|k-1}$ können der Erwartungswert als gewichtetes Mittel,

$$\hat{x}_k^- = \sum_{i=0}^{l-1} W_i \cdot \mathcal{X}_{i,k|k-1}, \quad (3-5)$$

sowie die Kovarianz als gewichtetes äußeres Produkt,

$$P_{xx,k}^- = \sum_{i=0}^{l-1} W_i (\mathcal{X}_{i,k|k-1} - \hat{x}_k^-) (\mathcal{X}_{i,k|k-1} - \hat{x}_k^-)^T \quad (3-6)$$

berechnet werden. Bei dieser Art der Transformation von statistischen Momenten stellen sich dabei direkt einige Fragen:

- Wieviele Sigma-Punkte \mathcal{X} sollen gewählt werden?
- Wo sollen die gewählten Sigma-Punkte \mathcal{X} liegen?
- Wie müssen die zugehörigen Gewichtungen gewählt werden?

Das Ziel der Sigma-Punkte-Berechnung ist die Erfassung von Informationen über die ersten beiden Momente der pdf $p(x_{k-1})$ der Zufallsvariablen x_{k-1} . Da kein praktisches und erst recht kein echtzeitfähiges Filter in der Lage ist, eine willkürliche Verteilungsfunktion mit beispielsweise Multimodalitäten in allen statistischen Momenten abzubilden, müssen immer Annahmen über die zugrundeliegende Verteilung getroffen werden. Eine Möglichkeit besteht in der Annahme, dass es sich bei $p(x_{k-1})$ um eine Gaußverteilung mit $p(x_{k-1}) = \mathcal{N}(\hat{x}_{k-1}, P_{xx,k-1}^-)$ handelt, was bereits in Abschnitt 2.2 zur Herleitung des Kalman-Filters angenommen wurde. Aus dieser Annahme folgt eine Wahrscheinlichkeitsdichteverteilung für x_{k-1} der Form [Kus67]

$$p(x_{k-1}) = \frac{1}{(2\pi)^{\frac{n}{2}} |P_{xx,k-1}|^{\frac{1}{2}}} \exp \left(-\frac{1}{2} (x_{k-1} - \hat{x}_{k-1}) P_{xx,k-1}^{-1} (x_{k-1} - \hat{x}_{k-1})^T \right), \quad (3-7)$$

mit $|P_{xx,k-1}|$ als Determinante von $P_{xx,k-1} \in \mathbb{R}^{n \times n}$ und n als der Systemordnung von $x_{k-1} \in \mathbb{R}^{n \times 1}$. Die zu berechnenden Sigma-Punkte müssen gewünschte Informationen über die Verteilung (3-7) des Zustandsvektors x_{k-1} erfassen, die in Form

von Nebenbedingungen vorliegen und beispielsweise die korrekte Erfassung des Erwartungswertes erzwingen:

$$c[\mathcal{S}_{k-1}, p(x_{k-1})] = 0. \quad (3-8)$$

In Abhängigkeit der verwendeten Anzahl an Sigma-Punkten sind die bis hier aufgestellten Bedingungen nicht in allen Fällen ausreichend, um einen Sigma-Punkte Satz \mathcal{S}_{k-1} nebst zugehörigen Gewichtungen W zu ermitteln [JU97a]. Hierzu wird eine Kostenfunktion $e[\mathcal{S}_{k-1}, p(x_{k-1})]$ eingeführt, die unerwünschte Effekte bestraft oder die bessere Erfassung von nicht exakt beschreibbaren Effekten der Verteilung von x_{k-1} durch \mathcal{S}_{k-1} belohnt. Allgemein kann daher geschrieben werden:

$$\min_{\mathcal{S}_{k-1}} e[\mathcal{S}_{k-1}, p(x_{k-1})] \quad \text{unter Nebenbedingung} \quad (3-8). \quad (3-9)$$

Welche gewünschten Informationen in der Beschränkungsfunktion c enthalten sein sollen und welche Effekte es durch e zu vermeiden gilt, hängt vom jeweiligen Anwendungsfall und der gewünschten Genauigkeit bzw. Rechenbelastung ab. In den folgenden Abschnitten werden die beiden in dieser Arbeit eingesetzten Verfahren zur Sigma-Punkte-Berechnung erläutert.

Bei dem ersten Verfahren handelt es sich um die Scaled Unscented Transform (SUT), die eine Weiterentwicklung der Second-Order Unscented Transform (SOT) darstellt. Da die Entwurfparameter sowohl aus der SUT als auch aus der SOT stammen, werden in den nächsten beiden Abschnitten beide Verfahren aufeinander aufbauend erläutert. Bei der zweiten genutzten Variante handelt es sich um den Spherical-Simplex-Punkte-Satz (SSPS), der auf den Minimal-Skew-Simplex-Punkte-Satz (MSSPS) aufbaut und nur in verkürzter Form vorgestellt wird, da das Vorgehen zur Bestimmung der Sigma-Punkte äquivalent zur SUT- bzw. SOT-Variante ist.

3.2.1 Second-Order Transformation

Die Anforderung an der Second-Order Transformation (SOT) ist die korrekte Erfassung der ersten beiden statistischen Momente einer Gaußverteilung der Zufallsvariablen x_{k-1} der Dimension n mit Erwartungswert \hat{x}_{k-1} und Kovarianz $P_{xx,k-1}$ [JU97a], was der A-posteriori-Schätzung des letzten Zeitschrittes in Kalman-Filter-Notation entspricht. Die Annahme einer Gaußverteilung hat hierbei den Vorteil, dass es sich um eine symmetrische Verteilung handelt und das dritte Moment der Schiefe gleich Null ist. Die Approximation von x_{k-1} mit beliebigem Erwartungswert und Kovarianz kann reduziert werden auf das Problem der Approximation einer Standard-Gaußverteilung z mit $E[z] = 0$ und $E[zz^T] = I$, da x_{k-1} als

$$x_{k-1} = \bar{x}_{k-1} + C_{k-1}z \quad (3-10)$$

dargestellt werden kann mit $C_{k-1} \in \mathbb{R}^{n \times n}$ als Matrixwurzel der Kovarianz $P_{xx,k-1}$ und $C_{k-1}C_{k-1}^T = P_{xx,k-1}$ [JU97a]. Der Term \bar{x}_{k-1} entspricht dabei dem tatsächlichen Erwartungswert des Zustandsvektors x_{k-1} , der im letzten Zeitschritt durch \hat{x}_{k-1} geschätzt wurde.

Bei der SOT kommen eine gerade Anzahl von $2n$ Sigma-Punkten zum Einsatz, die symmetrisch um den Erwartungswert $E[z] = 0$ positioniert werden, mit n als Systemordnung. Durch diese Symmetrie wird direkt die Symmetrie der zu erfassenden Gaußverteilung korrekt wiedergegeben, da pro Dimension der Zufallsvariablen z zwei Sigma-Punkte im gleichem Abstand zum Erwartungswert gesetzt werden, wie in Bild 3-7 dargestellt.

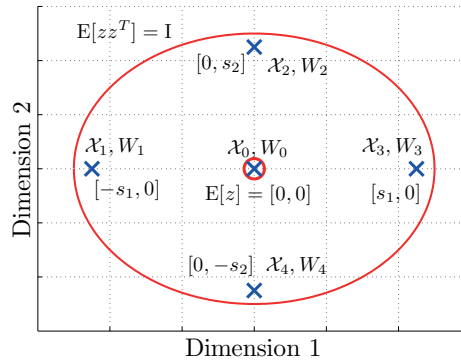


Bild 3-7: Sigma-Punkte-Verteilung der Second-Order Unscented Transformation einer zweidimensionalen Zufallsvariablen

Durch eine symmetrische Verteilung der Sigma-Punkte \mathcal{X}_i um den Erwartungswert $E[z] = 0$ in den Dimensionen der Zufallsvariablen wird die Symmetrie der Kovarianz von $E[zz^T] = I$ (siehe auch 3-12) sowie die Schiefe $E[zz^T]^3 = 0$ korrekt erfasst.

Bei einer symmetrischen Verteilung von Sigma-Punkten um den Erwartungswert ist dieser immer korrekt erfasst, allerdings kann kein weiterer Einfluss auf die höheren, geraden Momente der Verteilung, wie die Wölbung, genommen werden. Um diese beeinflussen zu können, wird neben den $2n$ Punkten gleichzeitig ein Sigma-Punkt in dem Erwartungswert positioniert: $\mathcal{X}_{0,k-1} = E[z]$, mit dessen Gewichtung die höheren Momente beeinflusst werden können. Eine erhöhte Gewichtung des Sigma-Punktes im Erwartungswert einer Verteilung entspricht einer erhöhten Wahrscheinlichkeit des Auftretens des Erwartungswertes und bestimmt somit die Wölbung einer Verteilung. Hieraus ergibt sich eine Gesamtzahl an Sigma-Punkten von $2n + 1$ [JU96].

Zur Erfassung der Standard-Normalverteilung z besitzen alle Punkte um den Erwartungswert die gleichen Abstände, wodurch $s_1 = s_2$ gilt. Zur Erfüllung der Forderung der korrekten Erfassung des Erwartungswertes bei gleichen Abständen

müssen alle symmetrisch um den Erwartungswert liegenden Sigma-Punkte die gleiche Gewichtung aufweisen, sodass $W_1 = W_2 = W_3 = W_4$ gilt.

Die Berechnung des Erwartungswertes $E[z] = \bar{z} = 0$ nach (3-5) ergibt sich daher zu:

$$E[z] = \sum_{i=0}^{2n} W_i \mathcal{X}_i = 0. \quad (3-11)$$

Die eingangs formulierte Forderung nach Approximation der Kovarianz $E[zz^T] = I$ (vgl. (3-6)) ergibt sich zu:

$$\begin{aligned} E[zz^T] &= \sum_{i=0}^{2n} W_i (\mathcal{X}_i - \bar{z})(\mathcal{X}_i - \bar{z})^T = I \\ &= 2W_1 s_1^2 \cdot I \end{aligned} \quad (3-12)$$

und die Normalisierungsbedingung nach (3-3) zu:

$$1 = \sum_{i=0}^{2n} W_i = W_0 + 2nW_1. \quad (3-13)$$

Die Ergebnisse von (3-12) unter Betrachtung nur einer Dimension² sowie (3-13) werden mittels der Zwangsfunktion (3-8) zusammengefasst und ergeben sich somit zu:

$$c[\mathcal{S}, p(z)] = \begin{bmatrix} 2W_1 s_1^2 - 1 \\ W_0 + 2nW_1 - 1 \end{bmatrix} = 0. \quad (3-14)$$

Zur Bestimmung der Gewichtungen W_i mit $i = 1, \dots, 4$ und des Abstandes s_1 wird (3-14) zu

$$\begin{aligned} s_1 &= \frac{1}{\sqrt{2W_1}} \\ W_0 &= 1 - 2nW_1 \end{aligned} \quad (3-15)$$

aufgelöst. Die antiproportionale Abhängigkeit der Sigma-Punkt-Abstände s_1 von den Gewichtungen W_1 , die nachfolgend festgelegt werden, um den Erwartungswert zeigt anschaulich, dass die Vorgabe eines Abstandes zu einer definierten Gewichtung führt. Je weiter die Sigma-Punkte vom Erwartungswert entfernt liegen, umso geringer wird ihre Gewichtung, d. h. umso geringer ist die Wahrscheinlichkeit, dass dieser Punkt im Zustandsraum wirklich auftritt. Umgekehrt: Je näher

²Da es sich um eine symmetrische Verteilung handelt, reicht die Betrachtung einer Dimension aus.

die Punkte an den Erwartungswert wandern, umso größer wird die Gewichtung, da die Wahrscheinlichkeit dieses Zustandsvektors immer größer wird. Zur Lösung der drei unbekannten W_0 , W_1 und s_1 führen Uhlmann und Julier in [JU97a] einen Entwurfsparameter κ ein und setzen die Gewichtungen der um den Erwartungswert verteilten Punkte zu $W_1 = \frac{1}{2(n+\kappa)}$, wodurch sich (3-15) zu

$$\begin{aligned} s_1 &= \frac{1}{\sqrt{2\frac{1}{2(n+\kappa)}}} = \sqrt{n+\kappa} \\ W_0 &= 1 - 2n\frac{1}{2(n+\kappa)} = \frac{\kappa}{n+\kappa} \end{aligned} \quad (3-16)$$

ergibt. Der Parameter κ stellt einen Entwurfsparameter dar, der direkt die Gewichtung des nullten Sigma-Punktes W_0 beeinflusst und mittels dem Einfluss auf die Wölbung einer Verteilung gewonnen werden kann, wie in Bild 3-8 dargestellt. Der als Stern gekennzeichnete Sigma-Punkt im Erwartungswert der zu erfassenden Verteilung geht mit unterschiedlichen Gewichtungen in die Berechnung ein, was im Bild als unterschiedliche betragsmäßige Höhen dargestellt ist. Die daraufhin berechnete Verteilung besitzt mit größerem κ eine immer größere Wölbung, flacht sich also ab, was bedeutet, dass im rechten Teil von Bild 3-8 Werte dicht am Erwartungswert mit höherer Wahrscheinlichkeit auftreten als im linken Teil. Zur Beibehaltung der korrekten Erfassung der Kovarianz werden die übrigen Sigma-Punkte zum Erwartungswert hin skaliert und verändern ihre Gewichtung.

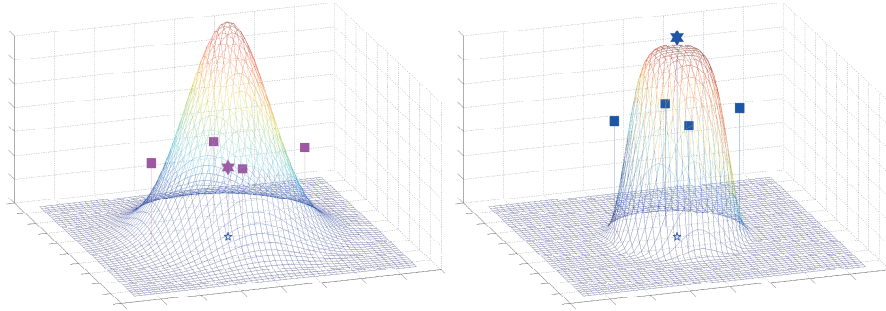


Bild 3-8: Wölbung zweier Verteilungen bei unterschiedlichen κ Werten

Durch Anwendung von (3-10) kann die Berechnung der Sigma-Punkte zur Erfassung von Erwartungswert und Kovarianz der Standard-Normalverteilung auf eine beliebige Verteilung übertragen werden, die im Sinne der Zustandsschätzung durch die A-posteriori-Schätzung des letzten Zeitschrittes gegeben ist. Nachfolgend ist das Sigma-Punkte-Verfahren der *Second-Order Unscented Transform* zusammen gefasst dargestellt, mit den als bekannt vorausgesetzten Größen \hat{x}_{k-1} ,

$P_{xx,k-1}$, die initial vorzugeben sind.

$$\begin{aligned}\mathcal{X}_{0,k-1} &= \hat{x}_{k-1} \\ \mathcal{X}_{i,k-1} &= \hat{x}_{k-1} + \gamma \left(\sqrt{P_{xx,k-1}} \right)_i \quad 1 \leq i \leq n \\ \mathcal{X}_{i,k-1} &= \hat{x}_{k-1} - \gamma \left(\sqrt{P_{xx,k-1}} \right)_i \quad n+1 \leq i \leq 2n\end{aligned}\tag{3-17}$$

mit

$$\begin{aligned}W_0 &= \frac{\kappa}{n + \kappa} \\ W_i &= \frac{1}{2(n + \kappa)} \quad 1 \leq i \leq 2n.\end{aligned}\tag{3-18}$$

Der Term $\left(\sqrt{P_{xx,k-1}} \right)_i$ steht dabei für die i -te Spalte der Wurzel der Kovarianzmatrix $P_{xx,k-1}$, mit $\gamma = \sqrt{(n + \kappa)}$. Die Matrixwurzel sollte mittels effektiver Verfahren wie der Cholesky-Zerlegung gebildet werden, wobei von [SDS⁺13] gezeigt wurde, dass die Wahl des Verfahrens zur Matrixwurzelbestimmung Einfluss auf die Schätzgüte haben kann. Der Abstand der Sigma-Punkte vom Erwartungswert sowie die relative Gewichtung der symmetrisch um den Erwartungswert liegenden Punkte zu dem auf dem Erwartungswert liegenden Sigma-Punkt werden dabei von der nicht beeinflussbaren Systemordnung n sowie von dem Entwurfparameter κ bestimmt. Der Einfluss der Wahl von κ wird nachfolgend anhand eines einfachen Beispielles verdeutlicht.

In Bild 3-9 ist das zu Beginn dieses Kapitels eingeführte Beispiel des Einmassenschwingers mit den Verteilungen zum Zeitschritt $k = 1$ und $k = 2$ dargestellt. Die Erfassung der beiden statistischen Momente erfolgt nach (3-17), wonach fünf Sigma-Punkten zu berechnen sind. Die Auswirkungen der Variation des Parameters κ ist durch die unterschiedlichen Sigma-Punkte-Sätze verdeutlicht, wobei die Lage durch die Punkte in der Zustandsebene und die Gewichtung jedes einzelnen Sigma-Punktes durch die betragsmäßige Höhe bestimmt sind. Die Auswirkungen der Parametervariation κ auf die Wölbung werden in dieser Darstellung aus Übersichtlichkeitsgründen nicht aufgeführt.

Gut zu erkennen ist, wie sich bei Änderungen von κ die Lage der Sigma-Punkte vom Erwartungswert sowie die Gewichtung, verbunden mit dem jeweiligen Sigma-Punkt, verändern. Bei einer Vergrößerung des Faktors κ entfernen sich die Sigma-Punkte vom Erwartungswert, und gleichzeitig erhöht sich die Gewichtung des nullten Sigma-Punktes. Die Verringerung des Faktors κ bewirkt das Gegenteil. Die Abstandsänderung bei unterschiedlichen Werten des Parameters κ ist durch den Term $\sqrt{(n + \kappa)}$ vor der Matrixwurzel $\sqrt{P_{xx,k-1}}$ und die erhöhte Gewichtung durch die Gewichtungsabhängigkeit von κ gegeben (siehe auch (3-17)). Durch die veränderte Gewichtung des Sigma-Punktes im Erwartungswert der Verteilung kann somit Einfluss auf das vierte Moment der Wölbung genommen und so eventuell vorhandene Informationen genutzt werden [JUDW95].

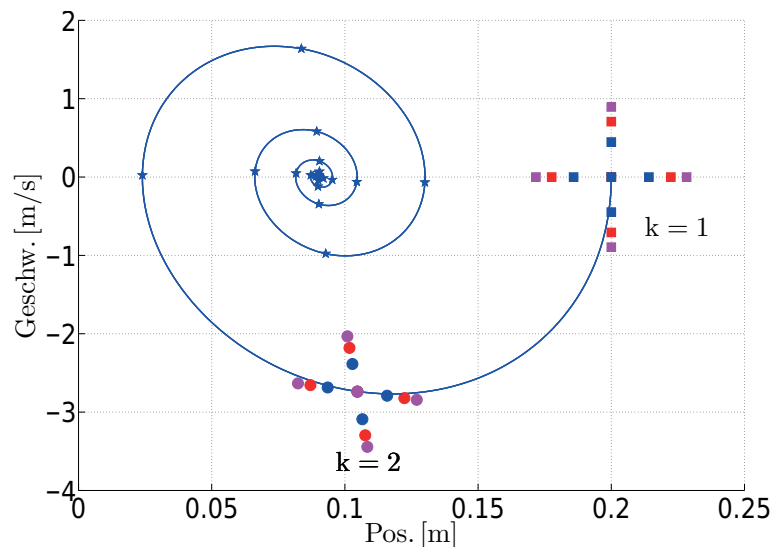
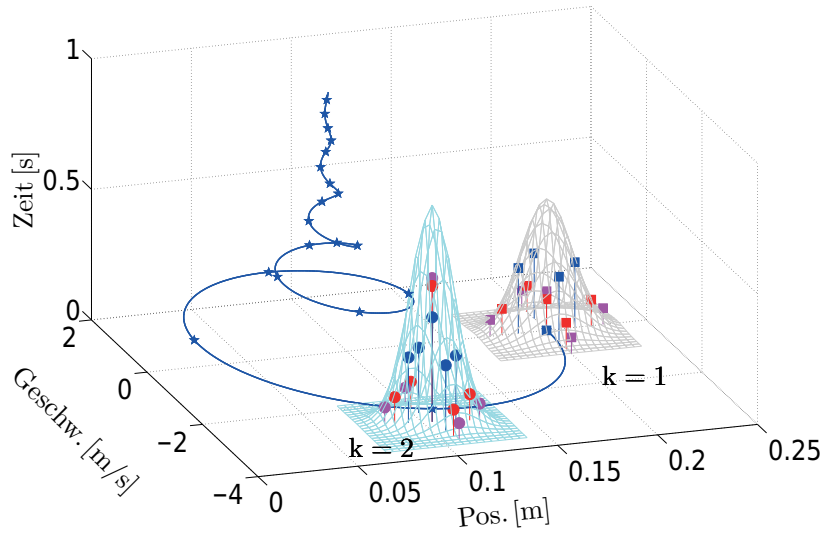


Bild 3-9: Berechnete und transformierte Sigma-Punkte bei unterschiedlicher Wahl von κ bei der SUT-Sigma-Punkte-Variante. Blau: $\kappa = 0$, Rot: $\kappa = 0.5$, Magenta: $\kappa = 1$

Wenn es sich bei der Zufallsvariablen um eine Gaußverteilung handelt, kann ein optimaler Wert für κ angegeben werden. Die ersten beiden Momente werden nach obiger Herleitung durch die Zwangsbedingungen erfasst. Da es sich bei der vorgestellten Wahl der Sigma-Punkte um einen symmetrischen Ansatz handelt, sind alle ungeraden Momente gleich Null und somit auch das dritte Moment, die Schiefe, korrekt wiedergegeben. Die Wahl von κ kann durch eine Minimierung des Fehlers im vierten Moment vorgenommen werden. Hierzu wird (3-12) erweitert, um das vierte Moment abzubilden. Dies erfolgt wiederum für die Standard-Gaußverteilung, deren Wölbung gleich 3 ist:

$$e[\mathcal{S}, p(z)] = [2W_1 s_1^4 - 3] = 0. \quad (3-19)$$

Durch Umformung und Einsetzen von (3-16) für s_1 und W_1 aus (3-17) in (3-19) ergibt sich die optimale Einstellung des Parameters κ bei vorliegender Gaußverteilung zu:

$$\begin{aligned} W_1 &= \frac{3}{2s_1^4} = \frac{3}{2(\sqrt{n+\kappa})^4} \\ \Rightarrow \kappa &= 3 - n. \end{aligned} \quad (3-20)$$

Nachteilig bei dieser Art der Sigma-Punkte-Bestimmung ist die Unbeschränktheit des Radius der Sigma-Punkte um den Erwartungswert. Nach (3-17) nimmt der Radius der Sigma-Punkte um den Erwartungswert mit $\sqrt{(n+\kappa)}$ zu. Dies kann schon bei moderaten Systemordnungen zur Berechnung von Sigma-Punkten führen, die weit vom Erwartungswert entfernt liegen. Systemtechnisch bedeutet dies die Auswertung eines Arbeitspunktes im Zustandsraum, der weit entfernt vom betrachteten liegt, wodurch Systemeigenschaften wie Nichtlinearitäten erfasst werden, die nichts mit dem eigentlichen Arbeitspunkt zu tun haben. Ein weiterer Nachteil besteht in der Wahl von $\kappa < 0$. Dies kann dazu führen, dass die Matrixwurzel nach (3-17) nicht berechenbar ist, da die positive Semidefinitheit nicht mehr gewährleistet werden kann, womit sich der nächste Abschnitt beschäftigt.

3.2.2 Scaled-Unscented Transformation

In diesem Abschnitt wird die Second-Order Unscented Transform zur Scaled-Unscented Transform erweitert, um den Nachteil der Unbeschränktheit des Sigma-Punkte-Radius \mathcal{X}_i mit $i = 1, \dots, 2n$ bei steigender Systemordnung zu umgehen und die Berechnung von nicht positiv semidefiniten Kovarianzmatrizen $P_{xx,k-1}$ zu verhindern. In [JUDW95] wurde vorgeschlagen, die Unbeschränktheit mit Hilfe von negativen Gewichtungen zu umgehen, was allerdings zu nicht positiv definiten Kovarianzmatrizen führen kann, aus denen keine reellwertige Matrixwurzel mehr berechnet werden kann. Die Verwendung negativer Gewichtungen stellt somit keine Alternative dar. In [Jul02] schlägt Julier eine Skalierung der Sigma-Punkte

um den Erwartungswert vor. Die Herleitung erfolgt wiederum an der Standard-Normalverteilung z , die zum Schluss dieses Abschnitts mittels der Abbildung aus (3-10) auf einen beliebigen Zustandsvektor übertragen werden kann.

Julier führt hierzu den Parameter α ein, der einen skalaren Wert mit $0 < \alpha \leq 1$ darstellt und einen bestehenden Satz an Sigma-Punkten durch die Transformation

$$\mathcal{X}'_i = \mathcal{X}_0 + \alpha(\mathcal{X}_i - \mathcal{X}_0) \quad \forall i = 1, \dots, 2n \quad (3-21)$$

zum Erwartungswert hin bzw. vom Erwartungswert weg skaliert, wodurch der Satz an Sigma-Punkten aus (3-2) zu

$$\mathcal{S}' = \{i = 0, 1, \dots, l-1 : \mathcal{X}'_i, W'_i\} \quad (3-22)$$

überführt wird. Zu ermitteln sind die Gewichtungen W'_i der skalierten Sigma-Punkte \mathcal{X}'_i mit $i = 0, \dots, 2n$, wobei eine Reihe von Anforderungen eingehalten werden müssen:

- Der Erwartungswert der Zufallsvariablen z muss weiterhin korrekt erfasst werden.
- Die Kovarianz der Zufallsvariablen z muss weiterhin korrekt erfasst werden.
- Die positive Semidefinitheit der Kovarianz muss sichergestellt sein.

Die Berechnung der Gewichtungsfaktoren W'_i mit $i > 0$ erfolgt durch die Erhaltung der Kovarianz. Im ursprünglichen Satz an Sigma-Punkten wird die Kovarianz P_{zz} durch die gewichtete Summe des äußeren Produktes dargestellt, mit $E[z] = 0 = \mathcal{X}_0$:

$$P_{zz} = \sum_{i=0}^{2n} W_i(\mathcal{X}_i - \mathcal{X}_0)(\mathcal{X}_i - \mathcal{X}_0)^T. \quad (3-23)$$

Diese Bedingung muss ebenso für den transformierten Satz an Sigma-Punkten mit $\mathcal{X}_0 = \mathcal{X}'_0$ gelten, da dieser auf dem Erwartungswert bleibt und nicht skaliert wird.

$$P_{zz} = \sum_{i=0}^{2n} W'_i(\mathcal{X}'_i - \mathcal{X}_0)(\mathcal{X}'_i - \mathcal{X}_0)^T \quad (3-24)$$

Durch Gleichsetzung von (3-23) mit (3-24) sowie Ersetzen von \mathcal{X}'_i durch $\mathcal{X}_0 + \alpha(\mathcal{X}_i - \mathcal{X}_0)$ aus (3-21) gelangt man zu folgender Darstellung:

$$\begin{aligned} \sum_{i=0}^{2n} W_i(\mathcal{X}_i - \mathcal{X}_0)(\mathcal{X}_i - \mathcal{X}_0)^T = \\ \sum_{i=0}^{2n} W'_i(\mathcal{X}_0 + \alpha(\mathcal{X}_i - \mathcal{X}_0) - \mathcal{X}_0)(\mathcal{X}_0 + \alpha(\mathcal{X}_i - \mathcal{X}_0) - \mathcal{X}_0)^T. \end{aligned} \quad (3-25)$$

Für $i > 0$ und unter Ausnutzung der Bedingung, dass $W_1 = W_2 = W_i$ gilt, mit $i = 3, \dots, 2n$, lässt sich (3-25) nach

$$W_1 \sum_{i=1}^{2n} (\mathcal{X}_i - \mathcal{X}_0)(\mathcal{X}_i - \mathcal{X}_0)^T = W'_1 \sum_{i=1}^{2n} (\alpha(\mathcal{X}_i - \mathcal{X}_0))(\alpha(\mathcal{X}_i - \mathcal{X}_0))^T \quad (3-26)$$

umformen. Da α eine Konstante ist, lässt sich (3-26) weiter nach

$$W_1 \sum_{i=1}^{2n} (\mathcal{X}_i - \mathcal{X}_0)(\mathcal{X}_i - \mathcal{X}_0)^T = W'_1 \alpha^2 \sum_{i=1}^{2n} (\mathcal{X}_i - \mathcal{X}_0)(\mathcal{X}_i - \mathcal{X}_0)^T \quad (3-27)$$

umstellen, woraus für die Berechnung der skalierten Gewichtungen $W'_1 = W'_i$ mit $i = 1, \dots, 2n$ folgt:

$$W'_i = \frac{W_i}{\alpha^2}. \quad (3-28)$$

Die Berechnung des Gewichtungsfaktors W_0 erfolgt über die Gleichsetzung der Normierungsbedingung (3-29) und Einsetzen der Lösung³ aus (3-28).

$$1 = \sum_{i=0}^{2n} W_i = W_0 + \sum_{i=1}^{2n} W_i \Rightarrow 1 = W_0 + \alpha^2 \sum_{i=1}^{2n} W'_i \quad (3-29)$$

Die Normalisierungsbedingung (3-3) des skalierten Sigma-Punkte-Satzes \mathcal{S}' lässt sich in gleicher Form durch

$$1 = \sum_{i=0}^{2n} W'_i = W'_0 + \sum_{i=1}^{2n} W'_i \Rightarrow \sum_{i=1}^{2n} W'_i = 1 - W'_0 \quad (3-30)$$

ausdrücken. Einsetzen von (3-30) in (3-29) führt auf die Lösung der Gewichtung W'_0 :

$$1 = W_0 + \alpha^2(1 - W'_0). \quad (3-31)$$

Zusammenfassend lassen sich die neuen Gewichtungen W'_i durch den Parameter α und die bekannten Gewichtungen W_i mit $i = 0, \dots, 2n$ in der Form

$$W'_i = \begin{cases} \frac{W_0}{\alpha^2} + 1 - \frac{1}{\alpha^2} & i = 0 \\ \frac{W_i}{\alpha^2} & 1 \leq i \leq 2n \end{cases} \quad (3-32)$$

³Dies ist gültig, weil $\sum_{i=1}^{2n} W_i = \sum_{i=1}^{2n} W'_i$ gilt.

berechnen. Bei einer Wahl von $\alpha = 1$ entsprechen die in (3-32) gezeigten Gewichtungen denen der SOT, was zeigt, dass es sich um eine reine Erweiterung des SOT-Verfahrens handelt und bei der Setzung von $\alpha = 1$ der Sigma-Punkte-Satz unverändert bleibt. Durch die Abbildungseigenschaft aus (3-10) kann die für die Standard-Normalverteilung hergeleitete Form auf die Erfassung von Verteilungen mit beliebigem Erwartungswert und Kovarianz erweitert werden. Die Wirkungsweise des Parameters α ist in Bild 3-10 anhand des bekannten Beispiels des Massenschwingers dargestellt. Die berechneten bzw. transformierten Sigma-Punkte sind als Quadrate bzw. Kugeln markiert.

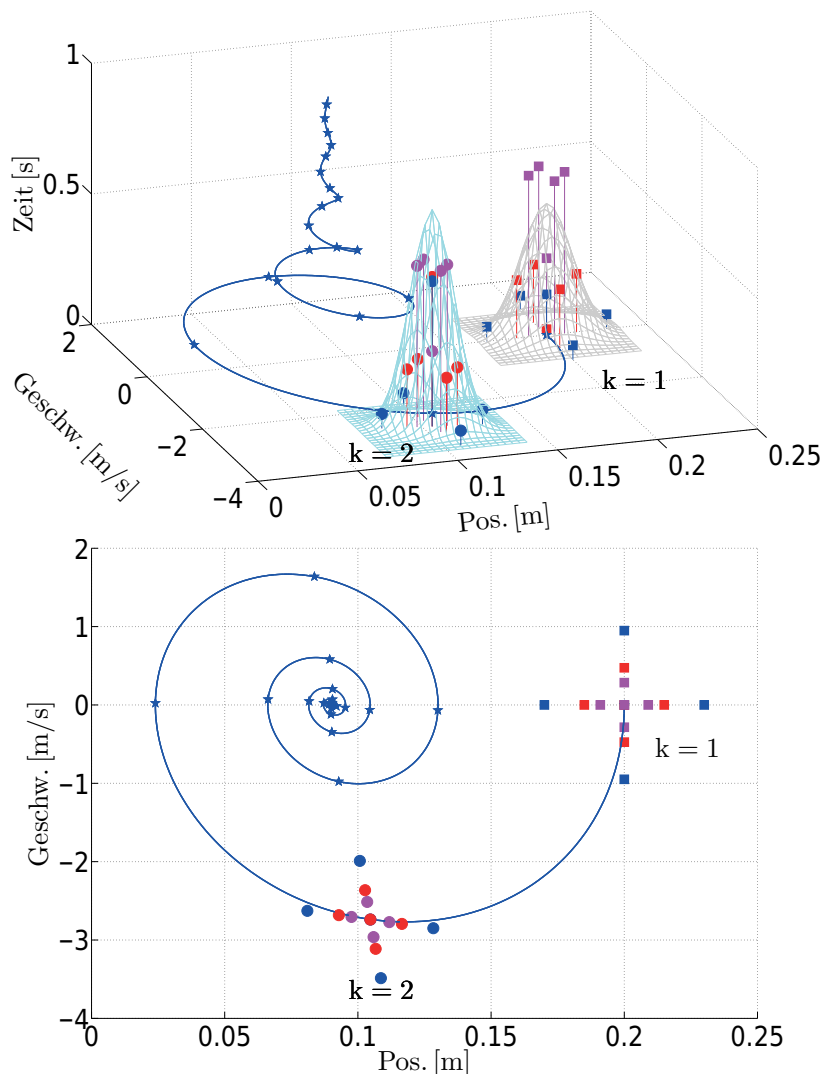


Bild 3-10: Berechnete und transformierte Sigma-Punkte bei unterschiedlicher Wahl von α . Rot: $\alpha = 0.5$, Blau: $\alpha = 1$; Magenta: $\alpha = 0.3$

Es ist gut zu erkennen, dass, je kleiner der Wert von α wird, die Sigma-Punkte näher an den Erwartungswert skaliert werden und die Gewichtung, dargestellt durch die betragsmäßige Höhe, zunimmt, da die Wahrscheinlichkeit des Auftre-

tens dieses Zustandsvektors erhöht wird. Erwartungswert und Kovarianz der zu bestimmenden cyanfarbenen Verteilung bleiben dabei erhalten. Im unteren Teil von Bild 3-10 ist die Draufsicht der Sigma-Punkte dargestellt, wobei die Punkte sich bei zunehmender Wahl von α entlang der Eigenvektoren der Kovarianz $P_{xx,k-1}$ und $P_{xx,k}^-$ vom Erwartungswert entfernen.

Bisher wurden alle Berechnungen durchgeführt, um den bestehenden Erwartungswert \hat{x}_{k-1} und die Kovarianz $P_{xx,k-1}$ einer Verteilung mittels eines Satzes an Sigma-Punkten $\mathcal{X}_{i,k-1}$ und zugehörigen Gewichtungen W_i mit $i = 0, \dots, 2n$ zu erfassen, wobei hier zur bekannten Kalman-Filter-Notation gewechselt wurde. Nach der Transformation der berechneten Sigma-Punkte $\mathcal{X}_{i,k-1}$ nach (3-4) können der A-priori-Erwartungswert \hat{x}_k^- und Kovarianz $P_{xx,k}^-$ mittels (3-6) und (3-5) berechnet werden.

In [Jul02] wird ein weiterer Gewichtungsfaktor β eingeführt, der auf die Gewichtung des nullten, transformierten Sigma-Punktes $\mathcal{X}_{0,k|k-1}$ bei der Kovarianzberechnung wirkt und mittels dem Einfluss auf die Wölbung der Verteilung **nach** der Transformation genommen werden kann. Der Unterschied zum Parameter κ besteht demnach darin, dass durch β die Wölbung der Verteilung $P_{xx,k}^-$ explizit verändert werden kann, die auf Basis der transformierten Sigma-Punkte $\mathcal{X}_{i,k-1}$ berechnet wurde, während mittels κ Informationen über die Wölbung der Verteilung $P_{xx,k-1}$ zur Berechnung des Sigma-Punkte-Satzes $\mathcal{X}_{i,k-1}$ eingebracht werden können. Für eine genaue Herleitung des Einflusses sei auf [Jul02] und [JU04] verwiesen, die auch gezeigt haben, dass bei Vorliegen einer Gaußverteilung der Fehler im vierten Moment durch die Wahl von $\beta = 2$ minimiert wird.

Die Wirkung von β ist in Bild 3-11 dargestellt. Je größer der Wert von β wird, umso höher wird der nullte, transformierte Sigma-Punkte $\mathcal{X}_{0,k|k-1}$ gewichtet, wodurch sich die Wölbung der cyanfarbenen Verteilung verändert, wie bereits in Bild 3-8 gezeigt. An dieser Stelle wird auf die Darstellung der unterschiedlichen Wölbungen aus Übersichtlichkeitsgründen verzichtet und nur der Fall von $\beta = 5$ dargestellt. Alle übrigen Sigma-Punkte $\mathcal{X}_{i,k}$, $i = 0, \dots, 2n$ und $\mathcal{X}_{i,k|k-1}$, $i = 1, \dots, 2n$ bleiben unverändert, ebenso wie die Gewichtungen, die in Tabelle 3-1 berechnet wurden.

Tabelle 3-1: Berechnete Gewichtungen bei unterschiedlichen Werten von β

Parametrierung	$\beta = 0$		$\beta = 2$		$\beta = 5$	
Gewichtung	$W_i^{(c)}$	$W_i^{(m)}$	$W_i^{(c)}$	$W_i^{(m)}$	$W_i^{(c)}$	$W_i^{(m)}$
W_0	0.6	0.6	2.6	0.6	5.6	0.6
$W_{1,\dots,4}$	0.1	0.1	0.1	0.1	0.1	0.1

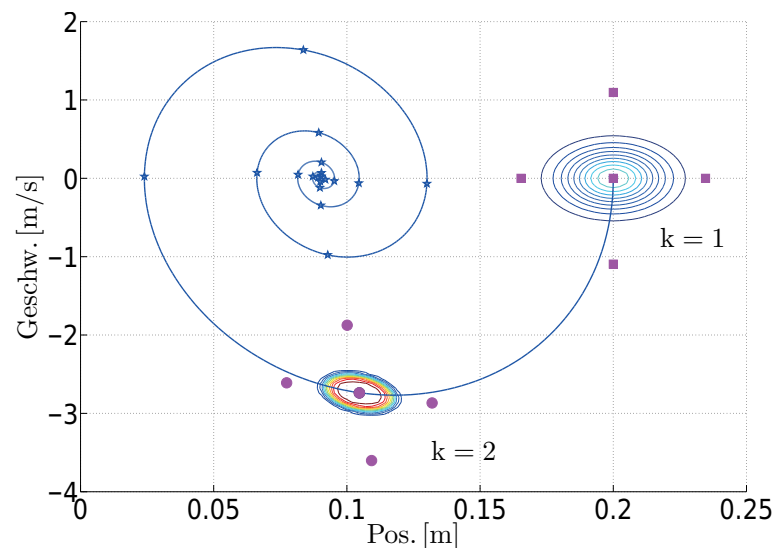
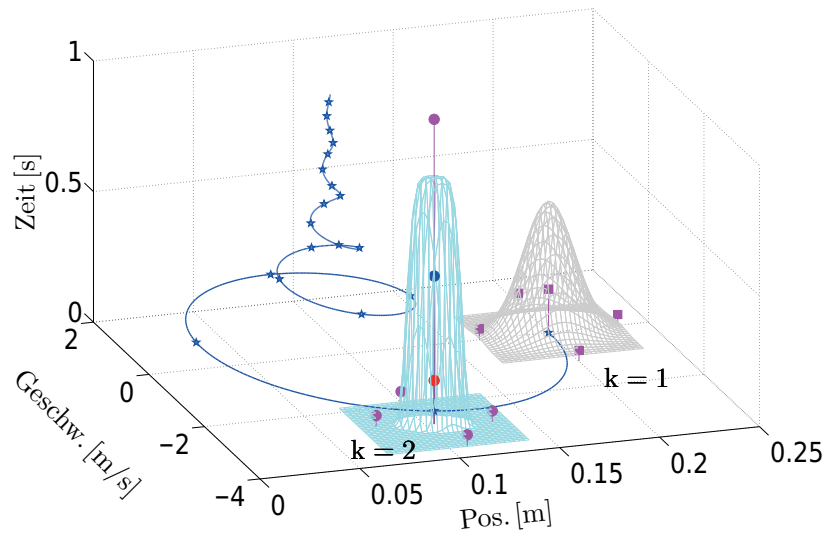


Bild 3-11: Berechnete und transformierte Sigma-Punkte bei unterschiedlicher Wahl von β . Rot: $\beta = 0$, Blau: $\beta = 2$; Magenta: $\beta = 5$

Zusammenfassend kann die SUT mit Einführung des Parameters $\lambda = \alpha^2(n+\kappa) - n$ als

$$\begin{aligned} \mathcal{X}_{0,k-1} &= \hat{x}_{k-1} & i &= 0 \\ \mathcal{X}_{i,k-1} &= \hat{x}_{k-1} + \left(\sqrt{(n+\lambda)P_{xx,k-1}} \right)_i & i &= 1, \dots, n \\ \mathcal{X}_{i,k-1} &= \hat{x}_{k-1} - \left(\sqrt{(n+\lambda)P_{xx,k-1}} \right)_i & i &= n+1, \dots, 2n \end{aligned} \quad (3-33)$$

dargestellt werden. Durch die Einführung des Parameters β wird die Gewichtung des nullten, transformierten Sigma-Punktes $\mathcal{X}_{0,k|k-1}$ zur Berechnung von Erwartungswert und Kovarianz verändert, wodurch die Gewichtungen aufgeteilt werden müssen. Die Gewichtungen zur Berechnung des Erwartungswertes werden im Folgenden mit $W_i^{(m)}$ bezeichnet und die Gewichtungen zur Berechnung der Kovarianz mit $W_i^{(c)}$, wobei in (3-34) nachrechenbar ist, dass sich die Gewichtungen gegen (3-32) nur hinsichtlich des additiven Parameters β unterscheiden:

$$\begin{aligned} W_0^{(m)} &= \frac{\lambda}{n+\lambda} \\ W_0^{(c)} &= \frac{\lambda}{n+\lambda} + 1 - \alpha^2 + \beta \\ W_i^{(m)} &= W_i^{(c)} = \frac{1}{2(n+\lambda)} & i &= 1, \dots, 2n. \end{aligned} \quad (3-34)$$

Das in (3-33) bzw. (3-34) hergeleitete Verfahren ist das erste der in dieser Arbeit genutzten, mit den Entwurfsparametern $\theta = [\alpha, \beta, \kappa]$.

3.2.3 Spherical-Simplex Transformation

In dieser Arbeit werden die Sigma-Punkte-Filter zur Zustands- und Parameterschätzung an Black-Box-Systemen mit der Zielrichtung der Echtzeitfähigkeit genutzt. Im letzten Abschnitt wurde bereits deutlich, dass jeder der berechneten $2n+1$ Sigma-Punkte, die zur Erfassung von \hat{x}_{k-1} und $P_{xx,k-1}$ genutzt werden, durch das als Black-Box vorliegende Modell transformiert werden muss. Dies stellt auch den rechenintensivsten Schritt dar, weshalb es nahe liegt, auch Verfahren anzuwenden, die mit weniger als $2n+1$ Sigma-Punkten auskommen. Hierzu werden im Folgenden die Eigenschaften der in dieser Arbeit genutzten Spherical-Simplex Transformation (SST) erläutert, die mit $n+2$ Sigma-Punkten auskommt und von Julier und Uhlmann in [Jul03] vorgestellt wurde.

Die grundsätzliche Verteilung der Sigma-Punkte zur Erfassung von Erwartungswert und Kovarianz einer Standard-Normalverteilung ist in Bild 3-12 links vergleichend gegenüber der SUT-Berechnung rechts dargestellt. Neben dem im Erwartungswert positionierten Sigma-Punkt \mathcal{X}_0 werden $n+1$ weitere um diesen verteilt.

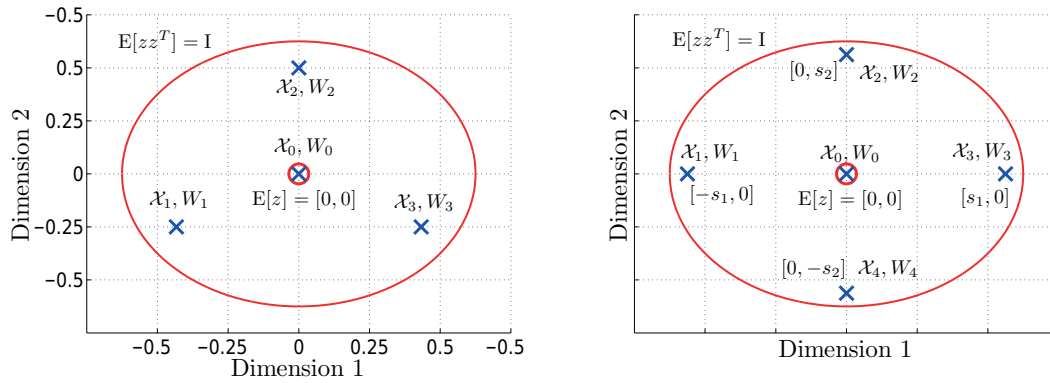


Bild 3-12: Sigma-Punkte-Verteilung zur Erfassung von Erwartungswert und Kovarianz einer Standard-Normalverteilung

Diese Verteilung von Sigma-Punkten bringt allerdings ein Problem mit sich: Die Verteilung ist nicht mehr symmetrisch, und damit ist der Term der Schiefe nicht mehr automatisch korrekt erfasst, wie im letzten Kapitel bei der Nutzung von $2n + 1$ Sigma-Punkten. Um dies zu verdeutlichen, wird das Verfahren von Mardia [KTO04]

$$\Psi = \frac{1}{l} \sum_{r,s}^l g_{r,s}^3 \quad (3-35)$$

angewandt, das eine Maßzahl Ψ der multivariaten Schiefe⁴ eines Stichprobenumfanges liefert, mit

$$g_{r,s} = (\mathcal{X}_r - \mathcal{X}_0) P_{zz}^{-1} (\mathcal{X}_s - \mathcal{X}_0) \quad r, s = 0, \dots, l \quad (3-36)$$

wobei l die Anzahl der Stichproben einer Verteilung darstellt, was in diesem Fall der Anzahl der Sigma-Punkte entspricht mit $l = n + 2$. Im Falle einer Normalverteilung, die in Bild 3-12 vorliegt, müsste $\Psi = 0$ gelten. Die Stichproben der beiden in Bild 3-12 dargestellten Sigma-Punkte-Verfahren sind in Tabelle 3-2 dargestellt mit der Berechnung der Maßzahl der Schiefe nach Mardia.

Hieraus ist erkennbar, dass durch die symmetrische Anordnung von 2 Sigma-Punkten pro Dimension der SUT keine Schiefe vorhanden ist, die Verteilung der SST allerdings einen Schiefe-Term aufweist. Hieraus kann allerdings ein Worst-case-Szenario resultieren, da bei der Nutzung der SST keine Aussage getroffen werden kann, wie sich die Schiefe ausbildet, d. h. ob sie bei einer eindimensionalen Verteilung rechts- oder linksschief ist, was in Bild 3-13 qualitativ als rote und grüne Verteilung dargestellt ist. Somit kann es passieren, dass durch das SST-Verfahren eine linksschiefe Verteilung folgt (roter Verlauf), in Realität allerdings eine rechtsschiefe Verteilung vorliegt (grüner Verlauf), was zu Fehlern in der Transformation führt.

⁴Schiefe einer mehrdimensionalen Zufallsvariablen

Tabelle 3-2: Schiefe-Anteile einer symmetrischen und einer asymmetrischen Verteilungserfassung

	Spherical-Simplex Transformation	Scaled-Unscented Transformation
\mathcal{X}_0	[0; 0]	[0; 0]
\mathcal{X}_1	[−0.6062; −0.3500]	[−0.7071; 0]
\mathcal{X}_2	[0; 0.7000]	[0; 0.7071]
\mathcal{X}_3	[0.6062; −0.3500]	[−0.7071; 0]
\mathcal{X}_4	—	[0; −0.7071]
Ψ	0.0417	0

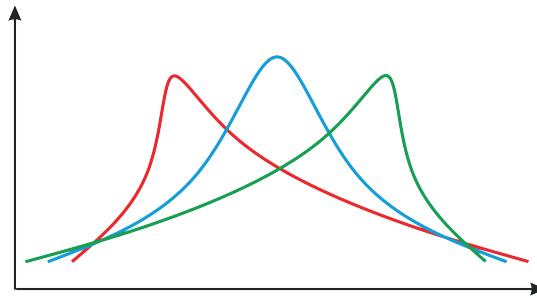


Bild 3-13: Sigma-Punkte-Verteilung zur Erfassung von Erwartungswert und Kovarianz einer Standard-Normalverteilung

In der in dieser Arbeit angewandten Kalman-Filter-Technik liegen allerdings keine Informationen über die Schiefe der Verteilung vor; es werden nur Erwartungswert und Kovarianz geschätzt und transformiert. Da keine Informationen über die Schiefe der gegebenen Verteilung des Zustandsvektors x_{k-1} zum Zeitpunkt k vorliegen, wird der Fehler bei der Nutzung des SST-Verfahrens minimal, wenn die Schiefe der SST-Verteilung minimal ist. Mit anderen Worten: Durch die Minimierung von Ψ nähert sich die Verteilung, die aus SST-Sigma-Punkten resultiert, der blauen Gaußverteilung aus Bild 3-13 an, wodurch der begangene Fehler bei realer rechts- oder linksschiefer Verteilung am kleinsten ist. In [Jul98] wird von Julier ein Verfahren vorgestellt, das die Schiefe einer Verteilung vollständig erfasst. Da dies allerdings mit einer erhöhten Anzahl an Sigma-Punkten einhergeht, wird es in der vorliegenden Arbeit nicht angewandt.

Um eine minimale Schiefe zu erreichen, wurde von Julier und Uhlmann in [JU02] das Minimal-Skew-Simplex-Verfahren eingeführt. Auf die Herleitung soll an dieser Stelle verzichtet werden, da das Vorgehen analog zu dem in den Abschnitten 3.2.1 und 3.2.2 ist. Die Zwangsbedingungen $c[\mathcal{S}, p(z)]$ setzen sich ebenfalls aus der Erfassung von Erwartungswert und Kovarianz sowie der Normalisierungsbedingung zusammen, und die zu minimierende Schiefe wird in der Kostenfunktion $e[\mathcal{S}, p(z)]$ berücksichtigt, woraus analog die Abstände der Sigma-Punkte vom Erwartungswert und die zugehörigen Gewichtungen folgen. Bei diesem Verfahren erhöht sich allerdings der Radius der Sigma-Punkte um den Erwartungswert mit

dem Faktor $2^{n/2}$, was in [Jul03] mittels der Erweiterung auf die SST auf den Faktor \sqrt{n} reduziert wurde. Da der Radius der Sigma-Punkte um den Erwartungswert bei zunehmender Systemordnung immer noch unbeschränkt ist, kann das bereits in Abschnitt 3.2.2 vorgestellte Verfahren der Skalierung hier ebenfalls angewandt werden und wird beispielsweise von [LCD⁺08] genutzt. Das abschließende Verfahren zur Berechnung der Sigma-Punkte und deren Gewichtungen mittels des SST-Verfahrens ist nachfolgend dargestellt. Die Matrix $\mathcal{Z} \in \mathbb{R}^{n \times (n+2)}$ stellt eine Hilfsmatrix dar, mit der eine Standard-Normalverteilung erfasst wird und die mittels

$$\mathcal{X}_{i,k-1} = \hat{x}_{k-1} + \left(\sqrt{P_{xx,k-1}} \right)_i \cdot \mathcal{Z} \quad i = 1, \dots, n+1 \quad (3-37)$$

den Erwartungswert und die Kovarianz in Kalman-Filter-Notation zum Zeitschritt k abbildet. \mathcal{Z} wird dabei initial iterativ berechnet, wobei \mathcal{Z}_i^j die j -te Dimension des i -ten Sigma-Punktes darstellt. Das Vorgehen zur Berechnung der Sigma-Punkte des SST-Verfahrens ist nachfolgend dargestellt.

- Initialisiere Vektorsequenz

$$\mathcal{Z}_0^1 = 0 \quad \mathcal{Z}_1^1 = -\frac{1}{\sqrt{2W_1^{(m)}}} \quad \mathcal{Z}_2^1 = \frac{1}{\sqrt{2W_1^{(m)}}} \quad (3-38)$$

- Erweitere Sequenz für die Dimensionen $j = 2, \dots, n$

$$\mathcal{Z}_i^j = \begin{cases} \begin{bmatrix} \mathcal{Z}_0^j \\ 0 \end{bmatrix} & i = 0 \\ \begin{bmatrix} \mathcal{Z}_i^{j-1} \\ -\frac{1}{j(j+1)W_1^{(m)}} \end{bmatrix} & i = 1, \dots, j \\ \begin{bmatrix} 0_{j-1} \\ \frac{j}{\sqrt{j(j+1)W_1^{(m)}}} \end{bmatrix} & i = j+1 \end{cases} \quad (3-39)$$

Die Gewichtungen können mit Vorgabe der Entwurfsparameter $0 < W_0 \leq 1$, $0 < \alpha \leq 1$ und $\beta > 0$ nach

$$\begin{aligned} W_0^{(m)} &= 1 + \frac{W_0 - 1}{\alpha^2} \\ W_0^{(c)} &= W_0^{(m)} + 1 - \alpha^2 + \beta \\ W_i^{(m)} &= W_i^{(c)} = \frac{1 - W_0}{\alpha^2(n+1)} \quad \forall i = 1, \dots, n+1 \end{aligned} \quad (3-40)$$

berechnet werden. Die Sigma-Punkte-Verteilung am bekannten Beispiel des Einmassenschwingers zur Vorhersage von A-priori-Zustandsvektor \hat{x}_k^- und Kovarianz $P_{xx,k}^-$ mit $k = 2$ auf Basis von \hat{x}_{k-1} und $P_{xx,k-1}$ ist in Bild 3-14 dargestellt. Der Einfluss der Entwurfparameter $\theta = [W_0, \alpha, \beta]$ ist äquivalent zu dem im Abschnitt 3.2.1 vorgestellten.

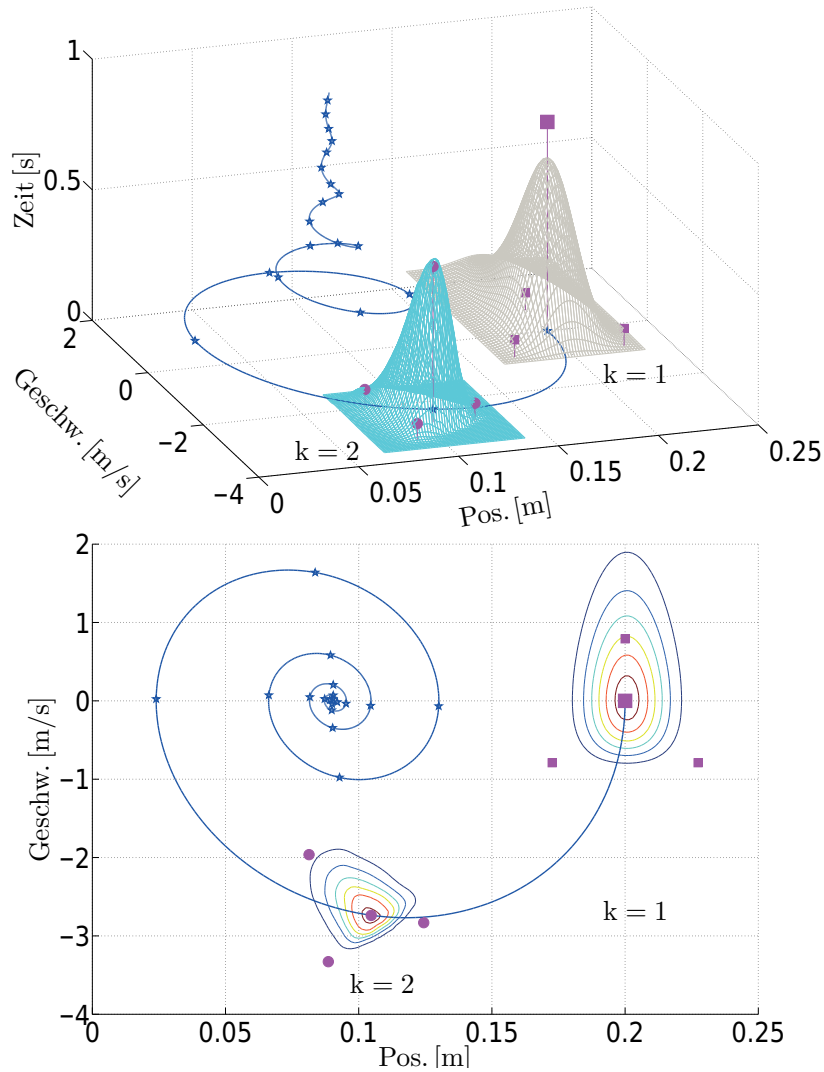


Bild 3-14: Berechnete Sigma-Punkte beim Einsatz der Spherical-Simplex Transformation am Einmassenschwinger zur Vorhersage von x_k^- und $P_{xx,k}$

3.2.4 Zusammenfassung der Sigma-Punkte-Verfahren

In den Abschnitten 3.2.1 und 3.2.3 wurden die beiden in dieser Arbeit verwendeten Sigma-Punkte-Berechnungsverfahren vorgestellt. Hieraus wird deutlich, dass hinsichtlich Genauigkeit und Echtzeitfähigkeit ein Zielgrößenkonflikt besteht. Neben den vorgestellten Varianten haben Julier in [JU97a] und Lerner in [Ler02] die

Fourth-Order Transformation vorgestellt, die Momente bis zur vierten Ordnung einer Verteilung erfasst und den Fehler im sechsten Moment minimiert. Tenne und Singh zeigten in [TS03], dass die Erfassung der Momente einer Gaußverteilten Zufallsvariablen x mit der Dimension n bis zur 8. Ordnung mit $5n$ -Sigma-Punkten möglich ist. Diese recht niedrige Anzahl an Sigma-Punkten erfordert allerdings eine Aufsplittung der Gewichtungen in zwei Gruppen, was mit einem höheren Aufwand in der Berechnung einhergeht. Durch eine weitere Aufsplittung in drei Gruppen ist die Erfassung einer Gaußverteilung bis zur 12. Ordnung möglich und wurde von Tenne als Higher-Order Transformation bezeichnet. In Bild 3-15 werden die genannten Verfahren hinsichtlich Genauigkeit und Echtzeitfähigkeit eingeordnet, woraus auch ersichtlich wird, warum in dieser Arbeit die Scaled-Unscented und die Spherical-Simplex Transformation zum Einsatz kommen.

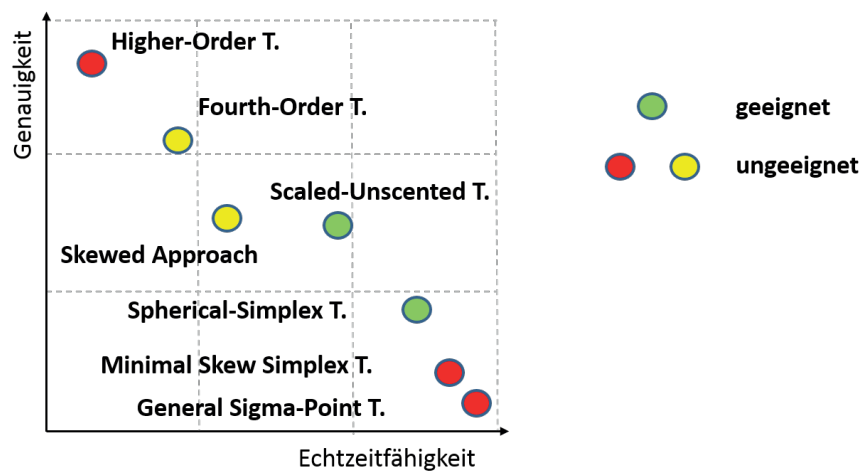


Bild 3-15: Einteilung der Sigma-Punkte-Verfahren hinsichtlich Echtzeitfähigkeit und Genauigkeit

3.3 Erweiterung des UKF-Algorithmus

Die Einbettung der beiden hergeleiteten Verfahren zur Sigma-Punkte-Berechnung in den Algorithmus eines Kalman-Filters führt direkt zum Unscented-Kalman-Filter, das nachfolgend vorgestellt wird. Die Grundstruktur entspricht dabei der bereits in Abschnitt 2.2 vorgestellten und besteht aus einer Filterinitialisierung sowie dem während des Betriebs laufenden Prädiktions- und Korrekturschritt. Welches der Verfahren zur Sigma-Punkte Berechnung genutzt wird, ist dabei vom Anwendungsfall abhängig, wobei das SST-Verfahren aufgrund seiner geringeren Anzahl an Sigma-Punkten zu bevorzugen ist. Sollte eine Schätzung mit dem SST-Verfahren keine zufriedenstellende Schätzgüte liefern, kann auf das SOT-Verfahren gewechselt werden, das eine höhere Güte aufweist, allerdings auch einen höheren Rechenaufwand bedeutet. In der Darstellung der folgenden Gleichungen wird das SOT-Verfahren eingesetzt.

Im ersten Schritt ist das Filter mit einem Startwert des Zustandsvektors und dessen Kovarianz zu initialisieren, wobei im Vorfeld die Gewichtungungen $W_0^{(m)}, W_0^{(c)}$ und $W_i^{(m)} = W_1^{(c)}$ mit $i = 1, \dots, 2n + 1$ auf Basis von $\theta = [\alpha, \beta, \kappa]$ zu berechnen sind. Der Startwert \hat{x}_0 wird in dieser Arbeit aus dem bestehenden Modell bezogen, was in Kapitel 5 erläutert wird. Die Startkovarianz kann aufgrund von mangelndem Wissen über den realen Zustand einfach als Einheitsmatrix angenommen werden:

$$\begin{aligned}\hat{x}_0 &= E[x_0] \\ P_{xx,0} &= E[(x_0 - \hat{x}_0)(x_0 - \hat{x}_0)^T].\end{aligned}\tag{3-41}$$

Nach der Initialisierung kann in die Betriebsphase für $k \in \{1, \dots, \infty\}$ übergegangen werden, wobei im ersten Schritt die in den letzten Abschnitten hergeleiteten Verfahren zur Berechnung der Sigma-Punkte zum Einsatz kommen. Auf Basis der A-posteriori-Schätzung \hat{x}_{k-1} und $P_{xx,k-1}$ bzw. der Initialisierung werden die Sigma-Punkte nach

$$\mathcal{X}_{k-1} = \begin{bmatrix} \hat{x}_{k-1} & \hat{x}_{k-1} + \gamma \left(\sqrt{P_{xx,k-1}} \right)_i & \hat{x}_{k-1} - \gamma \left(\sqrt{P_{xx,k-1}} \right)_i \end{bmatrix} \tag{3-42}$$

berechnet, mit $i = 1, \dots, n$, $\lambda = \alpha^2(n + \kappa) - n$, n als Systemordnung und $\gamma = \sqrt{n + \lambda}$ (vgl. (3-33)). Nach der Erfassung von Erwartungswert und Kovarianz mittels der Sigma-Punkte $\mathcal{X}_{k-1} \in \mathbb{R}^{n \times 2n+1}$ erfolgt die Transformation jedes Sigma-Punktes durch das Systemmodell, was in diesem Fall durch das als Black-Box vorliegende Modell gegeben ist. Die Transformation (vgl. (3-4)) wird dabei mittels eines in dieser Arbeit entwickelten Interfaces und unter Einwirkung des bekannten Systemeingangs u_{k-1} durchgeführt, das in Kapitel 5 vorgestellt wird und eine wesentliche Neuerung dieser Arbeit darstellt, mit dem zeitlich unkorrelierten Rauschprozess v_k .

$$[\mathcal{X}_{i,k|k-1}^*] = F(\mathcal{X}_{i,k-1}, v_k, u_{k-1}) \quad i = 1, \dots, 2n + 1 \tag{3-43}$$

Gleichzeitig bietet dies auch die Grundlage für eine Vielzahl weiterer Modellnutzungen wie zum Beispiel Model-Predictive-Control-Anwendungen. Die transformierten Sigma-Punkte $\mathcal{X}_{i,k|k-1}^*$ sind aus den vorherigen Abschnitten bereits bekannt. Im Anschluss kann auf Basis der transformierten Sigma-Punkte $\mathcal{X}_{i,k|k-1}^*$, $i = 0, \dots, 2n$ die A-priori-Vorhersage des Zustandsvektors \hat{x}_k^- (vgl. (3-5)) und dessen Fehlers $P_{xx,k}^-$ (vgl. (3-6)) mittels

$$\begin{aligned}\hat{x}_k^- &= \sum_{i=0}^{2n} W_i^{(m)} \mathcal{X}_{i,k|k-1}^* \\ P_{xx,k}^- &= \sum_{i=0}^{2n} W_i^{(c)} [\mathcal{X}_{i,k|k-1}^* - \hat{x}_k^-] [\mathcal{X}_{i,k|k-1}^* - \hat{x}_k^-]^T + R_{v,k}\end{aligned}\tag{3-44}$$

durchgeführt werden, wobei zur berechneten A-priori Fehlerkovarianz $P_{xx,k}^-$ die Kovarianzmatrix des Prozessrauschens $R_{v,k} = E[v_k v_k^T]$ als rein additiver Eingang

berücksichtigt wird, um die Unsicherheit in den vorhergesagten Zuständen zu verdeutlichen. Zur Einbeziehung dieser Unsicherheit in die Vorhersage der Messungen wird ein neuer Satz an Sigma-Punkten $\mathcal{X}_{k|k-1}$ nach

$$\mathcal{X}_{k|k-1} = \begin{bmatrix} \hat{x}_k^- & \hat{x}_{k-1} + \gamma \left(\sqrt{P_{xx,k}^-} \right)_i & \hat{x}_k^- - \gamma \left(\sqrt{P_{xx,k}^-} \right)_i \end{bmatrix} \quad (3-45)$$

gebildet, mit $i = 1, \dots, n$, $\lambda = \alpha^2 (n + \kappa) - n$, n als Systemordnung und $\gamma = \sqrt{n + \lambda}$ (vgl. (3-33)), die mittels des entwickelten Interfaces als Anfangszustände im Modell gesetzt werden. In bisherigen Arbeiten wie [VDMW⁺04] wird stets von einem Modell in Zustandsraumdarstellung ausgegangen, mit einer Trennung zwischen Dynamik und Messmodell. Durch die Nutzung von Black-Box-Modellen in dieser Arbeit existiert allerdings keine Trennung zwischen Mess- und Dynamikmodell. Beide Systeme sind in der Regel in einem Modell integriert. Zur Vorhersage der Messgrößen \hat{y}_k^- muss daher bei der nachfolgenden Transformation der Sigma-Punkte $\mathcal{X}_{k|k-1}$ beachtet werden, dass nur die Auswirkung des Prozessrauschens $R_{v,k}$ der Systemzustände auf die Messgrößen erfasst werden soll, mit dem zeitlich unkorrelierten Messrauschen n_k .

$$[\mathcal{Y}_{i,k|k-1}] = H(\mathcal{X}_{i,k|k-1}, n_k, u_{k-1}) \quad i = 1, \dots, 2n + 1 \quad (3-46)$$

Der Term $\mathcal{Y}_{i,k|k-1} \in \mathbb{R}^{n_y \times 2n+1}$ stellt die Ausgänge des als Black-Box vorliegenden Systemmodells dar, mit n_y als Anzahl der Systemausgänge. Pro transformiertem Sigma-Punkt $\mathcal{X}_{i,k|k-1}$ entsteht somit jeweils eine entsprechende „Messung“ des Modells, die im Folgenden zur Vorhersage der Messwerte \hat{y}_k^- genutzt werden.

Auf Basis der durch (3-46) vorhandenen Systemausgänge $\mathcal{Y}_{i,k|k-1}$ können die Vorhersage der Messungen \hat{y}_k^- und deren Kovarianz $P_{yy,k}$ erfolgen sowie die Kreuzkovarianz aus Zuständen und Messungen berechnet werden:

$$\begin{aligned} \hat{y}_k^- &= \sum_{i=0}^{2n} W_i^{(m)} \mathcal{Y}_{i,k|k-1} \\ P_{yy,k} &= \sum_{i=0}^{2n} W_i^{(c)} [\mathcal{Y}_{i,k|k-1} - \hat{y}_k^-] [\mathcal{Y}_{i,k|k-1} - \hat{y}_k^-]^T + R_{n,k} \\ P_{xy,k} &= \sum_{i=0}^{2n} W_i^{(c)} [\mathcal{X}_{i,k|k-1} - \hat{x}_k^-] [\mathcal{Y}_{i,k|k-1} - \hat{y}_k^-]^T. \end{aligned} \quad (3-47)$$

Der Term $R_{n,k} = E[n_k n_k'] \in \mathbb{R}^{n_y \times n_y}$ stellt die Kovarianzmatrix des Messrauschens dar, das ebenfalls in Abschnitt 2.2 vorgestellt wurde. Nach diesem Schritt liegen alle benötigten Größen für die Berechnung der Kalman-Verstärkung vor, über welche die Vorhersage des Zustandes \hat{x}_k^- nach dem Auftreten neuer Messungen y_k korrigiert werden kann und abschließend die vorhergesagte Fehlerkovarianz $P_{xx,k}^-$

angepasst wird:

$$\begin{aligned}\mathcal{K}_k &= P_{xy,k} P_{yy,k}^{-1} \\ \hat{x}_k &= \hat{x}_k^- + \mathcal{K}_k (y_k - \hat{y}_k^-) \\ P_{xx,k} &= P_{xx,k}^- - \mathcal{K}_k P_{yy,k} \mathcal{K}_k^T.\end{aligned}\tag{3-48}$$

Durch die Entwicklung eines Interfaces zur Transformation der Sigma-Punkte $\mathcal{X}_{i,k|k-1}$, $i = 0, \dots, 2n + 1$ und die Weiterentwicklung des Filteralgorithmus um Modelle, die Mess- und Dynamikmodell in sich vereinen, ist in dieser Arbeit die Einbindung von Black-Box-Modellen in ein Unscented-Kalman-Filter möglich. Gleichzeitig werden durch die aufgezeigten Vorteile von Sigma-Punkte-Filtern gegenüber etablierten Varianten wie dem EKF starke Nichtlinearitäten mit höherer Güte erfasst und somit die Schätzgüte einer Zustands- und Parameterschätzung erhöht.

Neben dieser Filtervariante wurden von van der Merwe in [VDMW⁺04] Varianten zur Reduzierung der Rechenintensität und Sicherstellung der positiven Definitheit eingeführt, wie das Square-Root-UKF, bei dem der rechenintensive Schritt der Bildung von Matrixwurzeln mittels Cholesky-Zerlegung durch ein Cholesky-Update reduziert wurde (siehe A.1).

3.4 Das Central-Difference-Kalman-Filter

Neben dem Unscented-Kalman-Filter kommen noch weitere Filtertypen für die Nutzung an Black-Box-Systemen mit starken Nichtlinearitäten in Frage, wie das Central-Difference-Kalman-Filter (CDKF), das von Ito und Xiong in [IX00] vorgestellt wurde. Das CDKF arbeitet sehr ähnlich zu der Funktionsweise des Unscented-Kalman-Filters und bettet ebenfalls eine Berechnungsmethode zur Transformation der statistischen Momente zweiter Ordnung eines Zustandsvektors mittels Sigma-Punkten in den Algorithmus eines Kalman-Filters ein. Die Berechnung der Sigma-Punkte erfolgt allerdings nicht über das in 3.2 vorgestellte Verfahren mittels Zwangs- und Strafbedingungen, sondern basiert auf Sterlings Polynom Interpolation [VDMW⁺04]. Aufgabe ist wiederum die Transformation einer Zufallsvariablen x_{k-1} mit Erwartungswert $\hat{x}_{k-1} = \bar{x}_{k-1}$ und Kovarianz $P_{xx,k-1}$ durch eine nichtlineare Funktion, wobei das in dieser Arbeit genutzte Black-Box-System durch $x_k = F(x_{k-1})$ gegeben ist. Diese Transformation kann ebenfalls ausgedrückt werden durch:

$$\begin{aligned}x_k &= F(x_{k-1}) \\ &= F(\bar{x}_{k-1} + \delta_{x_{k-1}}),\end{aligned}\tag{3-49}$$

wobei $\delta_{x_{k-1}}$ eine Zufallsvariable darstellt mit dem Erwartungswert 0 und der gleichen Kovarianz wie x_{k-1} . Die Transformation nach (3-49) kann dabei durch

eine Taylorreihenentwicklung um \bar{x}_{k-1} durchgeführt werden, was in Gleichung (3-50) mit einem Abbruch nach dem zweiten Term dargestellt ist:

$$y = F(\bar{x}_{k-1}) + D_{\delta_{x_{k-1}}} F + \frac{1}{2!} D_{\delta_{x_{k-1}}}^2 F + \dots \quad (3-50)$$

Der Term $D_{\delta_{x_{k-1}}}$ stellt dabei das totale Differential von F um \bar{x}_{k-1} dar. Die Bestimmung der notwendigen analytischen Ableitungen kann durch die Nutzung einer Interpolationsmethode umgangen werden, wobei eine Möglichkeit durch die Nutzung von Sterlings Polynominterpolation gegeben ist [VDMW⁺04]. (3-49) lässt sich dadurch zu

$$F(x_{k-1}) = F(\bar{x}_{k-1}) + \tilde{D}_{\Delta_{x_{k-1}}} F + \frac{1}{2!} \tilde{D}_{\Delta_{x_{k-1}}}^2 F \quad (3-51)$$

umformen, mit

$$\begin{aligned} \tilde{D}_{\Delta_{x_{k-1}}} F &= (x_{k-1} - \bar{x}_{k-1}) \frac{F(\bar{x}_{k-1} + \xi) - F(\bar{x}_{k-1} - \xi)}{2\xi} \\ \tilde{D}_{\Delta_{x_{k-1}}}^2 F &= (x_{k-1} - \bar{x}_{k-1})^2 \frac{F(\bar{x}_{k-1} + \xi) + F(\bar{x}_{k-1} - \xi) - 2F(\bar{x}_{k-1})}{\xi^2}. \end{aligned} \quad (3-52)$$

Der Parameter ξ stellt dabei eine Intervalllänge dar. Die Sterling-Interpolation kann nach (3-50), (3-51) und (3-52) als Taylorreihenentwicklung der nichtlinearen Funktion F um den Arbeitspunkt \bar{x}_{k-1} mit den zentralen Differenzenquotienten (3-52) angesehen werden, bei der die analytischen Ableitungen durch die zentralen Differenzenquotienten ersetzt werden [VDMW⁺04]. Die obige Angabe für den eindimensionalen Fall kann durch eine statistische Entkopplung auf den mehrdimensionalen Fall einer Zufallsvariablen übertragen und in [Sch97], [Sch11] nachvollzogen werden.

Van der Merwe zeigt in [VDMW⁺04], dass sich der Kern des Central Difference Filters, die Bildung von Differenzenquotienten um den Erwartungswert \bar{x}_{k-1} , auf

$$F(x_{k-1}) = F(x_{k-1}^- \pm \xi s_{x_{i,k-1}}). \quad (3-53)$$

abstrahieren lässt. Hierbei stellt $s_{x_{i,k-1}}$ die i -te Spalte der Matrixwurzel der Kovarianz $P_{xx,k-1}$ von x_{k-1} dar, die mittels Choleskyzerlegung gewonnen werden kann:

$$s_{x_{i,k-1}} = S_{x,k-1} e_i = (S_{x,k-1})_i = \left(\sqrt{P_{xx,k-1}} \right)_i \quad (3-54)$$

Der Term $\bar{x}_{k-1} \pm \xi s_{x_{i,k-1}}$ entspricht demnach von der Bedeutung her (3-10) bzw. (3-17) zur Berechnung des Sigma-Punkte-Satzes des UKF. Der einzige Unterschied besteht in dem Faktor ξ , mittels dessen die Gewichtungen der Sigma-Punkte bzw. der Abstand vom Erwartungswert eingestellt werden können. Die algorithmische Realisierung des CDKF ist in A.2 dargestellt.

Liu et al. stellen in [LWM11] das Central-Difference-Filter als dem UKF überlegen dar, da es neben einem problemspezifischen Vorteil mit nur einem Entwurfparameter $\theta_{CDF} = [\xi]$ auskommt, während das UKF mit $\theta_{UKF} = [\alpha, \beta, \kappa]$ parametrisiert werden muss. Dies stellt allerdings nur einen Vorteil dar, wenn man nicht in der Lage ist, die Parameter θ_{UKF} entsprechend der Verteilung der Zufallsvariablen einzustellen. Besitzt man Informationen über die höheren Momente der Zustandsverteilung x , entweder explizit über Expertenwissen oder implizit mittels Modellen und Anwendungsszenarien, so geht bei der Nutzung des CDKF die Möglichkeit verloren, diese Informationen gezielt einfließen zu lassen.

Neben den vorgestellten und implementierten Varianten des UKF und des CDKF existieren zahlreiche Implementierungsmöglichkeiten und problemspezifische Lösungen, die im Folgenden kurz angesprochen werden, deren Herleitung den Umfang der Arbeit allerdings bei Weitem übersteigen würde. Bei der Square-Root-Implementierung wird der Choleskyfaktor direkt durch das Sigma-Punkte-Satzes propagiert. Somit entfällt der rechenaufwändigste Schritt des UKF-Algorithmus: das Aktualisieren des Sigma-Punkte-Sets zu jedem Zeitschritt. Gleichzeitig bietet dieses Verfahren den großen Vorteil, dass die Matrix garantiert positiv semi-definit und dadurch numerisch stabil ist [VDMW01b], [ZMC07]. In [HXS⁺07] stellen Hao et al. das Rao-Blackwell UKF (RBUKF) vor, das entwickelt wurde, um dynamische Systeme zu behandeln, die neben dem linearen auch einen nichtlinearen Anteil besitzen. Das RBUKF kombiniert das UKF mit einem linearen KF, um den rechenintensiven Schritt der Sigma-Punkte-Transformation zu minimieren, was ebenfalls mit anderen Mitteln durch das Sequence-UKF (S-UKF) in [LXZ⁺08] erreicht wird. Li et al. stellen in [LPG⁺09] ein Strong-Tracking-Filter vor (STF-UKF), das durch zeitvariante Fading-Matrizen Anpassungen des Kalman-Gains und der Fehlerkovarianzmatrix vornimmt, um bei abrupten Änderungen des Zustandsvektors schneller auf diese reagieren zu können. Um auf plötzliche Änderungen im Messsignal reagieren zu können, wurde in [SH11] ein Verfahren vorgestellt, das die Sensorgrößen vor der Einbeziehung in den Update-Schritt auf Vertrauenswürdigkeit und Nutzbarkeit überprüft. Neben den echtzeitfähigen Anwendungen existieren ebenfalls Varianten, denen Messungen über den gesamten Betriebsraum zur Verfügung stehen und die in einem Offline-Verfahren arbeiten. Als bekanntester Vertreter ist das Rauch-Tung-Striebel UKF-Smoother [QXL⁺10] zu nennen, das in einem Vorwärts- und einem Rückwärtsfilter die Schätzgüte stark erhöhen kann.

In [AHH10] wird von Arasaratnam und Haykin das Cubature-Kalman-Filter (CKF) vorgestellt, das im Rahmen dieser Arbeit ebenfalls als eigenständiger Filtertyp zur Klasse der Sigma-Punkte-Filter gezählt wird. Der Ansatz des CKF ist sehr ähnlich zu dem des UKF und des CDKF. Eine erhöhte Genauigkeit gegenüber dem Extended-Kalman-Filter wird durch die Approximation eines Gauß-verteiltern Zustandsvektors mittels gewichteter Punkte gewonnen, die im Anschluss durch die nichtlineare Systemfunktion transformiert werden. Auf Basis der transformierten Punkte erfolgt die Berechnung der A-priori-Zustands- und

-Kovarianzschätzung. Die Berechnung der Punkte nebst Gewichtungen zur Approximation der Wahrscheinlichkeitsdichteverteilung unterscheidet sich von den Methoden des UKF und des CDKF und wird auf Basis von Cubature-Regeln durchgeführt, was die numerische Integration von mehrdimensionalen Integralen meint. Die Grundidee ist die Approximation der multidimensionalen Integrale über die Dichteverteilungen des Bayesschen Schätzers aus Abschnitt 2.2.3, der die allgemeine Form eines optimalen, rekursiven Schätzers darstellt. Zur Realisierung eines echtzeitfähigen Algorithmus gelingt dies allerdings nur unter einigen Annahmen, wie Gaußverteilten Zufallsvariablen. Als Resultat entwickelten Arasaratnam und Haykin ein Sigma-Punkte-Kalman-Filter, das mit $2n$ Punkten und einem Entwurfparameter $\theta_{CKF} = \alpha$ auskommt, mittels dem der Radius der symmetrisch um den Erwartungswert liegenden Punkte eingestellt werden kann. Arasaratnam und Haykin stellen das CKF als dem UKF und dem CDKF überlegen dar, was in dieser Arbeit nicht bestätigt werden kann, wie anhand des Filtervergleichs in 6.1.6 vorgestellt wird. Im Endeffekt basiert das CKF auf den gleichen Approximationen wie UKF und CDKF, um ein echtzeitfähiges, rekursives Filter zu realisieren. In [AHH10] wird das CKF mit dem UKF und der SOT-Sigma-Punkte-Berechnungsmethode verglichen, wobei der Entwurfparameter α als Skalierungsparameter bezeichnet wird, was nicht die Intention von Julier war. Die Skalierung erfolgt über die Anpassung der Gewichtung, da der Sigma-Punkt im Erwartungswert höher gewichtet wird, um das vierte Moment exakt zu erfassen. Die übrigen Sigma-Punkte müssen daher niedriger skaliert werden und sich vom Erwartungswert entfernen. Weiterhin hätte das CKF, wie in dieser Arbeit, mit der SST-Berechnungsmethode verglichen werden müssen. Abschließend kann das CKF mit dem CDKF verglichen werden, da es mit nur einem Entwurfparameter auskommt und daher genutzt werden sollte, falls keinerlei Informationen über das zugrundeliegende System vorliegen. Im Rahmen dieser Arbeit ist eine vollständig automatisierte Auslegung allerdings möglich, und Informationen für eine gezielte Auslegung des UKF mit seinen Entwurfparametern sind auf Basis von Szenarien ohne Expertenkenntnisse möglich.

3.5 Parameterschätzung mittels Sigma-Punkte-Filter

In den letzten Abschnitten wurde gezeigt, wie das Unscented-Kalman-Filter als Vertreter der Sigma-Punkte-Filter zur Zustandsschätzung eingesetzt werden kann und welche Erweiterungen und Anpassungen des Algorithmus für die Zustandsschätzung an Black-Box-Systemen in dieser Arbeit durchgeführt wurden. Grundsätzlich stehen eine Reihe von Verfahren und Implementierungsmöglichkeiten zur gleichzeitigen Schätzung von Zuständen und Parametern zur Verfügung, die nachfolgend als w_k bezeichnet werden und als zeitvariabel angesehen werden. Für einen guten Überblick sei an dieser Stelle auf [Nel00] verwiesen. In dieser Arbeit wurden das Verfahren der *Dual*-Schätzung und der *Joint*-Schätzung für die Anwendung an Black-Box-Systemen aufbereitet und hinsichtlich der Schätzgüte untersucht. Die Struktur der beiden Varianten ist in Bild 3-16 dargestellt.

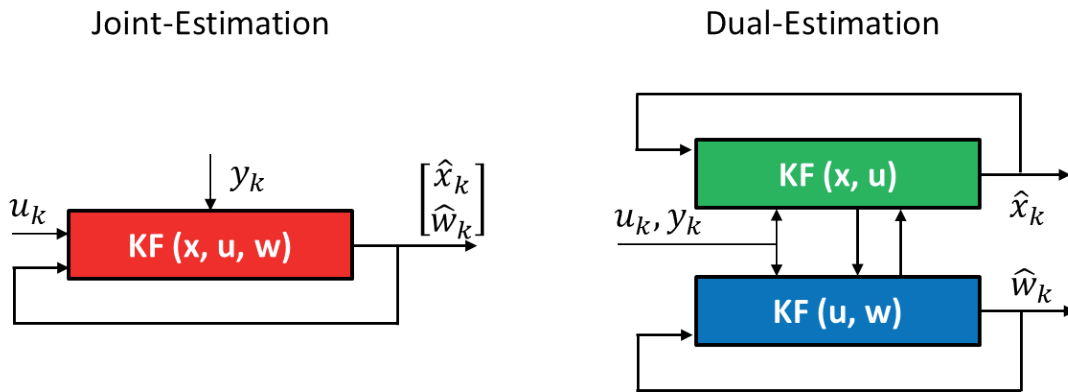


Bild 3-16: Unterschiedliche Varianten zur Realisierung einer Zustands- und Parameterschätzung mittels Kalman-Filterung. Links: Joint estimation mittels Augmentierung des Zustandsvektors. Rechts: Dual estimation mit Repräsentation der Parameterschätzung als eigene Zustandsraumdarstellung und Anwendung von zwei Filtern. Nach [VDM04]

Im Falle der *Joint*-Schätzung ist ein Filter für die gleichzeitige Vorhersage der Zustände und der unbekannten Parameter zuständig, während im Falle der *Dual*-Schätzung zwei Filter eingesetzt werden. Das erste Filter übernimmt die Aufgabe der Zustandsschätzung, während das zweite Filter die Schätzung des Parameters durchführt. Während jeden Schrittes tauschen die Filter ihre Zustands- bzw. Parameterschätzung aus, sodass beide Filter ihre jeweilige Aufgabe auf Basis der besten Schätzung des jeweils anderen Filters durchführen.

Das *Dual*-Verfahren hat dabei direkt mehrere Nachteile. Es resultiert ein höherer Implementierungsaufwand durch die Nutzung eines zweiten Filters, und es wird eine höhere Anzahl an Sigma-Punkten benötigt, die den größten Teil der Rechenzeit verursachen. Gleichzeitig hat die *Joint*-Schätzung für die in dieser Arbeit untersuchten Systeme durchweg die besseren Ergebnisse erbracht und wird daher im Folgenden hinsichtlich Funktionsweise und Anpassung zur Nutzung bei Black-Box-Systemen erläutert.

In diesem Fall ist neben dem zu bestimmenden unbekannten Zustandsvektor x_k zum Zeitschritt k ebenfalls der Parametervektor w_k zu ermitteln, wobei für beide Schätzvorgänge als Korrektur nur die verrauschten Messungen y_k zur Verfügung stehen.

3.5.1 Das Joint-Unscented-Kalman-Filter

Im Ansatz der *Joint*-Schätzung werden die zu bestimmenden Parameter $w_{k-1} \in \mathbb{R}^{n_p \times 1}$ zum Zeitpunkt $k-1$ als *Pseudo*-Zustände angesehen und der bisher zu schätzende Zustandsvektor x_{k-1} des Systems um diese augmentiert, mit n_p als Anzahl zu schätzender Parameter. Hieraus ergibt sich der erweiterte, zu bestimm-

mende Vektor der Zustände und Parameter zu

$$x_{w,k-1} = \begin{bmatrix} x_{k-1} \\ w_{k-1} \end{bmatrix}, \quad (3-55)$$

mit $x_{w,k-1} \in \mathbb{R}^{n+n_p \times 1}$, wobei die Parameter w_{k-1} als stückweise konstant in ihrer Dynamik angesehen werden. Durch diese Erweiterung erhöht sich die Dimension der Prozessrauschkovarianz $R^v \in \mathbb{R}^{n+n_p \times n+n_p}$ sowie die Dimension und Anzahl der Sigma-Punkte $\mathcal{X}_{k-1} \in \mathbb{R}^{n+n_p \times 2(n+n_p)+1}$. Die neue Fehlerkovarianz setzt sich durch die Erweiterung zu

$$\begin{aligned} P_{x_w x_w, k-1} &= E \left[(x_{w,k-1} - \hat{x}_{w,k-1}) (x_{w,k-1} - \hat{x}_{w,k-1})^T \right] \\ &= \begin{bmatrix} P_{xx,k-1} & P_{xw,k-1} \\ P_{wx,k-1} & P_{ww,k-1} \end{bmatrix} \end{aligned} \quad (3-56)$$

zusammen. Der Term $P_{xx,k-1}$ stellt in (3-56) die bereits bekannte Fehlerkovarianz der Zustände dar und der Term $P_{ww,k-1} \in \mathbb{R}^{n_p \times n_p}$ die Fehlerkovarianz der Parameter. Die Terme $P_{wx,k-1} = P_{xw,k-1}^T$ ergeben die Kreuzkovarianzen aus den Beziehungen der Zustände zu den Parametern und umgekehrt. Die Matrix der Sigma-Punkte setzt sich durch die Erweiterung um die Parameter w_{k-1} zu

$$\mathcal{X}_{k-1} = \begin{bmatrix} \mathcal{X}_{x,k-1} \\ \mathcal{X}_{w,k-1} \end{bmatrix} \quad (3-57)$$

zusammen, mit $\mathcal{X}_{x,k-1} \in \mathbb{R}^{n \times 2(n+n_p)+1}$, $\mathcal{X}_{w,k-1} \in \mathbb{R}^{n_p \times 2(n+n_p)+1}$, und werden mittels (3-43) transformiert. Hierbei ist zu beachten, dass die Sigma-Punkte $\mathcal{X}_{w,k-1}$ weiterhin einen Parameter darstellen und somit bei der Transformation nach (3-43) die Sigma-Punkte der Zustände $\mathcal{X}_{x,k-1}$ unter Einfluss der Sigma-Punkte der Parameter $\mathcal{X}_{w,k-1}$ transformiert werden. D. h.: Wie verändern sich die Zustände vom Zeitschritt $k-1$ zu k bei unterschiedlichen eingesetzten Parameterwerten, die in den Sigma-Punkten $\mathcal{X}_{w,k-1}$ berechnet wurden.

Wegen dieses Vorgehens führt die A-priori-Schätzung der Parameter nach (3-44) auf keine veränderten Werte für w_k , da $\mathcal{X}_{w,k-1} = \mathcal{X}_{w,k|k-1}$ gilt, wodurch die Schätzung von Parametern nur durch den A-posteriori-Schritt in (3-48) durchgeführt werden kann. Die Schätzung von Parametern wird somit stark durch die angesetzten Prozess- und Messrauschprozesse beeinflusst, da eine Korrektur nur über die Kalman-Verstärkung möglich ist, in das die Rauschkovarianzen als Vertrauensmaß eingehen.

Ein wichtiger Grund für die in dieser Arbeit aufgetretene bessere Schätzgüte des *Joint*-Verfahrens gegenüber der *Dual*-Schätzung liegt in der Beschaffenheit der Fehlerkovarianzmatrix $P_{x_w x_w, k-1}$. Bei der Nutzung eines *Dual*-Verfahrens arbeitet das Filter der Zustandsschätzung in dem bekannten Verfahren aus Abschnitt 3.3 mit der Fehlerkovarianz $P_{xx,k-1}$ aus (3-56), während das Filter der Parameterschätzung mit der Fehlerkovarianz $P_{ww,k-1}$ arbeitet. Durch diese Implementierung

besitzt das Filter nur $2n_p + 1$ Sigma-Punkte der Schätzparameter, deren Berechnung einzig auf $P_{ww,k-1}$ basiert. Die Erweiterung der Fehlerkovarianz zu $P_{xwxw,k-1}$ aus (3-56) bewirkt hingegen die Generierung von $2(n + n_p) + 1$ Sigma-Punkten der Schätzparameter, deren Werte auf den statistischen Zusammenhängen zu den Zuständen basieren und somit eine wesentlich höhere Schätzgüte bewirken.

3.5.2 Wahl der Kovarianz einer Parameterschätzung

Aufgrund der alleinigen Korrektur der zu schätzenden Parameter w_k im A-posteriori-Schritt stellt die Wahl des Prozessrauschens $R_{v,w}$ einen wichtigen Schritt im Entwurfsprozess dar. Die Prozessrauschmatrix $R_{v,xw}$ beim Einsatz des *Joint*-Verfahrens setzt sich dabei aus dem Prozessrauschen der Zustände $R_{v,x}$ und dem der Parameter $R_{v,w}$ nach

$$R_{v,xw} = \begin{bmatrix} R_v & 0 \\ 0 & R_{v,w} \end{bmatrix} \quad (3-58)$$

zusammen. Die Wahl von $R_{v,w}$ hat großen Einfluss auf das Konvergenzverhalten hinsichtlich Geschwindigkeit und Genauigkeit. Von van der Merwe wurde in [VDMW⁺04] gezeigt, dass eine beliebige Wahl von $R_{v,w}$ zu unerwünschtem Schätzverhalten führen kann und maßgeblichen Einfluss auf die Schätzgüte wie auch die Konvergenzzeit besitzt, was durch die Anwendungen in dieser Arbeit bestätigt werden konnte.

Demgemäß kann die Wahl des auf die zu schätzenden Parameter Einfluss nehmenden Prozessrauschens auf unterschiedlichen Wegen erfolgen. Die einfachste Möglichkeit besteht in einer willkürlichen Wahl, die bei fortschreitendem Zeitindex der Schätzung gegen Null „heruntergefahren“ wird. In [Nel00] wird von Nelson eine Fensterung der Daten nach

$$P_{ww,k}^- = \lambda_{RLS}^{-1} P_{ww,k-1} \quad (3-59)$$

vorgestellt, mit $P_{ww,k}^{-1} = R_{v,w}$ für $k = 0$. Der Term λ_{RLS} stellt einen Vergessensfaktor dar. Für Werte von $\lambda_{RLS} < 1$ werden Daten, die weiter in der Vergangenheit liegen, nicht mehr für die Berechnung der Parameterschätzung herangezogen. Die Begründung ist nach Nelson heuristisch: Die Schätzung der Zustände in einem Schätzverfahren sind zu Beginn stark fehlerhaft und sollte die Parameterschätzung zu späteren Zeitpunkten, bei vorliegenden verbesserten Schätzungen, nicht beeinflussen.

Eine weitere Möglichkeit besteht in der Schätzung der Kovarianz mittels der Robbins-Monro-Methode [RM51]. Bishop beschreibt in [BN06] den Einsatz der Robbins-Monro-Methode in der sequentiellen Schätzung, die von Wan und Nelson in [WN01] auf das Kalman-Filter erweitert wurde. Die von Wan und Nelson implementierte Variante ist durch

$$R_{v,w,k} = (1 - \alpha_{RM}) R_{v,w,k-1} + \alpha_{RM} \cdot \mathcal{K}_k (y_k - \hat{y}_k^-) (y_k - \hat{y}_k^-)^T \mathcal{K}_k^T \quad (3-60)$$

gegeben. Der Parameter α_{RM} ist ein positiver Wert < 1 und wird als Entwurfsparmeter angesehen [MHT09]. Im Wesentlichen ist die Robbins-Monro-Methode ein Werkzeug, um bei unbekannten oder stark gestörten Kovarianz-Matrizen eine Konvergenz zu einem “exakten” Wert zu ermöglichen.

Die Wahl der Methode zur Festlegung der Kovarianz zur Parameterschätzung hängt dabei von der Problemstellung ab. Nach [VDM04] bietet die Robbins-Monro-Methode die schnellste Konvergenzrate und die kleinsten Fehler, was durch die Anwendungen in dieser Arbeit bestätigt werden konnte und daher in den späteren Anwendungen eingesetzt wird.

3.6 Sigma-Punkte-Filter unter Nebenbedingungen

Bei der Nutzung von Sigma-Punkte-Filtern zur Schätzung von Zuständen und Parametern technischer Systeme sind oftmals Nebenbedingungen zu berücksichtigen, wodurch beispielsweise Wertebereiche eingeschränkt werden können. Insbesondere bei der Schätzung von Parametern kann es zur Berechnung von physikalisch nicht möglichen Sigma-Punkten kommen, etwa negativen Massen oder Federsteifigkeiten. Da diese Sigma-Punkte im Filterbetrieb als Parameter in das Modell gesetzt werden, kann es zu instabilem Modellverhalten kommen und damit zu Filterinstabilitäten.

Zur Verdeutlichung soll hier das Beispiel einer eindimensionalen Parameterschätzung der Masse eines technischen Systems gezeigt werden. Sofort ersichtlich ist, dass weder in der A-priori- noch in der A-posteriori-Schätzung Werte kleiner Null geschätzt werden dürfen. Neben der Beschränkung der A-priori- und A-posteriori-Schätzung sind auch die zu berechnenden Sigma-Punkte in die Betrachtung mit einzubeziehen. Einen wichtigen Beitrag dieser Dissertation stellen die Entwicklung und Einbeziehung von Nebenbedingungen zur Beschränkung von Parameterschätzungen bei unbekannten Systemmodellen unter Echtzeitbedingungen dar.

In Bild 3-17 sind die möglichen Fehlerfälle dargestellt, die im Algorithmus erkannt und angepasst werden müssen. Das Bild stellt eine eindimensionale Massenschätzung für einen symmetrischen Sigma-Punkte-Ansatz bei additivem Rauschen dar (vgl. (3-33)), wobei die grau hinterlegten Flächen die Schätzung beschränken, um keine negativen oder zu großen Massen zuzulassen. Zum Zeitpunkt 0 liegt die A-posteriori-Schätzung der Masse (blauer Kreis) mit zugehöriger Fehlerkovarianz (blaue Fehlerbalken) vor. Im darauf folgenden Zeitschritt 1 werden im ersten Schritt 1a drei Sigma-Punkte zur Erfassung von Erwartungswert und Kovarianz berechnet (rote Kreuze und roter Kreis), die durch das Modell von 1a zu 1b transformiert werden und die A-priori-Vorhersage der Masse mit Fehlerkovarianz zum Zeitpunkt 1b ermöglichen. Zur Einbindung des Prozessrauschens wird die berechnete A-priori-Fehlerkovarianz $P_{xx,k}^-$ additiv mit dem Prozessrauschen überlagert und neue Sigma-Punkte berechnet, die von 1c zu 1d transformiert werden. Nach dem Auftreten neuer Messungen zum Zeitschritt $k = 1d$ wird die Vorhersage der

Masse und deren Fehlerkovarianz angepasst. Diese A-posteriori-Schätzung steht daraufhin für den nächsten Zeitschritt $k = 2$ zur Verfügung.

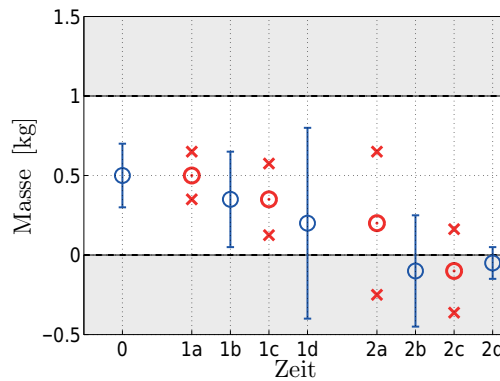


Bild 3-17: Fehlerfälle bei der Anwendung von Sigma-Punkte-Filtern am Beispiel einer eindimensionalen Massenschätzung. Fehlerhafter Sigma-Punkt $k = 2a$ und $k = 2c$. Fehlerhafte A-priori-Schätzung $k = 2b$. Fehlerhafte A-posteriori-Schätzung $k = 2d$.

Nach Beendigung des Zeitschrittes $k = 1$ werden die neuen Sigma-Punkte berechnet, transformiert und die A-priori-Schätzung zum Zeitpunkt $k = 2b$ mit zugehöriger Fehlerkovarianz durchgeführt. Sollte in $k = 1d$ die Fehlerkovarianz groß geworden sein, kommt es in $k = 2a$ zur Berechnung eines ungültigen Sigma-Punktes mit einer negativen Masse, die zu korrigieren ist, was einen Fehlerfall darstellt. Bei der nachfolgenden A-priori-Vorhersage der Masse und auch bei der späteren A-posteriori-Korrektur kann es vorkommen, dass die eigentliche Schätzung die gestellten Nebenbedingungen verletzt und somit zu korrigieren ist. Hierbei muss beachtet werden, dass die Massenschätzung nicht auf der Grenze der Nebenbedingung liegen darf, da ansonsten im nächsten Filterschritt Erwartungswert und Sigma-Punkte zusammenfallen, was zur Filterinstabilität führen kann. Hieraus folgt, dass zur Laufzeit des Filters sowohl die berechneten Sigma-Punkte als auch die A-priori- und die A-posteriori-Schätzung hinsichtlich vorhandener Nebenbedingungen zu kontrollieren und gegebenenfalls zu korrigieren sind. Nach Abschnitt 3.5 existiert bei einer Parameterschätzung die A-priori Schätzung zu den Zeitpunkten $k = 1b$ und $k = 2b$ nicht, da keine Transformation der zu schätzenden Parameter existiert. In diesem Beispiel wurde allerdings eine A-priori Schätzung angegeben, da die entwickelten Verfahren äquivalent auf die zu schätzenden Zustände angewandt wird, um beispielsweise beim Anschlag einer Masse bestimmte Wertebereiche auszublenden.

Die beiden entwickelten Verfahren zur Beschränkung der Sigma-Punkte und Korrektur der Parameterschätzung werden nachfolgend für den Fall eines Joint-Unscented-Kalman-Filters mit symmetrischem Sigma-Punkte Satz vorgestellt und lassen sich nach dem gleichen Vorgehen auf den asymmetrischen Sigma-Punkte Satz und weitere Sigma-Punkte-Filter wie das CDKF erweitern. Hierzu wird die

Menge \mathcal{M} definiert, die durch

$$\mathcal{M} = \{\mathcal{X}_{i,n_p,k-1}, \mathcal{X}_{i,n_p,k|k-1}, \hat{x}_{n_p,k}^-, \hat{x}_{n_p,k}\} \quad (3-61)$$

gegeben ist, mit n_p als der Anzahl der zu schätzenden Parameter und $i = 1, \dots, 2(n+n_p)+1$ als i -tem Sigma-Punkt. Im nächsten Schritt wird geprüft, ob Elemente m der Klasse \mathcal{M} die obere $NB_{P,o} \in \mathbb{R}^{n_p \times 1}$ bzw. die untere $NB_{P,u} \in \mathbb{R}^{n_p \times 1}$ verletzen⁵.

$$NB_{P,o} \geq m \geq NB_{P,u} \quad \forall \quad m \in \mathcal{M} \quad (3-62)$$

Sind in diesem Schritt Überschreitungen der A-priori- oder A-posteriori-Schätzungen erkannt worden, werden diese auf die Grenzen zurückgesetzt, um eine gültige Schätzung zu erlangen. Sollten ungültige Sigma-Punkte erkannt worden sein, erfolgt im ersten Schritt eine Neuordnung der Sigma-Punkte-Matrix \mathcal{X}_{k-1} zu $\mathcal{X}_{k-1,skal}$ in nicht zu korrigierende $\mathcal{X}_{k-1,\tau}$ und zu korrigierende Sigma-Punkte $\mathcal{X}_{k-1,\varsigma}$. Dieser Schritt ist äquivalent für den Fall einer Nebenbedingungsverletzung von $\mathcal{X}_{i,n_p,k|k-1}$ durchzuführen. Die Neusortierung ist zulässig, da im Folgenden die Sigma-Punkte einzeln transformiert werden und nur die Zuordnung der Gewichtungen zum Erwartungswert erhalten bleiben muss. Aus der Neuordnung ergibt sich

$$\mathcal{X}_{k-1,skal} = [\mathcal{X}_{0,k-1} \quad \mathcal{X}_{i,k-1,\tau} \quad \mathcal{X}_{j,k-1,\varsigma}], \quad (3-63)$$

wobei gilt: $i = 1, \dots, 2 \cdot (n+n_p)+1 - 2 \cdot n_z$ und $j = 1, \dots, 2 \cdot n_z$ mit n_z als Anzahl der zu korrigierenden Sigma-Punkte. Im nächsten Schritt erfolgt die Ermittlung des Sigma-Punktes, der die stärkste Verletzung der Nebenbedingungen bildet und im Folgenden als $\hat{\mathcal{X}}$ bezeichnet wird, mit⁶

$$\hat{\mathcal{X}} = \max(|NB_{P,u,o} - \mathcal{X}_{i,n_p,k-1,\varsigma}|) \quad \forall \quad i = 1, \dots, n_z. \quad (3-64)$$

Anhand der ermittelten maximalen Abweichung $\hat{\mathcal{X}}$ der Sigma-Punkte zeigt Bild 3-18 zwei Abbildungsvarianten von korrigierten Sigma-Punkten für den vereinfachten Fall eines einzelnen Filterschrittes bei der Sigma-Punkte-Berechnung. Als rote Verläufe sind die Dichtefunktionen als Gaußverteilungen der Zufallsvariablen dargestellt.

Der linke Plot stellt eine Variante zur Abbildung der Sigma-Punkte dar, die ähnlich zu der Berechnung der Scaled-Unscented Transform arbeitet. Da in diesem Fall die Sigma-Punkte einer Dimension symmetrisch um den Erwartungswert verteilt sind, bewirkt eine gleichsinnige Verschiebung aller Sigma-Punkte zum Erwartungswert die Beibehaltung der korrekten Erfassung von Erwartungswert und

⁵Da n_P auch mehrdimensional sein kann, erfolgt die Überprüfung der Nebenbedingungen komponentenweise.

⁶Bei einer Dimension von $n_P > 1$ muss diese Überprüfung pro Dimension von n_P durchgeführt und somit der Sigma-Punkt mit dem Parameter der maximalen Abweichung erfasst werden.

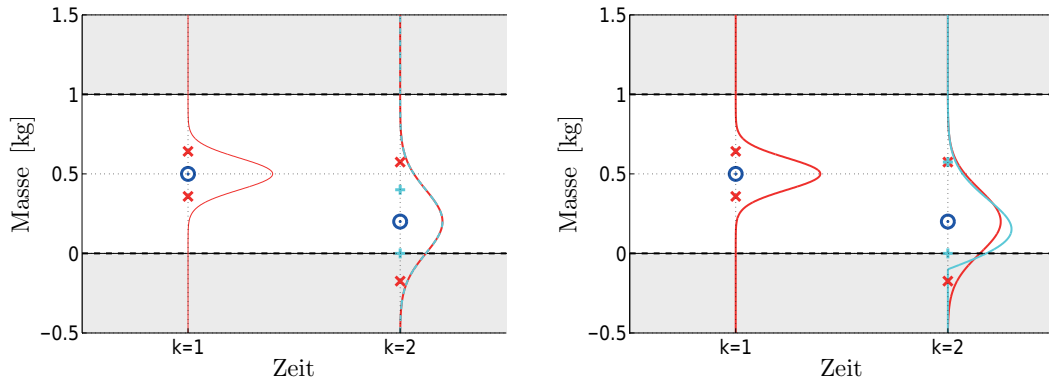


Bild 3-18: Links: Beibehaltung von symmetrischer Dichtefunktion bei symmetrischer Verschiebung von ungültigen Sigma-Punkten. Rechts: Schiefe Anteile bei asymmetrischer Verschiebung von ungültigen Sigma-Punkten

Kovarianz, wie im linken Plot von Bild 3-18 als roter und cyanfarbener Verlauf zu erkennen. Die Aufgabe liegt in diesem Fall in der Auffindung eines Skalierungsparameters γ' , der die ungültigen Sigma-Punkte in den gültigen Bereich abbildet und nachfolgend als *Skalierung symmetrischer Sigma-Punkte* bezeichnet wird.

Die Berechnung des nullten Sigma-Punktes und der nicht zu korrigierenden Sigma-Punkte erfolgt dabei mit der bereits aus Abschnitt 3.2.1 bekannten Methode:

$$\begin{aligned} \mathcal{X}_{0,k-1} &= \hat{x}_{k-1} \\ \mathcal{X}_{i,k-1,\tau} &= \hat{x}_{k-1} \pm \gamma \left(\sqrt{P_{xx,k-1}} \right)_i \quad i = 1, \dots, 2 \cdot (n + n_P) + 1 - 2n_z \end{aligned} \quad (3-65)$$

Die zu korrigierenden Sigma-Punkte $\mathcal{X}_{k-1,\varsigma}$ werden durch

$$\mathcal{X}_{j,k-1,\varsigma} = \hat{x}_{k-1} \pm \gamma' \left(\sqrt{P_{xx,k-1}} \right)_j \quad j = 1, \dots, 2n_z. \quad (3-66)$$

berechnet. Um eine symmetrische Verteilung beizubehalten, skaliert man allerdings nicht nur die Sigma-Punkte, welche die Nebenbedingungen verletzen, sondern ebenfalls die auf dem Eigenvektor der Matrixwurzel $\sqrt{P_{xx,k-1}}$ gegenüber liegenden Sigma-Punkte, was im linken Teil von Bild 3-18 zu erkennen ist.

Der Skalierungsparameter γ' für die zu korrigierenden Sigma-Punkte ist dabei durch den Sigma-Punkt $\hat{\mathcal{X}}$ gegeben, der am weitesten von der zu erfüllenden oberen oder unteren Nebenbedingung entfernt ist und sich mit

$$\gamma' = \left| \frac{NB_{P,u,o}}{\hat{\mathcal{X}}} \right| \quad (3-67)$$

berechnen lässt.

Um die Bedingungen, die in Abschnitt 3.2 formuliert wurden, nicht zu verletzen, müssen im Folgenden die Gewichtungen der Sigma-Punkte angepasst werden. Als Annahme wird hierbei festgesetzt, dass sich die Gewichtung des nullten

Sigma-Punktes nicht ändern soll, da dieser die Erfassung von höheren statistischen Momenten beeinflusst und unverändert bleiben soll. Zwingend erforderlich bleibt allerdings die Forderung nach der Beibehaltung der Kovarianz, wodurch die Gewichtungen der korrigierten Sigma-Punkte $\mathcal{X}_{j,k-1,\varsigma}$ angepasst werden müssen, im Folgenden als $W_{j,\varsigma}^{(m)'}$ und $W_{j,\varsigma}^{(c)'}$ bezeichnet werden und sich zu

$$W_{j,\varsigma}^{(m)'} = W_{j,\varsigma}^{(c)'} = \frac{1}{2 \cdot (n + \lambda')} \quad \forall \quad j = 1, \dots, 2 \cdot n_z \quad (3-68)$$

ergeben, wobei $\lambda' = (\alpha')^2 \cdot (n + \kappa) - n$ gilt. Der neue Wert für den Skalierungsparameter α' ergibt sich zu

$$\begin{aligned} \gamma' &= \sqrt{n + \lambda'} = \sqrt{n + (\alpha')^2 (n + \kappa) - n} = \alpha' \sqrt{n + \kappa} \\ \Rightarrow \alpha' &= \frac{\gamma'}{\sqrt{n + \kappa}}. \end{aligned} \quad (3-69)$$

Wegen der Neuberechnung der Gewichtungen der korrigierten Sigma-Punkte müssen die Gewichtungen der unkorrigierten Sigma-Punkte $\mathcal{X}_{i,k-1,\tau}$ ebenfalls angepasst werden, um der Normalisierungsbedingung (3-3) zu genügen. Diese werden im Folgenden mit $W_{i,\tau}^{(m)'} = W_{i,\tau}^{(c)'}$ für $i = 1, \dots, 2(n + n_P) + 1 - 2n_z$ benannt und ergeben sich nach

$$W_0^{(m)} + \sum_{i=1}^{2(n+n_P)+1-2n_z} W_{i,\tau}^{(m)'} + \sum_{j=1}^{2n_z} W_{j,\varsigma}^{(m)'} = 1 \quad (3-70)$$

und unter Berücksichtigung der Symmetrieeigenschaften, dass die Gewichtungen aller Dimensionen innerhalb der korrigierten bzw. der unkorrigierten Sigma-Punkte der Verteilung gleich sind, zu

$$W_0^{(m)} + (2(n + n_P) + 1 - 2n_z) \cdot W_{i,\tau}^{(m)'} + 2n_z \cdot W_{j,\varsigma}^{(m)'} = 1. \quad (3-71)$$

Hieraus ergeben sich die letztlich noch unbekannten Gewichtungen der nicht korrigierten Sigma-Punkte $W_{i,\tau}^{(m)'}$.

Das hergeleitete Verfahren zur Skalierung der Sigma-Punkte geht davon aus, dass sich der Skalierungsfaktor γ' aller zu korrigierenden Sigma-Punkte aus der maximalen Verletzung der Nebenbedingungen ergibt. Die obige Herleitung kann ohne Weiteres auf die Variante erweitert werden, in der jeder Sigma-Punkt, der eine Nebenbedingung verletzt, einen eigenen Skalierungsfaktor und eigene Gewichtungen erhält. In der vorliegenden Dissertation wurde dieses Verfahren allerdings nicht angewandt, da es neben der höheren Rechenbelastung keine signifikanten Vorteile in der Analyse zeigte.

Das vorgestellte Verfahren kann bei einer häufigeren oder andauernden Berechnung von Sigma-Punkten in der Nähe von $NB_{P,u,o}$ (was durchaus des öfteren auftritt, wie im weiteren Verlauf der Arbeit gezeigt) allerdings zu einer Beschneidung

der Filterdynamik führen, da die Sigma-Punkte sehr eng an den Erwartungswert transformiert werden, was zu einer zu hohen Gewichtung von lokalen Effekten führt. Um diesen Effekt zu vermeiden, hat man das Verfahren zur Erzeugung gültiger Sigma-Punkte dahingehend abgewandelt, dass nur der Sigma-Punkt, der die Nebenbedingungen verletzt, verschoben wird und der ihm „gegenüberliegende“ unverändert bleibt. Dieses Verfahren führt auf den rechten Teil von Bild 3-18. Bei den hergeleiteten Rechnungen müssen zur Erzeugung dieses Verhaltens nur die Dimensionen der zu korrigierenden Sigma-Punkte $\mathcal{X}_{j,k-1,\varsigma}$ von $2 \cdot n_z$ auf n_z reduziert werden. Somit können im Falle eines nahe an den Grenzen liegenden Erwartungswertes zumindest einige Sigma-Punkte nicht lokale Effekte erfassen und diese in die Transformation einfließen lassen. Als Nachteil aus dieser Methode ergibt sich allerdings ein asymmetrischer Satz an Sigma-Punkten, der die Fehlerkovarianz staucht, wie ebenfalls im rechten Teil von Bild 3-18 ersichtlich. Etwaige Fehler durch die Schiefeanteile der nicht mehr korrekt erfassten Gaußverteilung müssen dabei in Kauf genommen werden, was in der vorliegenden Arbeit allerdings als zielführender angesehen wird, da es nicht zu einer alleinigen Abbildung von Sigma-Punkten im Erwartungswert auf der Grenze kommen kann. Das Verfahren könnte hier noch um Aspekte der Minimierung von Schiefeanteilen erweitert werden, was ähnlich zum Vorgehen von Julier durchzuführen wäre und in Abschnitt 3.2.3 kurz angesprochen wurde.

Für die Bewertung beider Verfahren kann festgehalten werden, dass bei der Analyse die symmetrische Skalierung bei Schätzungen nahe an den Grenzen keine signifikant schlechteren Ergebnisse erzielte als die asymmetrische Sigma-Punkte-Skalierung und daher in dieser Arbeit weiter angewandt wird. Dies liegt daran, dass für Parameterschätzungen aufgrund der besseren Ergebnisse das Joint-Kalman-Filter verwendet wurde, bei dem der Parameter in jedem Sigma-Punkt eingeht und somit bei einer geringen Anzahl an zu korrigierenden Sigma-Punkten noch genügend weitere für die Aufrechterhaltung einer korrekten Schätzung zur Verfügung stehen.

3.7 Zusammenfassung der Sigma-Punkte-Filter

In diesem Kapitel wurden die theoretischen Grundlagen des UKF vorgestellt, das eine Erweiterung des klassischen Kalman-Filters darstellt. Durch die Implementierung der Unscented Transform in den Algorithmus eines Kalman-Filters ist das UKF in der Lage, eine höhere Schätzgüte bei nichtlinearen Systemen gegenüber dem EKF zu erzielen. Durch die Funktionsweise der Unscented Transform, die statistischen Momente einer Zufallsvariablen mittels einer geringen Anzahl an deterministisch berechneten Sigma-Punkten zu erfassen und diese durch das nichtlineare System zu transformieren, besteht ein wesentlicher Neuerungsbeitrag dieser Arbeit in der Anwendung auf Black-Box-Modelle, die nicht in analytischer Form vorliegen und stark nichtlinear sein können.

Gemäß dieser Idee wurden in diesem Kapitel Sigma-Punkte-Berechnungsverfahren hergeleitet, die den Anforderungen hinsichtlich Echtzeitfähigkeit und Schätzungsgüte nachkommen, und ein kurzer Überblick über weitere Verfahren gegeben. Als weiterer Vertreter der Sigma-Punkte-Filter wurde das CDKF vorgestellt und gezeigt, dass sich die unterschiedliche Berechnungsmethode der Sigma-Punkte auf das UKF zurückführen lässt. Dabei zeichnet es sich durch seine geringere Anzahl an Entwurfsparametern aus; eventuell vorhandene höhere statistische Informationen können allerdings nicht berücksichtigt werden, wodurch es äquivalent zur Scaled-Unscented Transform unter Vernachlässigung der Parameter β und κ ist. Als mögliche Variante zur Schätzung von Parametern wurde das *Joint*-Verfahren ausgewählt, vorgestellt und die Vorteile hinsichtlich der *Dual*-Schätzung aufgeführt sowie eine Möglichkeit zur automatischen Berechnung der Prozessrauschkovarianz bei der Parameterschätzung ausgewählt.

Nach der Grundidee der Anwendung von Sigma-Punkte-Filter auf nichtanalytisch vorliegende Modelle in Black-Box-Form wurde als weitere Neuerung dieser Arbeit der Algorithmus von Sigma-Punkte-Filter abgeändert, um mit den veränderten Bedingungen beim Einsatz an Black-Box-Systemen umgehen zu können, was beispielhaft am Algorithmus eines UKF gezeigt wurde.

Insbesondere bei der Schätzung von Parametern kann der Einsatz von Sigma-Punkte-Filtern zur Berechnung von ungültigen Sigma-Punkten führen. Werden diese ungültigen Sigma-Punkte in das als Black-Box vorliegende Modell eingesetzt, kann es zu unvorhersehbarem Filterverhalten kommen, etwa nicht ausführbaren Simulationen oder instabilem Modellverhalten. Eine wesentliche Neuerung dieser Arbeit besteht in der Entwicklung eines echtzeitfähigen Algorithmus zur Erkennung von ungültigen Sigma-Punkten und deren Abbildung in einen zulässigen Bereich, was als Filterung unter Nebenbedingungen bezeichnet wird. Ebenfalls werden mögliche falsche Schätzungen von Parametern erkannt und korrigiert, wodurch der Einsatz von Black-Box-Modellen zur Parameterschätzung möglich wird.

Im folgenden Kapitel wird eine Möglichkeit vorgestellt, die Entwurfsparameter des UKF und des CDF modellbasiert in einem Pre-processing-Schritt anhand von Anwendungsszenarien und Streckenmodellen auszulegen. In Kapitel 6 wird gezeigt, dass für ein komplexes Problem die Möglichkeit der Einbindung von statistischen Informationen über eine höhere Anzahl an Entwurfsparametern zu einer höheren Schätzungsgüte führt.

4 Wahl der Filter- und Entwurfparameter

Im letzten Kapitel wurden Sigma-Punkte-Kalman-Filter zur Zustands- und Parameterschätzung vorgestellt und gezeigt, wie diese für die Nutzung an Black-Box-Modellen in dieser Arbeit angepasst wurden. Da eine wichtige Anforderung in der Anwendung unter Echtzeitbedingungen besteht, wurden zwei Verfahren der Unscented Transform zur Sigma-Punkte-Berechnung vorgestellt, die in dieser Arbeit genutzt werden. Bei der Scaled-Unscented Transformation ist eine höhere Schätzgüte möglich, bei $2n + 1$ zu berechnenden und zu transformierenden Sigma-Punkten, während das Spherical-Simplex-Verfahren mit nur $n + 2$ Sigma-Punkten auskommt. Bei der Nutzung beider Verfahren sind allerdings jeweils drei Entwurfparameter zu bestimmen: $\theta_{SUT} = [\alpha, \beta, \kappa]$ und $\theta_{SST} = [W_0, \alpha, \beta]$, welche die Lage der Sigma-Punkte und deren Gewichtungen bestimmen und somit die Erfassung höherer statistischer Momente beeinflussen.

In [TR12b] wurde von Turner der Fall des *Sigma-point collapse* gezeigt. Hierbei wurde ein Sigma-Punkte-Satz in Abhängigkeit der Entwurfparameter θ_{SUT} gewählt, der die Dichtefunktion der Verteilung so ungünstig erfasst, dass ein Einsatz zur Zustands- und Parameterschätzung zu unvorhersehbaren Ergebnissen führen kann. In den Abschnitten 3.2.2 und 3.2.1 wurden optimale Werte für die Entwurfparameter β und κ bei einer vorliegenden Gaußverteilung hergeleitet. Wirkliche Gaußverteilungen werden allerdings nur auftreten, wenn das System, an dem die Zustandsschätzung durchzuführen ist, linear ist, was nicht das Anwendungsgebiet von Sigma-Punkte-Filtern ist und in dieser Arbeit wegen der Nutzung von Modellbibliotheken während der Modellierungsphase des als Black-Box vorliegenden Systems fast ausgeschlossen ist. Somit stellt die Annahme einer Gaußverteilung bei der Auslegung der Entwurfparameter bestenfalls den im Mittel geringsten Fehler dar, wenn keine Informationen über höhere statistische Momente zur Verfügung stehen.

Die Grundlage dieses Kapitels besteht in der Annahme, dass ein optimaler¹ Satz an Entwurfparametern für das jeweilige System und deren Anwendung existiert. Diese These wird durch zahlreiche empirische Empfehlungen in der Literatur für den Parametersatz $\theta_{SUT,SST}$ gestützt, die durch ihre spezielle Anwendung ein problemspezifisches Optimum darstellen. Van der Merwe vertritt in [VDMW⁺04] ebenfalls diese These und bestätigt die Vermutung durch unterschiedlich gewählte Parametersätze an seinen aufgeführten Beispielen. Gleichzeitig zeigt er auf, dass die Wahl der Entwurfparameter $\theta_{SUT,SST}$ wohl keinen starken Einfluss auf die Ergebnisse hat, solange diese zu einem Sigma-Punkte-Satz mit entsprechenden Gewichtungen führt, der sich numerisch *gut* verhält. Was dies bedeutet, wird allerdings offen gelassen.

¹Optimal im Sinne von: Es kann kein Parametersatz gefunden werden, der im Hinblick auf Schätzgüte, Konvergenzzeit und Stabilität bessere Ergebnisse liefert.

Das Problem des Findens eines optimalen Entwurfsparametersatzes könnte durch die Nutzung des vorgestellten Central-Difference- oder Cubature-Kalman-Filters vereinfacht werden, das mit nur einem Entwurfsparameter $\theta_{CDKF} = [\xi]$ bzw. $\theta_{CKF} = [\alpha]$ auskommt. Die Nutzung des CDKF bzw. CKF ist sicher anzuraten, wenn keine Informationen über die Momente einer Verteilung vorliegen. In dieser Arbeit liegen allerdings implizit Informationen über die zu erfassenden Momente in Form von physikalischen Verhaltensmodellen und deren Anregung vor, sodass A-priori-Informationen für die Filter angebbbar sind. Die Aufgabe besteht demnach in der Angabe optimaler, problemspezifischer Filterentwurfsparameter, wenn der Ingenieur kein Experte in dem zugrundeliegenden Filtermodell oder in den Algorithmen des zu nutzenden Filters ist.

Das Ziel dieses Kapitels besteht demnach in der Auffindung dieses problemspezifischen Entwurfsparametersatzes unter den Randbedingungen, dass das Modell nur als Black-Box vorliegt. Dadurch sind etwaige Nichtlinearitäten unbekannt, die maßgeblichen Einfluss auf die statistischen Momente besitzen und die es durch $\theta_{SUT,SST}$ zu erfassen gilt. Die Annahme zur Auffindung eines optimalen problemspezifischen Satzes an Entwurfsparametern besteht darin, dass Informationen über die statistischen Momente implizit im als Black-Box vorhandenen Modell enthalten sind. Eine optimale Erfassung der statistischen Momente mittels θ liegt demnach vor, wenn die Parameter θ eines Sigma-Punkte-Filters so angepasst wurden, dass bei der Anwendung zur Zustands- und Parameterschätzung keine höhere Schätzgüte durch einen anderen Satz an Entwurfsparametern θ erreicht werden konnte.

4.1 Optimaler Parametersatz zur Filterauslegung

Zur Erlangung eines optimalen Parametersatzes wird im Filterentwurfsverfahren ein Optimierungsproblem gelöst. Optimierungsparameter stellen die Entwurfsparameter des Filterverfahrens dar. Die Kostenfunktion wird durch die Schätzgüte bei der Zustands- und Parameterschätzung definiert. Hierzu wird nachfolgend ein Szenario definiert und die Kostenfunktion hergeleitet. Im Anschluss werden das genutzte Optimierungsverfahren und die darin enthaltenen Grundlagen und Parameter vorgestellt. Abschließend wird die Optimierungsmethodik erläutert.

4.1.1 Szenariodefinition

Die Definition des Szenarios [Mün12] stellt den Kern der Optimierung zur Auffindung eines optimalen Parametersatzes θ dar, wie in Bild 4-1 dargestellt.

Die in das Szenario eingehenden Parameter stellen die Optimierungsparameter dar, die hier beispielhaft mit θ zur optimalen Vorgabe der Entwurfsparameter bei Einsatz einer Scaled-Unscented Transformation angegeben sind. Ausgang des Szenarios ist ein Kostenfunktionswert, der die Güte des aktuellen Szenarios unter der Parametrierung θ angibt.

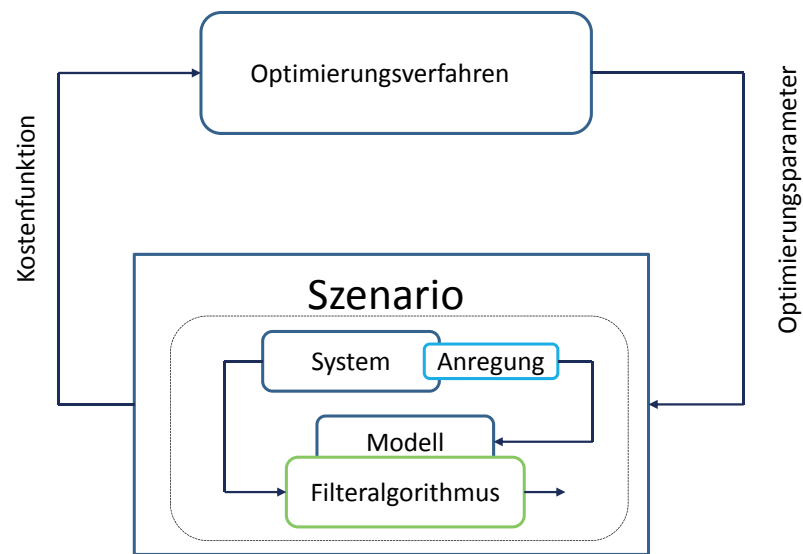


Bild 4-1: Zusammenspiel zwischen Szenario und Optimierungsverfahren

Das Szenario besteht vor dem Hintergrund dieser Arbeit aus dem Einsatz eines Sigma-Punkte-Filters zur Zustands- und Parameterschätzung an einem vorliegenden Modell. Eingangsgrößen des Modells müssen über Anregungsfälle gegeben sein und sind ebenfalls im Szenario eingeschlossen. Hierbei ist es wichtig, das System über die Anregungsfälle ausreichend anzuregen, da sonst etwaige im Modell vorhandene Nichtlinearitäten nicht angeregt werden. Das Filter würde, anders ausgedrückt, um den falschen Arbeitspunkt ausgelegt werden. Notwendige Messdaten müssen ebenfalls im Szenario enthalten sein und können aus realen Versuchen auf Basis der Anregungsfälle generiert werden, falls es sich beispielsweise um eine Weiterentwicklung handelt oder bereits Prototypen zur Verfügung stehen. Zwingend erforderlich für diesen Schritt ist allerdings eine Validierung des im Filter eingesetzten Streckenmodells mit einer hinreichenden Genauigkeit. Erschwerend bei dieser Art der Parameterbestimmung ist, dass das Streckenmodell immer Abweichungen zu dem realen System aufweist. Falls kein reales Experiment durchgeführt werden kann, besteht eine zweite Möglichkeit in der Generierung von Messdaten anhand eines Modells, das gegenüber dem Filtermodell verändert wurde. Die Definition des in dieser Arbeit eingesetzten Szenarios zur Ermittlung von optimalen Entwurfsparametern θ ist in Bild 4-2 dargestellt.

4.2 Ermittlung von Prozess- und Messrauschkovarianzen

Zur vollständigen Definition des Szenarios müssen bis auf die zu optimierenden Filterentwurfsparameter alle Parameter festgelegt sein. In den Kapiteln 2 und 3 wurden bereits die Prozess- und Messrauschkovarianzmatrizen eingeführt, mittels derer ein Vertrauensmaß (siehe auch Abschnitt 3.3) hinsichtlich Messung oder Modell gegeben ist. Da es sich bei den Sigma-Punkte-Filtern um die Einbettung einer

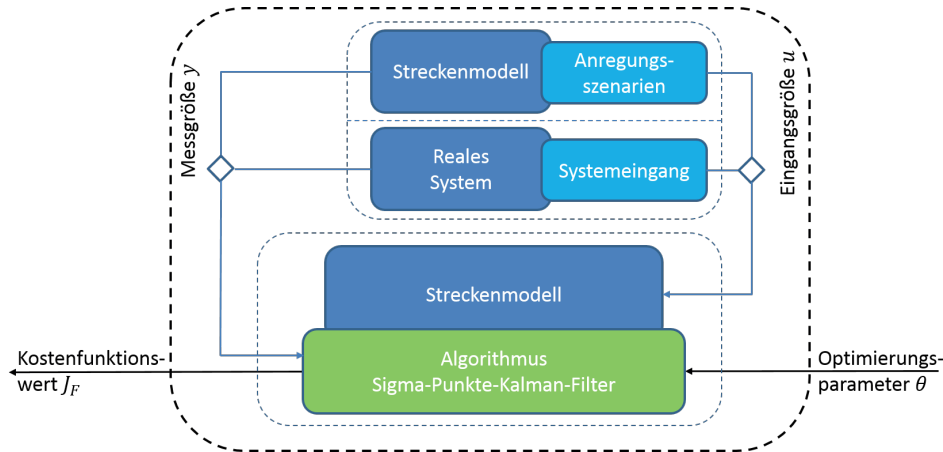


Bild 4-2: Szenariodefinition des Offline-Filtereinsatzes zur Optimierung der Entwurfsparameter eines UKF

Methode zur Transformation der statistischen Momente des Zustandsvektors in den Algorithmus eines Kalman-Filters handelt, erfolgt die Nutzung der Prozess- und Messrauschkovarianzen äquivalent. Bei der Bildung der Sigma-Punkte im Falle von additiven Rauschprozessen wird die Notwendigkeit einer geschickten Vorgabe des Prozessrauschens ersichtlich (vgl. (3-44), (3-45)), deren Berechnung nachfolgend nochmals dargestellt ist:

$$P_{xx,k}^- = \sum_{i=0}^{2n} W_i^{(c)} [\mathcal{X}_{i,k|k-1}^* - \hat{x}_k^-] [\mathcal{X}_{i,k|k-1}^* - \hat{x}_k^-]^T + R_v \quad (4-1)$$

$$\mathcal{X}_{i,k|k-1} = \begin{bmatrix} \hat{x}_k^- & \hat{x}_{k-1} + \gamma \left(\sqrt{P_{xx,k}^-} \right)_i & \hat{x}_k^- - \gamma \left(\sqrt{P_{xx,k}^-} \right)_i \end{bmatrix}.$$

Zur Bestimmung der A-priori Kovarianz wird die Prozessrauschkovarianz addiert, die geeignet vorzugeben ist. Verdeutlicht man sich allerdings, dass die Wahl der Prozessrauschkovarianz R_v bei Nutzung eines UKF mit additiven Rauschprozessen direkten Einfluss auf die Matrixwurzel $\sqrt{P_{xx,k-1}^-}$ hat, mit der die zur Simulation bestimmten Sigma-Punkte berechnet werden, so ist die Wahl von R_v eine wichtige Aufgabe. Insbesondere der Fall einer zu großen Wahl von R_v bewirkt in diesem Fall einen vergrößerten Radius der Sigma-Punkte um den Erwartungswert. Dies kann zur Erfassung von Modellnichtlinearitäten außerhalb des gewünschten Arbeitsbereiches führen oder zu nicht auswertbaren Sigma-Punkten. Als Beispiel kann der unterschiedliche Wertebereich zweier Zustände herangezogen werden. Variiert Zustand x_1 während einer gegebenen Anregung im Wertebereich von $[0, 1e-3]$ und x_2 im Wertebereich von $[0, 1e6]$, so wäre die Annahme der Einheitsmatrix für R_v nicht optimal. Die Hauptdiagonalelemente $R_{v,ii}$ mit $i = 1, \dots, n$ und n als Systemordnung spiegeln nach (4-1) und Kapitel 2 ein Vertrauensmaß in das Modell wider und erweitern die berechnete Fehlerkovarianz, also den Bereich, in dem sich der wahre Zustand befindet. Nach dieser Erklärung müssen sich

die Einträge der Prozessrauschmatrix R_v nach den Wertebereichen der Zustände richten.

Hinsichtlich der in dieser Arbeit genutzten Modelle und Szenarien ist die relative Änderung des Zustandes im Hinblick auf zu nutzende Anregungsszenarien entscheidend. Im linken Teil von Bild 4-3 ist ein qualitativer Verlauf des Zustandes x_j über einen definierten Zeitraum gezeigt, mit den diskreten Abtastungen x_j als roten Kreisen und $j = 1, \dots, l$ mit l als Anzahl der diskreten Abtastungen. Die in diesem entwickelten Verfahren getroffene Annahme besteht darin, dass es sich bei jeder Realisierung des Zustandes x_j um eine Gaußverteilung mit Erwartungswert und Varianz handelt. Der Erwartungswert ist direkt durch den als roten Kreis dargestellten Wert gegeben und die Varianz durch die Änderung des Zustandes zum nächsten Zeitschritt mittels

$$\text{var}[x_j] = \left| \frac{x_{j+1} - x_j}{\Delta t} \right| \quad (4-2)$$

angebar, mit Δt als Abtastschrittweite des Verlaufs, was im rechten Teil von Bild 4-3 als roter Vertrauensbereich dargestellt ist.

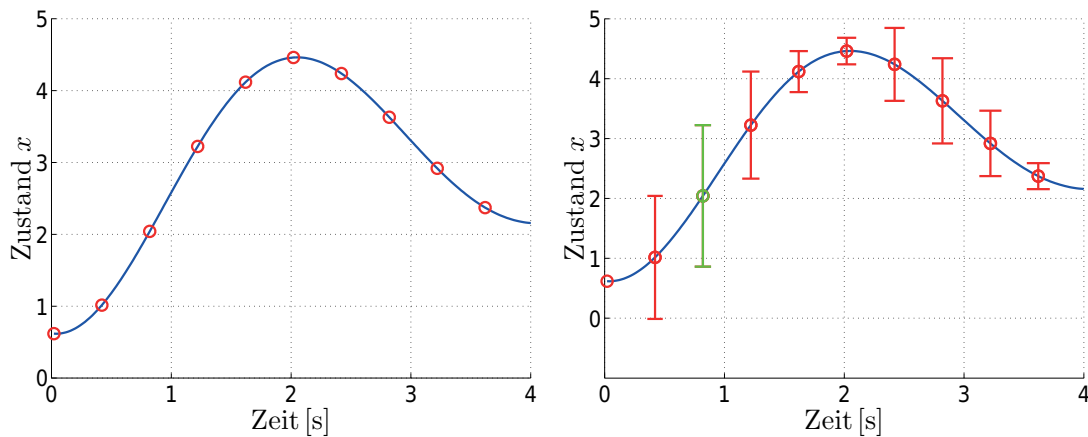


Bild 4-3: Automatisierte Ermittlung der Varianzen eines Zustandes als maximaler Differenzenquotient des zeitlichen Zustandsänderung

Das Prozessrauschen des i -ten Zustandes $R_{v0,ii}$ mit $i = 1, \dots, n$ und n der Systemordnung wird nach Auswertung aller Stützstellen als

$$R_{v0,ii} = \max \text{var}[x_i] \quad i = 1, \dots, n \quad (4-3)$$

angegeben und entspricht im rechten Teil von Bild 4-3 der grün markierten Varianz. Im Sinne der Zustandsschätzung bedeutet dies, dass, wenn die A-priori-Hauptdiagonalelemente der Fehlerkovarianz P_{xx}^- gegen Null gehen, wir demnach eine sehr gute Schätzung vorliegen haben und die restliche Unsicherheit durch $\max \text{var}[x_i]$ gegeben ist, was der maximalen zeitlichen Änderung der Zustände entspricht.

Dieses Vorgehen stellt eine weitere Neuerung dieser Arbeit dar und verringert in erheblichem Maße die Berechnung von ungültigen Sigma-Punkten, die ansonsten durch das in Abschnitt 3.6 entwickelte Verfahren zu korrigieren wären, und ist vollständig automatisiert durchführbar.

Um den vorgestellten Charakter eines Vertrauensmaßes in das Modell nicht zu verlieren, gehen wir nachfolgend von einer unkorrelierten Kovarianz aus, die mittels eines Faktors $0 < F_{R_v} < 1$ variiert wird und ein Maß für das Vertrauen in das Modell darstellt. Bei geringem Vertrauen in das Modell ermöglicht der Faktor durch einen großen Wert die Aufweitung der Kovarianz und somit die Erfassung von Sigma-Punkten, die weiter vom Erwartungswert entfernt liegen:

$$R_v = F_{R_v} \cdot R_{v0}. \quad (4-4)$$

Die Bestimmung der Messrauschkovarianz R_n erfolgt analog zu der vorgestellten Form anhand der Modellausgänge und führt auf die Basis des Messrauschens R_{n0} , das durch Multiplikation mit dem Faktor F_{R_n} das Messrauschen R_n ergibt.

4.3 Optimierung der Entwurfparameter

Der Kern des Optimierungsproblems ist durch das eingangs definierte Szenario gegeben, dessen Eingangsgrößen aus den Entwurfsparemtern θ des verwendeten Filter-Typs bestehen und dessen Ausgang eine die Filtergüte bewertende, noch zu definierende Kostenfunktion bildet. Um eine ausreichende Anregung des Systems zu gewährleisten, müssen allgemeingültige Anregungsfälle im Szenario definiert werden, die, wie in dieser Arbeit aufgetreten, mit teilweise langen Simulationszeiten einhergehen. Die Optimierungsparameter unterliegen dabei Beschränkungen, wie in Abschnitt 3.2.2 erläutert, und müssen in die Optimierungsstrategie aufgenommen werden.

Zur Lösung dieser Aufgabe kommen eine Vielzahl an Optimierungsmethoden in Betracht, etwa die Anwendung von genetischen Algorithmen, die mit einer geringen Anzahl an Entwurfsparemtern auskommen, Nebenbedingungen einbinden können und an einem als Black-Box vorliegenden System ableitungsfrei arbeiten. Allerdings erkaufte man sich dieses grundsätzlich für diesen Fall nutzbare Verhalten mit einer sehr hohen Anzahl an Funktionsaufrufen, was auch für evolutionäre Algorithmen und Particle-Swarm-Optimierungen gilt. Pattern-Search-Algorithmen sind ebenfalls in der Lage, an einem als Black-Box vorliegenden System ableitungsfrei nach Optima zu suchen, benötigen dafür allerdings auch eine hohe Anzahl an Funktionsaufrufen. In [TR12b] zeigte Turner, dass die Anwendung von gradientenbasierten Optimierungsverfahren keine zufriedenstellenden Lösungen liefert.

In Anlehnung an [Tur11] wird in dieser Arbeit die Gaußprozessoptimierung genutzt. Diese ist in der Lage, eine Optimierung an einem als Black-Box vorliegenden System ableitungsfrei durchzuführen und gleichzeitig nur eine geringe Anzahl

an Funktionsaufrufen zu benötigen, mit einer hohen Chance auf das globale Optimum. Hierzu wird nachfolgend der Gaußprozess, seine Funktionsweise und die enthaltenen Entwurfparameter erläutert.

Die Gaußprozessregression wurde durch ein von Daniel Krige im Jahr 1951 entwickeltes Verfahren in der Geologie bekannt, bei dem Krige den Ort mit der höchsten Wahrscheinlichkeit auf einen Goldfund vorhersagen wollte, wobei nur vereinzelte Bohrungen im Umkreis zur Verfügung standen. Im Gegensatz zu bisherigen Verfahren, in welchen die Messpunkte antiproportional zu ihrer Entfernung vom Auswertepunkt eingehen, nutzte Krige die Annahme einer Gaußverteilung der Messungen, wodurch jede Linearkombination dieser normalverteilten Messungen eine mehrdimensionale Gaußverteilung erzeugt [Kri51]. Somit ist die Vorhersage von unbekannten Funktionswerten² auf Basis von Stützstellen im Sinne einer Minimum-Varianz Schätzung möglich.

Dieses Verfahren der Vorhersage von Funktionswerten auf Basis von bekannten Stützstellen wird im maschinellen Lernen als Gaußprozessregression bezeichnet [RW96]. Hierdurch wird auch der Begriff „Prozess“ erklärt, da es oftmals das Ziel ist, auf Basis von Messungen eines Prozesses die nächsten Messungen vorherzusagen.

Im Sinne dieser Arbeit wird der Gaußprozess nicht zur Vorhersage eines zeitlichen Verlaufes genutzt, sondern zur Vorhersage des maximalen Wertes einer Kostenfunktion in Abhängigkeit der Filterentwurfparameter θ bei einer Szenarioauswertung. Im Folgenden wird dieser Zusammenhang zur Vereinfachung als $J_F = g(\theta)$ bezeichnet, wobei $g(\theta)$ das Szenario unter den Eingangsgrößen θ bildet und J_F den Ausgang des Szenarios als Kostenfunktionswert zur Beschreibung der Filtergüte³. Der Gaußprozess wirkt dabei als statistisches Modell des Szenarios und ist in der Lage, das Ein-/ Ausgangsverhalten des realen Szenarios abzubilden. Hierzu werden Stützstellen $\theta_1, \dots, \theta_N$ mit $\theta \in \mathbb{R}^{n_x \times 1}$ mit n_x als Dimension der Eingangsdaten am Szenario simuliert und die Ausgänge $J_{F,1}, \dots, J_{F,N}$ gebildet. Anhand dieser Stützstellen wird der Gaußprozess trainiert, wobei gilt: Je mehr Stützstellen, desto besser bildet der Gaußprozess die Funktion nach. Im Folgenden können Vorhersagen über das Szenarioverhalten getroffen, also auch der Wert θ vorhergesagt werden, der zu einem maximalen Ausgang J_F führt, was einer maximalen Filtergüte in Abhängigkeit der Entwurfparameter entspricht und die Anzahl der Auswertungen am Szenario reduziert. Bei der Anwendung des Gaußprozesses in der vorliegenden Arbeit handelt es sich demnach nicht um eine reine statistische Aufgabenstellung, vielmehr wird zur Lösung des Optimierungsproblems der Gaußprozess als Ersatzmodell verwendet.

Zur Erreichung dieses Ziels wird im folgenden Abschnitt der Gaußprozess formal eingeführt. Es wird gezeigt, wie anhand von grundlegender Statistik die Vorher-

²Für Krige war die unbekannte Funktion der Zusammenhang zwischen Bodenparametern und Goldgehalt.

³Die Herleitung der Größen J_F und $g(\theta)$ im Sinne der Vorgabe optimaler Parameter für ein UKF erfolgt im späteren Verlauf in Abschnitt 4.4.

sage von Funktionswerten möglich ist und wie Unsicherheiten der Messungen in den Gaußprozess eingebracht werden können. In Abschnitt 4.3.2 und 4.3.3 wird der Kern des Gaußprozesses, bestehend aus Kernel-Funktionen, sowie der Einfluss der freien Parameter des Kernels auf den Gaußprozess gezeigt und eine Methode zur Vorgabe der Kernel-Parameter zur optimalen Vorhersage von Funktionswerten vorgestellt. In Abschnitt 4.3.4 wird gezeigt, wie innerhalb des Gaußprozesses mit einfachen Mitteln ein Maximum gefunden werden kann, und somit die Brücke zur Gaußprozessoptimierung geschlagen.

In Abschnitt 4.4 wird das Verfahren der Gaußprozessoptimierung auf die Optimierung an Sigma-Punkte-Filtern angewandt. Hierzu wurde in dieser Arbeit ein vollständiges Vorgehensmodell für eine automatisierte Optimierung entwickelt und an Sigma-Punkte-Filtern angewandt, was eine weitere Neuerung in dieser Arbeit darstellt.

4.3.1 Gaußprozess

Ziel dieses Abschnittes ist die allgemeingültige Erläuterung der Funktionsweise eines Gaußprozesses zur Vorhersage von Gütefunktionswerten des Szenarios an der Funktion $J_F = g(\theta)$, wenn von dieser nur diskrete Messungen vorliegen und nur das Ein-/ Ausgangsverhalten bekannt ist. Hierzu ist im ersten Schritt der Gaußprozess mittels Stützstellen an die Funktion $J_F = g(\theta)$ anzupassen. Basis stellt die nachfolgend dargestellte multivariate Gaußverteilung einer Zufallsvariablen x dar, mit $|\Sigma|$ als Determinante der Kovarianzmatrix:

$$\begin{aligned} \text{1-Dim.: } x &\sim \mathcal{N}(\mu, \sigma^2), & x, \mu, \sigma &\in \mathbb{R} \\ p(x) &= \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \end{aligned} \quad (4-5)$$

$$\begin{aligned} \text{p-Dim.: } x &\sim \mathcal{N}_p(\mu, \Sigma) & x, \mu &\in \mathbb{R}^p, \Sigma \in \mathbb{R}^{p \times p} \\ p(x) &= \frac{1}{(2\pi)^{p/2} |\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)\right). \end{aligned} \quad (4-6)$$

Eine wichtige Eigenschaft ist dabei, dass sich die multivariate Verteilung komplett mittels Erwartungswert und Kovarianz definieren lässt [BN06]. Der Gaußprozess \mathcal{GP} nutzt diese Eigenschaft einer multivariaten Gaußverteilung, indem er jedes Stützstellenpaar $\theta_i, J_{F,i}$ einer Funktion $J_F = g(\theta)$ als Gaußverteilung auffasst, die eine Unterverteilung einer multivariaten Gaußverteilung bildet. Die Annahme einer Gaußverteilung für die als Zufallsvariable aufgefasste Stichprobe $J_{F,1} = g(\theta_1)$ führt auf die Darstellung

$$J_{F,1} \sim \mathcal{N}(\mathbb{E}[g(\theta_1)], \text{var}[g(\theta_1)]). \quad (4-7)$$

Bei Hinzukommen eines weiteren Stützstellenpaares lässt sich (4-7) auf eine mehrdimensionale Zufallsvariable mit Gaußverteilung erweitern und mit

$$\begin{bmatrix} J_{F,1} \\ J_{F,2} \end{bmatrix} = \mathcal{N}\left(\begin{bmatrix} \mathbb{E}[g(\theta_1)] \\ \mathbb{E}[g(\theta_2)] \end{bmatrix}, \begin{bmatrix} \text{var}[g(\theta_1)] & k(\theta_1, \theta_2) \\ k(\theta_2, \theta_1) & \text{var}[g(\theta_2)] \end{bmatrix}\right) \quad (4-8)$$

darstellen. Die interessanten Terme stellen die Kovarianzelemente $k(\theta_1, \theta_2) = k(\theta_2, \theta_1)'$ dar, mittels derer die Beziehung zwischen den Funktionsausgängen $J_{F,1}$ und $J_{F,2}$ auf Basis der Eingangsdaten θ_1 und θ_2 dargestellt wird. Anders ausgedrückt: Die Kovarianzelemente enthalten Informationen über das Funktionsverhalten g bei den Eingangsdaten θ_1, θ_2 . Diese Annahmen lassen sich sehr einfach für die Vorhersage von Funktionswerten nutzen. Nimmt man ein bestimmtes Verhalten für die Kovarianz an (welches Verhalten dies ist, wird im nächsten Abschnitt detailliert vorgestellt) und interessiert sich für den Funktionswert $J_{F,2}$ an der Stützstelle θ_2 , so lassen sich die Kovarianz $k(\theta_1, \theta_2)$ berechnen und eine Vorhersage von $J_{F,2}$ tätigen, unter der angenommenen Unsicherheit der Varianz.

Dieses Vorgehen stellt den Grundgedanken eines als Zufallsprozess aufgefassten Gaußprozess \mathcal{GP} dar, mit dem die Funktion $g(\theta)$ approximiert werden kann und die durch Erwartungswert und Kovarianz vollständig definiert ist:

$$g(\theta) \sim \mathcal{GP}(m(\theta), k(\theta, \theta')). \quad (4-9)$$

Hierbei entsprechen $m(\theta)$ der Erwartungswertfunktion und $k(\theta, \theta')$ der Kovarianzfunktion. In den meisten Anwendungen liegen a-priori keine Informationen über den Erwartungswert der Funktion $g(\theta)$ vor, wodurch ohne Einschränkung der Allgemeinheit der Erwartungswert des Gaußprozesses $m(\theta)$ zu Null gesetzt werden kann [BN06]. Die Annahme des Erwartungswertes zu Null ist zwar ein übliches Vorgehen, allerdings keinesfalls ein zwingendes. Es stellt allerdings keine drastische Beschränkung des Gaußprozesses dar, da der A-posteriori-Prozess nicht auf Null beschränkt ist [RW06], was im Folgenden auch noch gezeigt wird. Der Gaußprozess wird demnach allein durch die Kovarianzfunktion $k(\theta, \theta')$ definiert und beschreibt die Kovarianz zwischen einem Datenpaar der Ausgänge als Funktion der Eingänge [RW06].

Für die Vorhersage von Funktionswerten anhand des Gaußprozesses werden zu Beginn Trainingsdaten der Funktion $J = g(\theta)$ benötigt, die aus einer endlichen Zahl an Wertepaaren bestehen:

$$\begin{aligned} \text{Stützstellen } X &= [\theta_1, \dots, \theta_N]^T \\ \text{Messwerte } Y &= [J_{F,1}, \dots, J_{F,N}]^T, \end{aligned} \quad (4-10)$$

mit $X \in \mathbb{R}^{N \times n_x}$ und $Y \in \mathbb{R}^{N \times 1}$. Das Ziel ist die Vorhersage von Messwerten $J_{F,*}$ an den noch nicht ausgewerteten Stützstellen θ_* .

$$\begin{aligned} \text{Stützstellen } X_* &= [\theta_{1*}, \dots, \theta_{M*}]^T \\ \text{Messwerte } Y_* &= [J_{F,1*}, \dots, J_{F,M*}]^T \end{aligned} \quad (4-11)$$

Hierzu ist im ersten Schritt die Kovarianzfunktion $k(\theta, \theta')$ an allen Kombinationen der Trainingsdaten auszuwerten, mit $K = k(\theta_i, \theta_j), \forall (i, j) \in [1, \dots, N]$.

$$K = \begin{bmatrix} k(\theta_1, \theta_1) & k(\theta_1, \theta_2) & \cdots & k(\theta_1, \theta_N) \\ k(\theta_2, \theta_1) & k(\theta_2, \theta_2) & \cdots & k(\theta_2, \theta_N) \\ \vdots & \vdots & \ddots & \vdots \\ k(\theta_N, \theta_1) & k(\theta_N, \theta_2) & \cdots & k(\theta_N, \theta_N) \end{bmatrix} \quad K \in \mathbb{R}^{N \times N} \quad (4-12)$$

Im nächsten Schritt kann die Kovarianzfunktion an allen Kombinationen unter Einbeziehung der vorherzusagenden Stützstellen gebildet werden, mit $K_* = k(\theta_i, \theta_{j*}), K_{**} = k(\theta_{i*}, \theta_{j*}), \forall (i, j) \in \{[1, \dots, N] \times [1, \dots, M]\}$.

$$\begin{aligned} K_* &= [k(\theta_*, \theta_1), \dots, k(\theta_*, \theta_N)] & K_* &\in \mathbb{R}^{M \times N} \\ K_{**} &= [k(\theta_*, \theta_*)] & K_{**} &\in \mathbb{R}^{M \times M} \end{aligned} \quad (4-13)$$

Zur Vorhersage der Funktionswerte an den interessierenden Stützstellen X_* wird die Eigenschaft des Gaußprozesses aufgegriffen, dass jede beliebige, endliche Untermenge des Gaußprozesses wiederum eine Gaußverteilung darstellt:

$$\begin{bmatrix} Y \\ Y_* \end{bmatrix} \sim \mathcal{N} \left(0, \begin{bmatrix} K & K_*^T \\ K_* & K_{**} \end{bmatrix} \right) \quad (4-14)$$

Das Interesse liegt nun in der bedingten Wahrscheinlichkeit $p(Y_*|Y)$, wodurch ausgesagt werden kann, wie hoch die Wahrscheinlichkeit einer bestimmten Voraussage von Y_* unter den bekannten Daten Y ist, die ebenfalls einer Gaußverteilung folgt und deren Herleitung im Anhang von [RW06] nachvollzogen werden kann.

$$Y_*|Y \sim \mathcal{N} (K_* K^{-1} Y, K_{**} - K_* K^{-1} K_*^T) \quad (4-15)$$

Die beste Vorhersage des Funktionswertes Y_* an der Stelle X_* liegt demnach bei:

$$Y_* = K_* K^{-1} Y, \quad (4-16)$$

mit der zugehörigen Unsicherheit in der Schätzung, die durch die Varianz ausgedrückt wird:

$$\text{var}[Y_*] = K_{**} - K_* K^{-1} K_*^T. \quad (4-17)$$

Aus (4-16) ist zu erkennen, dass, trotz der oben erwähnten A-priori-Annahme der Erwartungswertfunktion des Gaußprozesses als Null (vgl. (4-14)), der A-posteriori-Erwartungswert durchaus ungleich Null ist.

Das Vorgehen sei im Anschluss kurz an einer eindimensionalen Funktion $g(\theta)$ gezeigt, die in Bild 4-4 links dargestellt ist und deren zu ermittelnder Funktionswert als cyanfarbenes Kreuz an der Stützstelle $\theta_* = 2$ markiert ist. Die grau hinterlegte Fläche stellt die Unsicherheit bzw. die Kovarianz als 95 %-Vertrauensbereich dar.

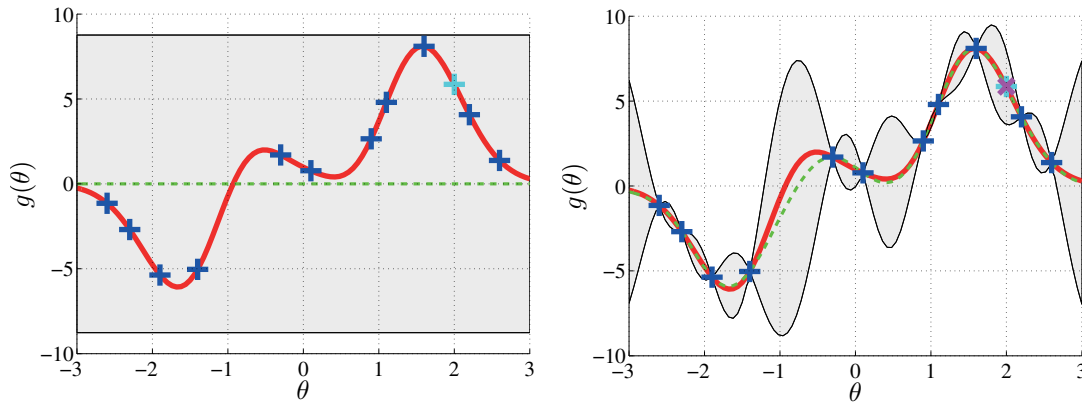


Bild 4-4: Unbekannte Zielfunktion (rot dargestellt) mit zu ermittelndem Funktionswert (cyanfarbenes Kreuz). Bestehende Messungen der Funktion (blaue Kreuze), initialer Gaußprozess definiert durch Erwartungswert (grüner Verlauf) und Unsicherheit (graue Fläche). Links: Initialer Gaußprozess. Rechts: Trainierter Gaußprozess und Vorhersage des Funktionswertes an $\theta_* = 2$ (magentafarbenes Kreuz)

Hierzu stehen Messwerte des Systems (blaue Markierungen) zur Verfügung, nach denen die Kovarianzmatrizen nach (4-12) und (4-13) berechnet werden können. Mit ihnen können der Erwartungswert und die zugehörige Unsicherheit nach (4-16) bzw. (4-17) bestimmt werden.

In Bild 4-4 ist der Erwartungswert des Gaußprozesses als grüne Funktion abgebildet. Da es sich um rauschfreie Messungen handelt, bildet der Gaußprozess diese vollständig ab und die Unsicherheit im Gaußprozess ist gleich Null. Der vorhergesagte Funktionswert ist in Bild 4-4 rechts als magentafarbenes Kreuz an der Stützstelle $\theta_* = 2$ dargestellt und weist dabei einen geringen Fehler zum realen Signalverlauf auf.

Im Normalfall treten die Messungen allerdings nicht rauschfrei auf, sondern liegen als verrauschte Sensordaten vor. Unter der Annahme von additivem, unabhängigem, normal verteiltem Rauschen mit der Varianz σ^2 ergibt sich

$$J_F = g(\theta) + \mathcal{N}(0, \sigma^2). \quad (4-18)$$

Der Aufbau der Kovarianzmatrizen im Vergleich zu rauschfreien Messungen (vgl. (4-14)) unterscheidet sich dabei nur durch die Addition der Diagonalmatrix der Rauschvarianzen [RW06]:

$$\begin{bmatrix} Y \\ Y_* \end{bmatrix} \sim \mathcal{N} \left(0, \begin{bmatrix} K + \sigma^2 I & K_*^T \\ K_* & K_{**} \end{bmatrix} \right). \quad (4-19)$$

Durch die Hinzunahme des Rauschprozesses als Maß der Unsicherheit verändert sich (4-17) bzw. (4-16) zu:

$$\begin{aligned} Y_*|Y &\sim \mathcal{N}(Y_*, \text{var}[Y_*]) \\ Y_* &= K_* [K + \sigma^2 I]^{-1} Y \\ \text{var}[Y_*] &= K_{**} - K_* [K + \sigma^2 I]^{-1} K_*^T. \end{aligned} \quad (4-20)$$

In Bild 4-5 ist die Vorhersage des Funktionswertes an der Stützstelle $\theta_* = 2$ dargestellt. Die Trainingsdaten des Gaußprozesses sind allerdings mit additivem Rauschen überlagert worden. Hierbei ist gut zu erkennen, dass die Messungen, dargestellt als blaue Kreuze, nicht mehr auf dem wahren Funktionsverlauf liegen. Durch die Hinzunahme der angenommenen Rauschvarianz in den Gaußprozess ist im Gegensatz zu Bild 4-4 rechts die Varianz an diesen Stützstellen im Gaußprozess nicht mehr Null. Es existiert demnach eine Unsicherheit im Gaußprozess selbst, wenn über die Stützstelle Messinformationen vorliegen. Die Vorhersage des Funktionswertes an der Stützstelle $\theta_* = 2$ stimmt hierdurch auch nicht mehr mit dem wahren Funktionswert überein.

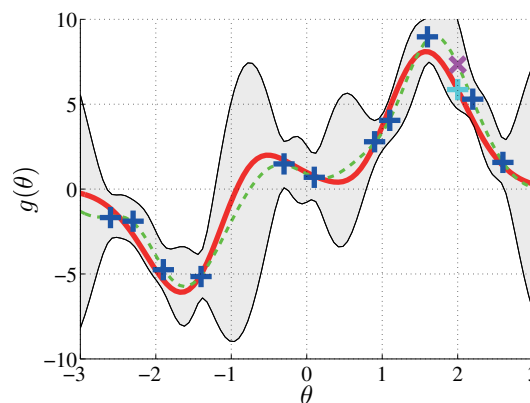


Bild 4-5: Vorhersage Funktionswert unter verrauschten Messungen. Blaue Kreuze: Stützstellen; Grün: Erwartungswert Gaußprozess; Grau: Varianz Gaußprozess; Magenta: Vorhersage; Rot: Ein-/ Ausgangsverhalten $g(\theta)$

Nach [RW96] handelt es sich beim Gaußprozess um ein nichtparametrisches Regressionsverfahren, da a priori keine Annahmen über die Natur der zu lernenden Funktion gemacht werden müssen, wie in anderen Regressionsverfahren. Die Berechnung der Einträge der Kovarianzmatrizen K , K_* und K_{**} erfolgt allerdings mittels Kernel-Funktionen, von denen eine Vielzahl verschiedener Vertreter existieren, die wiederum durch interne Hyperparameter⁴ an die jeweilige Problemstellung angepasst werden können. Somit besteht durchaus die Möglichkeit, mittels Annahmen über das zugrundeliegende Modell die Vorhersagequalität zu erhöhen,

⁴Hyperparameter meint in diesem Zusammenhang Parameter des statistischen Modells, um sie nicht mit Parametern der zu erlernenden Funktion $g(\theta)$ zu verwechseln.

da spezialisierte Kernel für beispielsweise die Anpassung an periodische Funktionen existieren. Hierzu wird im nächsten Abschnitt kurz die in dieser Arbeit genutzte Kernel-Funktion und der Einfluss der enthaltenen Hyperparameter vorgestellt.

4.3.2 Kernel-Funktionen und Hyperparameter

In Abschnitt 4.3.1 wurde gezeigt, dass der Gaußprozess einzig von der Kovarianzfunktion, auch Kernel genannt, abhängt. Die Wahl der Kovarianzfunktion bietet dabei einen Freiheitsgrad des Verfahrens, um a priori Wissen über das System einzubringen [RW06]. Die Grundidee der Kovarianzfunktion besteht in der Annahme, dass Systemeingänge θ , die eng beieinander liegen, zu ähnlichen Funktionswerten führen bzw. dass der Funktionswert eines Testpunktes, der dicht an einem Trainingspunkt liegt, ähnlich dem Funktionswert des Trainingspunktes ist [RW06].

Als häufigster Vertreter der Kovarianzfunktion kann die Squared Exponential (SE) genutzt werden. Je geringer bei dieser Funktion der euklidische Abstand zwischen zwei Punkten θ und θ' ist, desto größer wird die Abhängigkeit ihrer Funktionswerte:

$$k(\theta, \theta') = \sigma_f^2 \exp \left(-\frac{(\theta - \theta')^2}{2l_f^2} \right). \quad (4-21)$$

Die Parameter σ_f und l_f werden dabei als Hyperparameter bezeichnet und beeinflussen die Interpretation der zu approximierenden Daten. l_f wird als spezifische Längenskalierung bezeichnet, mit der die Abhängigkeit der Daten in Bezug zu ihren Entfernungen eingestellt werden kann. Für große Werte von l_f ergibt sich so eine Mittelung der Messwerte, wodurch sich die approximierte Funktion immer weiter einer Geraden annähert. Durch den Faktor σ_f lässt sich die maximale Varianz der Funktion einstellen. Die Wirkung der Faktoren wird im Folgenden anhand des bereits im letzten Abschnitt vorgestellten Beispiels gezeigt.

In Bild 4-6 sind die mittels der SE-Kovarianzfunktion erzeugten Verläufe bei konstantem Kovarianzparameter σ_f^2 und variablem Längenskalierungsfaktor l_f dargestellt. Der Gaußprozess kann aufgrund des sehr großen Längenskalierungsfaktors im linken Bild den Messungen nicht folgen und führt zu einer Ausgleichsfunktion. Im rechten Bild ist l_f zu klein gewählt, sodass der Gaußprozess zwischen den Stützstellen auf Null zurückgeht und eine zu große Unsicherheit im Gaußprozess besteht.

Die SE-Kovarianzfunktion ist dabei unendlich oft differenzierbar, was für die spätere Suche nach einem Maximum innerhalb des Gaußprozesses noch sehr wichtig wird, und stellt den am meisten genutzten Kernel [RW06] dar. Die in (4-21) dargestellte Kovarianzfunktion kann leicht auf höherdimensionale Probleme erweitert werden und bildet den *squared exponential with automatic relevance determination*-Kernel (SE-ARD), was für diese Arbeit auch notwendig ist, da die

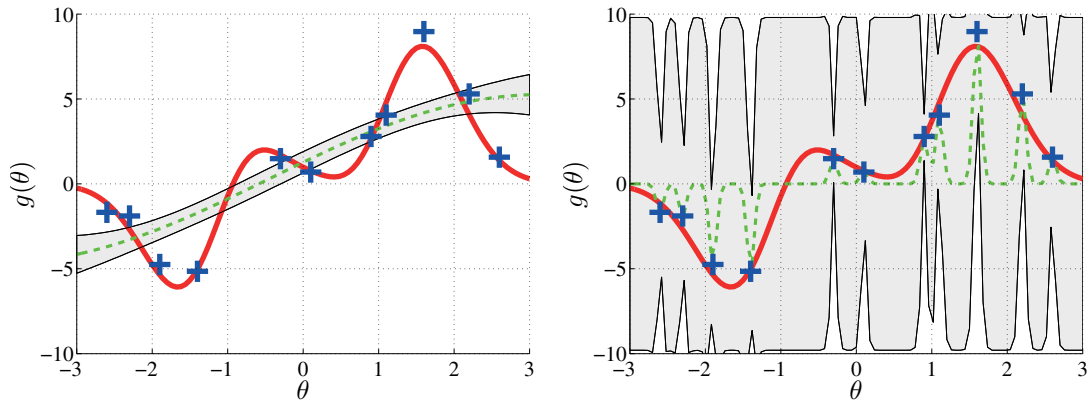


Bild 4-6: Gaußprozess unter verschiedenen Hyperparametern: Links: $l_f = 5, \sigma_f^2 = 1$ Rechts: $l_f = 0.1, \sigma_f^2 = 1$; Blaue Kreuze: Stützstellen; Grün: Erwartungswert Gaußprozess; Grau: Varianz Gaußprozess

Eingangsgröße θ aus den Entwurfsparametern der Filter besteht und beim UKF beispielsweise durch $\theta \in \mathbb{R}^{1 \times 3}$ gegeben ist.

$$k(\theta, \theta') = \sigma^2 \exp \left(-\frac{1}{2} (\theta - \theta')^T M (\theta - \theta') \right) \quad (4-22)$$

$$M = \text{diag}(l_{f,1}^{-2}, \dots, l_{f,n}^{-2})$$

Der Längenskalierungsfaktor kann dabei unabhängig für jede Dimension eingestellt werden [RW96]. Auf eine ausführliche Herleitung weiterer Kernel-Funktionen soll an dieser Stelle verzichtet und auf [RW06] verwiesen werden, da in dieser Arbeit nur der SE-ARD-Kernel verwandt wird.

In 4.3.1 wurde ausgeführt, dass es sich beim Gaußprozess um ein nichtparametrisches Verfahren handelt. Durch die vorgestellten Kovarianzfunktionen und deren Hyperparameter ist die Möglichkeit der Einbringung von A-priori-Wissen allerdings nicht nur gegeben, sondern zwingend erforderlich, da eine schlechte Wahl zu vollkommen falschen Interpretationen der Daten führen kann [RW06], [BN06]. Allerdings wurde von Rasmussen gezeigt, dass insbesondere die SE-Funktion exzellente Ergebnisse liefert, ohne Kenntnisse über die zugrundeliegende Funktion einfließen zu lassen. Von Ebdén wird allerdings in [Ebd08] angeführt, dass die Kovarianzfunktion trotzdem mit Vorsicht zu bestimmen ist und unterschiedliche Funktionen erprobt werden sollten.

Im hier vorliegenden Fall handelt es sich bei dem zugrundeliegenden Modell um das in Abschnitt 4.1.1 definierte Szenario, dessen Hauptbestandteil das topologieorientierte Modell mit unbekannten Nichtlinearitäten oder unbekannten Komponenten darstellt. Das Einbringen von A-priori-Informationen zur Bestimmung der Kernel-Funktion sowie der beinhaltenden Hyperparameter ist aufgrund des fehlenden Expertenwissens demnach nicht möglich. Die Ermittlung eines optimalen

Parametersatzes bei feststehender Kovarianzfunktion bei einem zugrundeliegenden Black-Box-Modell wird nachfolgend erläutert.

4.3.3 Optimale Vorgabe von Hyperparametern des Kernels

Nach Festlegung auf eine definierte Kernel-Funktion und der Rauschvarianz σ^2 müssen die jeweiligen Hyperparameter der Kernel-Funktion festgelegt werden, wodurch weiteres A-priori-Wissen über die zu approximierende Funktion eingebracht werden kann. Ein großer Vorteil der Gaußprozessoptimierung stellt die Möglichkeit dar, die Hyperparameter direkt aus den Trainingsdaten zu gewinnen [Sne07]. Der Grundgedanke besteht darin, die Hyperparameter so zu bestimmen, dass der Gaußprozess in der Lage ist, alle zur Verfügung stehenden Stützstellen wiederzugeben. Als Gütemaß allein die Abweichungen zwischen den Messungen an den Stützstellen und den vorhergesagten Funktionswerten heranzuziehen, ist nach Bild 4-6 nicht möglich, da der Gaußprozess im rechten Bild zwar sehr gut in der Lage ist, die Funktionswerte abzubilden, dennoch keine plausiblen Zwischenwerte generiert werden und die Kovarianz als Unsicherheit im Gaußprozess viel zu groß ist. Zur Erzeugung von optimalen Hyperparametern⁵ wird in der Literatur ([Ebd08], [Sne07], [RW06]) eine Maximierung der log-likelihood Funktion

$$\log(p(Y|X, [\sigma_f^2, l_f])) = -\frac{1}{2}Y^TKY - \frac{1}{2}\log|K| - \frac{p}{2}\log 2\pi \quad (4-23)$$

vorgenommen, mit $|K|$ als Determinante von K . Die likelihood Funktion stellt in diesem Fall die Wahrscheinlichkeit dar, dass die Funktionswerte Y bei vorhandenen X unter der Parametrierung $[\sigma_f^2, l_f]$ auftreten. Eine Maximierung dieser Funktion entspricht somit dem wahrscheinlichsten Parametersatz $[\sigma_f^2, l_f]$, unter dem Y als Funktionswerte von X auftreten. Die Logarithmierung der Likelihood-Funktion erleichtert dabei die Optimierung. Auf die Herleitung von (4-23) wird an dieser Stelle verzichtet; sie kann in [RW06] nachvollzogen werden.

Die beiden ersten Terme von (4-23) bilden gegenläufige Kostenfunktionen, um die in Bild 4-7 dargestellten Fälle zu unterbinden. Die Maximierung von $-\frac{1}{2}Y^TKY$ führt zu minimalen Abweichungen zwischen Gaußprozess und gemessenen Funktionswerten, wie im rechten Bild von 4-6 dargestellt, während eine Maximierung von $-\frac{1}{2}\log|K|$ zu einer schwach besetzten Kovarianzmatrix führt, wodurch Occam's razor implizit abgebildet ist und zu dem Verlauf führt, der im rechten Teil von Bild 4-6 dargestellt ist.

Da die partiellen Ableitungen der log marginal likelihood, mit $\zeta = [\sigma_f^2, l_f]$

$$\begin{aligned} \frac{\partial}{\partial \zeta_j} \log p(Y|X, \zeta) &= \frac{1}{2}Y^TK^{-1} \frac{\partial K}{\partial \zeta_j} K^{-1}Y - \frac{1}{2} \text{tr} \left(K^{-1} \frac{\partial K}{\partial \zeta_j} \right) \\ &= \frac{1}{2} \text{tr} \left(\left(K^{-1}YY^T (K^T)^{-1} - K^{-1} \right) \frac{\partial K}{\partial \zeta_j} \right) \end{aligned} \quad (4-24)$$

⁵Optimalität im Hinblick auf Erfassung der Stützstellen bei minimaler Varianz

analytisch gegeben sind, kann auf (4-23) und (4-24) direkt ein gradientenbasiertes Optimierungsverfahren angewandt werden [Ebd08].

Das globale Maximum kann dabei mit hoher Wahrscheinlichkeit durch mehrfaches Aufrufen eines gradientenbasierten Optimierers mit zufälligen Startwerten gefunden werden. Die Anwendung der log-likelihood Funktion und eine gradientenbasierte Optimierung auf das in den letzten Abschnitten gezeigte Beispiel der Abbildung der Funktion $g(\theta)$ führt auf die Hyperparameter $l_f = 0.68199, \sigma_f^2 = 3.8$, die den in Bild 4-7 dargestellten Gaußprozess bilden. Der Gaußprozess (grüner Verlauf) läuft zwar nicht durch jedes der blauen Kreuze, welche die verrauschten Messungen an der unbekannten roten Funktion darstellen, erfassen die Messungen im Mittel aber sehr gut bei einer minimalen Varianz, die den roten, zu bestimmenden Verlauf einschließt.

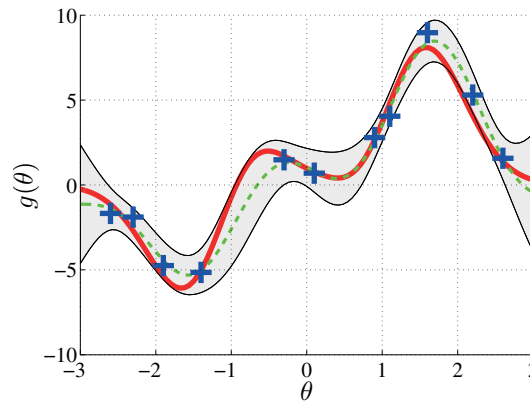


Bild 4-7: Gaußprozess unter optimierten Hyperparametern: $l_f = 0.68199, \sigma_f^2 = 3.8$; Blaue Kreuze: Stützstellen; Grün: Erwartungswert Gaußprozess; Grau: Varianz Gaußprozess; Rot: Ein-/ Ausgangsverhalten $g(\theta)$

4.3.4 Gaußprozessoptimierung

In den letzten Abschnitten wurde gezeigt, dass sich ein Gaußprozess mittels vorhandener Trainingsdaten, die aus Stützstellen X und Funktionsauswertungen Y bzw. Messungen an diesen Stellen bestehen, an die unbekannte Systemfunktion anpassen lässt. Im Anschluss können Funktionswerte am Gaußprozess der unbekannten Funktion, an noch nicht ausgewerteten Stützstellen X_* , vorhergesagt werden. Diese Vorhersage von Funktionswerten Y_* können interessierende Stützstellen sein oder die gezielte Ausgabe von Stützstellen, deren Funktionswerte gewünschte Merkmale aufweisen.

In der vorliegenden Arbeit wird in der Funktion $g(\theta)$, von der nur das Ein-/Ausgangsverhalten bekannt ist, das Maximum gesucht, das in Bild 4-8 mit einem cyanfarbenen Kreis bei $x = 1,6$ markiert ist. Das Ziel dieses Abschnittes ist die Suche der Stützstelle $\theta_* = 1,6$ bei der das Maximum der unbekannten

Systemfunktion $g(\theta)$ liegt. Durch die im letzten Abschnitt gezeigte analytische Ableitbarkeit des Gaußprozesses kann man sehr effizient gradientenbasierte Optimierungsverfahren einsetzen, um das Maximum im Gaußprozess zu finden, ohne Ableitungen der unbekannten Systemfunktion $g(\theta)$ durchführen zu müssen. Als Neuerung bei dieser Optimierung an einem Gaußprozess können neben den Informationen des vorhergesagten Funktionswertes J_* des Gaußprozesses an der Stützstelle θ_* mit $J_{F,*} = \mathcal{GP}(\theta_*) = K_* K^{-1} Y$ auch Informationen über die Unsicherheit des ermittelten Funktionswertes in Form der Varianz $\text{var}(\theta_*) = K_{**} - K_* K^{-1} K_*^T$ einbezogen werden. Die Funktionsweise soll an dem in den letzten Abschnitten bereits genutzten Anwendungsfall gezeigt werden, der in Bild 4-8 dargestellt ist. Im linken Plot ist die unbekannte Funktion $g(\theta)$ abgebildet, deren Maximum, markiert als cyanfarbener Kreis, zu ermitteln ist. Zu Beginn sind drei Stützstellen $X = [\theta_1 = -3, \theta_2 = 0, \theta_3 = 3]$ mit den zugehörigen Funktionswerten $Y = [J_{F,1}, J_{F,2}, J_{F,3}]$ bekannt und der Gaußprozess, dargestellt als grüner, gestrichelter Verlauf, an diesen trainiert. Die Unsicherheit des Gaußprozesses (graue Fläche) geht an den Stützstellen X auf das Rauschen σ^2 zurück und steigt auf das Maximum, je weiter man sich von den Stützstellen entfernt. Auf den Gaußprozess wird im Anschluss ein gradientenbasiertes Optimierungsverfahren angewandt, mit einem zufällig gewählten Startwert aus dem zulässigen Wertebereich der Eingangsgröße θ , wobei als Zielfunktion

$$\begin{aligned} J(\theta) &= J_F(\theta) + k_w \sqrt{\text{var}[J_F(\theta)]} \\ &= \mathcal{GP}(\theta) + k_w \sqrt{\text{var}[\mathcal{GP}(\theta)]} \end{aligned} \quad (4-25)$$

genutzt wird, die in Bild 4-8 als magentafarbener Verlauf abgebildet ist.

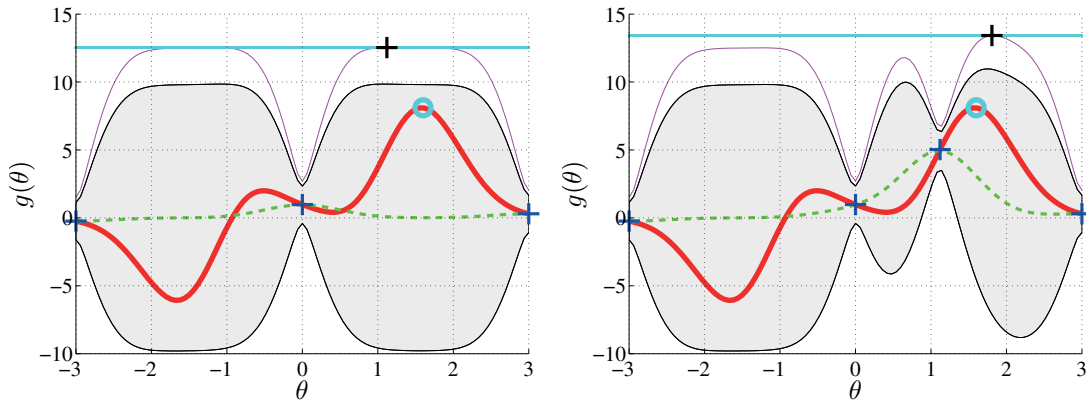


Bild 4-8: Optimierung am Gaußprozess: Schritt 1-2. Rot: Ein-Ausgangsverhalten $g(\theta)$; Grün: Erwartungswert Gaußprozess; Grau: Varianz Gaußprozess; Magenta: Kostenfunktion $J(\theta)$ (4-25); Blaue Kreuze: Stützstellen; Schwarzes Kreuz: Kandidat für Optimum

Die Einbeziehung der Varianz des Gaußprozesses in das Gütemaß des Optimierungsverfahrens bewirkt eine erhöhte Auswertung von bisher unbekannten Ge-

bieten im Gaußprozess. Anschaulich bedeutet dies, dass Stützstellen im Gaußprozess als Maximum der Funktion $g(\theta)$ in Betracht gezogen werden, wenn der Gaußprozess an diesen Stützstellen zwar einen sehr geringen Funktionswert vorhersagt, durch eine hohe Unsicherheit in der Vorhersage diese Stützstelle allerdings trotzdem als mögliches Maximum angegeben wird. Der Parameter k_w stellt einen dimensionslosen skalaren Gewichtungssparameter ein, mit dem das Verhältnis zwischen Erwartungswert und Varianz bei der Gütemaßbildung variiert werden kann. Ein großer Wert von k_w bewirkt eine explorative Auswertung, d. h. unbekannte Gebiete werden bekannten Gebieten mit hohem Funktionswert des Gaußprozesses bevorzugt und umgekehrt. Unbekannte Gebiete im Gaußprozess sind Wertebereiche, in denen nur wenige Stützstellen zur Verfügung stehen und somit die Vorhersagequalität des Gaußprozesses schlechter ist als in Gebieten mit einer Häufung an Stützstellen.

Im linken Plot von Bild 4-8 ist der auf Basis von (4-25) ermittelte Kandidat eines Maximums θ_* mit einem schwarzen Kreuz markiert. θ_* kann im ersten Schritt genauso links von $\theta = 0$ liegen, da die Kostenfunktion (4-25) um $\theta = 0$ im ersten Schritt symmetrisch ist. Der ermittelte Kandidat des Optimums θ_* wird daraufhin an der Funktion $J_{F,*} = g(\theta_*)$ ausgewertet und das Wertepaar dem bisherigen Gaußprozess zugeführt, wodurch sich der Gaußprozess dem erweiterten Datensatz $X = [-3, 0, 3, \theta_*]$, $Y = [J_{F,1}, J_{F,2}, J_{F,3}, J_{F,*}]$ anpasst und genauer wird, wie in Bild 4-8 rechts zu sehen ist, das den nächsten Schritt dieser iterativen Optimierung darstellt. Im rechten Teil von Bild 4-8 ist der neue Kandidat des Maximums gezeigt, der den maximalen Wert der magentafarbenen Gütefunktion markiert.

Dieses Vorgehen wird so lange wiederholt, bis ein gewähltes Abbruchkriterium erfüllt ist. Kriterien sind dabei durch eine maximale Anzahl an Iterationen gegeben oder durch den Fall, dass die Gaußprozessoptimierung mehrfach nacheinander einen Kandidaten für ein Maximum ausgibt, der sich in einem annähernd gleichen Wertebereich des letzten Kandidaten befindet, wobei „annähernd“ in diesem Fall 1 % des Gesamtwertebereiches des Optimierungsparameters θ entspricht. Die Fortführung der Optimierung aus Bild 4-8 ist in Bild 4-9 dargestellt, wobei in den letzten drei Bildern das Maximum gefunden wurde und der Kandidat des Maximums θ_* nur noch leicht variiert wird, was zum Beenden der Optimierung nach dem letzten Bild führt.

4.4 Anwendung am Unscented-Kalman-Filter

In diesem Abschnitt werden die im Vorfeld vorgestellten Methoden des Gaußprozesses zusammengeführt. Hieraus entstand als eine Neuerung im Rahmen dieser Arbeit ein Werkzeug, das beim Einsatz von Sigma-Punkte-Filtern zur Zustands- und Parameterschätzung in der Lage ist, ein in Abschnitt 4.1.1 definiertes Szenario aufzunehmen und einen optimalen Satz an Filterentwurfparametern θ auf Basis von definierten Anregungsfällen zu generieren. Die in den letzten Abschnitten

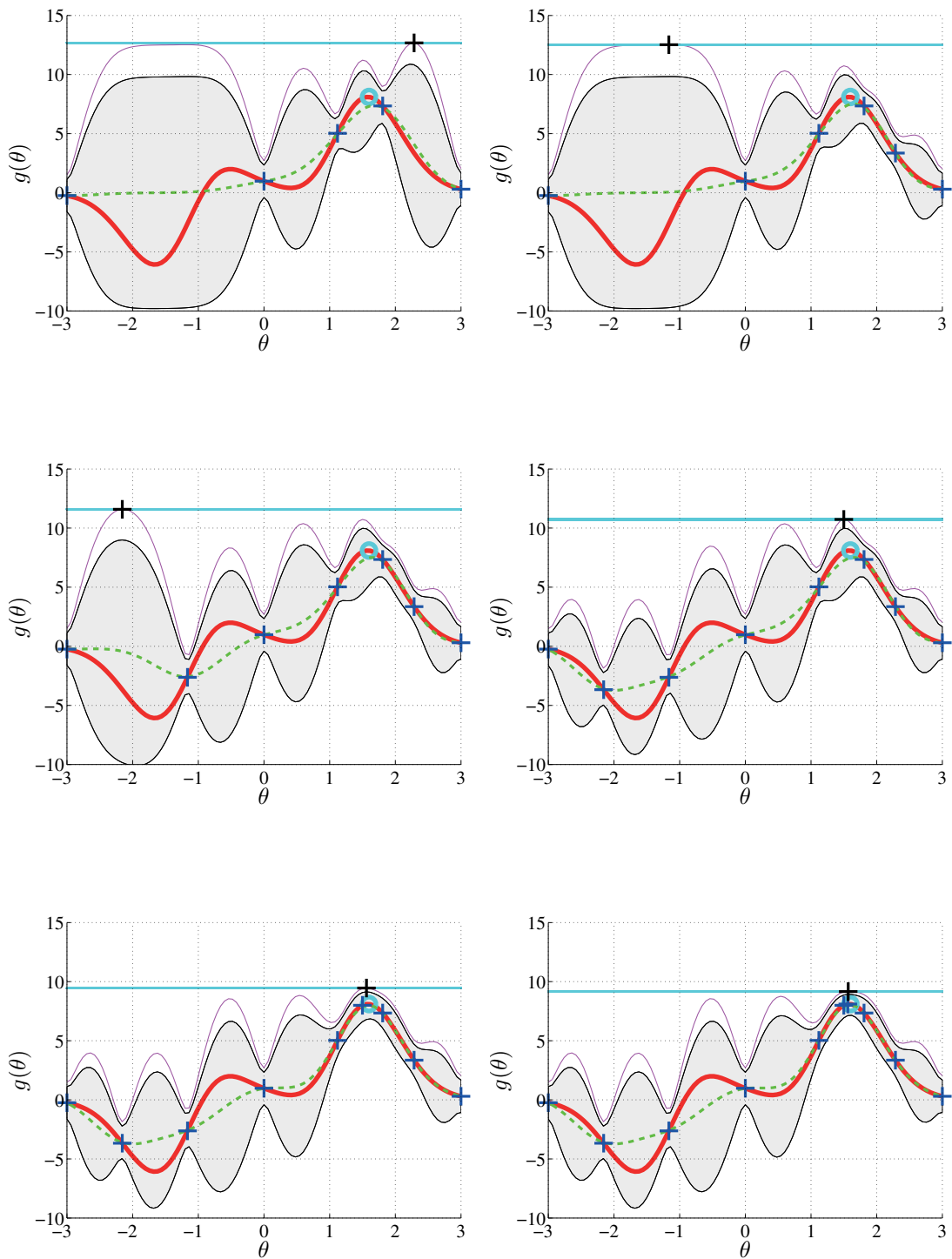


Bild 4-9: Optimierung am Gaußprozess: Schritt 5-8. Rot: Ein-/ Ausgangsverhalten $g(\theta)$; Grün: Erwartungswert Gaußprozess; Grau: Varianz Gaußprozess; Magenta: Kostenfunktion $J(\theta)$ (4-25); Blaue Kreuze: Stützstellen; Schwarzes Kreuz: Kandidat für Optimum

genutzte Funktion $g(\theta)$ stellt bei der Anwendung zur Ermittlung eines optimalen Filterentwurfsparametersatzes das vorgestellte Szenario dar, mit den Eingangsgrößen der filterspezifischen Entwurfparameter θ und der Kostenfunktion $J_F(\theta)$ zur Bewertung der Filtergüte, die zur Anwendung in der Gaußprozessoptimierung in der Form $J(\theta) = J_F(\theta) + k_w \sqrt{\text{var}[J_F(\theta)]}$ genutzt wird.

Als Kostenfunktion zur Bewertung der Filtergüte liegt es nahe, direkt den Fehler aus geschätzten und aus dem Szenario vorliegenden Zuständen zu nutzen und entsprechend

$$J(\theta) = \sum_{k=0}^T |x_k - \hat{x}_k| \quad (4-26)$$

über den zeitlichen Verlauf aufzusummieren, wobei x_k den Zustand des Systems zum Zeitschritt k darstellt und \hat{x}_k die Schätzung des Filters zum Zeitschritt k . Bei der Nutzung von realen Systemdaten ist allerdings eine Angabe der Zustände fast unmöglich oder nur mit unververtretbarem Aufwand. Als oftmals genutzte Fehlerfunktion zur Bewertung der Schätzungsgüte von Kalman-Filtern kann die likelihood Funktion des Prädiktionsschrittes des Filters angegeben werden, der in Analogie zu der in Abschnitt 4.3.3 vorgestellten likelihood-Funktion für jeden Zeitschritt ein Feedback über die Schätzungsgüte des Filters liefert, ohne auf Zustandsdaten des Systems zurückgreifen zu müssen, und durch

$$J(\theta) = \sum_{k=1}^{k_{end}} (y_k - \hat{y}_k)^T P_{yy,k}^- (y_k - \hat{y}_k) + \log |P_{yy,k}^-| \quad (4-27)$$

gegeben ist, mit y_k als Messungen zum Zeitschritt k , \hat{y}_k als geschätzte Messungen und $P_{yy,k}$ als Kovarianzmatrix der geschätzten Messungen. Eine Maximierung der Gütefunktion $J(\theta)$ aus (4-27) bewirkt eine Maximierung der Wahrscheinlichkeit, dass die vorhergesagten Messwerte \hat{y}_k , die auf der A-priori-Zustandsschätzung \hat{x}_k^- basieren, mit den gemessenen bzw. aus dem Szenario stammenden Messwerten y_k übereinstimmen.

4.4.1 Trainingsdaten

Zur Initialisierung des Gaußprozesses müssen einige Stützstellen X und Messwerte Y des Szenarios vorhanden sein, die im Folgenden als Trainingsdaten bezeichnet werden und aus Simulationen des Szenarios oder Versuchsdaten stammen.

Hierzu werden im ersten Schritt die Wertebereiche der zu optimierenden Parameter gesetzt, bei denen es sich um die Entwurfparameter des Sigma-Punkte-Verfahrens handelt, die in Abschnitt 3.1 vorgestellt wurden. Beispielhaft aufgeführt wird hier das SST-Verfahrens (vgl. Abschnitt 3.2.3). Als Optimierungsparameter müssen dabei $\theta_{SST} = [W_0, \alpha, \beta]$ bestimmt werden, wodurch sich die Dimension d des Optimierungsparametervektors zu $d = 3$ ergibt. Als zu ermittelnde

Trainingsdaten sollten eine möglichst geringe Anzahl an Simulationen durchgeführt werden, die den Optimierungsparameterraum gleichzeitig möglichst weitläufig erfassen. Hierzu werden Ränder und Mittelpunkte des (Hyper-)Würfel-Parameterraumes herangezogen, wie in Bild 4-10 dargestellt. Liegen keine realen Versuchsdaten vor, werden in dieser Arbeit die Wertebereiche in Abhängigkeit des gewählten Filters automatisch gesetzt, der daraus resultierende Parameterraum aufgebaut und das Szenario mit den Entwurfsparametern θ der Eck- und Mittelpunkte des Parameterraumes simuliert.

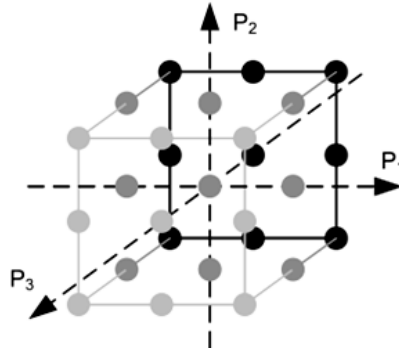


Bild 4-10: Würfelparameterraum bei dreidimensionalem Optimierungsparameterraum zur Ermittlung der Trainingsdaten zur Gaußprozessinitialisierung

Bei diesen Fällen kann es zu Parameterkonstellationen θ in Verbindung mit dem unbekannten Modell der Zustands- und Parameterschätzung kommen, die zu einem Filterabbruch führen. In diesen Fällen kann das Gütemaß nach (4-27) nicht gebildet werden, da es keine vergleichbare Basis zu den übrigen Auswertungen bietet. Als eine Neuerung in dieser Arbeit werden, um eine weitere Auswertung dieser Punkte im Parameterwürfel der Optimierung zu „bestrafen“, nach Abschluss der Ausführung aller 3^d Trainingsauswertungen die Gütefunktionswerte der abgebrochenen Szenarioauswertungen $J_{Error}(\theta)$ auf das Minimum der korrekten Auswertungen $J_{Min}(\theta)$ abzüglich 10 % der Spannweite der ermittelten Gütefunktionswerte gesetzt:

$$J_{error}(\theta) = J_{min}(\theta) - 0.1 \cdot (J_{max}(\theta) - J_{min}(\theta)). \quad (4-28)$$

Das Gütemaß $J_{max}(\theta)$ entspricht dabei dem maximalen Gütemaß der Szenarioauswertungen der Trainingsdaten, was deren bisherigem Optimum entspricht. Somit ist garantiert, dass Parameterkonstellationen, die zu instabilem Filterverhalten führen, für das Optimierungsverfahren uninteressant sind und gleichzeitig der Gaußprozess keine unnötig großen Wertebereiche der Trainingsdaten approximieren muss. Das hier beschriebene Vorgehen, das nachfolgend als Initialisierung benannt wird, ist in Bild 4-11 dargestellt.

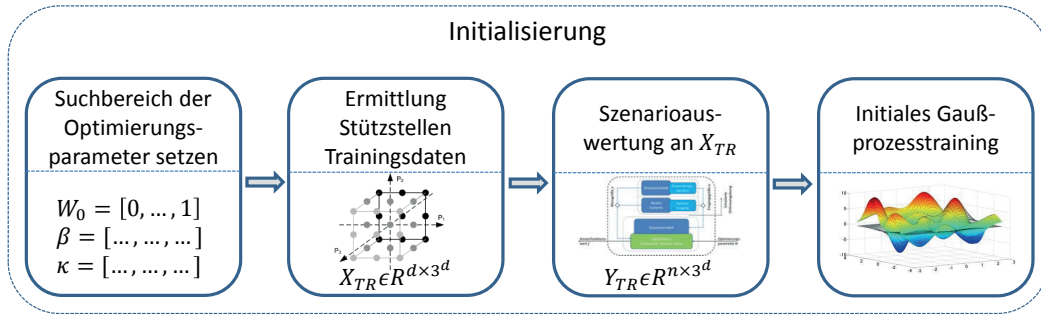


Bild 4-11: Vorgehen zur Initialisierung des Gaußprozesses. a) Festlegung der Optimierungsgrößen und ihrer Wertebereiche, b) Bestimmung der Ecken und Kanten bzw. Flächen- und Volumenmittelpunkte des Optimierungsparameterraumes, c) Simulation des Szenarios aus den vorab ermittelten Trainingsstützstellen, d) Initiales Training des Gaußprozesses mit den ermittelten Stützstellen und Funktionswerten

4.4.2 Optimierung zur Maximafindung und Abbruchbedingung

Nach der Initialisierung des Gaußprozesses kann die in Abschnitt 4.3.4 hergeleitete Optimierungsstrategie mittels der in Abschnitt 4.4 gezeigten Kostenfunktion des Kalman-Filters am Szenario gestartet werden. Ziel ist die Suche nach einem Kandidaten θ_* für ein mögliches Optimum des Szenarios, wobei der Benutzer den voreingestellten Wert der Gütefunktionsgewichtung k_w aus (4-8) verändern kann. Zur Erhöhung der Wahrscheinlichkeit, dass das gradientenbasierte Optimierungsverfahren nicht in einem lokalen Optimum „hängenbleibt“, wird das Verfahren c -mal mit Zufallsstartwerten aus dem Optimierungsparameterraum gestartet. Als Kandidat für ein wahrscheinliches Maximum θ_* wird die Stützstelle mit dem maximalen Gütefunktionswert der c -Iterationen ausgegeben. Der Kandidat θ_* wird daraufhin am Szenario ausgewertet und das daraus resultierende Wertepaar dem Gaußprozess zur Verfeinerung zugeführt. Dieses Vorgehen wird solange ausgeführt, bis die maximale Anzahl an Funktionsaufrufen M erreicht wurde, die vom Nutzer festzulegen ist, oder bis das Maximum durch mehrfaches, aufeinanderfolgendes Aufrufen eines ähnlichen⁶ Kandidaten für das Maximum gefunden wurde. Ist eine der Abbruchbedingungen erfüllt, wird der Parametersatz mit dem maximalen Funktionswert als Optimum des gewählten Filters unter den definierten Anwendungsszenarien ausgegeben. Das hier beschriebene Vorgehen wird nachfolgend mit dem Begriff *Optimierung* benannt und ist in Bild 4-13 dargestellt.

Die Anwendung der internen Optimierung des Gaußprozesses, um optimale Hyperparameter der Kovarianz σ_f^2 und der Längenskalierung l zu generieren, wird unter zwei Bedingungen gestartet, die beide erfüllt sein müssen. Zum Ersten müssen für eine Optimierung des Gaußprozesses genügend Stützstellen vorliegen, was in dieser Arbeit erzeugten Trainingsdaten von 3^d Stützstellen entspricht, womit

⁶Ähnlich heißt in diesem Fall: 1 % des Parameterraumes um den letzten Kandidaten.

nach der Initialisierung eine erste Optimierung der internen Hyperparameter gestartet wird.

Die zweite Bedingung zur Anwendung der internen Optimierung besteht in einem schlecht konditionierten Gaußprozess, was bedeutet, dass der Gaußprozess nicht in der Lage ist, die Gesamtheit der Stützstellen abzubilden, wie der linke Plot von Bild 4-12 zeigt.

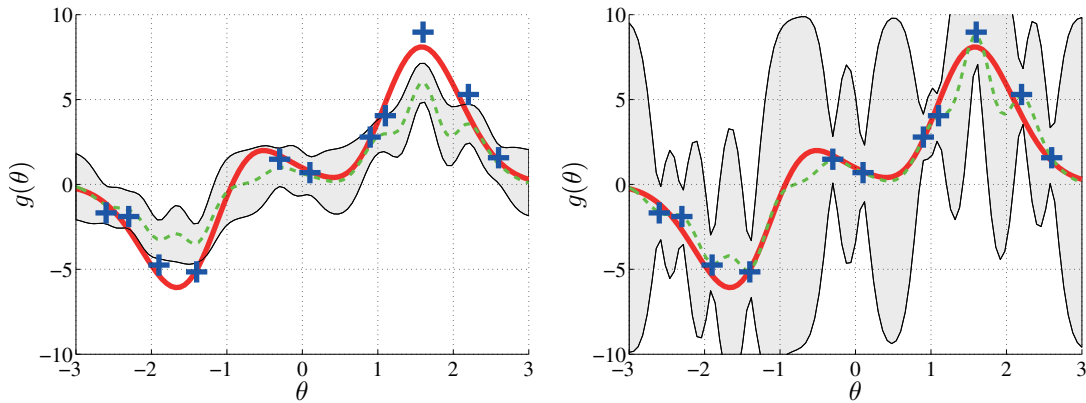


Bild 4-12: Gaußprozess unter verschiedenen Hyperparametern: Links: $l_f = 0.2, \sigma_f^2 = 1$; Rechts: $l_f = 0.2, \sigma_f^2 = 5$; Blaue Kreuze: Stützstellen; Grün: Erwartungswert Gaußprozess; Grau: Varianz Gaußprozess; Rot: Ein-/Ausgangsverhalten $g(\theta)$

Erkannt wird ein schlecht konditionierter Gaußprozess, wenn die Abweichung der Vorhersage von Funktionswerten des Gaußprozesses zu den gemessenen Werten außerhalb des 95 % Vertrauensbereichs liegt. Anders ausgedrückt: Wenn die gemessenen Funktionswerte durch den Gaußprozess nicht erfasst werden und gleichzeitig nicht in dem Unsicherheitsbereich liegen, was in Bild 4-12 links durch die grau hinterlegte Fläche gezeigt ist. Erfasst wird dieser Fall, wenn die Bedingung

$$|\mathcal{GP}(\theta_i) - y_i| > \varepsilon \quad i = 1, \dots, n_Z \quad (4-29)$$

mit n_Z als Anzahl der Stützstellen und

$$\begin{aligned} \varepsilon &\approx 1.96 \cdot \sqrt{\text{var}(\mathcal{GP}(\theta_i))} \\ &\approx 1.96 \cdot \sqrt{K(\theta_i)_{**} - K(\theta_i)_* K(\theta_i)^{-1} K(\theta_i)_*^T} \end{aligned} \quad (4-30)$$

erfüllt ist. Notwendig wird dieser Schritt, wenn durch neu hinzukommende Stützstellen ein Verhalten auftritt, das durch die vorher bestehenden noch nicht erfasst wurde. Die Erfassung des im rechten Plot von Bild 4-12 dargestellten Falls ist durch (4-29) nicht möglich, aber auch nicht nötig, da dies durch die Optimierungsmethodik der Hyperparameter aus Abschnitt 4.3.3 unterbunden wird.

Die Einbindung dieses Verfahrens in die Gesamtoptimierung wird als *Lernverfahren* bezeichnet, da es anhand immer neuer Daten, welche durch die Gaußprozessoptimierung geliefert werden, ein immer besseres Approximationsverhalten der wahren Funktion $g(\theta)$ erlernt, wodurch immer effizienter nach einem möglichen Maximum von $g(\theta)$ gesucht werden kann.

Das gesamte Verfahren der M -maligen Optimierung am Gaußprozess, mit der c -maligen internen Optimierung zur Suche eines Kandidats θ_* , ist in Bild 4-13 dargestellt, in dessen Zentrum das Lernverfahren zur Optimierung der Hyperparameter des Gaußprozesses eingebettet ist.

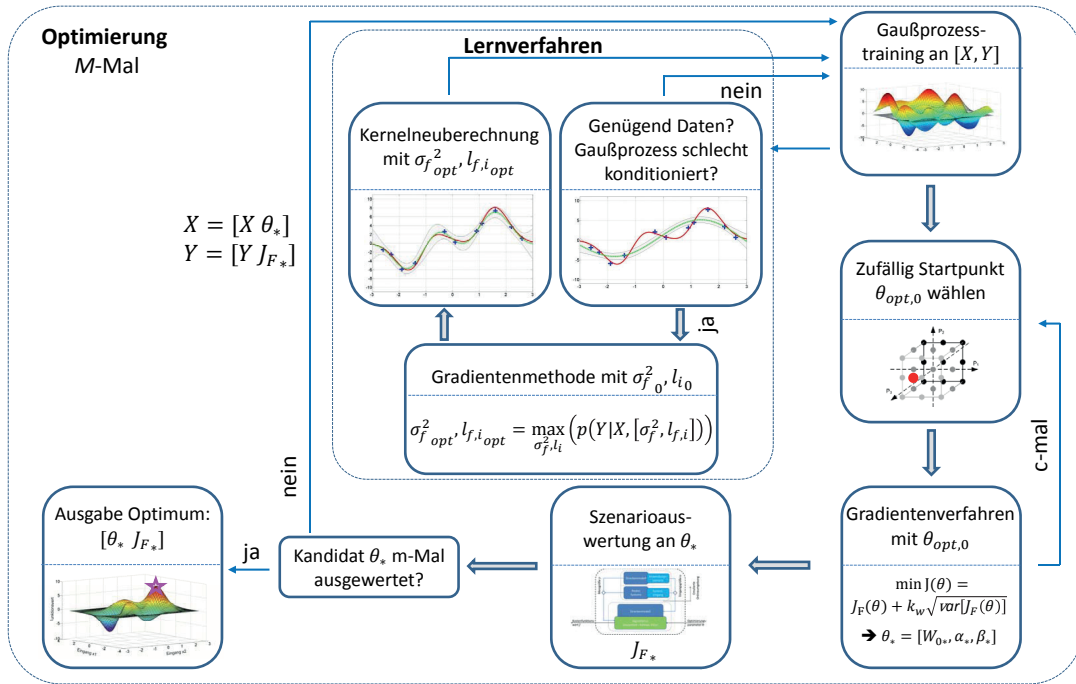


Bild 4-13: Vorgehen der Gaußprozessoptimierung zur Ermittlung eines optimalen Parametersatzes zum Filterentwurf

Zusammenfassend wurde in dieser Arbeit erstmalig ein Verfahren für die vollautomatische Auslegung der Entwurfsparameter von Sigma-Punkte-Filtern entwickelt, das auf einer Gaußprozessoptimierung und der Anwendung von Szenarien basiert. Der Benutzer muss im ersten Schritt nur das eingangs definierte Szenario laden, woraufhin Wertebereiche der Optimierungsparameter in Abhängigkeit des genutzten Sigma-Punkte-Filters gesetzt, Trainingsdaten generiert und ausgewertet werden, der Gaußprozess initial trainiert und im Anschluss die eigentliche Optimierung am Gaußprozess gestartet wird, unter fortlaufender Kontrolle der Approximationsgüte des Gaußprozesses und einer, falls erforderlich, Optimierung der Hyperparameter der Kernel-Funktion. Zum Abschluss werden die problemspezifischen, optimierten Entwurfsparameter ausgegeben und intern für die Weiterverwendung im Filter abgelegt. Der Benutzer muss zur Ermittlung dieser

Entwurfsparameter kein Expertenwissen über das Modell, die Filtertechnologie oder die Optimierungstechnik besitzen. Über veränderbare Wertebereiche oder Abbruchkriterien der Optimierung kann allerdings A-priori-Wissen eingebracht werden, ebenso über die Festsetzung der Faktoren F_{R_n} und F_{R_v} , die das Vertrauen in Messung oder Modell widerspiegeln. Sollten auch hierüber keine Kenntnisse vorliegen, können diese Faktoren neben den Entwurfsparametern ebenfalls der Optimierung übergeben werden.

5 Auslegungsmethodik

In diesem Kapitel wird eine Entwurfsmethodik für die Zustands- und Parameterschätzung an topologieorientiert erstellten Modellen vorgestellt. Hierzu werden die in den letzten Kapiteln erläuterten Verfahren in den Gesamtkontext eingeordnet und um notwendige Ergänzungen erweitert. Anforderungen stellen die bereits in Abschnitt 1.3 definierten Punkte dar. Die entwickelte Methodik lässt sich dabei in vier Phasen gliedern, die in Bild 5-1 dargestellt sind.

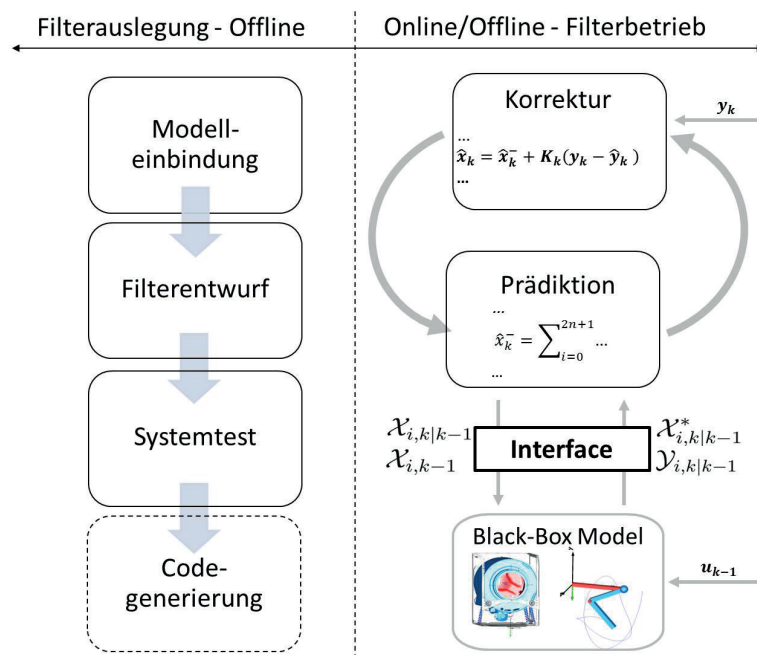


Bild 5-1: Viergeteilte Entwurfsmethodik zur Auslegung von Sigma-Punkte-Filtern an Black-Box-Modellen

Im linken Abschnitt ist der viergeteilte Filterentwurfsprozess abgebildet, der sich in die Phasen *Modelleinbindung*, *Filterentwurf*, *Systemtest* und *Codegenerierung* unterteilt. Der rechte Teil von Bild 5-1 stellt die Betriebsphase des Filters dar, die sowohl offline als auch online durchführbar ist. In den folgenden Abschnitten werden die einzelnen Abschnitte der Methodik erläutert und die vorgestellten Werkzeuge eingegliedert bzw. um weitere Verfahren erweitert.

Für alle Phasen der Entwurfsmethodik zu Auslegung, Test und Export von Sigma-Punkte-Filtern muss eine Schnittstelle bereit gestellt werden, die in der Lage ist, ein topologieorientiert erstelltes Streckenmodell, das in diesem Fall unter Dymola erzeugt wurde, einzubinden und die notwendigen Funktionen zur Filterfunktionalität bereitzustellen. Hierzu zählen:

- Ermittlung der Systemordnung, der Zustände und ihre modellinterne Bezeichnung.

- Ermittlung der Ein- und Ausgänge des Systems.
- Ermittlung von zu schätzenden Parametern im Modell, nebst zugehörigen Grenzbereichen.
- Neubesetzung des Zustandsvektors.
- Simulation des Modells um beliebige Zeitspannen.
- Auslesen des Zustandsvektors aus dem Modell.
- Echtzeitfähigkeit der Schnittstelle.

Als geeignete Schnittstelle und Austauschformat wird das Functional-Mockup-Interface (FMI) genutzt, das nachfolgend kurz vorgestellt wird.

5.1 Das Functional-Mockup-Interface

Bei dem Functional-Mockup-Interface (FMI) handelt es sich um ein Austauschformat für Modelle, das von der Daimler AG initialisiert und vorangetrieben wird und im MODELISAR-Projekt ITEA2 entwickelt wird. Der Grundgedanke besteht in der Schaffung eines offenen Standards zur Modellableitung von Zulieferern, deren Modelle in unterschiedlichen Entwicklungsumgebungen erstellt wurden, weshalb ihre Kopplung sich als schwierig erweist [BOA⁺11], [Cho13]. Für unterschiedliche Entwicklungsumgebungen existieren dabei bereits eine Vielzahl an Schnittstellen, wie die s-function für MATLAB/Simulink. Die Kopplung von Modellen mit solchen Schnittstellen ist allerdings nur mit großem Aufwand möglich. Die Functional-Mockup-Unit (FMU) beinhaltet das FMI und besteht aus einer modellbeschreibenden xml-Datei und einer die Funktion enthaltenden dll. Die Kopplung unterschiedlicher FMUs ist auf Basis der standardisierten Funktionen innerhalb der FMU möglich, die beispielsweise einen zentralen Datenaustausch unter den FMUs und deren Auswertung ermöglicht, wie in Bild 5-2 dargestellt.

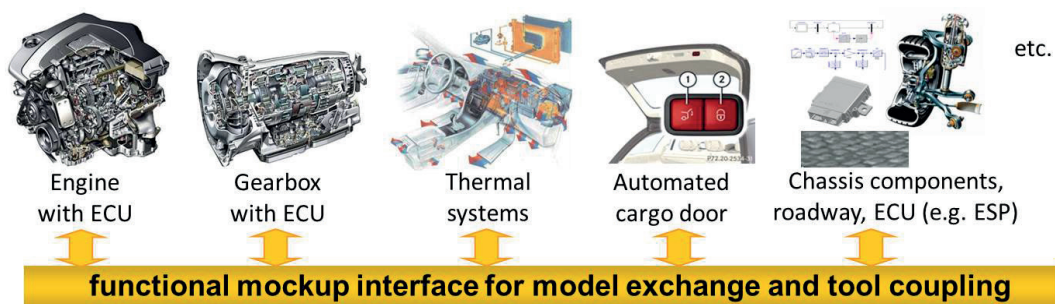


Bild 5-2: Möglichkeit zur Kopplung unterschiedlicher Modelle mittels des FMI-Standards [Quelle:Daimler AG]

Unterschiedliche Tools zur modellbasierten Entwicklung nutzen bereits den FMI-Standard wie Adams, CATIA, MapleSim genauso wie eine Reihe namhafter Un-

ternehmen: Atego, BOSCH, Daimler und Dassault Systèmes [BOA⁺11]. Der Modellexport aus dem Entwicklungstool kann dabei auf unterschiedliche Weise erfolgen: Zum Einen als *Model Exchange FMI* und zum Anderen als *Co-Simulation FMI*. Der Unterschied beider Austauschformate liegt in der Solveranwendung. Beim Export als *Co-Simulation FMI* wird ein geeigneter Solver direkt von der Modellentwicklungsumgebung in die FMU eingebettet. Bei der Einbindung der FMU durch ein Mastertool wird diese nur noch mit Eingangsdaten beaufschlagt und die Simulationsergebnisse zu definierten Zeitpunkten abgefragt (siehe Bild 5-3) [Con10a]. Bei dem Export als *Model Exchange FMI* wird die FMU ohne Solver exportiert, stellt allerdings alle notwendigen Schnittstellen zur Lösung mittels externer Solver bereit, die vom Mastertool genutzt werden. Zum Zeitpunkt dieser Arbeit steht die FMU-Version 1.0 zur Verfügung, die beim Export als *Model Exchange* das System als ordinary differential equation (ODE) beinhaltet [Con10b].

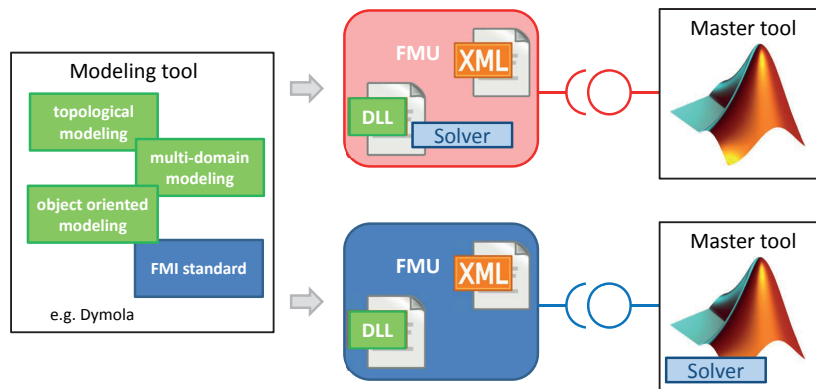


Bild 5-3: *Model Exchange FMI* und *Co-Simulation FMI* als Austauschformate der Modellentwicklungsumgebung

Das Anwendungstool Dymola unterstützt ebenfalls den Export von *Model Exchange*- und *Co-Simulation FMI*. Zum Zeitpunkt dieser Arbeit stellt das *Co-Simulation* Format nur einen Solver zur Verfügung. Bei diesem handelt es sich allerdings um den *Sundials*-Solver CVODE, der einen Multistep-Solver darstellt und somit keine echtzeitfähige Schnittstelle ermöglicht. Aus diesem Grund erfolgt der Export als *Model Exchange FMI*, das unter dem Mastertool MATLAB gelöst wird, was die Entwicklung eines Interfaces zur Einbindung von *Model Exchange*-FMUs unter MATLAB voraussetzt und im nächsten Abschnitt beschrieben wird.

5.1.1 MATLAB-Interface zur FMU-Ansteuerung

Die FMU-Einbindung unter MATLAB/Simulink erfolgt auf Basis eines Software Development Kit (SDK) von Qtronic [PC14]. Das SDK wurde dabei um einige Funktionen zur Nutzung unter MATLAB und filterspezifischen Anforderungen erweitert, wie in [STK13a] und [STK13b] gezeigt. Hierzu zählen die Identifikation

der Ein- und Ausgänge, der Identifikationsparameter, der Zustände des Systems sowie die Erweiterung um ein MATLAB-Interface, das die Ansteuerung der FMU aus MATLAB ermöglicht. Die Gesamtarchitektur der entwickelten Schnittstelle zur Einbindung von Modellen unter MATLAB, die als FMU vorliegen, ist in Bild 5-4 dargestellt.

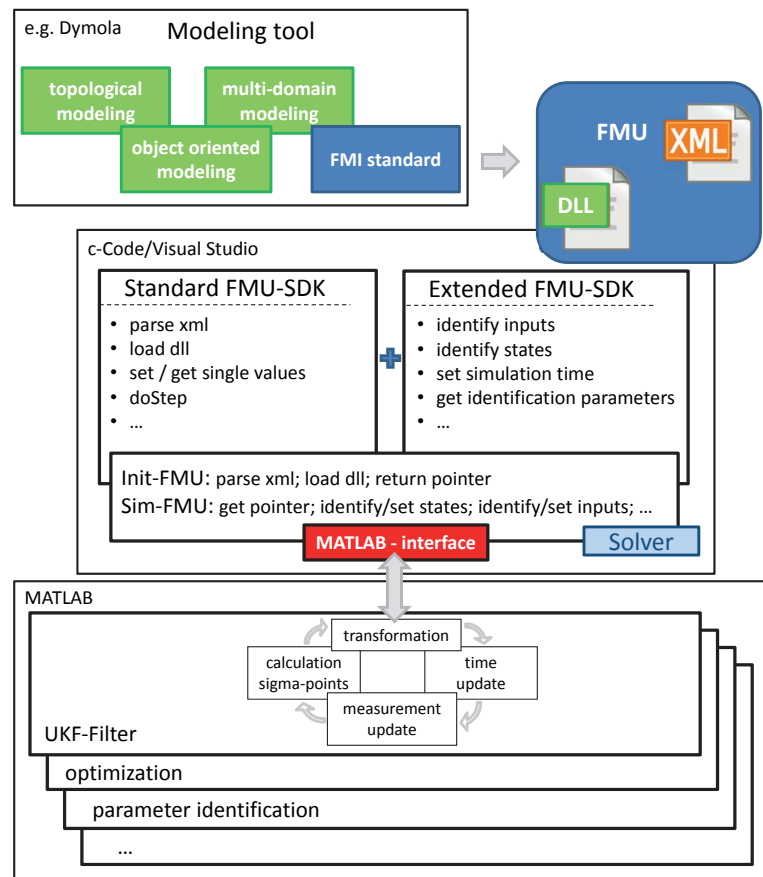


Bild 5-4: Entwickeltes Interface zur Einbindung von FMU-Modellen in MATLAB zur Nutzung in Filteralgorithmen oder weiteren modellbasierten Techniken

Ausgehend von dem Modellentwicklungstool, das den FMI-Standard beinhaltet und in diesem Fall durch Dymola gegeben ist, erfolgt die Ableitung der *Model Exchange* FMU, bestehend aus systembeschreibender *xml*- und modellenthaltender *dll*-Datei. Die Einbindung der erstellten FMU in das entwickelte Interface erfolgt durch den Funktionsaufruf aus MATLAB, woraufhin die *xml* analysiert und die *dll* geladen wird. Daraufhin stehen alle Funktionsaufrufe, die der FMI-Standard und die Erweiterung bieten, zur Verfügung.

Mittels der Analyse der systembeschreibenden *xml*-Datei können die benötigten Größen zur Filteranpassung gewonnen werden. In dieser Arbeit werden mittels eines entwickelten Interfaces in einem Pre-processing-Schritt die *xml*-Datei ana-

lysiert und alle für die Filterinitialisierung, die Filterauslegung und den Filterbetrieb notwendigen Größen ermittelt:

$$[n, nu, ny, np, np_{bounds}, init] = \text{Init FMU}(\text{Auswahl FMU}). \quad (5-1)$$

Der Parameter n bezeichnet die Systemordnung und die Bezeichner, nu bzw. ny die Anzahl Ein- und Ausgänge, np und np_{bounds} die Anzahl und die Bezeichner der Schätzparameter mit zugehörigen Parametergrenzen und $init$ die im Modell gewählten Initialisierungswerte des Zustandsvektors. Auf Basis der entwickelten Initialisierung ist die eingangs in Abschnitt 5 definierte Anforderung I erfüllt, sodass unbekannte Modelle eingebunden und alle notwendigen Informationen für die Filterauslegung und den Filterbetrieb automatisiert ermittelt werden können.

Die Transformation des Zustandsvektors entsprechend (3-43) und (3-46) stellt die Kernanforderung zur Realisierung der vorgestellten Sigma-Punkte-Filter dar und kann durch im FMI-Standard implementierte Funktionen realisiert werden. Zur Transformation der Sigma-Punkte wird innerhalb des Interfaces mittels der Funktion *fmiSetValue* der übergebene Sigma-Punkt als Zustandsvektor des Modells gesetzt und das Modell im Anschluss mit der Funktion *fmiDoStep* von der Startzeit t_{start} bis zur angegebenen Endzeit t_{end} simuliert. Der transformierte Sigma-Punkt wird im Anschluss mittels der Funktion *fmiGetValue* ausgelesen und abgespeichert. Der im Filteralgorithmus implementierte Interface-Aufruf erfolgt mittels

$$\begin{aligned} &\text{for } i = 1 : 2n + 1 \\ &\quad [\mathcal{X}_{i,k|k-1}^*] = \text{Transform}(\mathcal{X}_{i,k-1}, u_{k-1}, t_{start}, t_{end}) \\ &\text{end for;} \end{aligned} \quad (5-2)$$

und übergibt die zu transformierenden Sigma-Punkte $\mathcal{X}_{i,k-1}$ sowie die Start- und Endzeit t_{start} bzw. t_{end} . Den Rückgabewert stellt der transformierte Sigma-Punkt $\mathcal{X}_{i,k|k-1}^*$ dar.

Die Anforderung V nach Abschnitt 1.3 zur Nutzung in Echtzeitanwendungen ist zum Einen durch die Anwendung von Sigma-Punkte-Filtern erfüllt, bei denen in dieser Arbeit nur Sigma-Punkte-Varianten eingesetzt werden, die mit einer geringen Anzahl an Sigma-Punkten und damit an zu transformierenden Stützstellen der Verteilungen arbeiten. Zum Zweiten erfolgt die Integration der als ODE vorliegenden FMU mittels eines Explizit-Euler-Algorithmus mit einstellbarer Schrittweite, wodurch die Echtzeitfähigkeit seitens der Modellsimulation gewährleistet ist.

Das entwickelte Interface ist für weitere Anwendungen offen, wie beispielsweise eine halbautomatisierte Modellidentifikation und -validierung, die in [KST14] vorgestellt wurde, oder die Kopplung von als FMU vorliegenden Modellen aus unterschiedlichen Tools.

5.2 Modelleinbindung

Mit dem im letzten Abschnitt entwickelten Interface ist es möglich, eine FMU unter MATLAB einzubinden und alle für die Initialisierung des Filters notwendigen Größen zu ermitteln. Hierzu zählen die Anzahl der Zustände sowie die Anzahl der Ein- und Ausgänge nebst deren Bezeichnern, die im Modellierungsprozess festgelegt wurden. Ist das Modell für eine Parameterschätzung vorgesehen, werden die zu bestimmenden Parameter automatisch erkannt. Dies erfolgt mittels eines im Modellierungsprozess angegebenen Keywords, das bei der Initialisierung der FMU erkannt wird, woraufhin der zugehörige Parametername ausgelesen wird. Sollten Initialisierungswerte in der FMU abgelegt sein, werden diese direkt als Startwerte einer Parameterschätzung übergeben. Angegebene Parametergrenzen werden dabei ebenfalls dem Modell entnommen. Die Wertebereiche der zu schätzenden Parameter entsprechen dabei den Parametergrenzen und werden dem in 3.6 vorgestellten Verfahren zur Beschränkung der Sigma-Punkte mittels Reskalierung zur Verfügung gestellt. Vom Anwender sind für die Durchführung dieser Schritte nur die zu nutzende FMU und das in Abschnitt 4.1.1 definierte Szenario einzuladen. Anhand der erkannten Schätzparameter muss der Nutzer die Entscheidung treffen, ob diese auf den im Modell festgelegten Initialisierungswerten verbleiben sollen oder ob eine Parameterschätzung angewandt wird. Im Falle der Parameterschätzung werden alle Filtergrößen zur Anwendung des Joint-Kalman-Filters erweitert und die Nebenbedingungen eingebunden. Der abschließende Ablauf der Modelleinbindung ist in Bild 5-5 dargestellt.

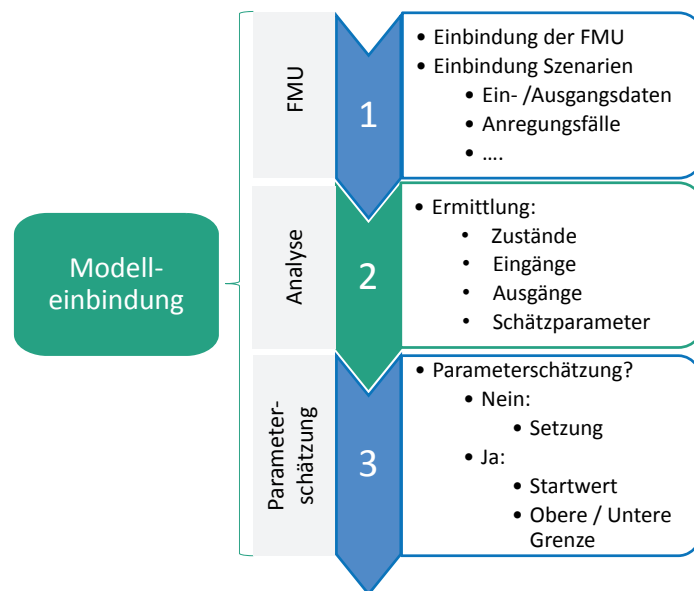


Bild 5-5: Einzelphasen in dem Entwurfsschritt der Modelleinbindung zur Filterauslegung

Automatisiert durchführbare Schritte sind grün gekennzeichnet. Abschnitte, in denen Nutzereingaben notwendig sind, erscheinen in Blau. Aus der Beschreibung der Modelleinbindung ist ersichtlich, dass beim Nutzer keine Expertenkenntnisse über das zugrundeliegende Modell vorhanden sein müssen. Alle filterimmanenten Größen werden mittels des entwickelten Interfaces automatisiert aus der FMU ermittelt und stehen für die weiteren Schritte der Auslegung und des Betriebs zur Verfügung. Die Anpassung des Filters an das in der FMU vorliegende Modell und an Anwendereingaben wie Parameterschätzungen erfolgt vollautomatisch. Die Anwendung der Klasse der Sigma-Punkte-Filter erfüllt in diesem Zusammenhang auch direkt die unter Abschnitt 1.3 gestellte Anforderung II, dass mit starken Nichtlinearitäten des Modells umgegangen werden können muss. Durch das im Rahmen dieser Arbeit entwickelte Verfahren der Erkennung und Korrektur von ungültigen Sigma-Punkten, das in Abschnitt 3.6 vorgestellt wurde, ist die Anforderung VI ebenfalls erfüllt, dass insbesondere im Rahmen einer Parameterschätzung keine ungültigen Sigma-Punkte berechnet werden, die bei der Simulation des Modells zu Instabilitäten und somit zum Abbruch des Filters führen.

5.3 Filterentwurf

Nachdem im letzten Abschnitt mittels des entwickelten Interfaces alle notwendigen Informationen automatisiert aus dem als FMU vorliegenden Modell gewonnen wurden, kann nachfolgend die Auslegung des Filters erfolgen. Im ersten Schritt erfolgt die Ermittlung der Basiskovarianzen des Prozess- und Messrauschens R_{v0} bzw. R_{n0} entsprechend Abschnitt 4.2 auf Basis der im Szenario definierten Anregungen. Vom Anwender sind im Anschluss die Faktoren F_{R_n} bzw. F_{R_v} zu definieren, durch die das Vertrauensmaß auf das Modell oder die Messungen gesetzt werden kann. Sind vom Anwender keine Eingaben zu den Faktoren möglich, werden diese der im späteren Verlauf des Filterentwurfs genutzten Gaußprozessoptimierung als weitere Entwurfsparameter übergeben.

Im nächsten Schritt erfolgt die Auswahl des zu nutzenden Filteralgorithmus des CDKF, des CKF bzw. des UKF mit seinen Varianten der Sigma-Punkte-Berechnungsmethoden der SUT oder SST. Wird keine Auswahl vorgenommen, wird das UKF mit dem SST-Verfahren genutzt, da dieses in den Anwendungen dieser Arbeit bessere Ergebnisse als CDKF und CKF erzielt hat und eine geringere Rechenbelastung als das SUT-Verfahren des UKF aufweist. Die jeweiligen Entwurfsparameter des entsprechenden Sigma-Punkte-Verfahrens werden automatisiert der Gaußprozessoptimierung zur Verfügung gestellt. Im Anschluss erfolgt, wie in Abschnitt 4.4.1 vorgestellt, die Berechnung der initialen Stützstellen, die im Anschluss an das Szenario ausgewertet werden und ein initiales Gaußprozesstraining ermöglichen. Nach dem initialen Training erfolgt die Suche des Maximums im Gaußprozess mittels einer gradientenbasierten Optimierung. Nach dem Auffinden eines vermeintlichen Maximums wird die Stützstelle im Szenario ausgewertet

und dem Gaußprozess zugeführt, wodurch dieser sich weiter dem Szenario annähert (siehe auch Abschnitt 4.4.2). Die Optimierung wird bis zum Erreichen der maximalen Funktionsaufrufe durchgeführt bzw. bis das Abbruchkriterium zum gefundenen Optimum im Gaußprozess greift, wie in Bild 4-13 dargestellt. Somit ist es möglich, ohne Eingriff eines Modellierungs- oder Regelungstechnikexperten eine optimale, problemspezifische Einstellung der Filterentwurfsparameter vorzunehmen, wie Bild 5-6 zeigt. Die Ermittlung von problemspezifischen, optimalen Entwurfsparametern des Filters kann durch den Anwender übersprungen und die Entwurfsparameter manuell angegeben werden.

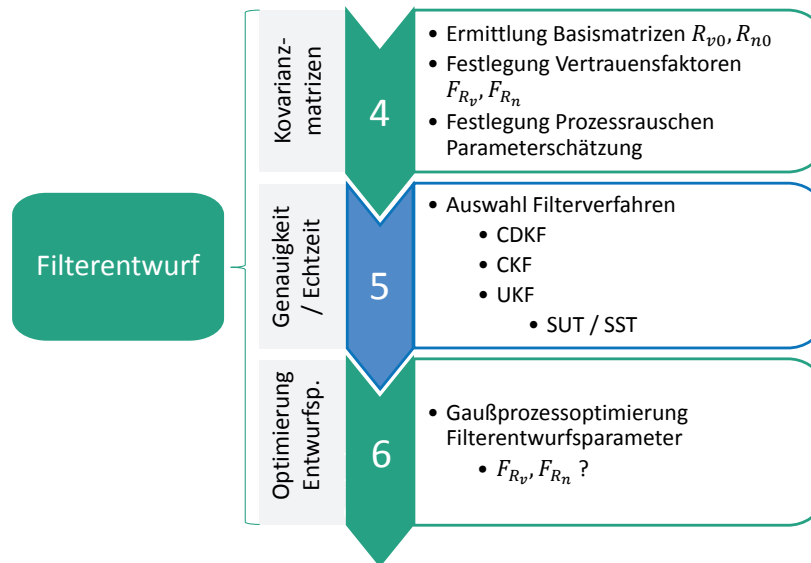


Bild 5-6: Einzelphasen in dem Entwurfsschritt der Filterauslegung mit Ermittlung der Prozess- und Messrauschkovarianzen R_v bzw. R_n ; Auswahl des Sigma-Punkte- bzw. Filterverfahrens und Optimierung der Filterentwurfsparameter θ

Während der gesamten Anwendung des in Kapitel 4 entwickelten modellbasierten Optimierungsverfahrens sind keine Eingriffe vom Anwender vorzunehmen, wodurch die Forderung nach einer maximalen Automatisierung des Entwurfsprozesses und Anforderung III aus Abschnitt 1.3 erfüllt sind. Durch die optimale Einstellung der Entwurfsparameter zur Erfassung von Informationen über höhere statistische Momente wird ebenfalls Anforderung II untermauert.

5.4 Systemtest

Nach der automatisierten Einbindung des Modells über das entwickelte Interface in MATLAB und der daraufhin stattfindenden Auslegung und Ermittlung optimaler Parametrierungen muss die ermittelte Parametrierung hinsichtlich der Filtergüte getestet werden, deren sequentieller Ablauf in Bild 5-7 links dargestellt ist.

Hierzu ist im ersten Schritt der Test des ausgelegten Filters gegen die Anregungsszenarien durchzuführen, die in dem Optimierungsszenario definiert wurden, was dem einfachsten Test der Filterfunktionalität entspricht und eine ausreichende Schätzgüte aufweisen muss. Sollte in diesem Schritt keine zufriedenstellende Filtergüte erreicht worden sein, können verschiedene Gründe vorliegen. Im einfachsten Fall liegen in dem Modell Nichtlinearitäten vor, für die ein Verfahren mit einer höheren Anzahl an Sigma-Punkten erforderlich ist. Da man die Nichtlinearitäten nicht kennt, zeigt sich dies erst nach dem Durchlaufen der Optimierung der Entwurfsparameter, wie im letzten Abschnitt beschrieben. Eine weitere Ursache kann in der Gaußprozessoptimierung selbst liegen. Wie in Abschnitt 4.3.2 vorgestellt, werden im Gaußprozess zur Abbildung des Ein-/ Ausgangsverhaltens des Optimierungsszenarios Kernel-Funktionen genutzt. Das in dieser Anwendung genutzte Verfahren stellt das squared-exponential-Kernel dar, das eine breite Klasse an Modellen abbilden kann. Allerdings kann es passieren, dass trotz modellbasiertem Lernen der Hyperparameter die Kernel-Funktion nicht in der Lage ist, das System zu erfassen, weshalb keine zufriedenstellende Optimierung der Entwurfsparameter des Filters durchgeführt werden kann. Zur Lösung dieses unwahrscheinlichen Falls ist eine adaptive Auswahl von Kernel-Funktionen notwendig, was nicht Gegenstand dieser Arbeit ist. Eine weitere, näherliegende Erklärung liegt in der Beobachtbarkeit des Systems. Zur Erreichung einer zufriedenstellenden Filtergüte müssen die zu schätzenden Zustände des Systems und die zu schätzenden Parameter auf Basis des Messmodells beobachtbar sein, was in der vorliegenden Arbeit nicht automatisiert wurde und in einer Weiterentwicklung zu berücksichtigen ist.

Bei bestandenem Systemtest gegenüber dem Optimierungsszenario kann zu Robustheitstests übergegangen werden. Falls Modelle höherer Güte vorhanden sind, kann dies in Form von Anwendungen des Filters gegen diese Modelle durchgeführt werden, da das im Filter implementierte Modell gegenüber der Realität reduziert ist, um den Berechnungsaufwand für Echtzeitanwendungen in einem vertretbaren Rahmen zu halten. Weitere Testfälle können veränderte Rauschprozesse sein oder eine stark veränderte Parametrierung des Streckenmodells gegenüber dem Filtermodell oder ein Signalausfall bei Messungen. Handlungsempfehlungen bei nicht bestandenem Test sind in Bild 5-7 dargestellt.

Abschließend ist das Filter bei Vorliegen eines Prototyps gegen reale Messdaten zu testen. Treten in diesem Abschnitt Fehler auf, ist im ersten Schritt aufgrund unkorrekt erfasster Nichtlinearitäten im Modell ein höherwertiger Sigma-Punkte-Satz zu wählen bzw. ein Modell höherer Güte, was in beiden Fällen mit erhöhtem Rechenaufwand einhergeht. Führen diese Wege nicht zum Erfolg, ein Filter zur Zustands- und Parameterschätzung mit zufriedenstellender Güte bei gesicherter Beobachtbarkeit mittels des gewählten Messmodells auszulegen, so muss die Modellvalidierung hinterfragt und beispielsweise auf ein automatisiertes Modellidentifikationsverfahren zurückgegriffen werden, wie in [KST14] beschrieben. Im Anschluss muss der Entwurfsprozess erneut durchlaufen werden.

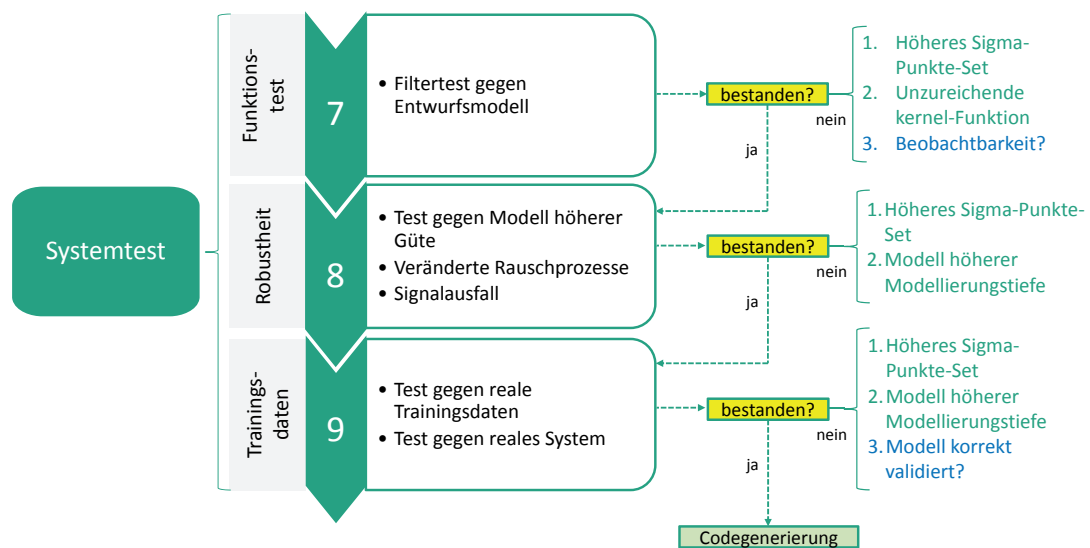


Bild 5-7: Einzelphasen in dem Entwurfsschritt des Systemtests mit Handlungsempfehlungen bei Nichtbestehen einzelner Schritte

5.5 Zusammenfassung Auslegungsmethodik

In diesem Kapitel wurde eine Entwurfsmethodik für Sigma-Punkte-Filter zur Zustands- und Parameterschätzung an Black-Box-Modellen vorgestellt, die eine wesentliche Neuerung dieser Arbeit darstellt. Vom Anwender sind zu Beginn nur das Filtermodell und das Szenario zu laden. Der nachfolgende Entwurfsprozess des Filters erfolgt mit einem maximalen Automatisierungsgrad. Alle notwendigen Parameter für die Filterauslegung und den Filterbetrieb werden automatisiert aus dem Filtermodell gewonnen: Anzahl der Zustände, Eingänge und Ausgänge. Sollten im Modell Schätzparameter erkannt worden sein, die in der Modellierungsphase markiert wurden, so werden diese nebst Initialwerten und Ober- bzw. Untergrenze dem Anwender angezeigt und können als zu schätzende Parameter festgelegt werden. Die Basiskovarianzmatrizen des Prozess- und Messrauschens werden automatisiert auf Basis von definierten Szenarien ermittelt und sind vom Anwender nur noch hinsichtlich Modell- oder Messvertrauen zu skalieren. Sollte vom Anwender keine Aussage getroffen werden, wird das Vertrauensmaß als weiterer Entwurfparameter angesehen. Der zu nutzende Filteralgorithmus kann im Anschluss vom Anwender ausgewählt werden und verbleibt ansonsten auf dem UKF-Filter mit der SST-Sigma-Punkte-Berechnungsmethode. Im nächsten Schritt erfolgt vollautomatisiert die Auslegung der Filterentwurfsparameter anhand der definierten Szenarien mittels einer Gaußprozessoptimierung, die für die Anwendung an ein Sigma-Punkte-Filter angepasst wurde. Abschließend kann die Schätzgüte des entworfenen Filters im Rahmen eines Robustheitstests bewertet werden und bei unzureichender Schätzgüte Handlungsempfehlungen ausgegeben werden. Die gesamte Kommunikation des Filters mit dem Filtermodell erfolgt

auf Basis eines entwickelten Interfaces, mit dem Modelle, die als FMU vorliegen, eingelesen werden können [KST14] [STK13a]. Somit stehen die entwickelten Methoden und Algorithmen nicht nur für die in dieser Arbeit genutzten Dymola-Modelle zur Verfügung, sondern ebenfalls für Modelle aus zahlreichen Werkzeugen wie SimulationX, MapleSim, MATLAB/Simulink, Adams und vielen mehr.

Die in dieser Arbeit entstandene Filterentwurfsmethode, die angepassten Filteralgorithmen und deren Weiterentwicklungen sowie die entwickelte Gaußprozessoptimierung zur Entwurfparameterbestimmung sind in einer MATLAB-Toolbox zusammengefasst, die in A.3 kurz vorgestellt wird. Diese bietet ein Schritt-für-Schritt-Vorgehen von der Einladung des Modells und des Szenarios bis zum Test des entworfenen Filters in Form einer Robustheitsanalyse.

6 Anwendungsbeispiele

In diesem Kapitel werden die in den letzten Abschnitten gezeigten Filteralgorithmen und Entwurfsverfahren zur Zustands- und Parameterschätzung an drei Beispielen erprobt. Abschnitt 6.1 behandelt eine Zustands- und Parameterschätzung an einem Doppelpendelmodell, das sich durch seine geringe Anzahl von sechs Zuständen und seiner starken Nichtlinearität sehr gut eignet, um die entwickelten Verfahren zu erproben und die Funktionsweise darzustellen. In 6.2 wird eine Zustandsschätzung an dem Modell eines Waschautomaten mit unbekannter zu ermittelnder Unwucht gezeigt. Das Modell besitzt eine höhere Systemordnung und ist erheblich komplexer als das Doppelpendel und zeigt so die korrekte Funktionsweise aller entwickelten Werkzeuge auch bei komplexen Systemen. Abschließend wird in 6.3 eine Zustandsschätzung an dem leicht gepanzerten Fahrzeug Dingo 2 durchgeführt, wobei die Anregung des Szenarios aus realen Messungen besteht. Hierdurch wird gezeigt, dass das entwickelte Verfahren und die Anpassung der Sigma-Punkte-Filter auch bei realen Systemen einsetzbar sind.

Alle in diesem Kapitel genutzten Modelle entstammen dem Modellierungstool Dymola. Für eine Einbindung in die Filteralgorithmen und die Auslegung werden allerdings keine Informationen über den inneren Aufbau genutzt und nur auf die in Kapitel 5 ermittelten Größen, die in der FMU hinterlegt sind, zurückgegriffen.

6.1 Doppelpendel

An dem einfachen Beispiel eines Doppelpendels werden in diesem Abschnitt für den Einsatz eines UKF mit SUT-Sigma-Punkte-Berechnungsverfahren gezeigt, welche Auswirkungen ein schlecht gewählter Satz von Entwurfsparametern θ auf die Schätzungsgüte hat und wie ohne Kenntnisse des Modells, der Filteralgorithmen oder der Optimierungsstrategie das in Kapitel 4 entwickelte Verfahren zur Ermittlung von optimalen Entwurfsparametern eine sehr hohe Schätzungsgüte erzielt werden kann. Für die Parameterschätzung wird unter Einsatz eines Joint-UKF mit SUT-Sigma-Punkte-Verfahren die Auswirkung der in Abschnitt 3.5.2 vorgestellten Verfahren zur adaptiven Anpassung der Prozessrauschkovarianzen erprobt. In Abschnitt 6.1.5 wird die in Abschnitt 3.6 entwickelte Erkennung und Korrektur von ungültigen Sigma-Punkten bei der Parameterschätzung vorgestellt und Schätzungen ohne Einbindung von Nebenbedingungen durchgeführt. Abschließend werden die unterschiedlichen Sigma-Punkte-Kalman-Filter aus Kapitel 3: Das UKF mit seinen Varianten der Sigma-Punkte-Berechnung in Form der SUT und SST sowie das CDKF und CKF hinsichtlich Schätzungsgüte bewertet.

6.1.1 Modellbildung Doppelpendel

Das unter Dymola erstellte Modell bildet das stark vereinfachte, dynamische Verhalten eines Doppelpendels auf einem Schlitten ab und besitzt eine Systemordnung von sechs. Die Struktur ist in Bild 6-1 dargestellt:

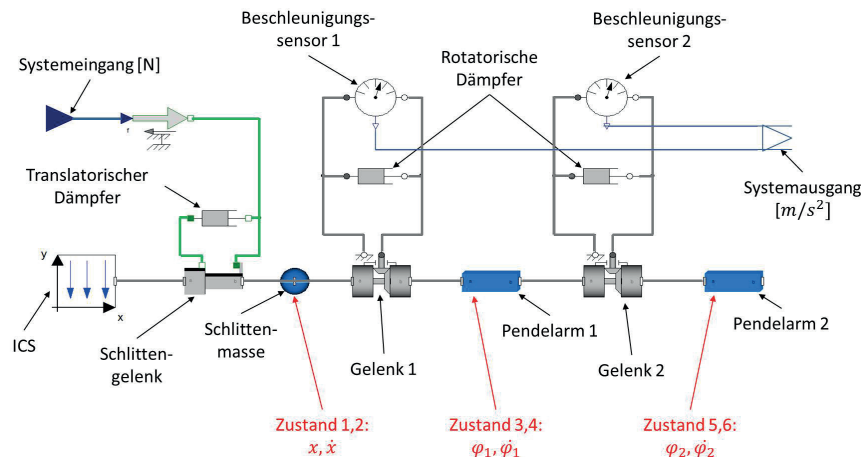


Bild 6-1: Struktur Dymolamodell eines vereinfachten Doppelpendels, bestehend aus ICS, Schlittenmasse und zwei Pendelarmmassen

Das System besteht dabei aus dem Inertialsystem (ICS), an dem das translatorische Gelenk, versehen mit einer viskosen Dämpfung und einem Krafteingang, mit der Schlittenmasse angebracht ist. Die Pendelarme 1 und 2 sind jeweils über rotatorische Gelenke und eine viskose Dämpfung aneinander bzw. an den Schlitten angebunden. Als Messgrößen werden zwei Beschleunigungssignale, welche die Relativrotation in den rotatorischen Gelenken erfassen, genutzt sowie eine Wegmessung des Schlittens.

Zum Zeitpunkt $t = 0$ befinden sich die Pendelarme in der Horizontalen und starten aus dieser Anfangsposition die Bewegung wie in Bild 6-2 zu den Zeitpunkten $t = 0$, $t = 0,9$ und $t = 1,8$ dargestellt. In Blau abgebildet ist dabei die Trajektorie des unteren Pendelarmpunktes.

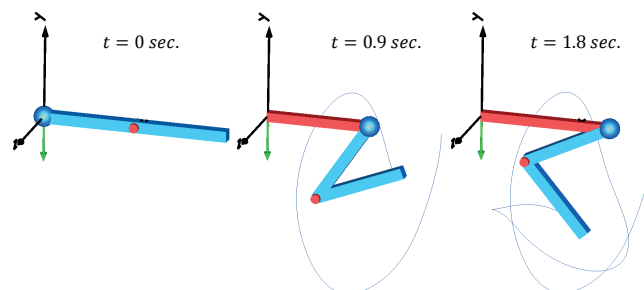


Bild 6-2: Links: Doppelpendel, horizontale Ausgangslage; Mitte: Pendel nach 0,9 Sekunden; Rechts: Pendel nach 1,8 Sekunden

Die translatorische wie auch die rotatorische Dämpfung sind dabei sehr gering, was ein Überschlagen der Pendelarme ermöglicht und somit starke Nichtlinearitäten und das Durchlaufen großer Winkel erzeugt. Die in den folgenden Abschnitten gezeigten Ergebnissen basieren auf simulierten Messwerten. Es werden demnach keine realen Messungen eines Doppelpendels genutzt, sondern Simulationswerte eines Doppelpendelmodells. Um in nachfolgenden Abschnitten die gleichen Bezeichnungen nutzen zu können, werden auch im Falle von simulierten Messwerten diese in Abbildungen als Messung bezeichnet.

6.1.2 Zustands- und Parameterschätzung unter idealen Bedingungen

Die korrekte Funktionsweise der unter Kapitel 5 gezeigten Schnittstelle zur Einbindung des Filtermodells in den Algorithmus, die Übergabe der im Filter berechneten Sigma-Punkte \mathcal{X} als Zustandsvektoren in das als FMU bereitstehende Modell, die Transformation der Sigma-Punkte \mathcal{X} zu \mathcal{X}^* und die Übergabe der transformierten Sigma-Punkte an das Filter werden anhand einer Zustands- und Parameterschätzung unter idealen Bedingungen gezeigt. Ideale Bedingungen heißt, dass sowohl das Streckenmodell vollständig bekannt und somit identisch ist mit dem Filtermodell als auch die Systemausgänge ideal ohne überlagerte Störungen zur Verfügung stehen. Die einzige Unbekannte stellt der Parameter der Dichte des unteren Pendelarmes dar. Bild 6-3 zeigt die korrekte Funktionsweise der entwickelten Einbindungsmethodik und der Schnittstelle anhand der Messungsrekonstruktion und der Parameterschätzung.

Das Schätzsignal der Pendelarmdichte ist nach 0,3 Sekunden eingeschwungen, und somit ist das Filter in der Lage, die zur Verfügung stehenden Systemausgänge ideal zu reproduzieren, was exemplarisch anhand der Winkelbeschleunigung des ersten Gelenkes gezeigt ist.

6.1.3 Optimierter Satz von Entwurfsparametern

Nachfolgend wird die Notwendigkeit der Nutzung eines optimalen Entwurfsparametersatzes für das Sigma-Punkte-Filter gezeigt. Hierzu wird im Gegensatz zum letzten Abschnitt das Streckenmodell gegenüber dem Modell, aus dem die Systemausgänge stammen, hinsichtlich Parametrierung und Anfangszustand unter Dymola abgeändert. Somit stehen zwei Modelle mit unterschiedlichem Verhalten zur Verfügung, wobei ein Modell in das UKF eingebunden wird und das andere zur Erzeugung der Messgrößen genutzt wird. Bei realen Anwendungen entsprechen diese Parameteränderungen immer auftretenden Modellungenauigkeiten, da eine Modellierung nie in der Lage ist ein technisches System vollständig zu beschreiben. In den Systemausgängen spiegelt sich die unterschiedliche Parametrierung in den Beschleunigungssignalen nach ca. 2 Sekunden wider, ebenso in der Messung der Wagenposition, wie in Bild 6-4 ohne Anwendung von Filtertechniken dargestellt.

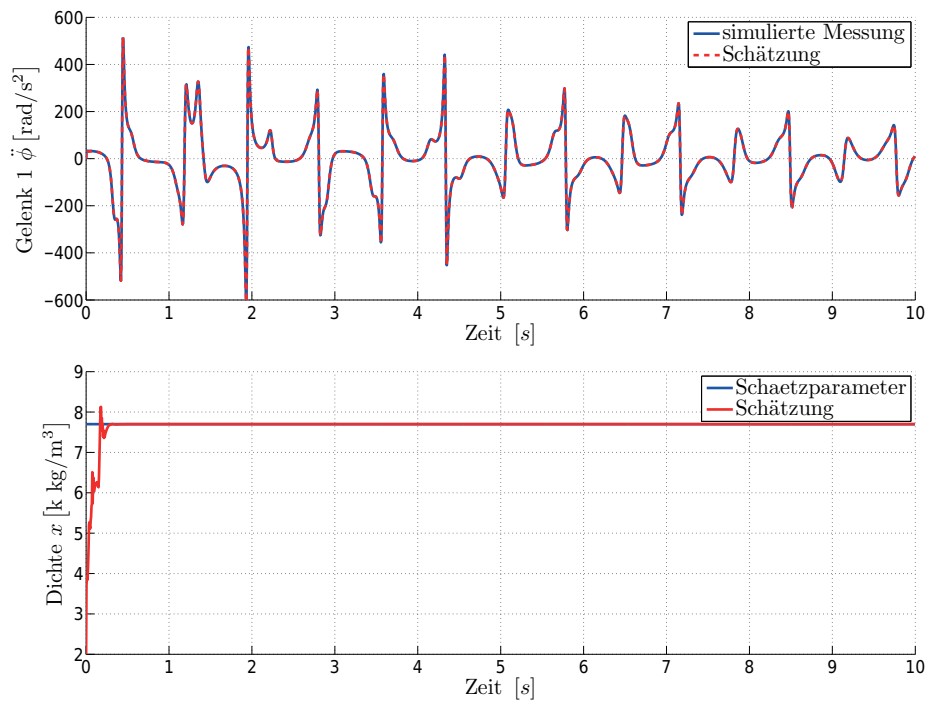


Bild 6-3: Oben: Rekonstruktion eines Systemausganges mittels UKF unter idealen Bedingungen. Unten: Schätzung des unbekannten Systemparameters als Dichte des unteren Pendelarms

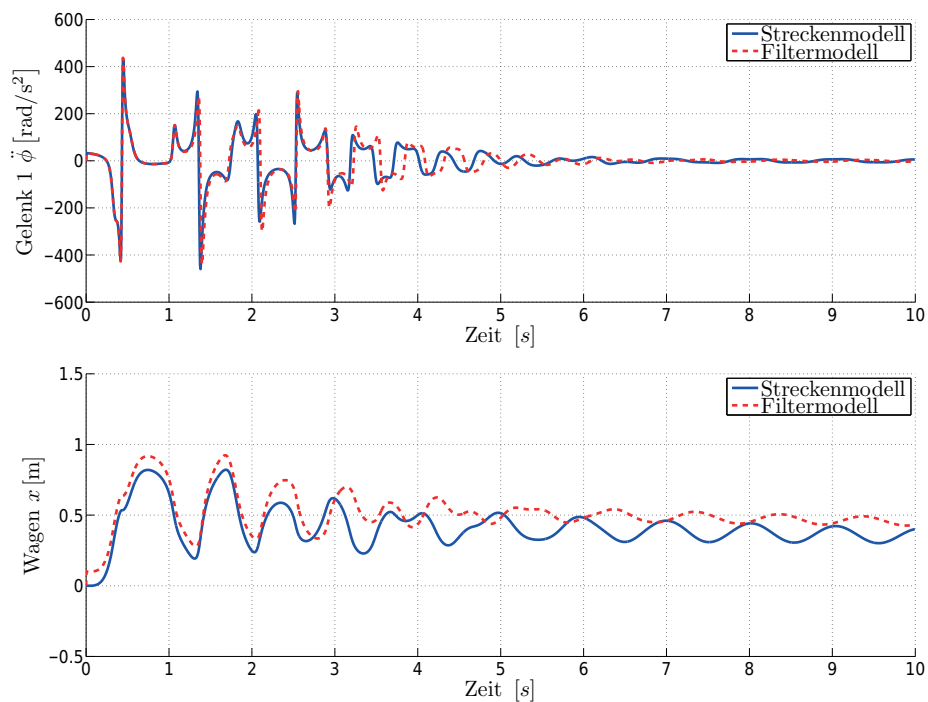


Bild 6-4: Systemausgänge von Streckenmodell (blau) und Filtermodell (rot) aufgrund unterschiedlicher Parametrierung und Startzustandsvektor

Tabelle 6-1: Empfohlene und optimierte Parametrierung eines UKF mit SUT Methode am Doppelpendel

	α	β	κ
θ_a	1e-4	0	-3
θ_b	0.99	2	0
θ_{opt}	0.435	2.14	0.311

Zusätzlich werden die Systemausgänge des Streckenmodells mit additivem weißem Rauschen überlagert. Die Notwendigkeit des Findens eines optimalen Satzes von Entwurfparametern wird an drei unterschiedlichen Parametrierungen eines Unscented-Kalman-Filters mit einem symmetrischen Sigma-Punkte-Verfahren gezeigt. Hierzu werden zwei in der Literatur oft genutzte und empfohlene Parametrierungen θ_a , θ_b und ein optimierter Satz θ_{opt} verglichen, der nach dem in Kapitel 4 entwickelten modellbasierten Optimierungsverfahren ausgelegt wurde und deren Parametrierungen in Tabelle 6-1 dargestellt sind.

Die Unterschiede der Filtergüte zeigt Bild 6-5. In dem oberen Plot ist in Blau der Systemausgang des Streckenmodells markiert und in Grau der Systemausgang, der mit weißem Rauschen überlagert wurden und dem Filter zur Verfügung gestellt wurde. Die Rekonstruktionen des Systemausganges unter den Parametrierungen $\theta_{a,b,opt}$ sind Cyan, Lila und Rot. Der untere Plot zeigt die Schätzung der unbekannten Dichte des unteren Pendelarmes: Blau den zu schätzenden Wert von 7700kg/m^3 und Lila, Cyan und Rot die Schätzungen des UKF unter den angegebenen Parametrierungen. Die empfohlenen Parametrierungen $\theta_{a,b}$ sind hierbei nicht in der Lage, eine brauchbare Schätzung zu liefern bzw. führen bei der Nutzung von θ_b zu einem instabilen Filterverhalten, was zum Abbruch bei ca. 5 Sekunden führt. Trotz der Stabilität des Parametersatzes θ_a und einer guten Parameterschätzung zeigen sich in der Rekonstruktion des verrauschten Systemausganges große Fehler, die sich ebenso in der Schätzung der Systemzustände niederschlagen, wie in Bild 6-6 dargestellt. Eine vollständige Darstellung der rekonstruierten Systemausgänge und der Schätzung der Zustände ist im Anhang in Bild 1-7 bzw. 1-8 zu finden.

Aus diesen Ergebnissen ist einerseits die in Kapitel 4 aufgezeigte Motivation zur Ermittlung eines problemspezifischen Satzes von Entwurfparametern sowie die korrekte Funktionsweise des entwickelten Optimierungsverfahrens zur Filterentwurfparameterbestimmung ersichtlich. Das Filter ist selbst unter den Bedingungen eines nur ungenügend bekannten Systemmodells und verrauschter Messungen in der Lage, den unbekannten Systemparameter der Pendelarmdichte in ausreichender Genauigkeit zu bestimmen. Zwar ist im Gegensatz zur Schätzung unter idealen Bedingungen eine ausreichende Güte erst nach ca 1,5 Sekunden vorhanden, aber trotzdem können Systemausgänge und Zustände nachfolgend mit hoher Güte rekonstruiert bzw. geschätzt werden, im Gegensatz zu den Fällen der Nutzung von θ_a und θ_b .

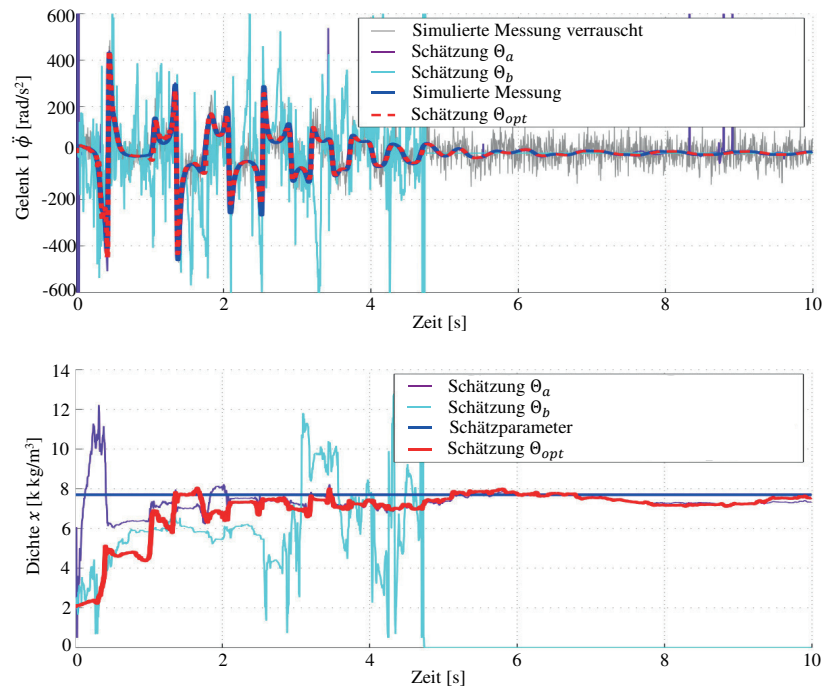


Bild 6-5: Oben: Systemausgang Streckenmodell (blau); Filter zur Verfügung gestellte, verrauschte Systemausgänge (grau), Rekonstruktion Systemausgang UKF mit $\theta_{a,b,opt}$ (lila, cyan, rot). Unten: Schätzung unbekannte Dichte des unteren Pendelarmes durch UKF mit $\theta_{a,b,opt}$ (lila, cyan, rot)

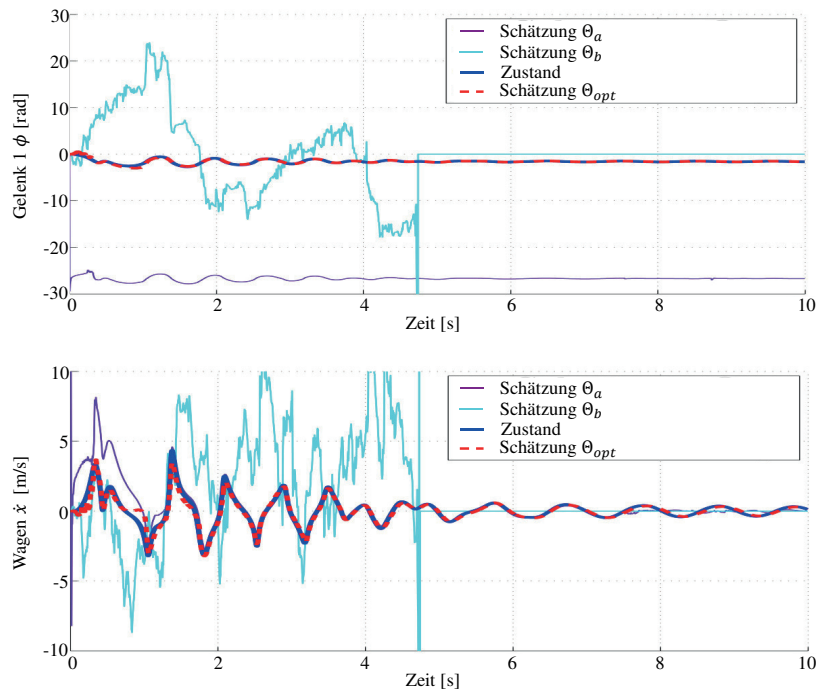


Bild 6-6: Schätzung der Zustände des Doppelpendels bei unterschiedlicher Parametrierung des eingesetzten UKF mit SUT-Sigma-Punkte-Berechnung

6.1.4 Wahl von Prozess- und Messrauschkovarianz

Bei der Nutzung eines Unscented-Kalman-Filters mit symmetrischem oder sphärischem Sigma-Punkte-Ansatz und der Annahme additiver Rauschprozesse gehen das Prozess- und das Messrauschen nach

$$\begin{aligned}
 P_{xx,k}^- &= \sum_{i=0}^{2n} W_i^{(c)} [\mathcal{X}_{i,k|k-1}^* - \hat{x}_k^-] [\mathcal{X}_{i,k|k-1}^* - \hat{x}_k^-]^T + R_v \\
 P_{yy,k} &= \sum_{i=0}^{2n} W_i^{(c)} [\mathcal{Y}_{i,k|k-1} - \hat{y}_k^-] [\mathcal{Y}_{i,k|k-1} - \hat{y}_k^-]^T + R_n
 \end{aligned} \tag{6-1}$$

in den Algorithmus ein (vgl. (3-44), (3-47)) und haben direkten Einfluss auf die Bestimmung der Sigma-Punkte $\mathcal{X}_{k|k-1}$ und die Berechnung der Kalman-Verstärkung \mathcal{K}_k . Verdeutlicht man sich diesen Zusammenhang und schaut sich insbesondere die Systemausgänge in Bild 6-5 an, so erklärt sich das in Abschnitt 4.2 vorgestellte Problem, dass bei einer pauschalen Festlegung der Kovarianzmatrix keine zufriedenstellende Schätzgüte erreicht werden kann. Eine individuelle Bestimmung jeder einzelnen Kovarianz ist allerdings aufgrund der Modellunkenntnis nicht möglich, weshalb das in der vorliegenden Arbeit vorgestellte systematische und automatisierte Verfahren notwendig ist.

Die Systemausgänge der Beschleunigungen liegen in Wertebereichen von $[-400, 400]$ rad/s^2 , wobei die Wagenposition im Wertebereich von $[0, 1]m$ liegt. Die verein-

fachte Annahme von Einheitsmatrizen für das Prozess- oder das Messrauschen R_v und R_n hätte demnach bei den Messungen der Wagenposition einen erheblichen Einfluss, jedoch bei der Messung der Beschleunigungen einen vollkommen zu vernachlässigenden. Bei den Zuständen stellt sich das gleiche Problem, da für die Zustände der Winkelgeschwindigkeiten andere Größenordnungen zu setzen sind als bei der Position des Wagens, was in Bild 6-6 deutlich wird. Die Anwendung der Optimierungsstrategie aus Kapitel 4 auf dieses Problem mit der Erweiterung der UKF-Entwurfsparameter um die Vertrauensfaktoren zur Multiplikation von R_v und R_n lieferte keine zufriedenstellende Schätzgüte. Eine Aufteilung der einzelnen Varianzen, also eine problemspezifische Bestimmung der einzelnen Hauptdiagonalelemente von R_v und R_n , ist allerdings bereits bei moderaten Systemgrößen nicht mehr von Hand durchführbar und stellt für Optimierungsalgorithmen eine große Herausforderung dar.

Die Anwendung des in Abschnitt 4.2 entwickelten Verfahrens zur Bestimmung einer Basis R_{v0} bzw. R_{n0} , die auf der Hauptdiagonalen den maximalen Gradienten der Systemausgänge und Zustände über einen Abtastschritt entspricht, führt auf folgende Matrizen:

$$R_{n0} \approx \begin{bmatrix} 262 & 0 & 0 \\ 0 & 198 & 0 \\ 0 & 0 & 0.431 \end{bmatrix} \quad (6-2)$$

$$R_{v0} \approx \begin{bmatrix} 0.07 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1.23 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.14 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2.27 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.02 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.53 \end{bmatrix}. \quad (6-3)$$

Insbesondere die Matrix R_{n0} verdeutlicht nochmals den erläuterten Zusammenhang. Die Nutzung dieser sehr einfach automatisiert zu ermittelnden Matrizen lässt die Nutzung von einfachen Vertrauensfaktoren F_{R_n} und F_{R_v} in der Form

$$R_n = R_{n0} \cdot F_{R_n} \quad (6-4)$$

$$R_v = R_{v0} \cdot F_{R_v} \quad (6-5)$$

zu, wodurch die Möglichkeit gegeben ist, in einer einfachen Form Vertrauen in Messung oder Streckenmodell anzugeben. Neben dieser einfachen Form besteht weiterhin die Möglichkeit, Einfluss auf einzelne Einträge der ermittelten Basismatrizen zu nehmen, um zusätzlich differenzierte Gewichtungen bei gleichen Wertebereichen, aber unterschiedlichen Störungen vornehmen zu können.

Der Wahl der Kovarianz einer Parameterschätzung ist nach Abschnitt 3.5.2 besondere Aufmerksamkeit zu widmen, da insbesondere im Falle der Annahme eines stückweise konstanten zu schätzenden Parameter die definierte Prozessrauschkovarianz die einzige Möglichkeit für das Filter ist, Einfluss auf die Schätzung zu

nehmen. Aber auch die konstante Annahme einer Prozessrauschkovarianz erzielt nach der Begründung in 3.5.2 keine optimalen Ergebnisse und führte zur Vorstellung der Fensterfunktion nach Nelson und der Robbins-Monro Methode, deren Ergebnisse in Bild 6-7 dargestellt sind. Die besten Ergebnisse wurden bei der Nutzung der Robbins-Monro-Methode mit einem Faktor von $\alpha_{RM} = 0,9$ erzielt, was einer starken Gewichtung der Messresiduen unter Einbeziehung der aktuellen Kalman-Verstärkung entspricht (siehe (3-60)). Die Nutzung der Fensterfunktion mit unterschiedlichen Vergessensfaktoren λ_{RLS} führte zu keiner Verbesserung der Schätzgüte, ebenso wenig wie die Nutzung einer konstanten Prozessrauschkovarianz für den Schätzparameter oder der Robbins-Monro-Methode mit $\alpha_{RM} = 0,1$. Aus diesem Grund wird für die nachfolgenden Vorstellungen der Anwendungsbeispiele die Robbins-Monro-Methode mit einem Vergessensfaktor von $\alpha_{RM} = 0,9$ für die Bestimmung der Parameterkovarianz eingesetzt.

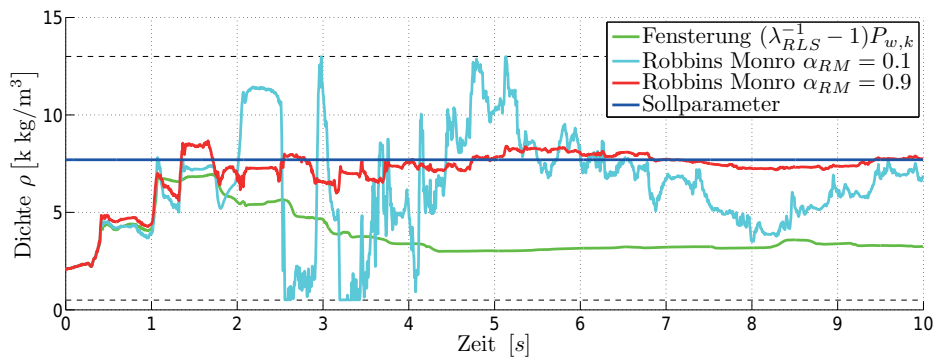


Bild 6-7: Parameterschätzungen der Pendelarmdichte beim Einsatz unterschiedlicher Kovarianzbestimmungsverfahren; Fensterung $(\lambda_{RLS}^{-1} - 1)P_{w,k}$ (grün), Robbins-Monro $\alpha_{RM} = 0,1$, Robbins-Monro $\alpha_{RM} = 0,9$, Sollparameter (blau)

6.1.5 Einbeziehung von Nebenbedingungen

In Abschnitt 3.6 wurde eine Möglichkeit entwickelt, Nebenbedingungen in die Algorithmen von Sigma-Punkte-Filtern einfließen zu lassen. Die Notwendigkeit und die Funktionalität werden nachfolgend gezeigt.

Insbesondere bei der Schätzung von Systemparametern liegen oftmals Informationen über den zu bestimmenden Parameter vor. Wenn es sich um die Bestimmung von Masse, Federsteifigkeit, Dichte oder ähnlichen Größen handelt, kann zumindest ausgeschlossen werden, dass der Parameter Werte im negativen Bereich annimmt. Weiter kann oftmals eine Beschränkung der Wertebereiche angegeben werden. So kann beispielsweise ungefähr abgeschätzt werden, welche Masse bzw. welche Dichte ein Körper im System besitzt oder was für eine Steifigkeit für beispielsweise ein Fahrwerk angenommen werden kann. Diese Nebenbedingungen lassen sich für die A-priori- und A-posteriori-Schätzung wie auch für die zu

transformierenden Sigma-Punkte in der Form

$$\begin{aligned} \mathcal{M} &= \{\mathcal{X}_{i,n_p,k-1}, \mathcal{X}_{i,n_p,k|k-1}, \hat{x}_{n_p,k}^-, \hat{x}_{n_p,k}\} \\ NB_{P,o} &> m > NB_{P,u} \quad \forall \quad m \in \mathcal{M} \end{aligned} \quad (6-6)$$

angeben, mit n_p als der Anzahl der zu schätzenden Parameter und $i = 1, \dots, 2(n + n_p) + 1$ als i -tem Sigma-Punkt. In den letzten Abschnitten wurde gezeigt, dass ein optimal gewählter Satz von Entwurfsparametern des Schätzalgorithmus sehr gute Ergebnisse in der Schätzung des Parameters und der Zustände sowie der Rekonstruktion der Messungen liefert. Ist dies allerdings nicht gewährleistet, beispielsweise durch das Auftreten von veränderten Rauschprozessen oder Anregungsszenarien, die im Vorfeld nicht trainiert wurden, kann nicht garantiert werden, dass der gefundene Entwurfsparametersatz optimal ist. In diesem Anwendungsfall wurde das Optimierungsverfahren zur Ermittlung der Entwurfsparemeter mit einem Anregungsszenario durchgeführt, bei dem das Doppelpendel nur kurz vor seiner stabilen Ruhelage losgelassen wird und somit wesentlich kleinere Winkel und Beschleunigungen durchläuft. Die Anwendung der Parameter auf das bereits bekannte Anwendungsszenario mit Überschlagen der Pendelarme wird hier in Form einer Robustheitsanalyse zur Verdeutlichung der Parametergrenzen herangezogen.

Bild 6-8 zeigt die Rekonstruktion eines verrauschten Systemausganges sowie die Schätzung des unbekannten Systemparameters der Dichte des unteren Pendelarmes. Bei einer vollständigen Vernachlässigung der Parametergrenzen sind die Ergebnisse in Lila abgebildet. Hierbei ist gut zu erkennen, dass nach 2,5 Sekunden das Filter divergiert und abbricht, da keine positive Definitheit der Kovarianzmatrix mehr gegeben ist, was auf die Berechnung von negativen Massen der Sigma-Punkte basiert und zu einer Instabilität des Filtermodells geführt hat. Unter A.4 in Bild 1-9 ist in Cyan die Schätzung bei einer unteren Nebenbedingung dargestellt, die besagt, dass die Dichte des unteren Pendelarmes nicht kleiner als $0,5 \text{ kg/m}^3$ sein kann. Hierbei bleibt das Filter zwar stabil, allerdings ist es nicht in der Lage, die Messung in ausreichender Güte zu rekonstruieren, was aus der im unteren rechten Plot dargestellten Schätzung des Parameters resultiert. Diese kann aufgrund der Nebenbedingung zwar keine nichtphysikalischen Werte in Form von negativen Dichten annehmen, allerdings steigt die Schätzung auf über 30 kg/m^3 und ist nicht in der Lage, eine vernünftige Bestimmung der unbekannten Dichte zu ermöglichen. In Rot dargestellt ist die Schätzung bei eingebundenen unteren und oberen Nebenbedingungen. Im Bereich von ca. 1,8 bis 4 Sekunden gerät die Schätzung in die obere und untere Parametergrenze und wird in dieser gehalten, was im Folgenden ab ca. 4 Sekunden zu einer stabilen Schätzung des Systemparameters führt, wie im unteren rechten Plot von Bild 6-8 ersichtlich ist.

Neben der Begrenzung der Schätzung, die als Erwartungswert zur Berechnung der Sigma-Punkte dient, ist ebenso eine Begrenzung der Sigma-Punkte selbst notwendig, da diese als Zustandsvektoren in das Systemmodell eingesetzt werden und als Startwert der Simulation dienen. Eine Besetzung von nicht-physikalischen

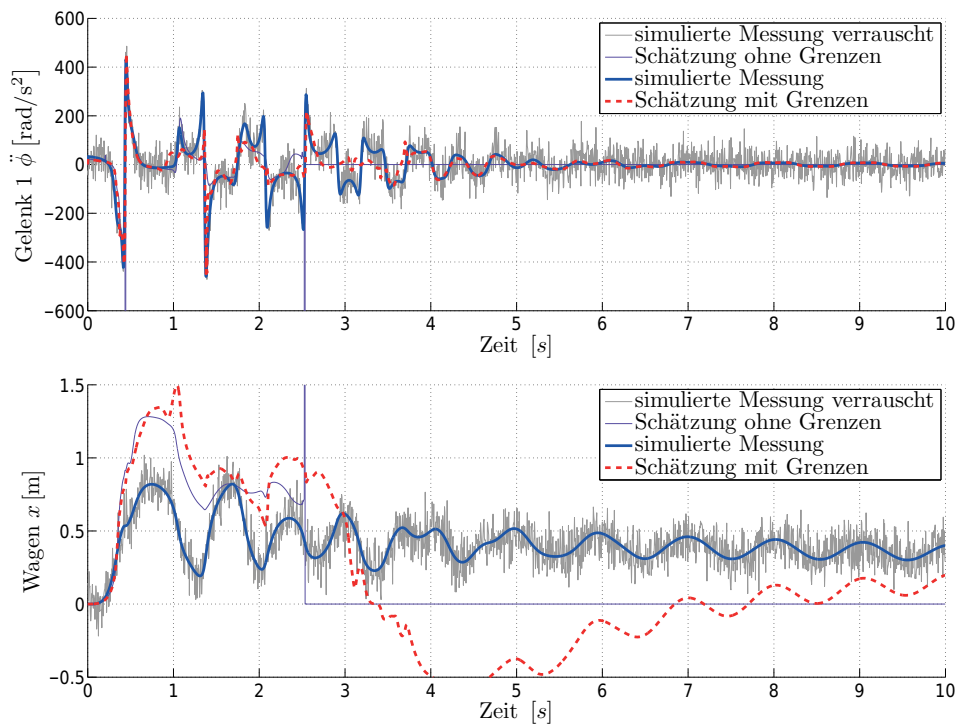


Bild 6-8: Systemausgänge des Streckenmodells (blau); dem Filter zur Verfügung gestellte Systemausgänge, mit Rauschen überlagert (grau), Filterrekonstruktion des Systemausganges durch UKF ohne Parameternebenbedingungen (lila) und vollen Nebenbedingungen (rot).

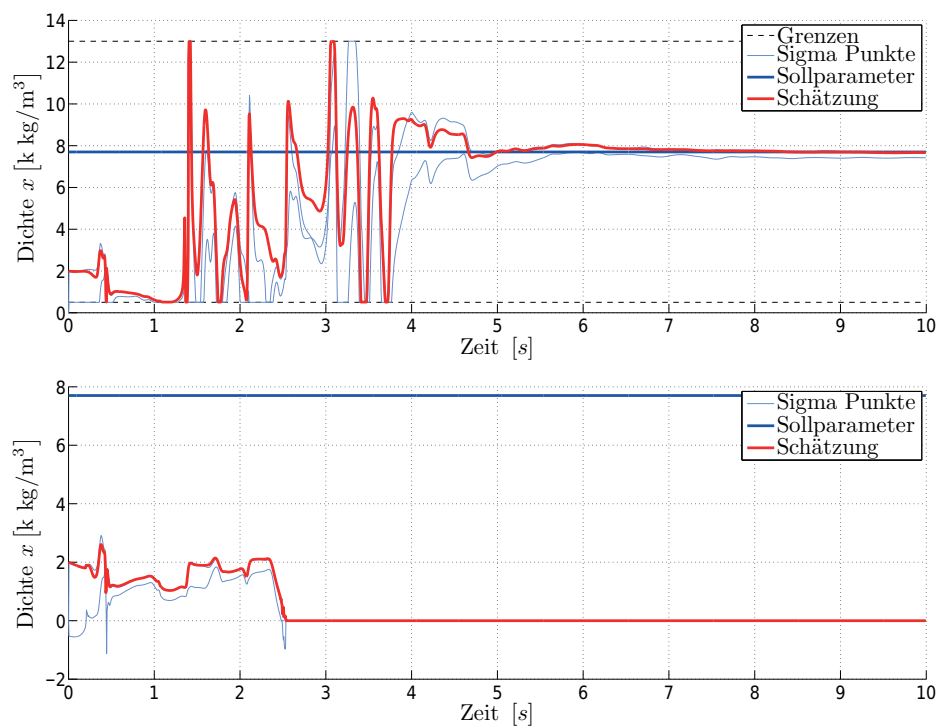


Bild 6-9: Schätzung des unbekannten Systemparameters (roter Verlauf) der Pendelarmdichte (dunkelblauer Verlauf) bei oberer und unterer Nebenbedingung (oben) und ohne Nebenbedingung (unten) mit Einbeziehung der berechneten Sigma-Punkte (hellblauer Verlauf)

Werten wie negativen Dichten kann hier zu unabsehbaren Ergebnissen führen. In Bild 6-9 sind die Schätzungen mit Nebenbedingungen (oben) und ohne Nebenbedingungen (unten) abgebildet. In Rot ist die Schätzung des Filters abgebildet, in Dunkelblau die Sollwerte der Schätzung, in Hellblau die Sigma-Punkte-Varianzen unter Vernachlässigung der Kovarianzen, was im dargestellten Fall

$$\begin{aligned}\mathcal{X}_{k-1,7} &= \hat{x}_{k-1} + \gamma \left(\sqrt{P_{xw,k}} \right)_7 \\ \mathcal{X}_{k-1,15} &= \hat{x}_{k-1} - \gamma \left(\sqrt{P_{xw,k}} \right)_7\end{aligned}\tag{6-7}$$

entspricht. Unter Einbeziehung der Nebenbedingungen ist im oberen Bild gut zu erkennen, dass neben der Begrenzung der eigentlichen Schätzung auch die Sigma-Punkte $\mathcal{X}_{k-1,7}$ und $\mathcal{X}_{k-1,15}$, in Hellblau dargestellt, begrenzt werden und es somit zu keiner Simulation von negativen Dichten kommt. Im unteren Bild ist die Schätzung ohne Begrenzung dargestellt. Zum Zeitpunkt von ca. 2,5 Sekunden wird das Modell aufgrund der berechneten Sigma-Punkte und der negativen Massenschätzung instabil, was zum Abbruch des Filters führt. Im Anhang ist in Bild 1-10 nur eine eingebundene untere Nebenbedingungen dargestellt, sodass eine Schätzung von ca. 60000 kg/m³ entsteht und Sigma-Punkte bis zu 80000 kg/m³ berechnet werden.

6.1.6 Vergleich unterschiedlicher Filtervarianten

Im Kapitel 3 wurden unterschiedliche Vertreter von Sigma-Punkte-Filtern vorgestellt, wie das Unscented-Kalman-Filter mit dem symmetrischen und dem spärlichen Satz von Sigma-Punkten, die in dieser Arbeit verwendet wurden. Als weitere Vertreter der Klasse der Sigma-Punkte Filter wurden das Central-Difference- sowie das Cubature-Filter implementiert, die neben der Nutzung einer veränderten Berechnungsvorschrift der Sigma-Punkte mit nur einem Entwurfsparameter auskommen und somit die Suche nach einem optimalen Satz vereinfachen.

Schätzgüte und Rechendauer der in dieser Arbeit genutzten vier Filtervarianten werden im Folgenden anhand der Parameterschätzung am Doppelpendel verglichen. Als Grundlage dienen unterschiedliche Strecken- und Filtermodelle, die in Abschnitt 6.1.3 vorgestellt wurden und deren, dem Filter zur Verfügung gestellte Messungen ebenfalls mit Rauschprozessen überlagert wurden.

Bild 6-10 zeigt die Schätzung der unbekannten Dichte des unteren Pendelarmes. Anhand der Verläufe der unterschiedlichen Filterimplementierungen ist zu erkennen, dass das Unscented-Kalman-Filter mit symmetrischem Sigma-Punkte-Satz die besten Resultate nach Anwendung des modellbasierten Lernverfahrens liefert. Die Verläufe zeigen, dass durch die Möglichkeit der Einbringung von Wissen über die Verteilungen der Zustände ein besseres Ergebnis erzielt werden kann. Dies ist im Fall des Unscented-Kalman-Filters durch die Entwurfsparameter κ und β möglich, die im entwickelten Lernverfahren anhand von Anwendungsszenarien

Tabelle 6-2: Standardabweichung und Rechenzeit beim Einsatz unterschiedlicher Filtervarianten

	Standardabweichung	Rechenzeit [s]
UKF symmetrisch	1.268	3.82
UKF asymmetrisch	1.801	2.75
Central-Difference-KF	3.260	3.91
Cubature-KF	1.727	3.87

trainiert wurden. Bei dem Einsatz des Cubature- als auch des Central-Difference-Filters steht nur ein Freiheitsgrad über einen Entwurfparameter zur Verfügung, über den die Streuung der Sigma-Punkte um den Erwartungswert beeinflusst, allerdings kein Einfluss auf höhere Momente der Verteilung gewonnen werden kann. Aufgrund der ähnlichen Struktur und der daraus resultierenden Ergebnisse wird im folgenden auf eine explizite Darstellung der Ergebnisse des Cubature-Filters verzichtet. Die vereinfachte Auslegung dieser Filter ist bei der Anwendungsmöglichkeit eines automatisierten Optimierungsverfahrens der Filterentwurfparameter, wie in Kapitel 4 vorgestellt, nicht mehr gegeben und zeigt zudem schlechtere Ergebnisse. Die Schätzung mittels eines Unscented-Kalman-Filters und der SST Sigma-Punkte-Berechnungsmethode ist in Rot dargestellt und liefert leicht schlechtere Ergebnisse als das Cubature-Filter, wie auch in Tabelle 6-2 anhand der Standardabweichung der Parameterschätzung zu sehen ist. Dies resultiert trotz Anwendung des Optimierungsverfahrens aus der asymmetrischen Verteilung der Sigma-Punkte und dem damit verbundenen Fehler in der Erfassung des dritten Momentes einer Verteilung. Allerdings ist die Standardabweichung der Parameterschätzung nur geringfügig schlechter als die des Cubature-Filters, wobei jedoch die Berechnungszeit für das dargestellte Szenario um annähernd 30 % reduziert werden konnte.

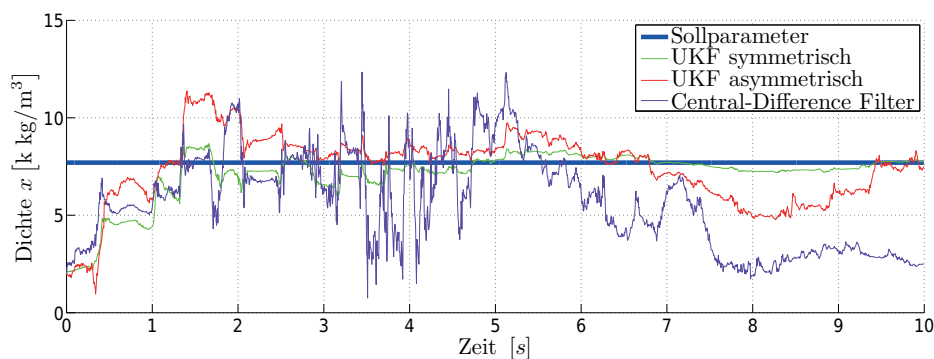


Bild 6-10: Vergleich der unterschiedlichen Sigma-Punkte-Kalman-Filter bei der Parameterschätzung am Doppelpendel. Grün: UKF-SUT, Rot: UKF-SST, Lila: CDKF

Das Ergebnis der unterschiedlichen Güte in der Parameterschätzung ist auch in Bild 6-11 ersichtlich. Insbesondere in der Rekonstruktion der Messung der Wagenposition zeigt sich eine starke Abweichung des Cubature-Filters, der erst für eine längere Simulationszeit wieder dem realen Verlauf entspricht. Das Central-Difference-Filter sowie das Unscented-Kalman-Filter mit asymmetrischer Sigma-Punkte-Verteilung zeigen Schwächen in der Rekonstruktion. Das Unscented-Kalman-Filter mit symmetrischem Sigma-Punkte-Ansatz zeigt auch hier die besten Resultate, was sich auch in der Schätzung der Zustände widerspiegelt, wie in Bild 6-12 ersichtlich ist.

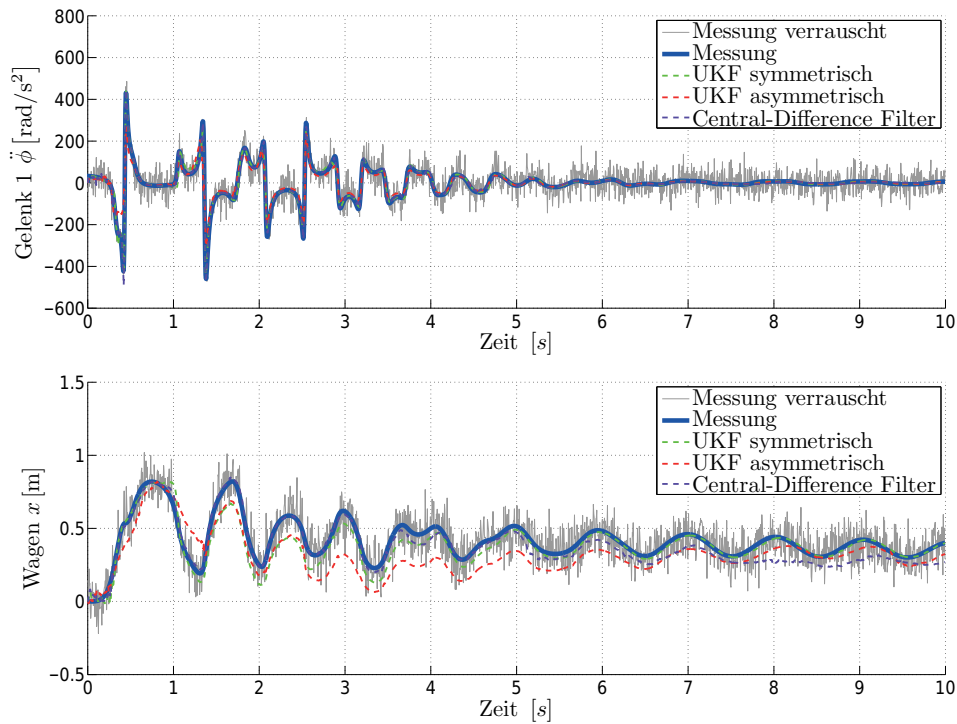


Bild 6-11: Vergleich der unterschiedlichen Sigma-Punkte-Kalman-Filter bei der Rekonstruktion verrauschter Messungen am Doppelpendel. Grün: UKF-SUT, Rot: UKF-SST, Lila: CDKF

Als ein Ergebnis dieser Arbeit bleibt festzuhalten, dass die Parameter- und Zustandsschätzung mit Hilfe von Sigma-Punkte-Filtern anhand des gezeigten Beispiels die besten Resultate durch den Einsatz eines Unscented-Kalman-Filters mit symmetrischem Sigma-Punkte-Satz lieferte. Die Möglichkeit, durch eine erhöhte Anzahl von Entwurfparametern Einfluss auf die Erfassung der Dichtefunktion zu nehmen, ist durch das entwickelte modellbasierte Lernverfahren gegeben und zeigt sehr gute Ergebnisse. Der Entwurfsprozess ist hierdurch trotz unbekanntem Systemmodell und ohne Expertenkenntnisse über das Filter möglich und bringt somit keine Nachteile gegenüber Filtervarianten mit einer geringeren Anzahl von Entwurfparametern. Das Unscented-Kalman-Filter mit asymmetrischem Sigma-Punkte-Satz zeigt nach Tabelle 6-2 fast gleichwertige Filterergebnisse wie das

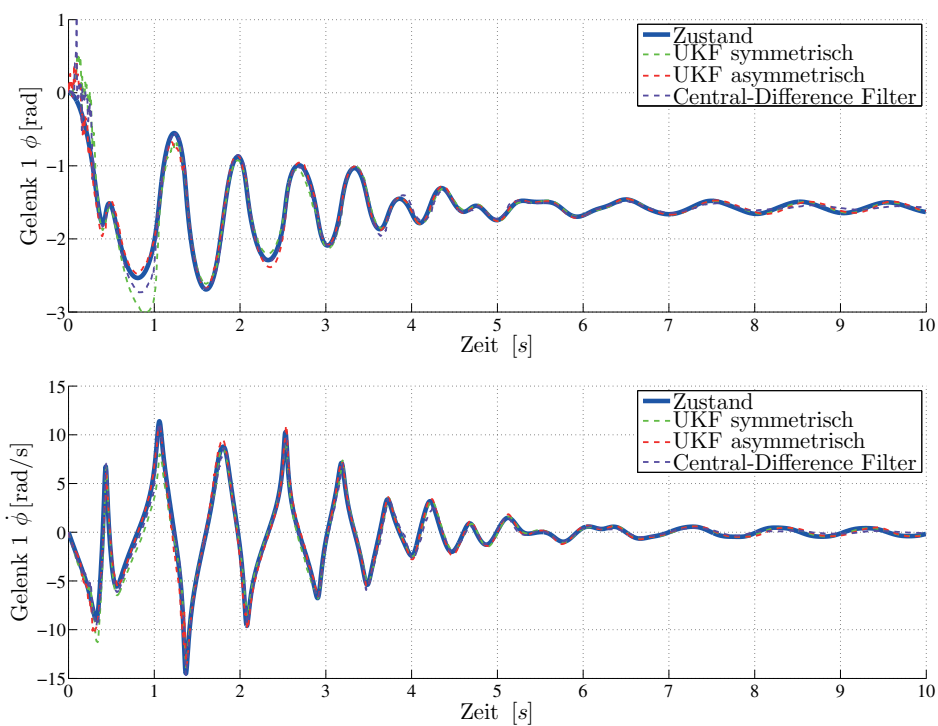


Bild 6-12: Vergleich der unterschiedlichen Sigma-Punkte-Kalman-Filter bei der Zustandsschätzung am Doppelpendel. Grün: UKF-SUT, Rot: UKF-SST, Lila: CDKF

Central-Difference und das Cubature-Filter bei einer allerdings um 30 % verkürzten Rechendauer.

Im Folgenden werden die Ergebnisse dieses Abschnittes anhand zweier weiterer Anwendungsfälle validiert, wobei auf die Darstellung der Ergebnisse des CDKF und des CKF verzichtet wird, da qualitativ die gleichen Resultate wie beim vorgestellten Doppelpendel entstehen und das UKF mit SUT Sigma-Punkte-Methode die höchste Schätzungsgüte liefert.

6.2 Schätzung zeitvarianter Unwucht in Waschautomaten

Ein komplexeres Anwendungsbeispiel ist durch die Zustands- und Parameterschätzung an einem Waschautomaten gegeben. Dieser besteht in seinen Hauptkomponenten aus einem Gehäuse und einem schwingungsfähig gelagerten Aggregat im Gehäuse. Das Aggregat besteht wiederum aus den einzelnen Elementen des Laugenbehälters, der die Waschlauge beinhaltet, und der Trommel, in welcher die Wäscheposten eingebracht werden. Die Verbindung zwischen Trommel und Laugenbehälter erfolgt durch ein Rotationsgelenk im Laugenbehälterboden, das idealisiert nur die Relativrotation um die Längsachse von Trommel und Laugenbehälter zulässt und somit die Waschrhythmen und den Schleuderprozess ermöglicht. Die Schleuderphase schließt nach dem eigentlichen Waschen an und dient der Reduktion der Restfeuchte in der Wäsche. Hierbei legt sich bei hohen Trommeldrehzahlen von bis zu 1600 rpm die Wäsche an die Trommelwand an, und der Feuchtegrad kann aufgrund der hohen Fliehkräfte reduziert werden.

In der Schleuderphase kann es durch eine ungleichmäßige Verteilung der Wäsche innerhalb der Trommel, die aus etwa ungünstiger Beladung mit wenigen Einzelstücken oder der Bildung von Wassertaschen resultiert, zur Ausbildung einer Unwucht kommen. Beim Durchlaufen der Resonanzen, die im unteren Nutzdrehzahlbereich der Schleuderphase liegen, kann es durch eine Unwucht zum Anschlagen des Aggregates an das Gehäuse kommen, wodurch Bauteile beschädigt werden oder im schlimmsten Fall der Waschautomat versetzt wird. Im hohen Drehzahlbereich treten geringere Schwingungsamplituden des Aggregates auf, dafür allerdings große, dämpfungsbedingte Kräfte, die zu einem Wandern des Automaten führen können, mit der Gefahr des Abreißen von Wasserleitungen. Aus diesem Grund muss im Entwurfsprozess der Lagerung des Aggregates innerhalb des Gehäuses hohe Aufmerksamkeit gewidmet werden. Als Lagerungsmöglichkeiten existieren eine Vielzahl an Varianten am Markt, wie Zugfedern an der Oberseite und/oder an den Seiten sowie Dämpfern an der Unterseite des Aggregates. Im vorliegenden Testsystem wird das Aggregat mittels drei Federbeinen an der Laugenbehälterunterseite im Gehäuse gelagert und die Momentenaufnahme der Federbeine in den Koppelpunkten durch Elastomere realisiert. Die Auslegung der dynamischen Eigenschaften der Federbeinkomponenten wie Steifigkeit, Dämpfungskraft, Elastomersteifigkeit und -dämpfung, stellt dabei einen Zielgrößenkonflikt zwischen minimalen Aggregatsauslenkungen innerhalb des Gehäuses,

um maximalen Bauraum für Zusatzkomponenten zu ermöglichen und minimalen Bodenkkräfte dar, die zu einem Wandern des Gesamtautomaten führen können. Die Abstimmung der dynamischen Koppelemente stellt dabei einen sehr wichtigen Schritt des Entwurfs dar, da die resultierenden Resonanzen des Aggregates im Nutzfrequenzbereich der Schleuderphase liegen und somit bei vorliegender Unwucht vollständig angeregt werden. Erschwert wird die Auslegung der dynamischen Komponenten durch stark variierende Parameter des schwingungsfähigen Systems. Die Masse der Beladung schwankt von einigen wenigen kg bis zu 30 kg mit unterschiedlichsten Verteilungen innerhalb der Trommel, die den Trägheitstensor und damit das gesamte Schwingungsbild des Aggregates beeinflussen. Weitere Effekte resultieren aus der Abdichtung des Laugenbehälters gegenüber der Umgebung an der Vorderseite durch den Türdichtring, der ein Elastomerbauteil darstellt und neben seinen frequenz- und amplitudenabhängigen Steifigkeits- und Dämpfungseigenschaften stark ausgeprägte Formfaktoren aufweist, die das dynamische Verhalten durch Aufklappvorgänge beeinflussen.

Erkennung und Bestimmung einer aufgetretenen Unwucht können in diesem Fall als Anhaltspunkt genutzt werden, ob ein Schleuderprozess abgebrochen werden sollte oder als Regelgröße für aktive Verfahren genutzt werden kann. Nachfolgend wird die Erkennung einer bestehenden Unwucht an einem Modell vorgenommen, da in diesem die Unwucht beliebig eingestellt werden kann und die Zustände zur Verfügung stehen, womit die Filtergüte gezeigt werden kann.

6.2.1 Modellbildung Waschautomat

Das unter Dymola erstellte Modell des Waschautomaten besitzt eine Systemordnung von 18 und besteht im Wesentlichen aus den Starrkörperelementen des Gehäuses, des Laugenbehälters und der Trommel. Das Gehäuse ist in diesem Fall fest mit der Umwelt verbunden und stellt das Inertialsystem dar. Der Laugenbehälter ist mittels dreier Federbeine gelagert, wie in Bild 6-13 dargestellt, wobei die Momentenaufnahme der Federbeinkopplungspunkte zum Gehäuse und Laugenbehälter durch Elastomere erfolgt, die stark nichtlineare frequenz- und amplitudenabhängige Steifigkeiten und Verlustwinkel aufweisen und als auf Messungen basierende Kennfelder im Modell realisiert wurden. Die Federbeine bestehen aus einer Parallelschaltung von Feder und Dämpfer und beinhalten weitere Nichtlinearitäten in Form von Freilauf und Haftung im Dämpferzweig sowie abheben-den Federn bei starker Anregung des Aggregates. Der Türdichtring besteht aus einer Kennlinie des verwendeten Elastomers und zeigt ebenfalls stark frequenz-, amplituden- und bewegungsformabhängiges dynamisches Verhalten. Die in das System eingebrachte Unwucht stellt durch ihr zeitvariantes Verhalten, das die Ausprägung und das Entwässern darstellt, ebenfalls eine starke Nichtlinearität im System dar.

Laugenbehälter und Trommel werden vereinfachend als Starrkörper angesehen, wobei die Kopplung mittels eines einfachen Rotationsgelenkes realisiert wird. So-

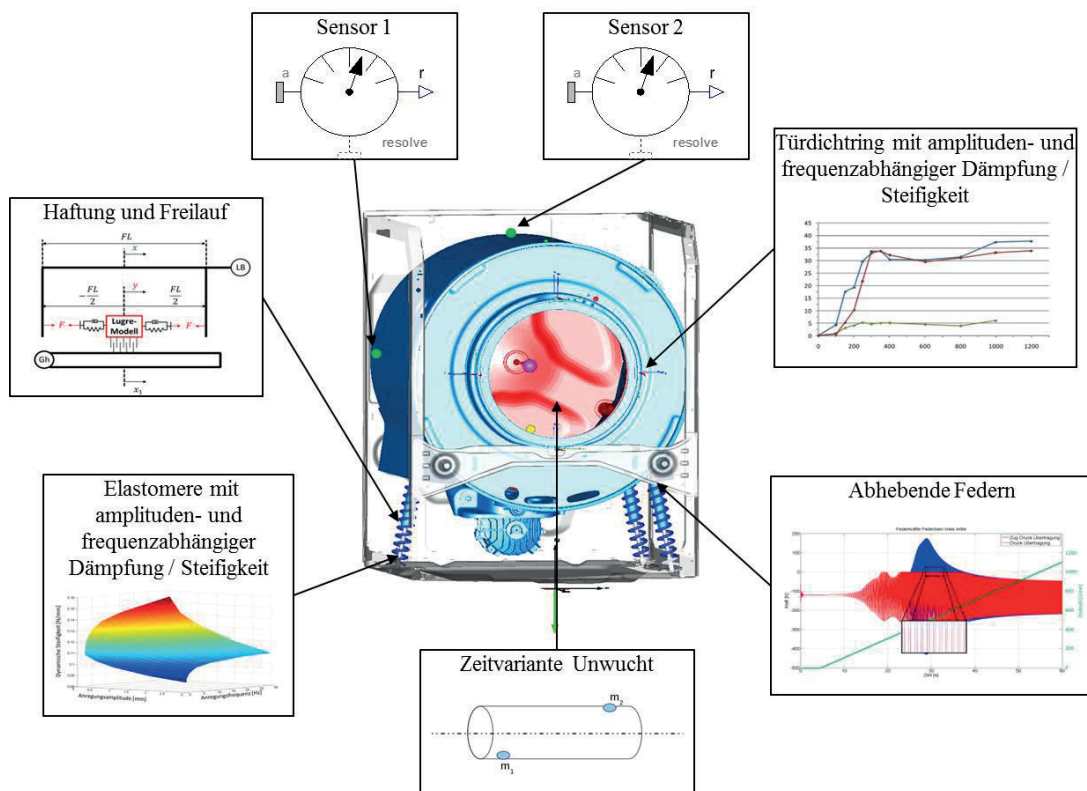


Bild 6-13: Strukturbild des unter Dymola modellierten Waschautomaten mit angegebenen Nichtlinearitäten und Sensoren

mit besitzt der Laugenbehälter alle sechs Freiheitsgrade: die der Translation in x-, y- und z-Richtung und die der Rotation um alle Achsen. Die Trommel besitzt nur noch den Freiheitsgrad der Relativrotation zum Laugenbehälter um ihre Längsachse. Hieraus resultieren insgesamt 14 Zustände, die der Translation und der Rotation von Laugenbehälter und Trommel im Schwerpunkt sowie die Relativrotation der Trommel zum Laugenbehälter. Weitere drei Zustände entstammen der Modellierung des Freilaufes der Federbeine und ein weiterer der Drehzahlregelung, die als PI-Regler realisiert ist und somit die Gesamtordnung von 18 ergibt.

Die Messung von Systemgrößen erfolgt mittels zwei 3-axialen Beschleunigungsaufnehmern, die auf dem Mantel des Laugenbehälters am hinteren Ende auf 12 und 9 Uhr-Position angebracht sind. Zusätzlich wird die Drehzahl der Trommel gemessen, da diese zur Realisierung der Drehzahlregelung bekannt sein muss.

6.2.2 Unwuchtschätzung im Schleuderprozess

Zur Bestimmung der eingesetzten Unwucht wird diese als stückweise konstanter Parameter angesehen und die Joint-Estimation-Variante eines Unscented-Kalman-Filters mit symmetrischem Sigma-Punkte-Ansatz genutzt. Das entwickelte Optimierungsverfahren zur Bestimmung der Filterentwurfsparameter α, β, κ wird dabei mit einem Anregungsszenario angewandt, das aus einem linearen Schleuderhochlauf bis 800 1/min in 10 Sekunden besteht. Filtermodell und Streckenmodell, die zur Generierung der Messungen bzw. der Systemausgänge genutzt werden, unterscheiden sich dabei nur in geringer Form, um die korrekte Filterfunktionalität bei komplexen Modellen zu verdeutlichen. Die aus dem Streckenmodell resultierenden Messungen werden mit normalverteiltem Rauschen überlagert.

Die Bestimmung des unbekannten Parameters der Unwuchtmasse nach Anwendung des modellbasierten Lernverfahrens ist in Bild 6-14 als roter Verlauf dargestellt und zeigt eine sehr hohe Schätzgüte im Vergleich zur eingesetzten Unwucht (blauer Verlauf). Insbesondere am Anfang, nachdem die Schätzung mit einem Startwert von 1 kg initialisiert wurde, treten noch größere Abweichungen im Schätzverlauf auf, die aus den verrauschten Messungen und den leichten Modellabweichungen stammen. Dies resultiert aus der Tatsache, dass die Unwucht zu Beginn aufgrund der noch sehr geringen Drehzahl noch schwer beobachtbar ist, da ihre Auswirkungen quadratisch mit der Drehzahl zunehmen. Im weiteren Verlauf erhöht sich die Schätzgüte, sodass selbst starken Unwuchtänderungen im Bereich von 4 bis 5 Sekunden und bei 6 Sekunden ebenso wie Plateaus und langsameren Unwuchtänderungen sehr gut gefolgt werden kann. Zum Ende des Szenarios kann die Unwucht annähernd perfekt bestimmt werden, und es tritt kaum noch Rauschen in der Schätzung auf, was auf den Einsatz der Robbins-Monro-Methode zurückzuführen ist. Durch diese wird die Prozessrauschkovarianz in Abhängigkeit der Abweichungen reduziert und somit im Update-Schritt keine Korrektur des Schätzparameters mehr durchgeführt.

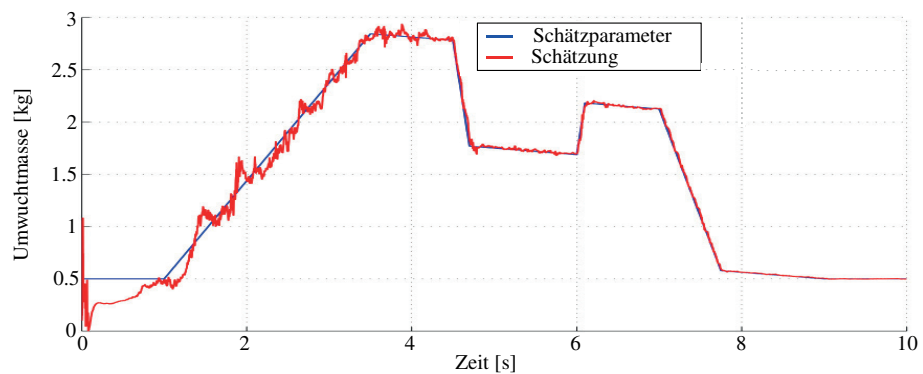


Bild 6-14: Schätzung einer zeitvarianten Unwucht in einem Waschautomaten

Bild 6-15 zeigt exemplarisch die Rekonstruktion der verrauschten Messungen anhand der x-Richtung des Sensors 1 (siehe auch Bild 6-13) sowie des Drehzahl-sensors und eine vergrößerte Darstellung ausgewählter Bereiche. Hierbei ist gut zu erkennen, dass das Filter durch die Schätzung der unbekannten Unwucht sehr gut in der Lage ist, das Rauschlevel zu unterdrücken und eine Rekonstruktion der Messungen zu realisieren. Die Darstellung der y-Richtung des Sensors 1 sowie der x-Richtung des Sensors 2 ist im Anhang A.5 zu finden.

Exemplarisch für die sehr gute Schätzung der Zustände sind in Bild 6-16 die Schätzungen von translatorischer Bewegung und Rotationsgeschwindigkeit des Laugenbehälters im Schwerpunkt in y-Richtung bzw. um die x-Achse dargestellt. Die Geschwindigkeit des Laugenbehälters im Schwerpunkt in die y-Richtung, die Rotationsgeschwindigkeit des Laugenbehälters um die x-Achse sowie der Rotationswinkel und die Drehzahl der Trommel sind in Anhang A.5 dargestellt. In Blau sind die Zustände des Streckenmodells abgebildet und in Rot die Rekonstruktion mittels des Unscented-Kalman-Filters. Bis auf einen Einschwingvorgang zu Beginn der Schätzung und den Bereich bis ca. 1 Sekunde in dem die Unwucht aufgrund der mangelnden Beobachtbarkeit noch nicht erkannt wurde, zeigt das Filter sehr gute Ergebnisse.

6.2.3 Robustheit gegenüber Modellungenauigkeiten

Im Sinne der in Abschnitt 5.3 vorgestellten Entwurfsmethodik erfolgt im nächsten Schritt ein Robustheitscheck des ausgelegten Filters unter Einbeziehung der durch das Optimierungsverfahren zur Filterentwurfparameterbestimmung ermittelten Parameter. In diesem Fall werden das bereits bekannte Anregungsszenario des Hochlaufes auf 800 1/min und der aus dem vorigen Abschnitt vorgestellte Unwuchtverlauf genutzt. Das Modell wird allerdings gegenüber dem Streckenmodell, das die Messungen bzw. Systemausgänge bereitstellt, stark verändert hinsichtlich Parametrierung von Massen und Tensoren der Starrkörper, Steifigkeiten

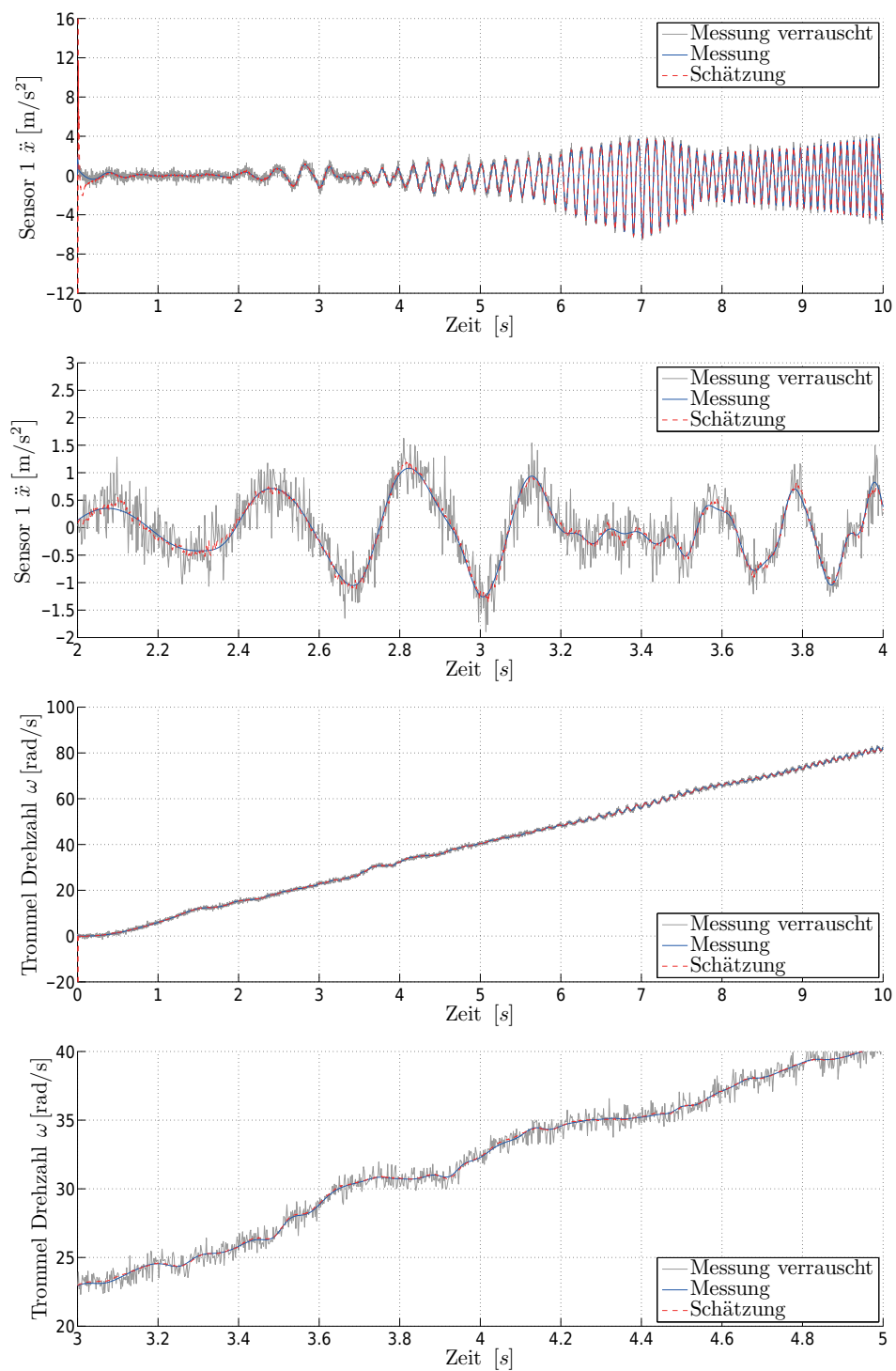


Bild 6-15: Rekonstruktion verrauschter Messungen eines Waschautomaten bei der Schätzung einer unbekannten, zeitvarianten Unwucht. Der Filterausgang wird als "Schätzung" bezeichnet und basiert auf dem Filter zur Verfügung gestellten Signal "Messung verrauscht", welches dem Signal "Messung" entspricht und mit additivem Rauschen überlagert wurde.

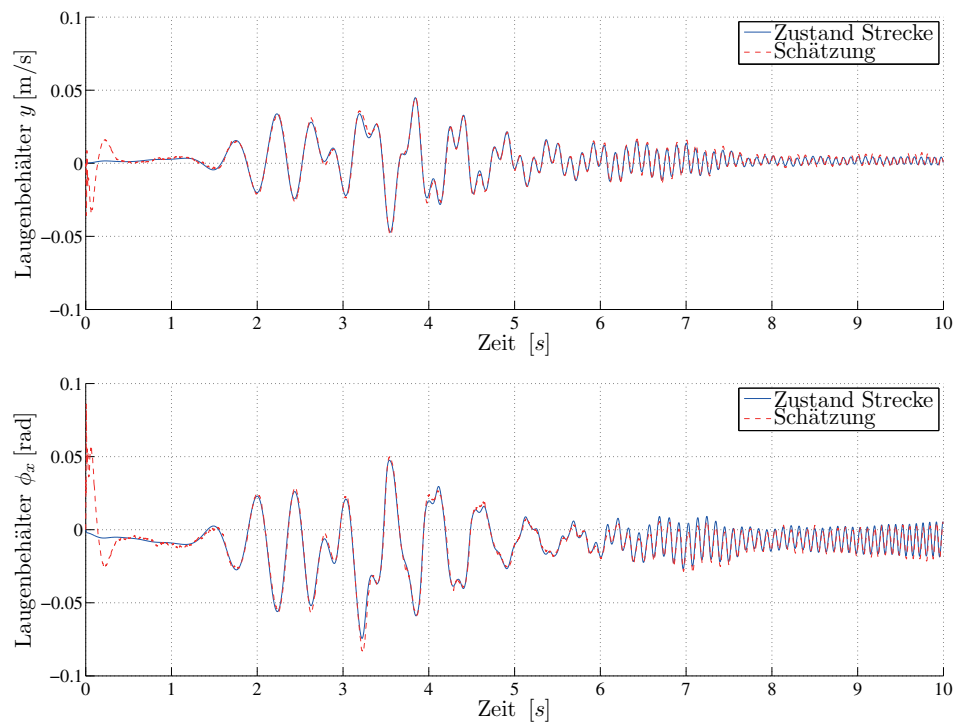


Bild 6-16: Zustandsschätzung eines Waschautomaten bei der Schätzung einer unbekannten, zeitvarianten Unwucht. Mit "Zustand Strecke" ist der zu schätzende Zustand der Strecke bezeichnet, der aufgrund der Erzeugung an einem Modell bekannt ist. Mit "Schätzung" ist die Zustandsschätzung des Filters gekennzeichnet.

und Dämpfungen der Federbeine und des Türdichtringes sowie der Position der Federbein im Gehäuse.

Die Auswirkungen der Systemparameteränderungen sind aus den Bildern 6-18 und 6-19 ersichtlich. Darin sind die Zustände bzw. Messungen des Streckenmodells in Blau dargestellt und werden mit *Messung*, bzw. *Zustand Strecke* bezeichnet. Sie stellen die Referenzsignale der zu bestimmenden Messungen und Zustände dar. Wird das Anregungsszenario, bestehend aus dem Drehzahl- und dem Unwuchtverlauf, ohne Anwendung von Filtertechniken auf das Filtermodell aufgeprägt, ergeben sich die cyanfarbenen Verläufe der Zustände und Systemausgänge, die mit *Systemausgang Modell*, bzw. *Zustand Modell* bezeichnet sind. Hierbei wird deutlich, dass sowohl in den Zuständen als auch in den Messungen starke Abweichungen in der Antwortfrequenz als auch der Verstärkung auftreten, was die Auswirkungen der unterschiedlichen Parametrierungen zwischen Strecken- und Filtermodell verdeutlicht und einen Robustheitstest des entworfenen Filters darstellt, da beim realen Einsatz auch niemals das eingesetzte Filtermodell in der Lage sein wird, die real vorliegende Strecke vollständig zu beschreiben.

Bild 6-17 zeigt die Schätzung der unbekannten Unwucht gegenüber dem realen, blauen Verlauf, der mit *Schätzparameter* bezeichnet ist. Das Filter ist trotz der starken Unterschiede zwischen Filter- und Streckenmodell in der Lage, die unbekannte Unwucht in einer absolut zufriedenstellenden Güte zu bestimmen. Die Rekonstruktion ausgewählter Zustände ist in Bild 6-18 als roter Verlauf dargestellt. Im höheren Drehzahlverlauf ab ca. 6 Sekunden treten Fehler in der Schätzung auf, die auf die Unterschiede in den Modellen zurückzuführen sind. Eine Darstellung weiterer Zustände ist in Anhang A.5 zu finden. In Bild 6-19 ist insbesondere in den vergrößerten Darstellungen ausgewählter Bereiche die sehr gute Rekonstruktion der Messungen ersichtlich. Das Filter ist in der Lage, sowohl das Rauschen zu reduzieren als auch mittels der Schätzung der unbekannten Unwucht den Messungen zu folgen. Insbesondere zu Beginn der Schätzung, was in diesem Fall der Initialisierung des Filters entspricht, sind teilweise starke Schwankungen der Schätzung zu erkennen, was aus der hohen Startkovarianz resultiert, die in den ersten Schritten durch das Filter abgebaut wird. Diese können durchaus auch kleiner gewählt werden, sollten allerdings in diesem Fall wegen der relativ hohen Unsicherheit in der Kenntnis der Anfangszustände auch groß gewählt werden. Die Rekonstruktion weiterer Messsignale ist im Anhang unter A.5 dargestellt.

Aus den erzielten Ergebnissen bei diesem komplexen, stark nichtlinearen System mit einer hohen Systemordnung ist erkennbar, dass sowohl das entwickelte Optimierungsverfahren zur Ermittlung eines optimalen Satzes von Entwurfsparametern als auch das entwickelte Interface und die Methodik zur Modelleinbindung, Filterauslegung und Filtertest sehr gut funktionieren.

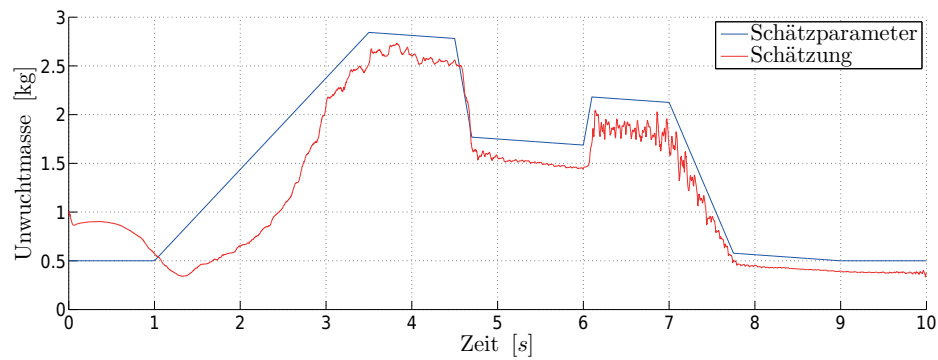


Bild 6-17: Schätzung einer zeitvarianten Unwucht in einem Waschautomaten mit starken Abweichungen zwischen dem Filtermodell und den dem Filter zur Verfügung stehenden Messungen.

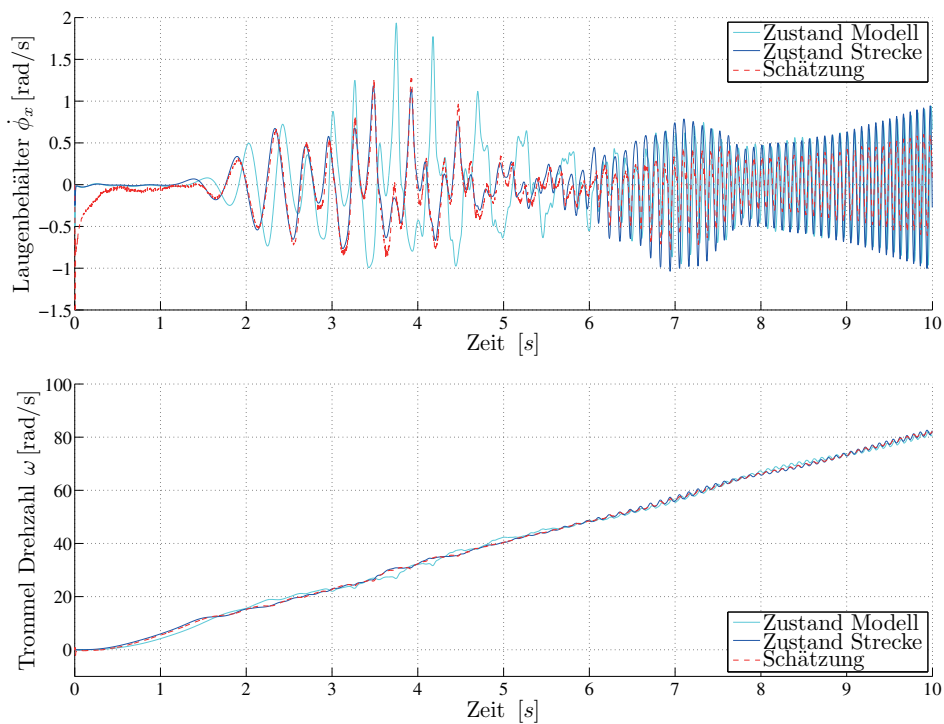


Bild 6-18: Zustandsschätzung eines Waschautomaten bei der Schätzung einer unbekannten, zeitvarianten Unwucht mit starken Abweichungen zwischen dem Filtermodell und den dem Filter zur Verfügung stehenden Messungen.

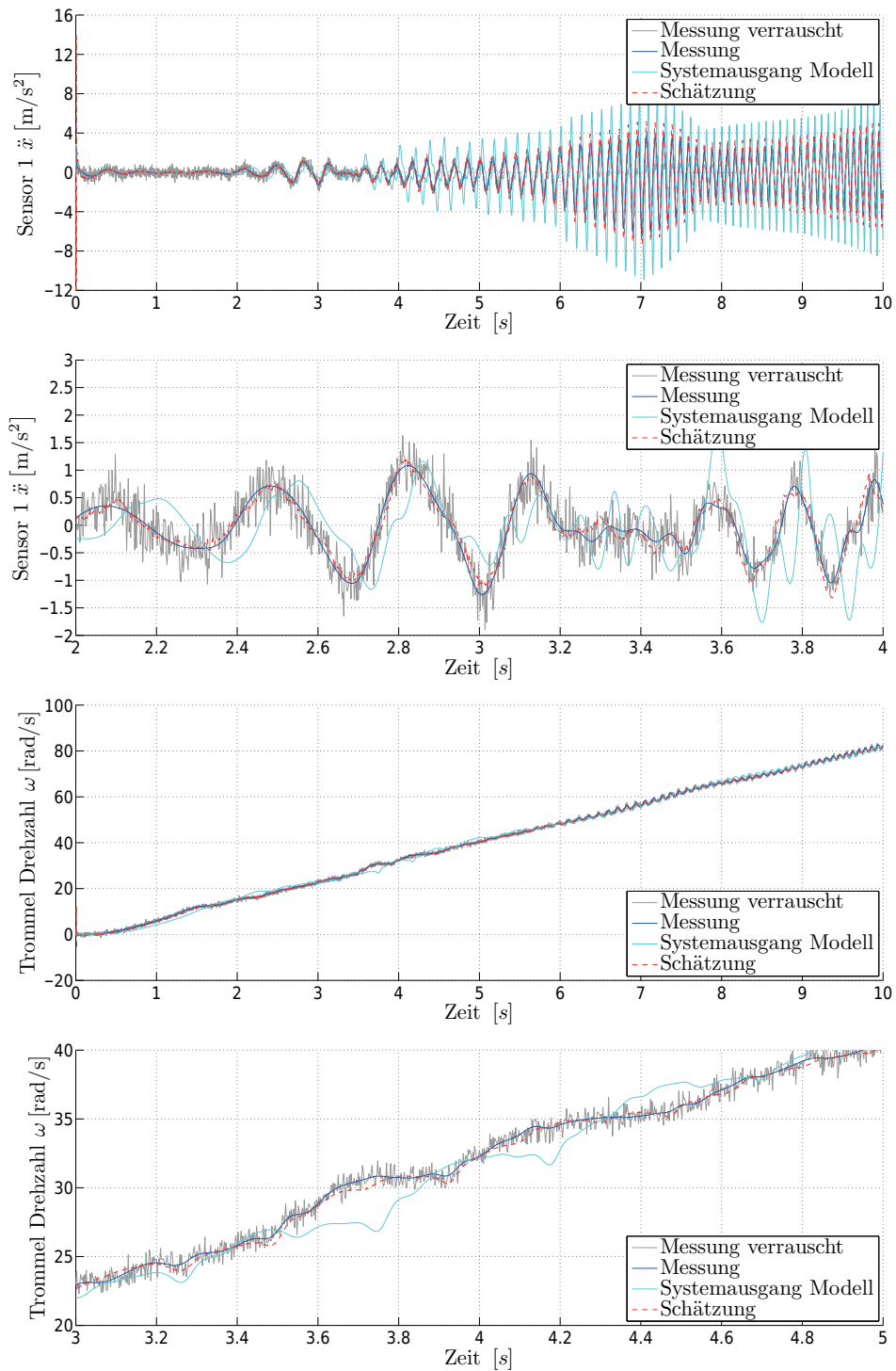


Bild 6-19: Rekonstruktion verrauschter Messungen eines Waschautomaten bei der Schätzung einer unbekannten, zeitvarianten Unwucht mit starken Abweichungen zwischen dem Filtermodell und den dem Filter zur Verfügung stehenden Messungen

6.3 Zustandsschätzung Dingo 2

Der Dingo 2 ist ein leicht gepanzertes, minengeschütztes, leicht bewaffnetes Radfahrzeug, das von Krauss-Maffei-Wegmann entwickelt wurde, und stellt die Verbindung zwischen den leichten ungepanzerten Transportfahrzeugen und den leichten Panzerfahrzeugen her (siehe Bild 6-20 links). Das Fahrzeug kann neben seiner zweiköpfigen Besatzung sechs weitere Personen transportieren und wird als Führungs-, Verbindungs-, Sanitätsfahrzeug etc. eingesetzt. Der Dingo 2 zeichnet sich durch gute Geländegängigkeit aus. Hauptkomponenten sind das Fahrgestell, die Schutzzelle/Kabine, der Stauraum/Rahmen und der Motorraum. Die Schutzzelle ist dabei über zwei elastische Lager I und II (siehe Bild 6-20 rechts) mit dem Motorraum und über ein elastisches Lager III mit dem Stauraum verbunden. Durch diesen Aufbau ist eine sehr große Verwindung des Fahrzeugs möglich. Er sorgt allerdings für Schwierigkeiten bei der Realisierung einer Aufbauregelung [Kru09].

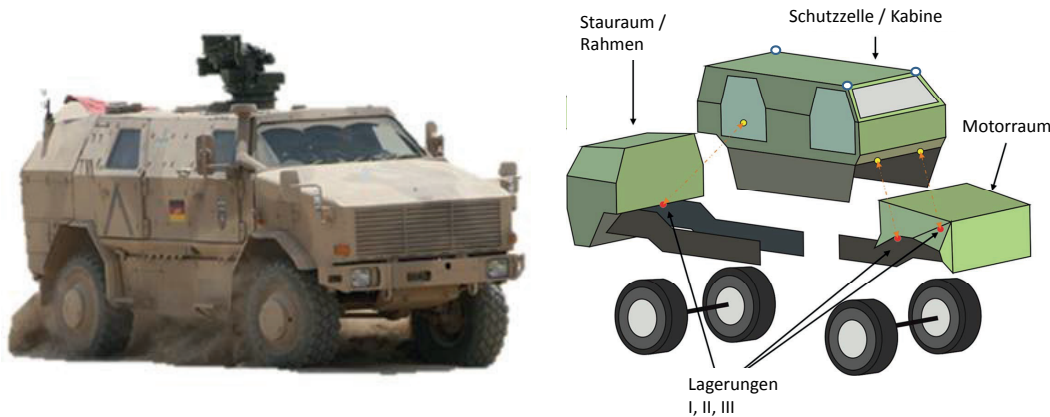


Bild 6-20: Links: Dingo 2 [Quelle: www.kmweg.de]. Rechts: Modularität Dingo 2 mit Ansicht der eingesetzten Sensoren nach [Kru09]

In früheren Projekten wurden Aufbauregelstrategien über die Realisierung von Eingrößenreglern vorgenommen. Um eine Strategie mittels Zustandsregelung zu realisieren, müssen alle Zustände des Systems bekannt sein, was im Folgenden durch den Einsatz von Sigma-Punkte-Filtern und der entwickelten Entwurfsstrategie geschieht. Ein weiterer Vorteil ist die durch das Filter bestehende Sensordatenfusion, durch die Sensoren eingespart werden konnten, die vormals für die Realisierung der Eingrößenregelungen notwendig waren.

6.3.1 Modellbildung Dingo

In [Kru09] wurde ein detailliertes MKS-Modell des Dingo 2 unter Dymola zur Aufbaureglerauslegung entwickelt. Die oben bereits vorgestellten Module des Aufbaus werden als Starrkörper realisiert und mittels Gelenken und Kraftelementen, wie Elastomere, miteinander verbunden. Die Massen des Motors, des Getriebes,

der gepanzerten Motorhaube und der vorderen Rahmenhälfte sind in der Masse *Rahmen vorn* zusammengefasst. Tank, Kraftstoff, Abschleppstange und Ladung inklusive Beladung ergeben den *Rahmen hinten*. Die gepanzerte Fahrgastzelle sowie die Massen der Insassen und der sonstigen Zuladung werden zur Kabine zusammengefasst. Eine Verwindung des Rahmens vorn und hinten gegeneinander wie auch eine geringe Durchbiegung werden durch die Verbindung mittels eines Gelenkes mit zwei rotatorischen Freiheitsgraden realisiert.

Insgesamt besitzt das Modell eine Systemordnung von 26 mit vier Eingängen, bestehend aus den Anregungen der vier Reifen, und vier Ausgängen, bestehend aus Beschleunigungssignalen des Rahmens und der Kabine.

Die Anwendungsszenarien entstammen aus den Teststrecken der Wehrtechnischen Dienststelle in Trier, die aus Einzelhindernissen, Sinusbahn und einer Schweizerbahn bestehen. Nachfolgend wird zur Verdeutlichung der korrekten Funktionsweise der entwickelten Auslegung die Fahrt über versetzte Einzelhindernisse mit 10 km/h als Anwendungsszenario gewählt, das in 6-21 dargestellt ist.

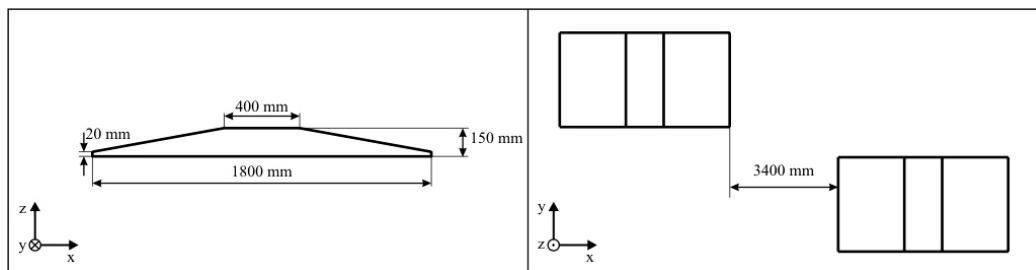


Bild 6-21: Versetzte Trapezhindernisse der im Szenario enthaltenen Anregungen [Kru09]

6.3.2 Zustandsschätzung bei Hindernisfahrt

Besonders erschwerend bei dieser Zustandsschätzung ist die stark unterschiedliche Anregung des Systems, da das Szenario aus der Fahrt über ebenen Untergrund besteht, in der nur eine minimale Reifenanregung stattfindet, und das Filter somit kaum Eingangsgrößen besitzt, um bestehende Unsicherheiten in der Initialisierung auszugleichen. Als Folge hieraus müssen die Fehlerkovarianzen sehr klein werden, um ein Driften des Filters zu unterbinden. Beim Erreichen der Hindernisse erfährt das System plötzlich eine sehr starke Anregung, auf die das Filter reagieren muss. Weiter erschwerend kommt in diesem Fall hinzu, dass das Modell nur eine stark vereinfachte Abbildung des realen Systems ist, von dem die Messwerte stammen.

Nachfolgend sind in Bild 6-22 die realen Messungen von zwei Sensoren, die von der Wehrtechnischen Dienststelle in Trier aufgenommen wurden, in Blau abgebildet, die Ausgänge des im Filter verwendeten parallel mitlaufenden Modell allerdings ohne Schätzfehlerückkopplung in Cyan und die rekonstruierten Messungen bei Anwendung eines Unscented-Kalman-Filters mit einem symmetrischen

Sigma-Punkte-Satzes in Rot, nach der Anwendung des Optimierungsverfahrens zur Filterparameterbestimmung. In Bild eins und drei ist der Verlauf über die gesamten 12 Sekunden des Manövers dargestellt, in denen zu erkennen ist, dass das Filter die verrauschten Messungen rekonstruieren kann. Sowohl in den Bereichen vor bzw. nach den Hindernissen von 0 bis ca. 3,5 Sekunden bzw. 7 bis 12 Sekunden als auch während der Hindernisse von 3.5 bis 7 Sekunden ist das Filter in der Lage, bei nur geringer Anregung stabil zu bleiben und dennoch beim Auftreten der Hindernisse zu reagieren.

Im zweiten und dritten Bild 6-22 sind vergrößerte Ausschnitte der Messungsrekonstruktion zu sehen. Beim Sensor 1 ist gut zu erkennen, dass sich das Filter von einem anfangs falschen Wert schnell auf die Messungen einschwingt und die Grundfrequenz der Messung annimmt. Der Modellausgang ohne Anwendung von Filtertechniken in Cyan zeigt ein starkes Überschwingen und kann die Grundfrequenz nicht nachbilden. Dieses Verhalten resultiert aus der Vernachlässigung eines nichtlinearen Dämpfungsverhaltens im aktiven Fahrwerk, das in dem vereinfachten Filtermodell durch einen linearen Anteil ersetzt wurde. Dieser sorgt allerdings dafür, dass bei geringer Anregung das System stärker gedämpft ist als in der Realität, was zu einem schnellen Abklingen der Grundschrwingungen führt. Der gleiche Effekt wird in der vergrößerten Darstellung des zweiten Sensors deutlich, der in 1-15 dargestellt ist und den Bereich nach Überfahren der Hindernisse darstellt. Wie erwartet, schwingt sich das reine Modellverhalten in Cyan sehr schnell ein, obwohl die reale Messung noch eine Grundfrequenz von ca. 2 Hz aufweist. Der Filterausgang in Rot ist hier geeignet diese Grundfrequenz zu rekonstruieren und gleichzeitig das Rauschlevel deutlich zu reduzieren.

In den vergrößerten Bereichen von Sensor 3 und 4 in Bild 6-22 und Bild 1-15 sind Bereiche beim Überfahren der Hindernisse dargestellt. Der reine Ausgang des Modells zeigt hier insbesondere beim Sensor 3 eine vergrößerte Amplitude, die durch Anwendung des Filters deutlich reduziert werden kann und dabei das Sensorrauschen unterdrückt. Dieses Verhalten ist auch im Ausschnitt des Sensors 4 zu erkennen, wobei auch stark fehlerhafte Sensorwerte, etwa der Messpeak bei ca. 5,8 Sekunden, unterdrückt werden und die Grundfrequenz der Messung sehr gut rekonstruiert wird.

Bild 6-23 zeigt die Zustandsschätzung beispielhaft für den Zustand 4 der Winkelgeschwindigkeit um die y-Achse des vorderen Rahmens und den Zustand 1 der Absolutlage in der Vertikalen des vorderen Rahmens.

Wie schon bei der Rekonstruktion der Messungen zu erkennen war, werden vom Filter die Grundschrwingungen vor und nach der Hindernisfahrt in die Zustände eingebracht (roter Verlauf), die vom reinen Modellzustand (blau) ohne Anwendung von Filtertechniken nicht erfasst werden können. Die Abweichungen zwischen reinem Modellausgang und Messung sind hier in den Unterschieden zwischen Schätzung und Modellzustand bei Überfahren der Hindernisse zu erkennen. In der Darstellung der aktuellen vertikalen Lage wird die fehlerhafte Initialisie-

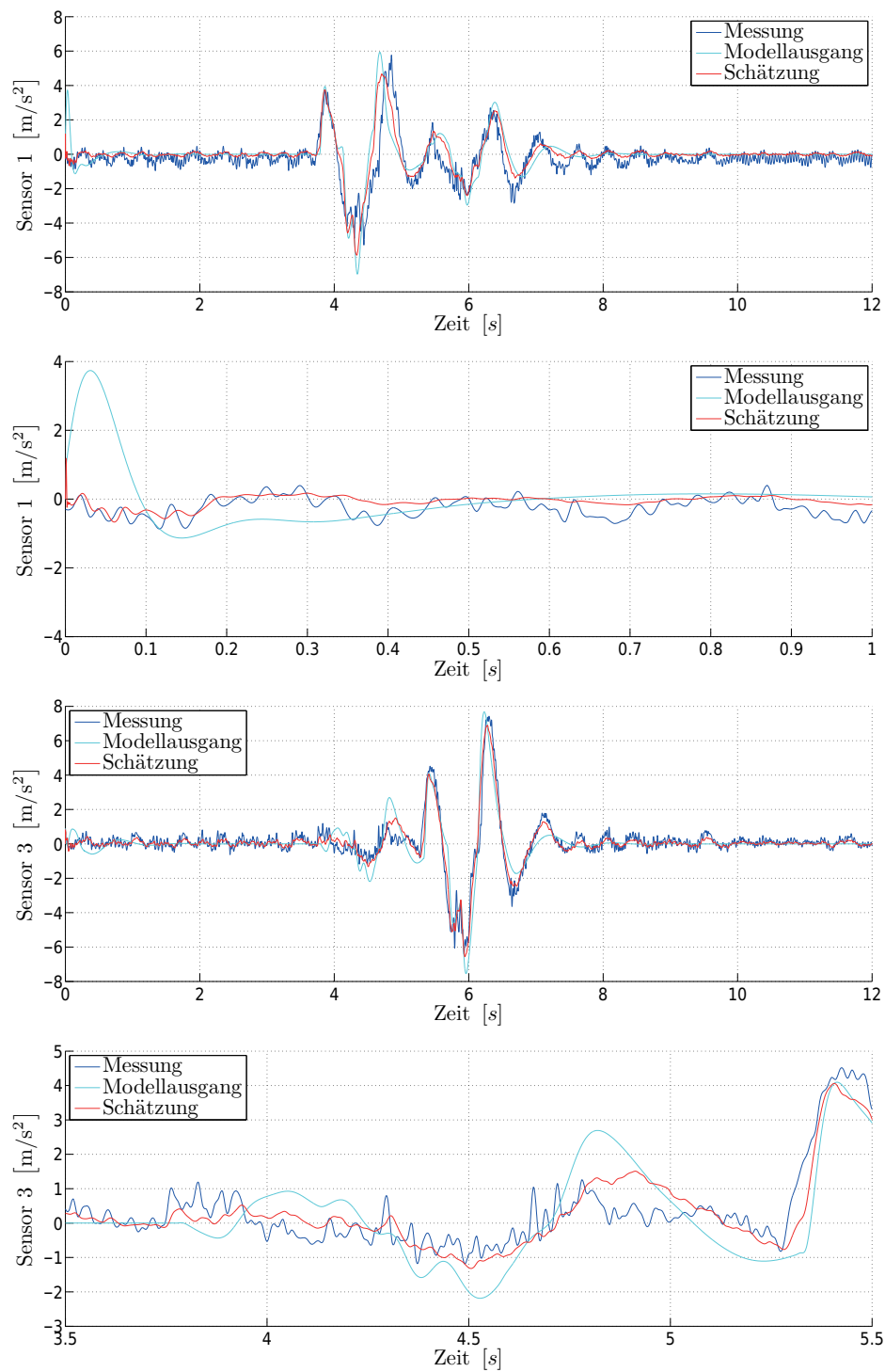


Bild 6-22: Rekonstruktion der realen Messungen eines Dingo 2 bei Überfahren von Trapezhindernissen

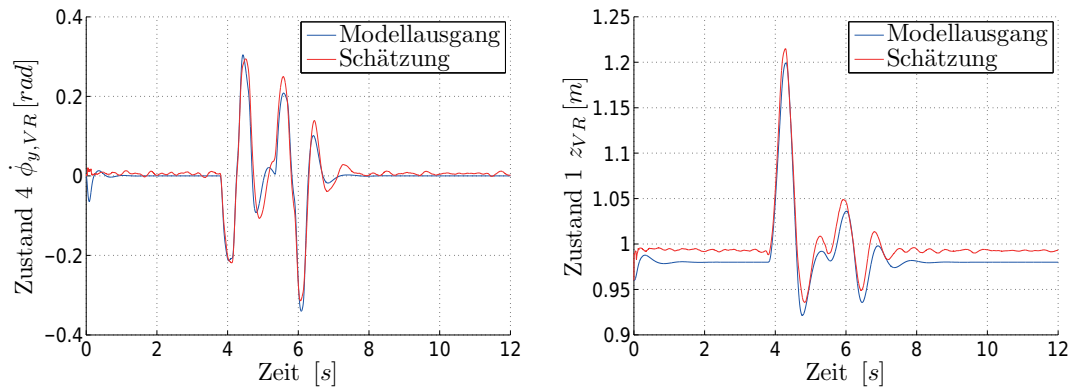


Bild 6-23: Zustandsschätzung des Dingo 2 bei Überfahren von Trapezhindernissen

rung der Zustände deutlich, da das Filter zum reinen Modellzustand einen Offset zeigt.

6.4 Zusammenfassung Anwendungsbeispiele

In diesem Kapitel konnte gezeigt werden, dass die in dieser Dissertation erarbeiteten Algorithmen und Methoden sehr gute Ergebnisse bei der Zustands- und Parameterschätzung an nicht-analytischen, unbekannten Systemmodellen liefern.

Hierzu wurden die Modelle als Function-Mockup-Unit mittels des entwickelten Interfaces in MATLAB importiert und alle notwendigen Modellinformationen für eine Filterauslegung und den Filterbetrieb automatisch bereitgestellt. Im Anschluss wurden Anwendungsszenarien importiert, die als beliebig generierte Datenstruktur vorliegen können.

Bei der Anwendung an einem einfachen, nichtlinearen System des Doppelpendels wurde in 6.1 gezeigt, dass die Ermittlung eines optimalen Satzes von Entwurfparametern sowie die Generierung von Basismatrizen für das Prozess- und Messrauschen zwingend erforderlich sind und mittels des Verfahrens zur Optimierung der Entwurfparameter sehr gute Resultate erzielt werden konnten. Im Anschluss konnte anhand eines stark veränderten Filtermodells die Robustheit des entworfenen Filters gezeigt werden. Bei der Anwendung der Parameterschätzung wurde gezeigt, dass die Robbins-Monro-Methode die besten Ergebnisse liefert, sodass für weitere Anwendungsfälle diese Methode zur Gewinnung einer Parameterkovarianz genutzt wurde. Die Notwendigkeit und korrekte Funktionsweise der Einbindung von Nebenbedingungen in die Schätzung von Parametern wurden in 6.1.5 vorgestellt. Abschließend wurde ein Vergleich hinsichtlich Schätzgüte und Geschwindigkeit bei unterschiedlichen Sigma-Punkte-Filtern vorgenommen, bei dem gezeigt werden konnte, dass durch die Anwendung des entwickelten Optimierungsverfahrens zur Filterentwurfparameterbestimmung die höhere Anzahl an Entwurfparameter vollautomatisch ausgelegt werden konnte, ohne Kenntnisse über das zugrundeliegende Modell zu erfordern. Die Schätzgüte übertrifft dabei

die Qualität des Cubature- oder des Central-Difference-Kalman-Filters, die mit nur einem Entwurfparameter auskommen. Durch die Anwendung des Unscented-Kalman-Filters mit einem Spherical Sigma-Punkte-Satz konnte die Rechenzeit um ein Drittel reduziert werden, bei gleichzeitiger Reduktion der Schätzungsgüte.

In Abschnitt 6.2 wurde die entwickelte Methodik mit einem Unscented-Kalman-Filter und einem symmetrischem Sigma-Punkte-Satz am Modell eines Waschautomaten zur Zustandsschätzung und zur Bestimmung einer zeitvarianten Unwucht in der Trommel angewandt. Es wurde gezeigt, dass das entworfene Filter bei dem zur Verfügung stehenden Anwendungsszenario automatisiert ausgelegt werden konnte und nahezu perfekte Ergebnisse bei der Bestimmung der unbekannten Unwucht zeigt. Bei der nachgeschalteten Robustheitsanalyse durch Nutzung eines stark veränderten Filtermodells, wurde gezeigt, dass die Unwucht qualitativ sehr gut geschätzt werden konnte, allerdings Fehler durch das veränderte Modell entstehen.

Übertragbarkeit und Anwendbarkeit der entwickelten Algorithmen und Methoden auf real vorliegende Messdaten wurden bei der Zustandsschätzung in Abschnitt 6.3 an dem leicht gepanzerten Fahrzeug Dingo 2 gezeigt. Modellunsicherheiten konnten durch das Filter ausgeglichen und gleichzeitig die Güte im Gegensatz zum Modellausgang ohne Anwendung von Filtertechniken bei einer Unterdrückung des Sensorrauschens erhöht werden. Die Zustände des Systems konnten geschätzt werden und stehen somit für die Implementierung einer Zustandsregelung zur Verfügung.

7 Resümee und Perspektiven

Die Entwicklung von komplexen mechatronischen Systemen steht unter immer größerem Kosten- und Zeitdruck. Gleichzeitig wird kundenseitig eine immer höhere Funktionalität gefordert, die das Zusammenwirken der unterschiedlichen Disziplinen Mechanik, Elektronik, Informatik und Regelungstechnik zwingend erfordert, was neue Herausforderungen an Entwicklungswerkzeuge und Methoden stellt. Die Einbindung von physikalisch motivierten Verhaltensmodellen zur Auslegung und Eigenschaftsabsicherung gewinnt dabei zunehmend an Bedeutung. Die Modelle der unterschiedlichen Disziplinen wurden bisher oftmals getrennt voneinander entwickelt oder im Rahmen der zunehmenden Aufteilung in der Entwicklung komplexer Systeme von unterschiedlichen Abteilungen eines Unternehmens oder sogar von externen Zulieferern bereitgestellt. Gleichzeitig wird der Modellierungsprozess immer stärker von Toolunterstützung durchdrungen, wie beispielsweise Dymola, das auf der offenen Sprache Modelica basiert. Die Modellbildung erfolgt nicht mehr auf ODE-Ebene, wie bei der Modellierung unter MATLAB/Simulink, sondern direkt auf der physikalischen Ersatzschaltbildebene, wobei auf zahlreiche vorgefertigte Elemente aus Bibliotheken von Fachdisziplinexperten, wie Zulieferern, Universitäten und Forschungseinrichtungen, zurückgegriffen wird.

Aus diesen Gründen steht für die modellbasierte Auslegung in der Fachdisziplin Regelungstechnik im Anschluss an die Modellbildungsphase eine Abbildung des Systems zur Verfügung, in dem der Regelungstechniker kein Experte mehr sein muss. Verschärft wird diese Problematik durch neue Entwicklungssystematiken wie ENTIME, in der das Auslegungsmodell teil- oder vollautomatisiert auf Basis von Lösungselementen und Anforderungen unter Dymola erstellt wird. Weder die innere Struktur des Modells noch enthaltene Nichtlinearitäten sind dem Regelungstechniker bekannt. Gleichzeitig liegt das Modell nur noch als Black-Box, mit bekanntem Ein- und Ausgangsverhalten vor, und nicht, wie in der für Regelungstechniker bekannten und gebräuchlicheren Form, als nichtlinearer Zustandsraum, wodurch die Anwendung von regelungstechnischen Methoden wie Zustands- und Parameterschätzung erschwert wird.

Der Schwerpunkt dieser Arbeit lag in der Entwicklung einer Methode und aller notwendigen Werkzeuge zur Auslegung von echtzeitfähigen Zustands- und Parameterschätzern an Black-Box-Systemen, die mittels moderner Entwicklungswerkzeuge wie Dymola erzeugt wurden. Hierzu wurden in einem ersten Schritt Sigma-Punkte-Filter als zu nutzende Filteralgorithmen identifiziert, da diese eine höhere Schätzgüte bei starken Nichtlinearitäten bieten, von denen in einem unbekannten Modell ausgegangen werden muss, was eine Neuerung dieser Arbeit darstellt. Gleichzeitig sind keine Ableitungen in Form von Jacobi- oder Hessematrizen des als Black-Box vorliegenden Modells notwendig, anders als bei der Nutzung eines Extended-Kalman-Filters und seinen Erweiterungen. Die höhere Genauigkeit

der Sigma-Punkte-Filter resultiert aus einer Erfassung der Wahrscheinlichkeitsdichteverteilung des Zustandsvektors mittels einer geringen Anzahl an deterministisch zu berechnenden Sigma-Punkten, die durch das vollständige nichtlineare Modell transformiert werden können. Zur Schätzung von unbekannten Systemparametern wurde als weitere Neuerung ein Verfahren zur Erkennung und Korrektur von ungültigen Schätzungen und Sigma-Punkten, die zu Instabilitäten des Filtermodells führen, entwickelt und implementiert. Zur Realisierung dieser in der vorliegenden Arbeit entwickelten Methode wurde ein Interface entwickelt, das die aus Dymola stammenden Modelle als Black-Boxes einbindet, alle notwendigen Informationen für den Filterbetrieb bereitstellt und die Möglichkeit der Sigma-Punkte-Transformation bietet. Zum Austausch der Modelle wurde das FMI-Format gewählt, das einerseits durch enthaltene Funktionen die Einbindung und den Filterbetrieb unterstützt und andererseits ein zur Zeit dieser Arbeit intensiv vorangetriebenes Austauschformat von Modellen darstellt, das von einer Vielzahl an Modellierungswerkzeugen unterstützt wird.

Für die Auslegung der in den Sigma-Punkte-Filtern enthaltenen Entwurfsparameter wurde in dieser Arbeit eine Optimierungsstrategie entwickelt, die auf einer Gaußprozessoptimierung basiert. Hierzu wird das unbekannte Modell, der zu nutzende Sigma-Punkte-Filteralgorithmus sowie Anregungsfälle als ein Szenario definiert, das dem Optimierungsverfahren übergeben wird. Somit ist es möglich, ohne Expertenkenntnisse des zugrundeliegenden Filtermodells noch des Filteralgorithmus oder der Anregungsfälle einen problemspezifischen, optimalen Satz von Filterentwurfsparametern zu bestimmen, was eine weitere Neuerung dieser Arbeit darstellt. Die Bestimmung von Basiskovarianzmatrizen des Prozess- und Messrauschens wird in einem Pre-Processingschritt vor der Optimierung anhand der definierten Anregungsfälle des Szenarios vollautomatisiert durchgeführt, um den Effekt unterschiedlicher Wertebereiche von Messgrößen und Zuständen zu eliminieren. Um die Interpretation der Prozess- und Messrauschkovarianzen als Vertrauensfaktoren nicht zu verlieren, kann der Anwender einen Faktor der Rauschprozesse wählen, der das Modell- oder Messvertrauen abbildet und zu einer höchstgradig automatisierten Filterauslegung beiträgt, was eine weitere Neuerung darstellt.

Abschließend wurde in dieser Arbeit erstmalig ein Vorgehensmodell für die Auslegung, den Test und den Export von Zustands- und Parameterschätzern an als Black-Boxes vorliegenden Filtermodellen mittels Sigma-Punkte-Filtern entwickelt. Der Fokus des Vorgehensmodells liegt auf einem maximalen Automatisierungsgrad der Filterauslegung, da der Ingenieur kein Experte mehr in dem zu nutzenden Filtermodell sein muss. Gleichzeitig wird an wichtigen Stellen des Filterentwurfsprozesses die Möglichkeit zum Eingriff durch den Anwender ermöglicht, um mögliches Expertenwissen einfließen zu lassen. Bei ungenügender Schätzgüte werden im Anschluss Handlungsempfehlungen definiert, die Veränderungen im Entwurfsprozess entsprechen, um eine höhere Schätzgüte zu ermöglichen.

Die korrekte Funktionsweise der entwickelten Methoden und Algorithmen wurde anhand von drei Beispielen vorgestellt. Beim ersten Beispiel handelte es sich um eine Zustands- und Parameterschätzung an einem mathematischen Pendel, an dem unterschiedliche Sigma-Punkte-Filter, sowie automatisierte Vorgaben des Parameterprozessrauschens, der entwickelten Sigma-Punkte-Korrektur und der vollautomatischen Filterentwurfsparameterauslegung erprobt wurden. Die Übertragung auf ein komplexes mechatronisches System in Form eines Waschautomaten wurde im zweiten Beispiel gezeigt, wobei auch sehr gute Ergebnisse bei der Robustheitsanalyse erzielt wurden, indem ein stark abgeändertes Modell in das Filter eingebunden wurde. Im letzten Anwendungsfall wurde das Filter an dem Modell eines Dingo 2 ausgelegt, wobei als Anregungsfälle im Szenario reale Messungen der Bundeswehr zur Verfügung standen. Sowohl die Filterauslegung als auch der anschließende Betrieb zeigten, dass die Anwendung auf reale Messungen ebenfalls sehr gute Ergebnisse zeigt und sich somit für den Onlineeinsatz an realen Systemen eignet.

Der nächste Schritt bei der Auslegung und dem Betrieb von Zustands- und Parameterschätzern für Black-Box-Systeme stellt die automatisierte Codegenerierung auf eine Zielplattform dar. Da alle genutzten Filtertechniken und die Weiterentwicklungen zur Transformation der Sigma-Punkte und die Erkennung und die Korrektur von ungültigen Sigma-Punkten echtzeitfähig sind, kann dies ohne Einschränkungen durchgeführt werden.

Im Rahmen der vollautomatisierten Optimierung der Filterentwurfsparameter wurde bereits angesprochen, dass eine Reihe von Kernel-Funktionen der genutzten Gaußprozesse existieren. In dieser Arbeit wurde das SE-ARD-Kernel genutzt, das die in dieser Arbeit gezeigten Szenarien darstellen konnte und zu sehr guten Ergebnissen geführt hat. Trotzdem kann durch die Einbringung einer Kernel-Bibliothek und mittels einer Vorabanalyse oder im späteren Optimierungsverlauf beim Erkennen einer schlechten Abbildbarkeit des Szenarios auf einen besser geeigneten Kernel gewechselt werden. Das Optimierungsverfahren wäre somit auch geeignet, abseits der Anwendung in dieser Arbeit für beliebige Optimierungsverfahren eingesetzt zu werden.

Ein weiterer wichtiger Punkt bei der Verbesserung besteht in einer automatisierten Überprüfung der Beobachtbarkeit von Zuständen und zu schätzenden Parametern. Zur Zeit wird dies nicht überprüft und in den Handlungsempfehlungen bei unzureichender Schätzgüte als letzte Möglichkeit der Überprüfung angegeben. Durch die Einbringung solch einer automatisierten Überprüfung könnten ein eventueller Filterentwurfsprozess frühzeitig abgebrochen und neue Anforderungen an den Entwicklungsprozess des Systems gestellt werden.

Eine Reduktion der Rechenbelastung kann durch die Anwendung eines reduzierten UKF erreicht werden. Hierzu muss das System hinsichtlich linearer Anteile analysiert werden. Nach der Identifikation eines beispielsweise linearen Ausgangssystems könnte dieses aus der Sigma-Punkte-Transformation ausgenommen werden und mittels eines linearen Kalman-Filter-Ansatzes eingebunden werden.

Das in einer Kooperation entwickelte Interface zur Einbindung von FMUs eignet sich für eine Vielzahl weiterer regelungstechnischer Funktionen. Zum Einen können neben den für einen Onlineeinsatz gedachten Sigma-Punkte-Filtern ebenfalls weitere Techniken zur Zustands- und Parameterschätzung realisiert werden, wie Moving-Horizon oder Particle-Filter, die sich bei eher langsamer Systemdynamik und hoher Rechenleistung für den Offlineeinsatz eignen. In [KST14] wird die Schnittstelle für eine teilautomatisierte Modellidentifikation und Validierung genutzt. Ein weiteres Anwendungsgebiet kann die Hinzunahme von Techniken des maschinellen Lernens sein, um die physikalische Modellierung der FMUs um Neuronale Netze zur Abbildung von weiteren Effekten zu erweitern.

Das entwickelte Interface sollte im nächsten Schritt auf die Einbindung von FMUs 2.0 erweitert werden, da diese bei der Generierung aus Dymola schon einen Fixed-step-Solver enthalten und somit die externe Integration des Modells unter MATLAB entfällt. Die Möglichkeit der Einbindung von Fixed-Step-Solvern besteht allerdings erst seit der Dymola-Version 2014 FD01, die zum Zeitpunkt der Arbeit noch nicht zur Verfügung stand.

8 Literaturverzeichnis

- [AHH10] ARASARATNAM, I. ; HAYKIN, S. ; HURD, T.: Cubature Kalman filtering for continuous-discrete systems: theory and simulations. In: *IEEE Transactions on Signal Processing* Volume 58 (2010), S. 4977–4993
- [BN06] BISHOP, C. ; NASRABADI, N.: *Pattern recognition and machine learning*. Bd. 1. Springer, New York, 2006
- [BOA⁺11] BLOCHWITZ, T. ; OTTER, M. ; ARNOLD, M. u. a.: The functional mockup interface for tool independent exchange of simulation models. In: *Proceedings of the 8th International Modelica Conference*. Dresden : Linköping University Press, 2011, S. 20–22
- [BOZ11] BREMBECK, J. ; OTTER, M. ; ZIMMER, D.: Nonlinear Observers based on the Functional Mockup Interface with Application to Electric Vehicles. In: *Proceedings of the 8th International Modelica Conference*. Dresden : Linköping University Press, 2011, S. 474–483
- [BS94] BRAMMER, Karl ; SIFFLING, Gerhard ; FOELLINGER, Otto (Hrsg.) ; SARTORIUS, Hans (Hrsg.) ; KREBS, Volker (Hrsg.): *Kalman-Bucy-Filter*. 4. R. Oldenbourg Verlag, 1994
- [CF56] CARLTON, A. ; FOLLIN, J.: Recent developments in fixed and adaptive filtering. In: *Proceedings of the second guided missiles seminar: guidance and control*, North Atlantic Treaty Organization, Paris, 1956
- [Cho13] CHOMBART, P.: Multidisciplinary modelling and simulation speeds development of automotive systems and software / Dassault Systèmes. 2013. – Forschungsbericht
- [CHZ⁺13] CHEN, Y. ; HU, J. ; ZHANG, F. u. a.: EMG-based estimation of knee joint angle under functional electrical stimulation using an artificial neural network. In: *32nd Chinese Control Conference (CCC)*. Xian, China, 2013, S. 4661–4665
- [Con10a] CONSORTIUM, MODELISAR: Functional Mock-up Interface for Co-simulation / MODELISAR. 2010. – Forschungsbericht
- [Con10b] CONSORTIUM, MODELISAR: Functional Mock-up Interface for Model Exchange / MODELISAR. 2010. – Forschungsbericht
- [Ebd08] EBDEN, M.: Gaussian Process for Regression: A Quick Introduction / University of Oxford. 2008. – Forschungsbericht

- [Ehr09] EHRENSPIEL, K.: *Integrierte Produktentwicklung–Denkabläufe, Methodeneinsatz, Zusammenarbeit*, vol. 4. Bd. 5. Carl Hanser Verlag, München, 2009
- [EKT08] ECKEY, H. ; KOSFELD, R. ; TÜRCK, M.: *Deskriptive Statistik: Grundlagen-Methoden-Beispiele*. Bd. 5. Gabler Verlag, Wiesbaden, 2008
- [Elm78] ELMQVIST, H.: *A Structured Model Language for Large Continuous Systems*, Department of Automatic Control, Lund University, Sweden, Diss., Mai 1978
- [Föl78] FÖLLINGER, O.: *Regelungstechnik. Einführung in die Methoden und ihre Anwendung*. Bd. 10. VDE VERLAG, Berlin, 1978
- [Föl13] FÖLLINGER, O. ; KONIGORSKI, U. (Hrsg.) ; LOHMANN, B. (Hrsg.) ; ROPPENECKER, G. (Hrsg.) ; TRÄCHTLER, A. (Hrsg.): *Regelungstechnik. Einführung in die Methoden und ihre Anwendung*. Bd. 11. VDE VERLAG, Berlin, 2013
- [FP03] FELDKAMP, L. ; PROKHOROV, D.: Recurrent Neural Networks for State Estimation / Research and Advanced Engineering Ford Motor Company, Dearborn, USA. 2003. – Forschungsbericht
- [Frä05] FRÄNKEN, D.: *Dynamische Zustandsschätzung*. Vorlesungsskript, Paderborn, 2005. – Studienmodell Kommunikationstechnik, Wahlpflichtfach and Fachprüfung
- [GA10] GREWAL, M. ; ANDREWS, A.: Applications of Kalman Filtering in Aerospace 1960 to the Present [Historical Perspectives]. In: *Control Systems, IEEE* 30 (2010), Nr. 3, S. 69–78
- [Gau57] GAUSS, C. ; DAVIS, Charles H. (Hrsg.): *Theory of the motion of the heavenly bodies moving about the sun in conic sections: a translation of Carl Frdr. Gauss "Theoria motus"*. Little, Brown and Company, Boston, 1857
- [Gib11] GIBBS, B.P.: *Advanced Kalman filtering, least-squares and modeling: a practical handbook*. Bd. 1. John Wiley & Sons, New York, 2011
- [GKS⁺12] GRÄBER, M. ; KIRCHES, C. ; SCHARFF, D. u. a.: Using Functional Mock-up Units for Nonlinear Model Predictive Control. In: *Proceedings of the 9th International MODELICA Conference*. München : Linköping University Press, 2012, S. 781–790
- [GSS93] GORDON, N.J. ; SALMOND, D.J. ; SMITH, A.F.: Novel approach to nonlinear/non-Gaussian Bayesian state estimation. In: *IEE Proceedings-F Radar and Signal Processing*, 140 (1993), Nr. 2, S. 107–113

- [GTS14] GAUSEMEIER, J. ; TRÄCHTLER, A. ; SCHÄFER, W.: *Semantische Technologien im Entwurf mechatronischer Systeme: Effektiver Austausch von Lösungswissen in Branchenwertschöpfungsketten*. Carl Hanser Verlag, München, 2014
- [HE04] HU, X. ; ERSSON, T.: Active state estimation of nonlinear systems. In: *Automatica* 40 (2004), Nr. 12, S. 2075–2082
- [HXS⁺07] HAO, Y. ; XIONG, Z. ; SUN, F. u. a.: Comparison of unscented-kalman-filters. In: *International Conference on Mechatronics and Automation (ICMA), 2007*. Harbin, China, 2007, S. 895–899
- [IKS10] IMSLAND, L. ; KITILSEN, P. ; SCHEI, T.: Model-Based Optimizing Control and Estimation Using Modelica Model. In: *Modeling, Identification and Control* 31 (2010), Nr. 3, S. 107–121
- [Ise95] ISERMANN, R.: Mechatronische Systeme. In: *at - Automatisierungstechnik* Volume 43 (1995)
- [Ise02] ISERMANN, R.: *Mechatronische Systeme: Grundlagen*. Bd. 2. Springer, Berlin Heidelberg New York, 2002
- [IX00] ITO, K. ; XIONG, K.: Gaussian filters for nonlinear filtering problems. In: *IEEE Transactions on Automatic Control* 45 (2000), Nr. 5, S. 910–927
- [JL07] JULIER, S.J. ; LAVIOLA, J.J.: On Kalman filtering with nonlinear equality constraints. In: *IEEE Transactions on Signal Processing* Volume 55 (2007), Nr. 6, S. 2774–2784
- [JU96] JULIER, S.J. ; UHLMANN, J.K.: A general method for approximating nonlinear transformations of probability distributions / University of Oxford. 1996. – Forschungsbericht
- [JU97a] JULIER, S.J. ; UHLMANN, J.K.: Consistent debiased method for converting between polar and Cartesian coordinate systems. In: *The Proceedings of AeroSense: The 11th International Symposium on Aerospace/Defense Sensing, Simulation and Controls*, 1997, S. 110–121
- [JU97b] JULIER, S.J. ; UHLMANN, J.K.: New extension of the Kalman filter to nonlinear systems. In: *The Proceedings of AeroSense: The 11th International Symposium on Aerospace/Defense Sensing, Simulation and Controls*, 1997, S. 182–193
- [JU02] JULIER, S.J. ; UHLMANN, J.K.: Reduced sigma point filters for the propagation of means and covariances through nonlinear transformations. In: *Proceedings of the 2002 American Control Conference* Bd. 2 IEEE, 2002, S. 887–892

- [JU04] JULIER, S.J. ; UHLMANN, J.K.: Unscented filtering and nonlinear estimation. In: *Proceedings of the IEEE* Volume 92 (2004), Nr. 3, S. 401–422
- [JUDW95] JULIER, S.J. ; UHLMANN, J.K. ; DURRANT-WHYTE, H.F.: A new approach for filtering nonlinear systems. In: *Proceedings of the American Control Conference* Bd. 3, 1995, S. 1628–1632
- [Jul98] JULIER, S.J.: Skewed approach to filtering. In: *Signal and Data Processing of Small Targets*, Society of Photo Optical, 1998, S. 271–282
- [Jul02] JULIER, S.J.: The scaled unscented transformation. In: *Proceedings of the American Control Conference* Bd. 6, 2002, S. 4555–4559
- [Jul03] JULIER, S.J.: The spherical simplex unscented transformation. In: *Proceedings of the American Control Conference* Bd. 3, 2003, S. 2430–2434
- [Kal60] KALMAN, R.E.: A new approach to linear filtering and prediction problems. In: *Journal of basic Engineering* Volume 82 (1960), Nr. 1, S. 35–45
- [KHP⁺14] KAAKINEN, M. ; HUTTUNEN, S. ; PAAVOLAINEN, L. u.a.: Automatic detection and analysis of cell motility in phase-contrast time-lapse images using a combination of maximally stable extremal regions and Kalman filter approaches. In: *Journal of microscopy* Volume 253 (2014), Nr. 1, S. 65–78
- [KHW12] KITTILSEN, P. ; HAUGER, S.O. ; WASBØ, S.: Designing models for online use with Modelica and FMI. In: *Proceedings of the 9th International Modelica Conference*. München : Linköping University Press, 2012, S. 197–204
- [KKS14] KIM, J. ; KIM, K. ; SOHN, H.: Autonomous dynamic displacement estimation from data fusion of acceleration and intermittent displacement measurements. In: *Mechanical Systems and Signal Processing* Volume 42 (2014), Nr. 1, S. 194–205
- [Kre80] KREBS, Volker ; FOELLINGER, Otto (Hrsg.) ; SARTORIUS, Hans (Hrsg.): *Nichtlineare Filterung*. R. Oldenbourg Verlag, 1980 (Methoden der Regelungstechnik)
- [Kri51] KRIGE, D. G.: *A statistical approach to some mine valuation and allied problems on the Witwatersrand*, University of the Witwatersrand, Johannesburg, Diss., 1951
- [Kru09] KRUSE, D.: *Entwurf und Auslegung der Aufbauregelung für ein schweres Geländefahrzeug mit aktiver Federung*, Universität Paderborn, Diplomarbeit, 2009

- [KST14] KRUSE, D. ; SCHWEERS, C. ; TRÄCHTLER, A.: Methodolgy for a partly automated parameter identification to validate multi-domain models. In: *Proceedings of the ASME 2014 International Mechanical Engineering Congress & Exposition IMECE 2014*. Montreal, Québec, Canada, November 2014
- [KTO04] KANKAINEN, A. ; TASKINEN, S. ; OJA, H.: On Mardias tests of multinormality. In: *Statistics for Industry and Technology*. Birkhäuser Verlag Basel, 2004, S. 153–164
- [Kus67] KUSHNER, H: Approximations to optimal nonlinear filters. In: *IEEE Transactions on Automatic Control* Volume 12 (1967), Nr. 5, S. 546–556
- [LBG⁺12] LI, J. ; BARILLAS, J. ; GUENTHER, C. u. a.: A Comparative Study of State of Charge Estimation Algorithms for LiFePO Batteries used in Electric Vehicles. In: *Journal of Power Sources* Volume 230 (2012), S. 244–250
- [LCD⁺08] LOZANO, J.G. ; CARRILLO, L. ; DZUL, A. u. a.: Spherical simplex sigma-point Kalman filters: a comparison in the inertial navigation of a terrestrial vehicle. In: *Proceedings of the American Control Conference*. Seattle, USA, 2008, S. 3536–3541
- [Ler02] LERNER, U.N.: Hybrid Bayesian networks for reasoning about complex systems / Stanford University. 2002. – Forschungsbericht
- [Lin09] LINDEMANN, U.: *Methodische Entwicklung technischer Produkte: Methoden flexibel und situationsgerecht anwenden*. Bd. 3. Springer, 2009
- [LMW14] LEVANT, J. ; MA, J. ; WOHR, M.E.: Stock Return and Inflation: An Analysis Based on the State-Space Framework. In: *Recent Advances in Estimating nonlinear Models*. Springer, 2014, S. 1–13
- [LOT14] LOCHBICHLER, M. ; OESTERSÖTEBIER, F. ; TRÄCHTLER, A.: Dynamic Behavior Models and their Modeling Depth in the Design Process of Mechatronic Systems. In: *Proceedings of the ASME 2014 International Mechanical Engineering Congress & Exposition IMECE 2014*. Montreal, Québec, Canada, November 2014
- [LPG⁺09] LI, Z. ; PLI, Z. ; GAO, D. u. a.: An Improved Unscented Kalman Filter Based on STF for Nonlinear Systems. In: *2nd International Congress on Image and Signal Processing*. Tianjin, China, 2009
- [LWM11] LIU, G. ; WORGOTTER, F. ; MARKELIC, I.: Nonlinear estimation using central difference information filter. In: *Statistical Signal Processing Workshop (SSP)*. Nice, France, 2011, S. 593–596
- [LXZ⁺08] LI, H. ; XU, D. ; ZHANG, F. u. a.: Sequence unscented Kalman filtering algorithm. In: *3rd IEEE Conference on Industrial Electronics and Applications*. Singapore, 2008, S. 1374–1378

- [Mac98] MACKAY, D.: Introduction to monte carlo methods. In: *Learning in graphical models*. A Bradford Book, Cambridge, USA, 1998, S. 175–204
- [May79] MAYBECK, P.S.: *Stochastic models, estimation, and control, Volume 1*. Academic Press New York, San Francisco, London, 1979
- [MHT09] MUNROE, M. ; HENCH, T. ; TIMMONS, M.: Sales Rate and Cumulative Sales Forecasting Using Kalman Filtering Techniques / Intel Corporation. 2009. – Forschungsbericht
- [Mün12] MÜNCH, E.: *Selbstoptimierung verteilter mechatronischer Systeme auf Basis paretooptimaler Systemkonfigurationen*, Universität Paderborn, Diss., 2012
- [Nel00] NELSON, A.T.: *Nonlinear estimation and modeling of noisy time series by dual Kalman filtering methods*, Oregon Graduate Institute of Science and Technology, Diss., 2000
- [Ott93] OTTER, M.: *Objektorientierte Modellierung mechatronischer Systeme am Beispiel geregelter Roboter*, Ruhr-Universität Bochum, Diss., 1993
- [PBF⁺06] PAHL, G. ; BEITZ, W. ; FELDHUSEN, J. u. a.: *Konstruktionslehre*. Bd. 7. Springer, Berlin, Heidelberg, New-York., 2006
- [PC14] PAVLOV, I. ; CLARK, J.: *FMU SDK 1.0.2*. Online. https://resources.qtronic.de/fmusdk/FmuSdk_reference.html. Version: 2014
- [PU92] PSICHOGIOS, D.C. ; UNGAR, L.H.: A hybrid neural network-first principles approach to process modeling. In: *American Institute of Chemical Engineers, AiChE Journal* 38 (1992), Nr. 10, S. 1499–1511
- [QXL⁺10] QU, C. ; XU, Z. ; LI, N. u. a.: An Improved UKF algorithm based on Rauch-Tung-Striebel smoother. In: *3rd International Congress on Image and Signal Processing (CISP)* Bd. 8. Yantai, China, 2010, S. 4020–4024
- [RC04] ROBOTHERS, P. ; CASELLA, G.: *Monte Carlo statistical methods*. Springer, Berlin Heidelberg New York, 2004
- [RM51] ROBBINS, H. ; MONRO, S.: A stochastic approximation method. In: *The Annals of Mathematical Statistics* 22 (1951), Nr. 3, S. 400–407
- [Ros06] ROSS, S.M.: *A first course in probability*. Springer, Berlin Heidelberg New York, 2006
- [RW96] RASMUSSEN, C.E. ; WILLIAMS, C.: Gaussian Process for Regression. In: *Advances in Neural Information Processing Systems* 8 Bd. 8, The MIT Press, Cambridge, London, 1996, S. 514–520

- [RW06] RASMUSSEN, C.E. ; WILLIAMS, C.: *Gaussian Process for Machine Learning*. The MIT Press, Cambridge, London, 2006
- [Sch97] SCHEL, T.S.: A finite-difference method for linearization in nonlinear estimation algorithms. In: *Automatica* 33 (1997), Nr. 11, S. 2053–2058
- [Sch09] SCHRÖDER, D.: *Elektrische Antriebe-Regelung von Antriebssystemen*. Bd. 3. Springer, Berlin Heidelberg New York, 2009
- [Sch11] SCHNEIDER, P.: *Sigma-Punkt Kalman-Filter mit Ungleichungsnebenbedingungen*, Universität des Saarlandes, Saarbrücken, Diss., 2011
- [SDS⁺13] STRAKA, O. ; DUNIK, J. ; SIMANDL, M. u. a.: Aspects and comparison of matrix decompositions in unscented Kalman filter. In: *Proceedings of the American Control Conference* Bd. Volume 7. Washington, USA : Wiley Online Library, 2013, S. 3075–3080
- [SEF04] SHISHKO, R. ; EBBELER, D.H. ; FOX, G.: NASA technology assessment using real options valuation. In: *Systems Engineering* 7 (2004), Nr. 1, S. 1–13
- [SH11] SOOTHERS, E. ; HAJIYEV, C.: REKF and RUKF development for pico satellite attitude estimation in the presence of measurement faults. In: *5th International Conference on recent Advances in Space Technologies*. Istanbul, Türkei, 2011, S. 891–896
- [Sim06] SIMON, D.: *Optimal state estimation: Kalman, H infinity, and nonlinear approaches*. John Wiley & Sons, New York, 2006
- [Sim10] SIMON, D: Kalman filtering with state constraints: a survey of linear and nonlinear algorithms. In: *Control Theory & Applications, IET* Volume 4 (2010), Nr. 8, S. 1303–1318
- [SLPP⁺13] SANCHEZ-LOPEZ, J.L. ; PESTANA, J. ; PUENTE, P. de l. u. a.: Visual Quadrotor Swarm for IMAV 2013 Indoor Competition. In: *ROBOT2013: First Iberian Robotics Conference, Universidad Politécnica de Madrid*, Springer International Publishing, 2013, S. 55–63
- [SM85] SCHMIDT, Stanley F. ; MCGEE, Leonard A.: Discovery of the Kalman Filter as a Practical Tool for Aerospace and Industry / NASA, Ames Research Center, Moffett Field California. 1985. – Forschungsbericht
- [Sne07] SNELSON, E.L.: *Flexible and efficient Gaussian process models for machine learning*, University of London, Diss., 2007
- [Sor70] SORENSON, H. W.: Least-Squares estimation: from Gauss to Kalman. In: *IEEE Spectrum* 7 (1970), S. 63–68

- [STK13a] SCHWEERS, C. ; TRÄCHTLER, A. ; KRUSE, D.: Automated Design of an Unscented Kalman Filter for State- and Parameter Estimation on unknown Models. In: *IEEE International Conference on Control, Automation, Robotics & Embedded Systems*. Jabalpur, Indien, 2013
- [STK13b] SCHWEERS, C. ; TRÄCHTLER, A. ; KRUSE, D.: Entwurf eines Unscented-Kalman Filters zur Zustands- und Parameterschätzung an Dymola-Modellen. In: *Tagungsband VDI Mechatroniktagung*. Dresden, 2013
- [TCT⁺09] TEIXEIRA, B. ; CHANDRASEKAR, J. ; TÔRRES, L. u.a.: State estimation for linear and non-linear equality-constrained systems. In: *International Journal of Control* 82 (2009), Nr. 5, S. 918–936
- [Trä12a] TRÄCHTLER, A.: *Modellbildung in der Mechatronik*, Universität Paderborn. 2012. – Vorlesungsskript, Universität Paderborn
- [TR12b] TURNER, R. ; RASMUSSEN, C.E.: Model based learning of sigma points in unscented Kalman filtering. In: *Neurocomputing* 80 (2012), S. 47–53
- [TS03] TENNE, D. ; SINGH, T.: The higher order unscented filter. In: *Proceedings of the American Control Conference* Bd. 3. Denver, USA, 2003, S. 2441–2446
- [Tur11] TURNER, R.D.: *Gaussian Process for State Space Models and Change Point Detection*, University of Cambridge, Diss., 2011
- [VDI04] *VDI 2206 Entwicklungsmethodik für mechatronische Systeme*. VDI Gesellschaft Produkt- und Prozessgestaltung, Beuth Verlag GmbH, 2004
- [VDM04] VAN DER MERWE, R.: *Sigma-point Kalman filters for probabilistic inference in dynamic state-space models*, University of Stellenbosch, Diss., 2004
- [VDMW01a] VAN DER MERWE, R. ; WAN, E.A.: Efficient derivative-free Kalman filters for online learning. In: *European Symposium on Artificial Neural Networks*. Bruges, Belgien, 2001, S. 205–210
- [VDMW01b] VAN DER MERWE, R. ; WAN, E.A.: The square-root unscented Kalman filter for state and parameter-estimation. In: *IEEE International Conference on Acoustics, Speech, and Signal Processing* Bd. 6. Salt Lake City, USA, 2001, S. 3461–3464
- [VDMW⁺04] VAN DER MERWE, R. ; WAN, E.A. u.a.: Sigma-point Kalman filters for nonlinear estimation Estimation and Sensor-Fusion - Applications to Integrated Navigation -. In: *Proceedings of the AIAA Guidance, Navigation, and Control Conference*. Providence, USA, 2004, S. 1735–1764

- [WB06] WELCH, G. ; BISHOP, G.: An introduction to the Kalman Filter / Department of Computer Science, University of North Carolina. 2006. – Forschungsbericht
- [WBT05] WIGREN, T. ; BRUS, L. ; TAYAMON, S.: MATLAB software for recursive identification and scaling using a structured nonlinear black box model revision 2 / Systems and Control, Department of Information Technology, Uppsala University, Schweden. 2005. – Forschungsbericht
- [Wie64] WIENER, N.: *Extrapolation, interpolation, and smoothing of stationary time series: with engineering applications*. Bd. 8. MIT Press, Cambridge, USA, 1964
- [WN01] WAN, E.A. ; NELSON, A.T.: Dual extended Kalman filter methods. In: *Kalman Filtering and Neural Networks*. John Wiley & Sons, New York, USA., 2001
- [WVDM00] WAN, E.A. ; VAN DER MERWE, R.: The unscented Kalman filter for nonlinear estimation. In: *Adaptive Systems for Signal Processing, Communications, and Control Symposium*, IEEE, New York City, 2000, S. 153–158
- [WVDM01] WAN, E.A. ; VAN DER MERWE, R.: The unscented Kalman filter. In: HAYKIN, S. (Hrsg.): *Kalman Filtering and Neural Networks*. John Wiley & Sons, New York, USA., 2001, S. 221–280
- [WVDMN99] WAN, E.A. ; VAN DER MERWE, R. ; NELSON, A.T.: Dual Estimation and the Unscented Transformation. In: *In Neural Information Processing Systems*, MIT Press, Cambridge, USA, 1999, S. 666–672
- [XLD⁺13] XU, L. ; LI, X. ; DUAN, Z. u. a.: Modeling and State Estimation for Dynamic Systems with Linear Equality Constraints. In: *IEEE Transactions on Signal Processing* Volume 61 (2013), Nr. 11, S. 2927–2939
- [XSC⁺14] XIONG, K. ; SUN, F. ; CHEN, Z. u. a.: A data-driven multi-scale extended Kalman filtering based parameter and state estimation approach of lithium-ion olymer battery in electric vehicles. In: *Applied Energy, Elsevier, Amsterdam* Volume 113 (2014), S. 463–476
- [YS06] YADAIAH, N. ; SOWMYA, G.: Neural network based state estimation of dynamical systems. In: *International Joint Conference on Neural Networks*. Vancouver, Kanada, 2006, S. 1042–1049
- [ZMC07] ZHANG, L. ; MA, H. ; CHEN, Y.: Square-root unscented Kalman filter for vehicle integrated navigation. In: *International Conference on Machine Learning and Cybernetics* Bd. 1. Hong Kong, China, 2007, S. 556–561

- [ZZH⁺14] ZHONG, L. ; ZHANG, C. ; HE, Y. u. a.: A method for the estimation of the battery pack state of charge based on in-pack cells uniformity analysis. In: *Applied Energy* 113 (2014), S. 558–564

A Anhang

A.1 Square-Root Unscented Kalman Filter

Die rekursive Berechnung der Matrixwurzel $S_k S_k^T = P_{xx,k}$ stellt den neben der Transformation der Sigma-Punkte den rechenintensivsten Schritt des UKF dar, und gleichzeitig ist positive Definitheit nicht garantiert. Die nachfolgend dargestellte Implementierung ersetzt die rekursiv durchzuführende Berechnung der Matrixwurzel durch ein Cholesky-Update und ist [VDMW01b] entnommen. Diese hat neben der Reduktion der Berechnungszeit noch den wichtigen Vorteil, dass positive Definitheit garantiert wird, solange die durch das Modell transformierten Sigma-Punkte nicht unzulässig sind, wie beispielsweise wegen einer Modellinstabilität.

Initialisierung Filter

Startmomente:

$$\begin{aligned}\hat{x}_0 &= E[x_0] \\ S_0 &= \text{chol} \left(E \left[(x_0 - \hat{x}_0)(x_0 - \hat{x}_0)^T \right] \right)\end{aligned}$$

Betriebsphase Zeitschritt $k \in \{1, \dots, \infty\}$:

Berechnung Sigma-Punkte:

$$\mathcal{X}_{k-1} = [\hat{x}_{k-1} \quad \hat{x}_{k-1} + \gamma S_k \quad \hat{x}_{k-1} - \gamma S_k]$$

Zeit-Update

$$\begin{aligned}\mathcal{X}_{k|k-1}^* &= F[\mathcal{X}_{k-1}, u_{k-1}] \\ \hat{x}_k^- &= \sum_{i=0}^{2L} W_i^{(m)} \mathcal{X}_{i,k|k-1}^* \\ S_k^- &= \text{qr} \left(\begin{bmatrix} \sqrt{W_1^{(c)}} & (\mathcal{X}_{1:2n,k|k-1} - \hat{x}_k^-) & \sqrt{R_v} \end{bmatrix} \right) \\ S_k^- &= \text{cholupdate} \left(S_k^-, \mathcal{X}_{0,k}^* - \hat{x}_k^-, W_0^{(c)} \right) \\ \mathcal{X}_{k|k-1} &= [\hat{x}_k^- \quad \hat{x}_k^- + \gamma S_k^- \quad \hat{x}_k^- - \gamma S_k^-] \\ \mathcal{Y}_{k|k-1} &= h[\mathcal{X}_{k|k-1}] \\ \hat{y}_k^- &= \sum_{i=0}^{2n} W_i^{(m)} \mathcal{Y}_{i,k|k-1}\end{aligned}$$

Mess-Update

$$\begin{aligned}
S_{\tilde{y}_k} &= \text{qr} \left(\begin{bmatrix} \sqrt{W_1^{(c)}} & (\mathcal{Y}_{1:2n,k} - \hat{y}_k) & \sqrt{R_{n,k}} \end{bmatrix} \right) \\
S_{\tilde{y}_k} &= \text{cholupdate} \left(S_{\tilde{y}_k} \quad \mathcal{Y}_{0,k} - \hat{y}_k \quad W_0^{(c)} \right) \\
P_{xy,k} &= \sum_{i=0}^{2L} W_i^{(c)} [\mathcal{X}_{i,k|k-1} - \hat{x}_k^-] [\mathcal{Y}_{i,k|k-1} - \hat{y}_k^-]^T \\
\mathcal{K}_k &= P_{xy,k} P_{yy,k}^{-1} \\
\hat{x}_k &= \hat{x}_k^- + \mathcal{K}_k (y_k - \hat{y}_k^-) \\
P_{xx,k} &= P_{xx,k}^- - \mathcal{K}_k P_{yy,k} \mathcal{K}_k^T
\end{aligned}$$

A.2 Central-Difference-Filter

Nachfolgend ist der Algorithmus des Central-Difference-Kalman-Filters für den Fall additiven Rauschens dargestellt. Das Central-Difference-Filter kann genau wie das UKF als Square-Root-Variante implementiert werden.

Initialisierung Filter

Startmomente:

$$\begin{aligned}
\hat{x}_0 &= E[x_0] \\
P_{xx,0} &= E[(x_0 - \hat{x}_0)(x_0 - \hat{x}_0)^T]
\end{aligned}$$

Betriebsphase Zeitschritt $k \in \{1, \dots, \infty\}$:

Berechnung Sigma-Punkte:

$$\mathcal{X}_{k-1} = \begin{bmatrix} \hat{x}_{k-1} & \hat{x}_{k-1} + \xi \sqrt{P_{xx,k-1}} & \hat{x}_{k-1} - \xi \sqrt{P_{xx,k-1}} \end{bmatrix}$$

Zeit-Update

$$\begin{aligned}
\mathcal{X}_{k|k-1}^* &= g[\mathcal{X}_{k-1}, u_{k-1}] \\
\hat{x}_k^- &= \sum_{i=0}^{2n} W_i^{(m)} \mathcal{X}_{i,k|k-1}^* \\
P_{xx,k}^- &= \sum_{i=1}^{2n} W_i^{(c_1)} [\mathcal{X}_{i,k|k-1}^* - \mathcal{X}_{n+i,k|k-1}^*]^2 \\
&\quad + W_i^{(c_2)} [\mathcal{X}_{i,k|k-1}^* + \mathcal{X}_{n+i,k|k-1}^* + 2\mathcal{X}_{0,k|k-1}^*]^2 + R_v
\end{aligned}$$

(Erneuern der Sigma-Punkte)

$$\begin{aligned}
\mathcal{X}_{k|k-1}^* &= \begin{bmatrix} \hat{x}_k^- & \hat{x}_k^- + h \sqrt{P_{x,k}^-} & \hat{x}_k^- - h \sqrt{P_{x,k}^-} \end{bmatrix} \\
\mathcal{Y}_{k|k-1} &= h[\mathcal{X}_{k|k-1}^*] \\
\hat{y}_k^- &= \sum_{i=0}^{2n} W_i^{(m)} \mathcal{Y}_{i,k|k-1}
\end{aligned}$$

Mess-Update

$$\begin{aligned}
P_{yy,k} &= \sum_{i=1}^{2n} W_i^{(c_1)} [\mathcal{Y}_{i,k|k-1} - \mathcal{Y}_{n+i,k|k-1}]^2 \\
&\quad + W_i^{(c_2)} [\mathcal{Y}_{i,k|k-1} + \mathcal{Y}_{n+i,k|k-1} + 2\mathcal{Y}_{0,k|k-1}]^2 + R_n \\
P_{xy,k} &= \sqrt{W_1^{(c_1)} P_{xx,k}^-} [\mathcal{Y}_{1:n,k|k-1} - \mathcal{Y}_{n+1:2n,k|k-1}]^T \\
\mathcal{K}_k &= P_{xy,k} P_{yy,k}^{-1} \\
\hat{x}_k &= \hat{x}_k^- + \mathcal{K}_k (y_k - \hat{y}_k^-) \\
P_{xx,k} &= P_{xx,k}^- - \mathcal{K}_k P_{yy,k} \mathcal{K}_k^T
\end{aligned}$$

Der Parameter ξ stellt den in 3.4 vorgestellten Parameter zur Gewichtung der Sigma-Punkte und deren Berechnung dar, mit $\xi > 1$. Der Term n entspricht der Ordnung der Zufallsvariablen x und somit der Anzahl der Zustände. Mittels der Matrizen R_v bzw. R_n werden Prozessrauschkovarianz bzw. die Messrauschkovarianz in den Algorithmus eingebracht. Die Notation \square^2 steht für das Vektorprodukt $[a]^2 = aa^T$. Die Gewichtungsterme entsprechen: $W_0^{(m)} = \frac{\xi^2 - n}{\xi^2}$; $W_i^{(m)} = \frac{1}{2\xi^2}$; $W_i^{(c_1)} = \frac{1}{4\xi^2}$; $W_i^{(c_2)} = \frac{\xi^2 - 1}{4\xi^4}$ mit $i = 1, \dots, 2n + 1$. Die Herleitung der Berechnung von Erwartungswert und Kovarianz auf Basis von Sigma-Punkten, die mittels Sterling-Interpolation bestimmt wurden, sowie der entsprechenden Gewichtungen kann in [VDMW⁺04], [Sch11] nachvollzogen werden.

A.3 MATLAB-Toolbox

Die in dieser Dissertation erarbeiteten Methoden und Algorithmen sind in eine unter MATLAB erstellte Toolbox eingeflossen, die es Anwendern ermöglicht mit einem maximalen Automatisierungsgrad Sigma-Punkte-Filter für als FMU vorliegende Modelle zu erstellen. Die Toolbox ist so angelegt, dass sie dem in Kapitel 5 vorgestellten Vorgehensmodell entspricht. Im ersten Schritt ist vom Anwender das Modell als FMU einzuladen, woraufhin die in der *.xml* abgelegten Informationen in Form von Systemordnung, Zustandsbezeichnern, Ausgängen, Eingängen, Schätzparametern und deren Parametergrenzen angezeigt werden, wie in Bild 1-1 gezeigt.

Im Anschluss kann das Szenario eingeladen und bestehende Messdaten den Ein- und Ausgängen der Toolbox zugewiesen werden (vgl. Bild 1-2).

Im Anschluss kann eine Überprüfung des Modells vorgenommen werden, indem das Filtermodell mit den im Szenario hinterlegten Eingängen simuliert wird, und ein Vergleich der simulierten Modellausgänge kann mit den im Szenario hinterlegten, realen Messungen durchgeführt werden. Weichen simulierte Ausgänge und reale Messungen zu stark voneinander ab, so ist im Vorfeld der Filterauslegung eine Parameteridentifikation ratsam, wie in [KST14] vorgestellt, oder vielleicht ein

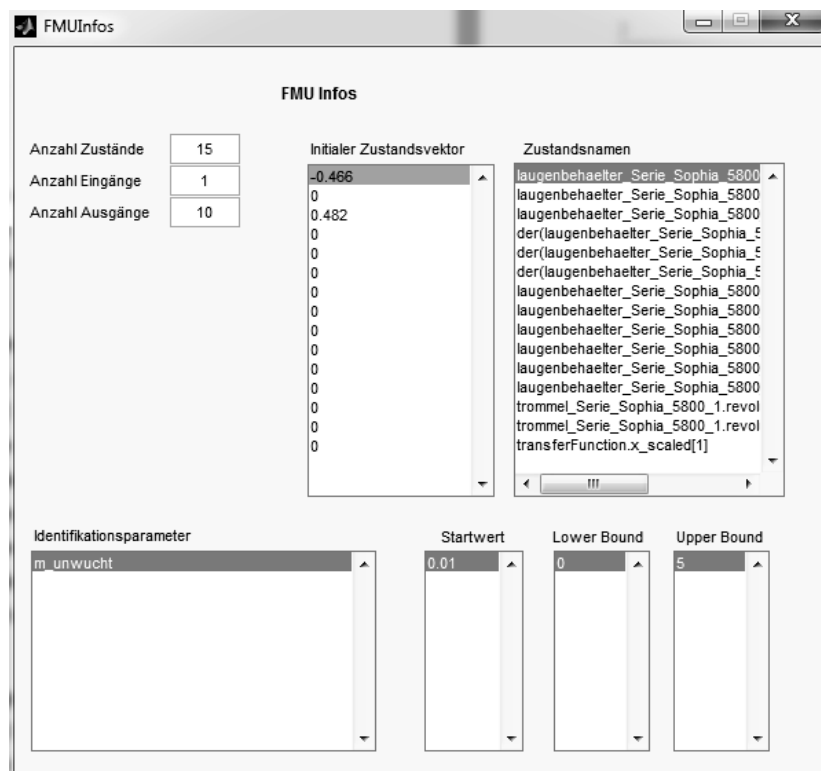


Bild 1-1: Modellinformationen nach Auslesen der FMU-xml

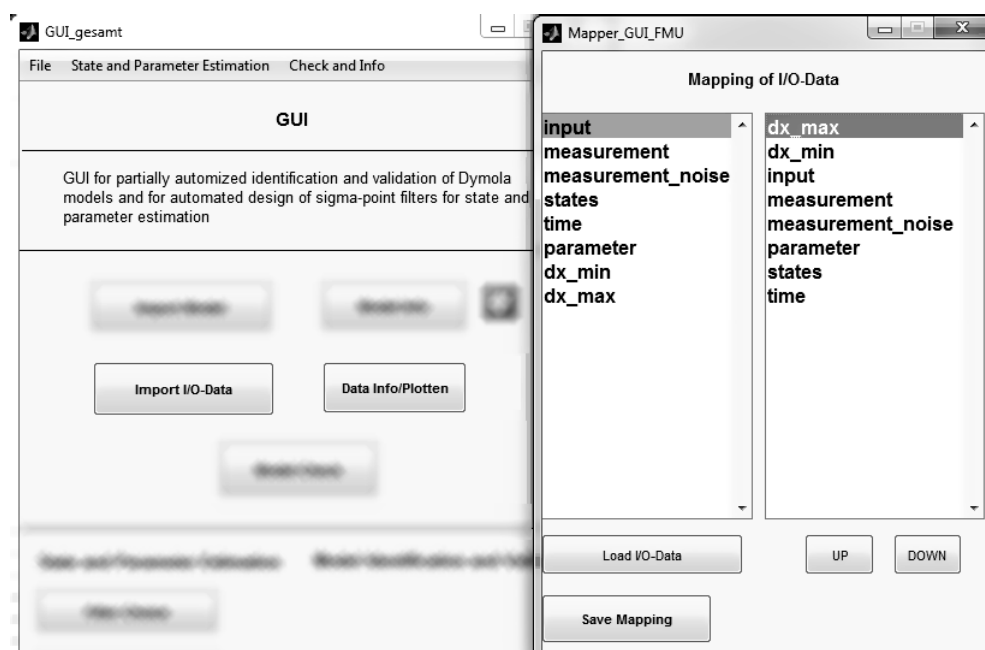


Bild 1-2: Einlesen von Szenarien und Zuordnung von Ein- und Ausgängen

Modell höherer Komplexität, wie in [LOT14] gezeigt. Bei zufriedenstellender Modellgüte können im Anschluss die Prozess- und Messrauschkovarianzen nach Bild 1-3 manuell oder automatisiert gewählt werden. Bei der Wahl einer automatisierten Vorgabe wird das in Abschnitt 4.2 vorgestellte Verfahren ausgeführt, und vom Nutzer ist nur noch die Angabe der Vertrauensfaktoren F_{R_v} und F_{R_n} notwendig, was durch einen Schieberegler in der Toolbox realisiert wurde.

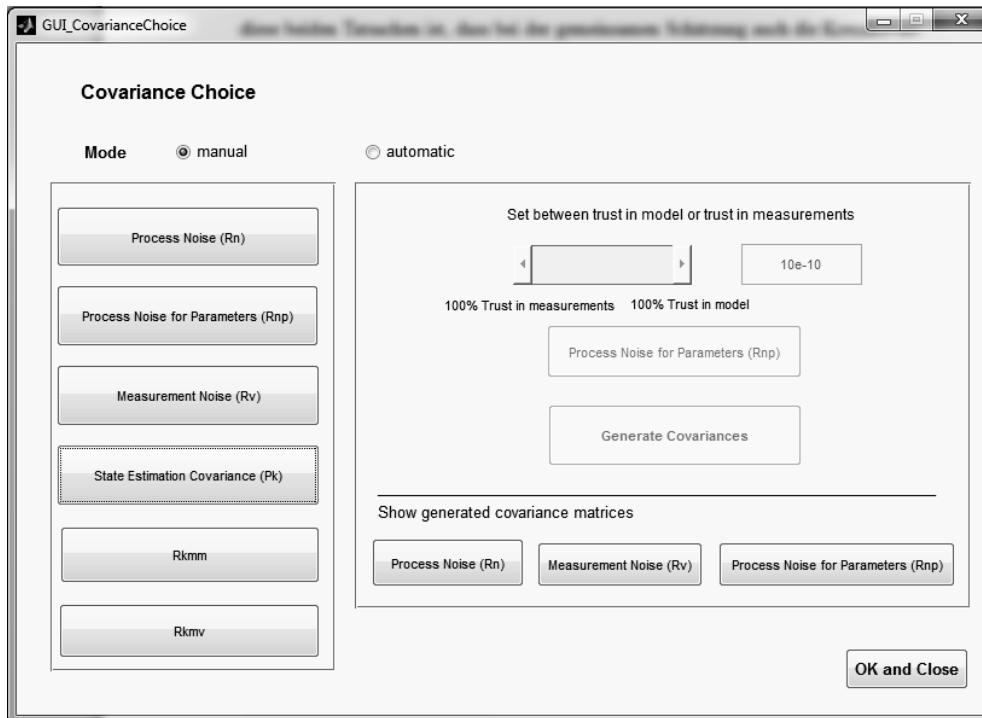


Bild 1-3: Automatisierte oder manuelle Generierung von Kovarianzmatrizen des Prozess- und Messrauschens

Falls in der FMU Schätzparameter mittels einer Markierung im Modellierungsprozess erkannt wurden, kann im nächsten Schritt vom Benutzer gewählt werden, ob die Parameter zur Schätzung mittels Joint- oder Dual-Verfahren freigegeben werden oder auf den Initialparametern verbleiben. Im Falle der Schätzung werden die im Dymolamodell hinterlegten Parametergrenzen direkt übernommen, können vom Anwender aber manuell verändert werden, wie in Bild 1-4 dargestellt. Abschließend erfolgt vom Benutzer noch die Wahl des zu verwendenden Filteralgorithmus in Form des UKF mit SUT oder des SST-Sigma-Punkte-Berechnungsverfahrens, des CDKF oder CKF.

Die Auslegung der im ausgewählten Filteralgorithmus enthaltenen Entwurfsparameter wird im nächsten Schritt durch die automatisierte Übergabe des Szenarios und der vom Nutzer getroffenen Einstellungen an das in Kapitel 4 vorgestellte Optimierungsverfahren durchgeführt. Der Benutzer hat an dieser Stelle noch die Möglichkeit, Abbruchbedingungen, den Gewichtungsfaktor k_w und Anzeigeeigenschaften einzustellen, wie in Bild 1-5 dargestellt. Die generierten, optimierten

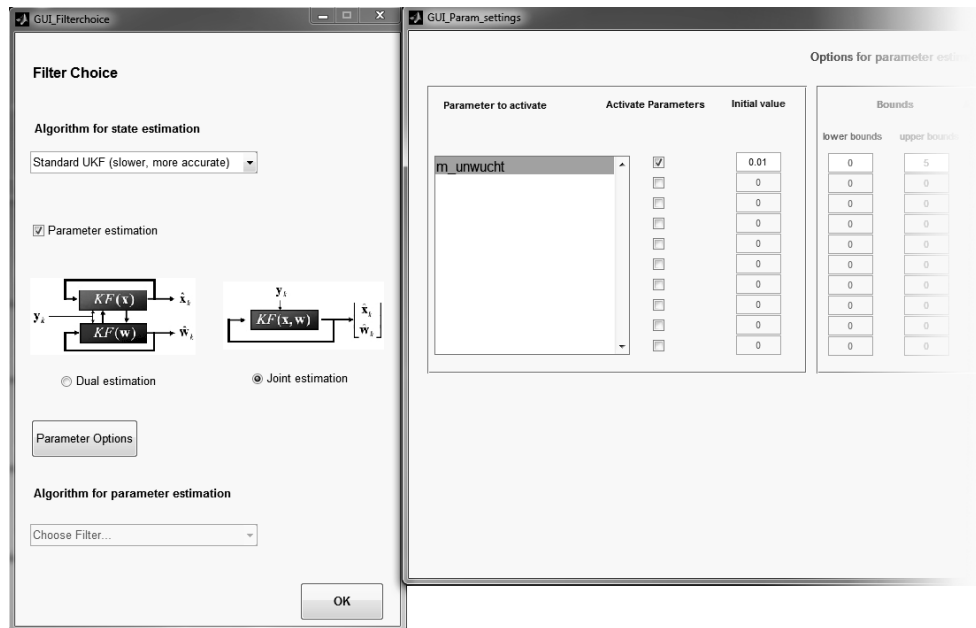


Bild 1-4: Auswahl des Parameterschätzverfahrens, Freigabe der zu schätzenden Parameter und Einstellung der Parametergrenzen

Werte der Entwurfsparameter werden angezeigt und intern für die weitere Verarbeitung abgelegt.

Nach Durchlaufen der Optimierung kann das entworfene Filter im Rahmen einer Robustheitsanalyse erprobt und die Schätzgüte beurteilt werden. Die Toolbox bietet für den gesamten Entwurfsprozess ein Schritt-für-Schritt-Vorgehen an, mit der abschließenden, noch zu entwickelnden Möglichkeit des Filterexports auf unterschiedliche Echtzeitsysteme oder auch Offlineumgebungen. Der gesamte Ablauf der Toolbox ist in Bild 1-6 dargestellt.

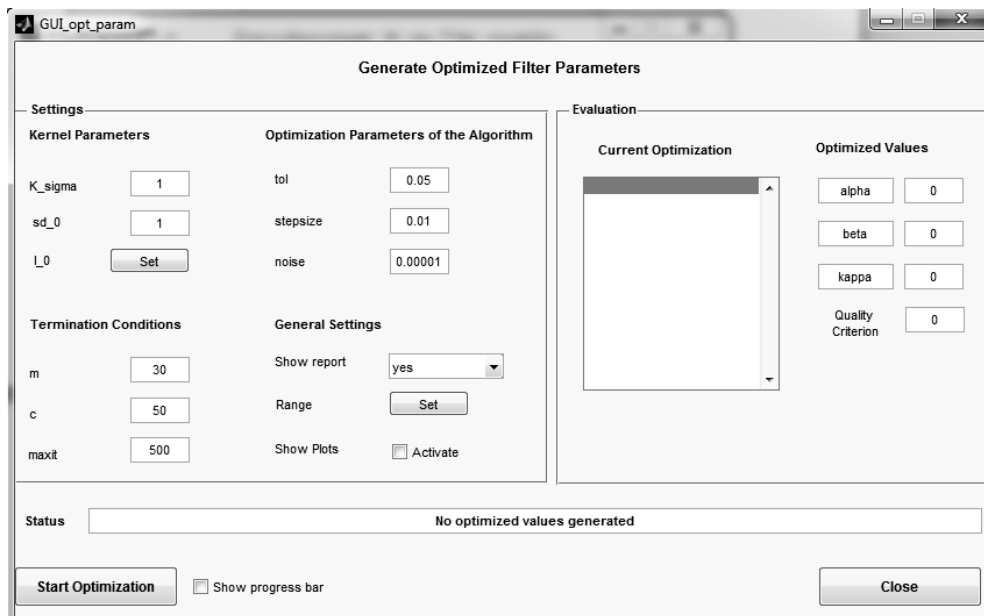


Bild 1-5: Optimierung der Filterentwurfparameter

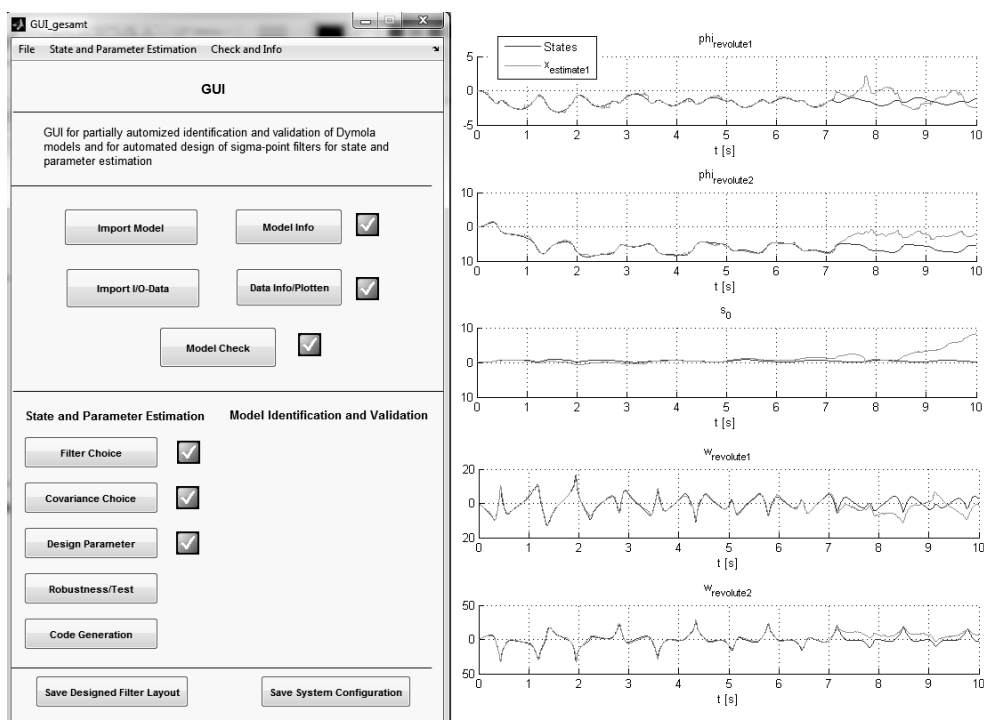


Bild 1-6: Optimierung der Filterentwurfparameter

A.4 Doppelpendel

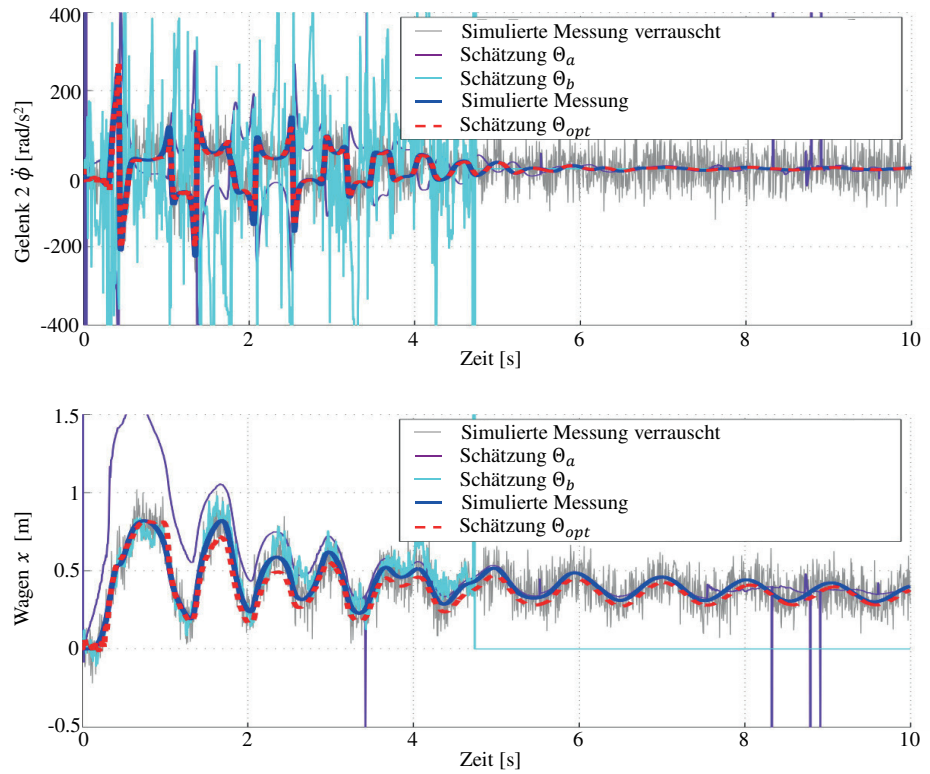


Bild 1-7: Oben, unten links: Systemausgang des Streckenmodells (blau); dem Filter zur Verfügung gestellte Systemausgänge, mit Rauschen überlagert (grau), Filterrekonstruktion des Systemausganges durch UKF mit $\theta_{a,b,opt}$ (lila, cyan, rot). Unten rechts: Schätzung der unbekannten Dichte des unteren Pendelarmes durch UKF mit $\theta_{a,b,opt}$ (lila, cyan, rot)

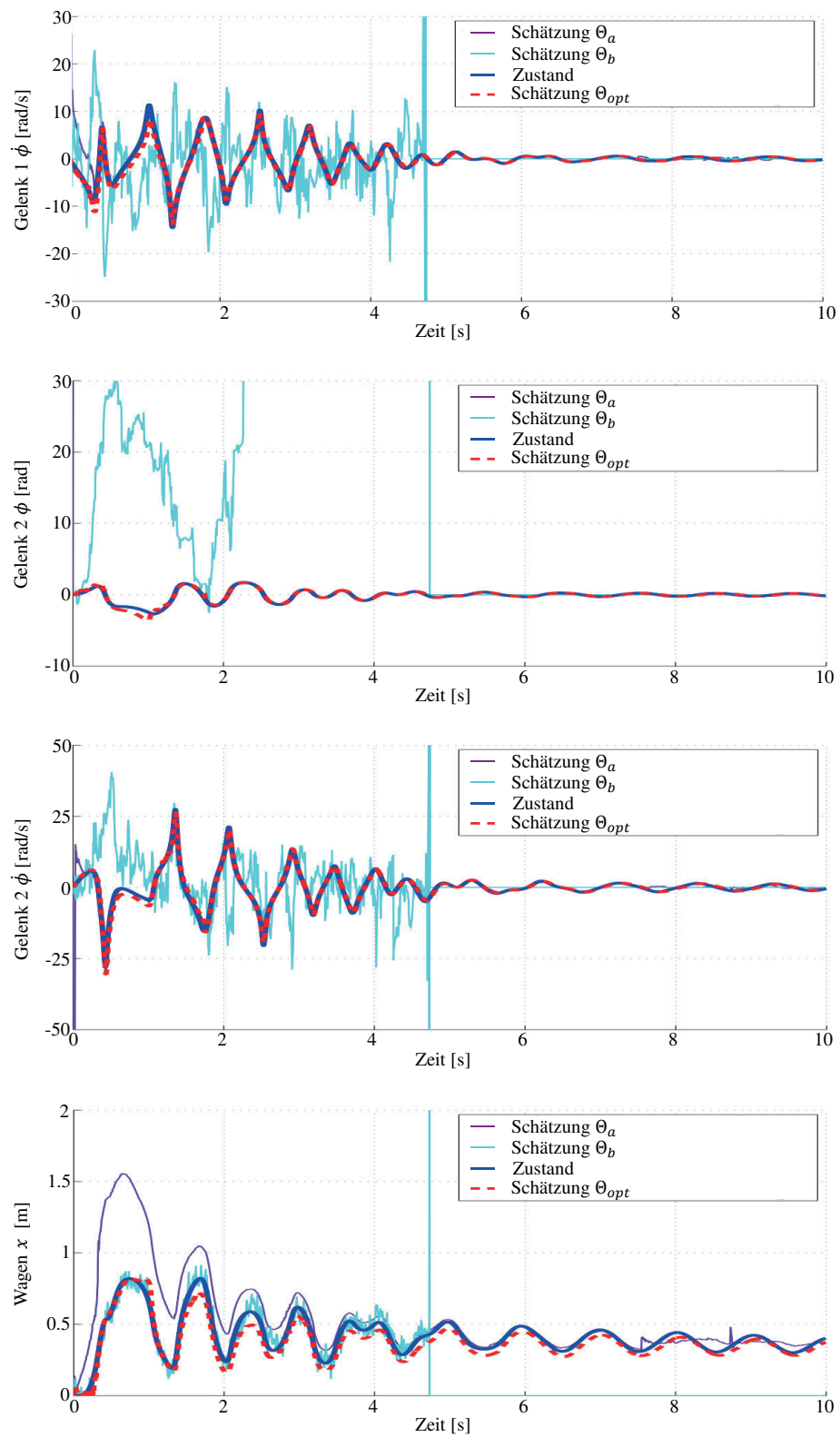


Bild 1-8: Schätzung der Zustände des Doppelpendels bei unterschiedlicher Parametrierung des eingesetzten Unscented-Kalman-Filter mit SUT-Sigma-Punkte-Berechnung

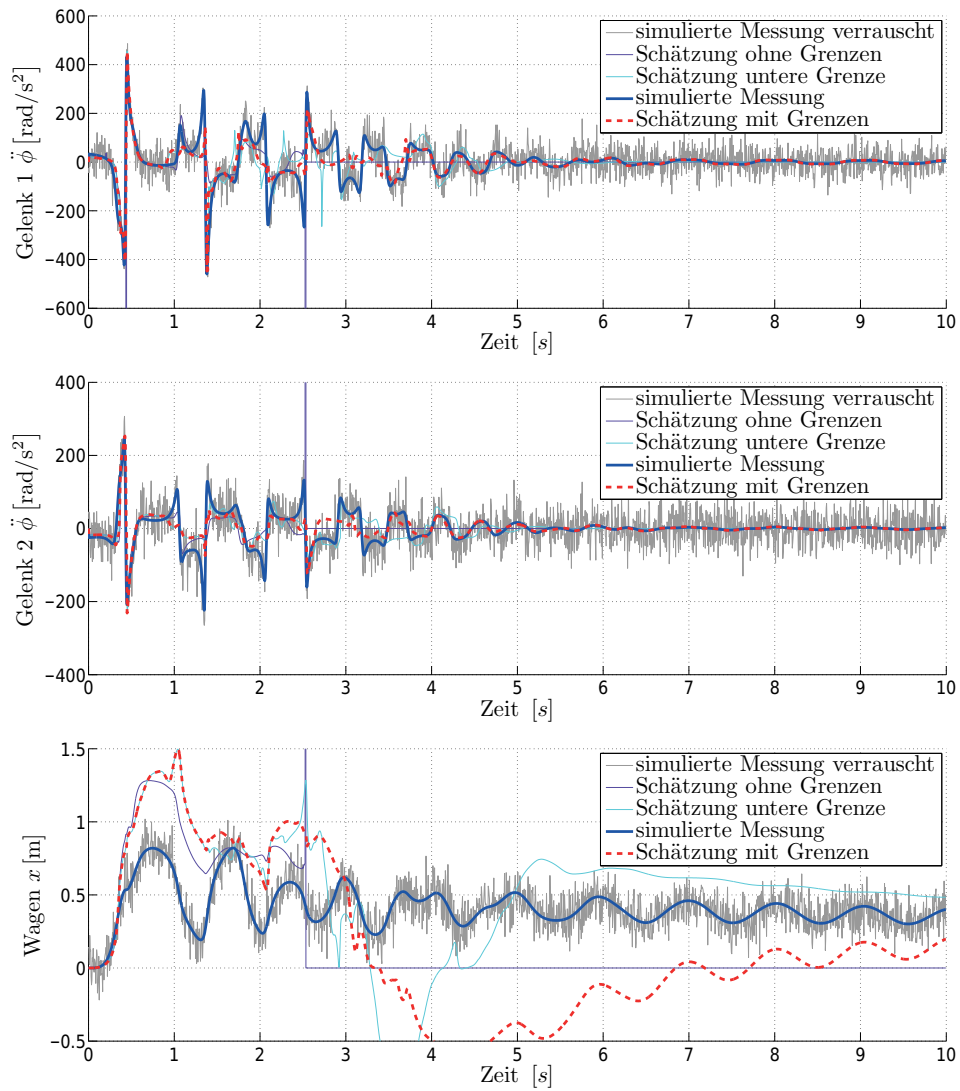


Bild 1-9: Systemausgänge des Streckenmodells (blau); dem Filter zur Verfügung gestellte Systemausgänge, mit Rauschen überlagert (grau), Filterrekonstruktion des Systemausganges durch UKF ohne Parameternebenbedingungen (lila), mit unterer Begrenzung (cyan) und vollen Nebenbedingungen (rot).

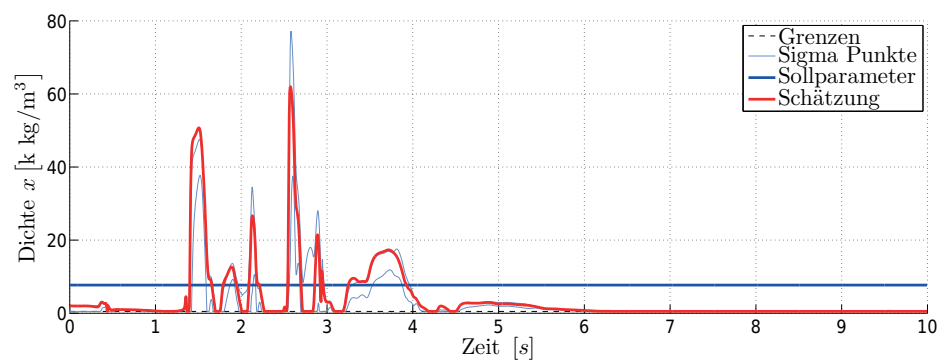


Bild 1-10: Schätzung des unbekannten Systemparameters (roter Verlauf) der Pendelarmdichte (dunkelblauer Verlauf) bei unterer Nebenbedingung mit Einbeziehung der berechneten Sigma-Punkte (hellblauer Verlauf)

A.5 Schätzung zeitvarianter Unwucht im Waschautomaten

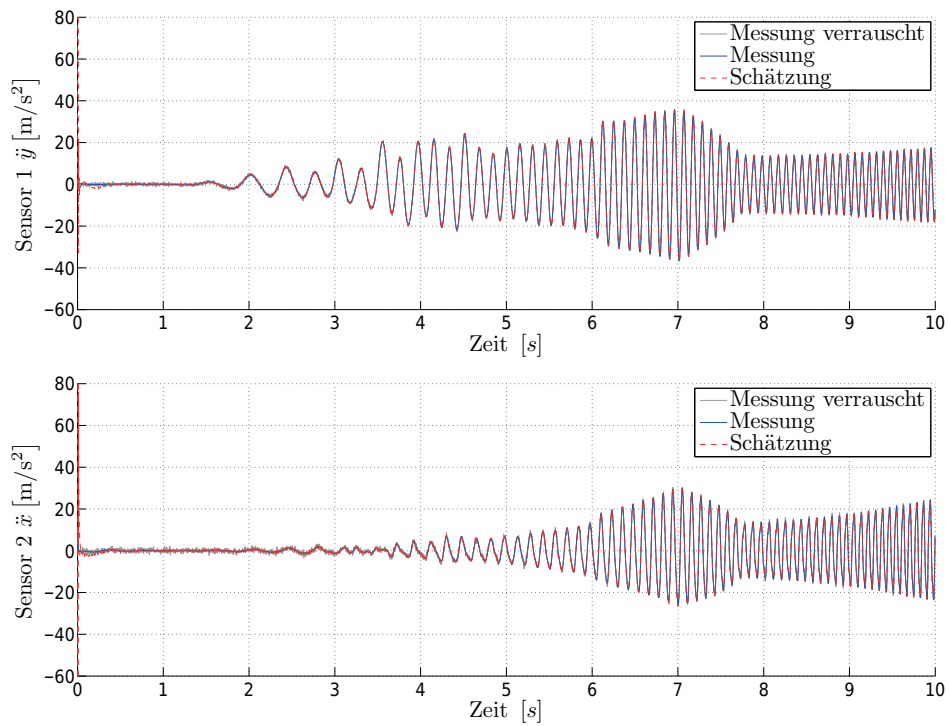


Bild 1-11: Rekonstruktion verrauschter Messungen eines Waschautomaten bei der Schätzung einer unbekannten, zeitvarianten Unwucht

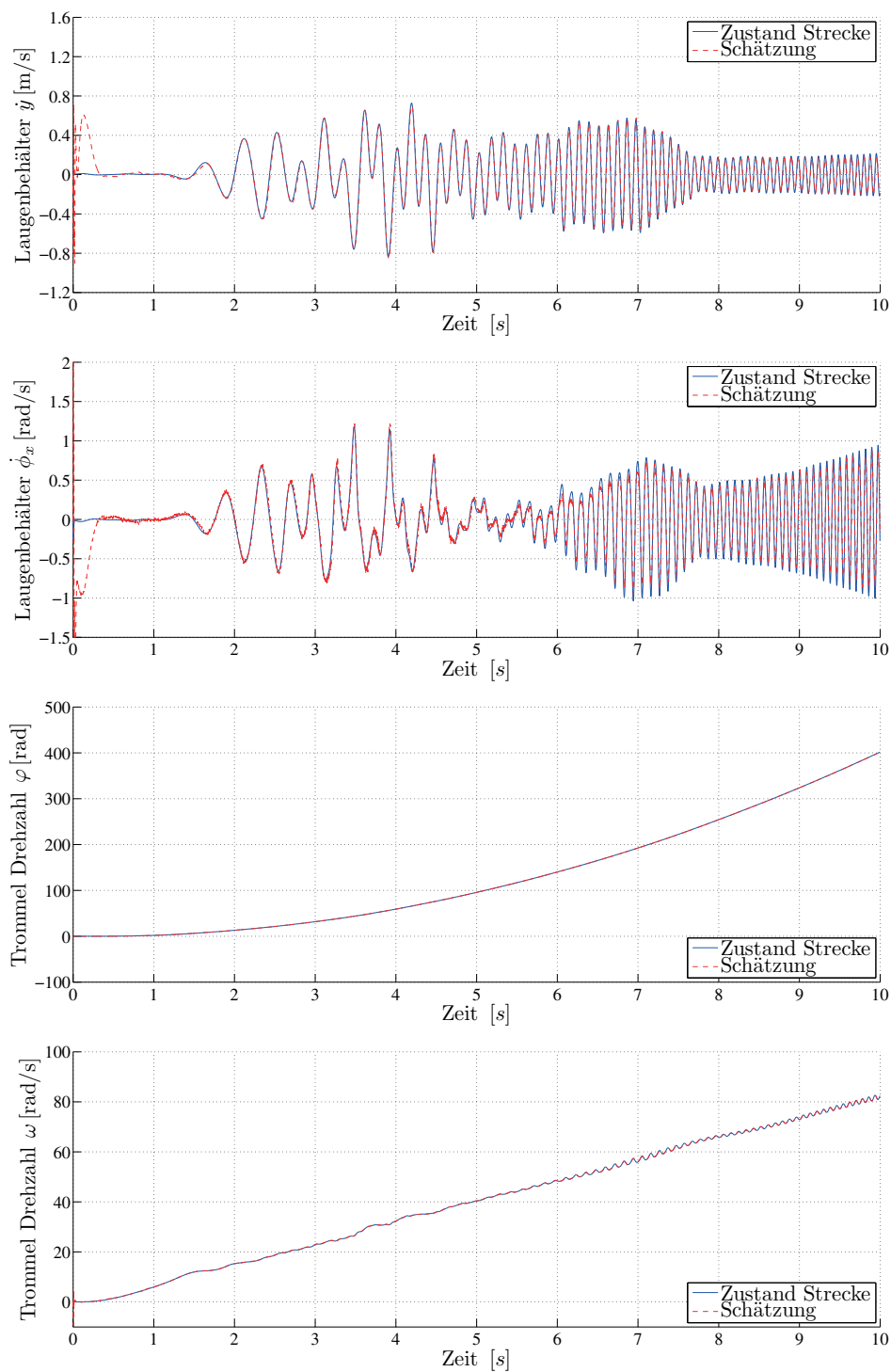


Bild 1-12: Zustandsschätzung eines Waschautomaten bei der Schätzung einer unbekannten, zeitvarianten Unwucht

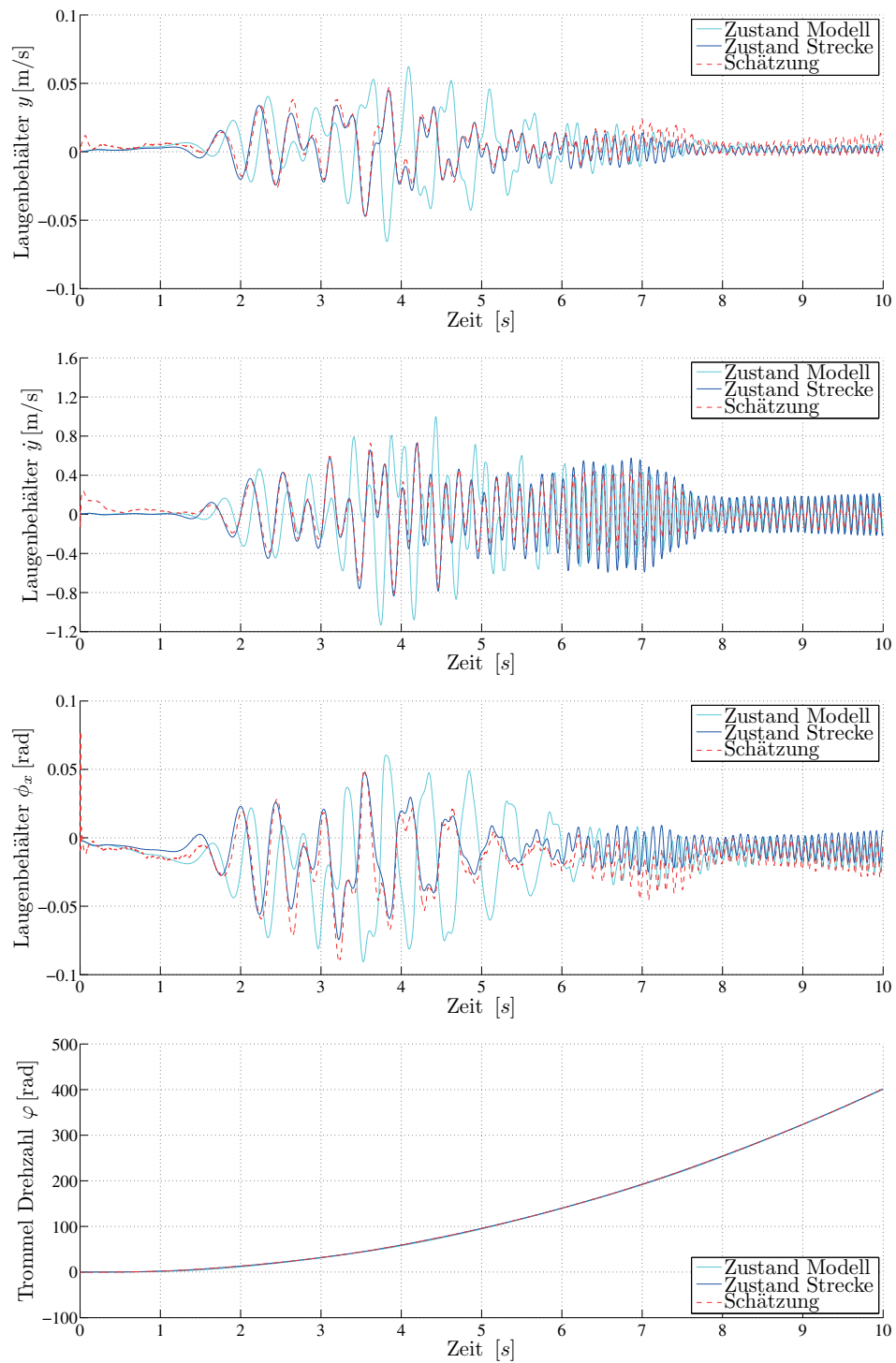


Bild 1-13: Rekonstruktion verrauschter Messungen eines Waschautomaten bei der Schätzung einer unbekannten, zeitvarianten Unwucht mit starken Abweichungen zwischen dem Filtermodell und den dem Filter zur Verfügung stehenden Messungen

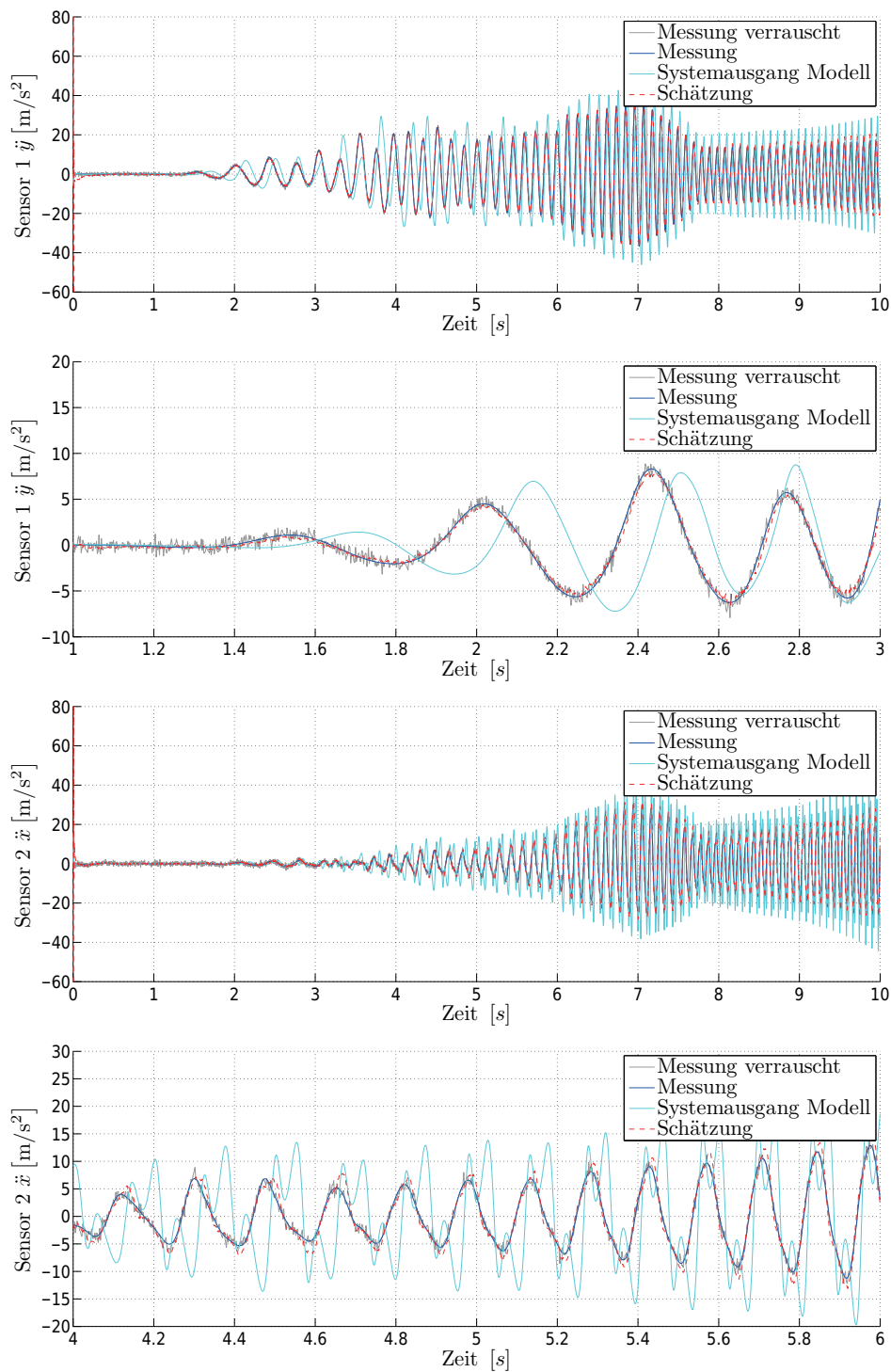


Bild 1-14: Zustandsschätzung eines Waschautomaten bei der Schätzung einer unbekannten, zeitvarianten Unwucht mit starken Abweichungen zwischen dem Filtermodell und den dem Filter zur Verfügung stehenden Messungen

A.6 Zustandsschätzung Dingo 2

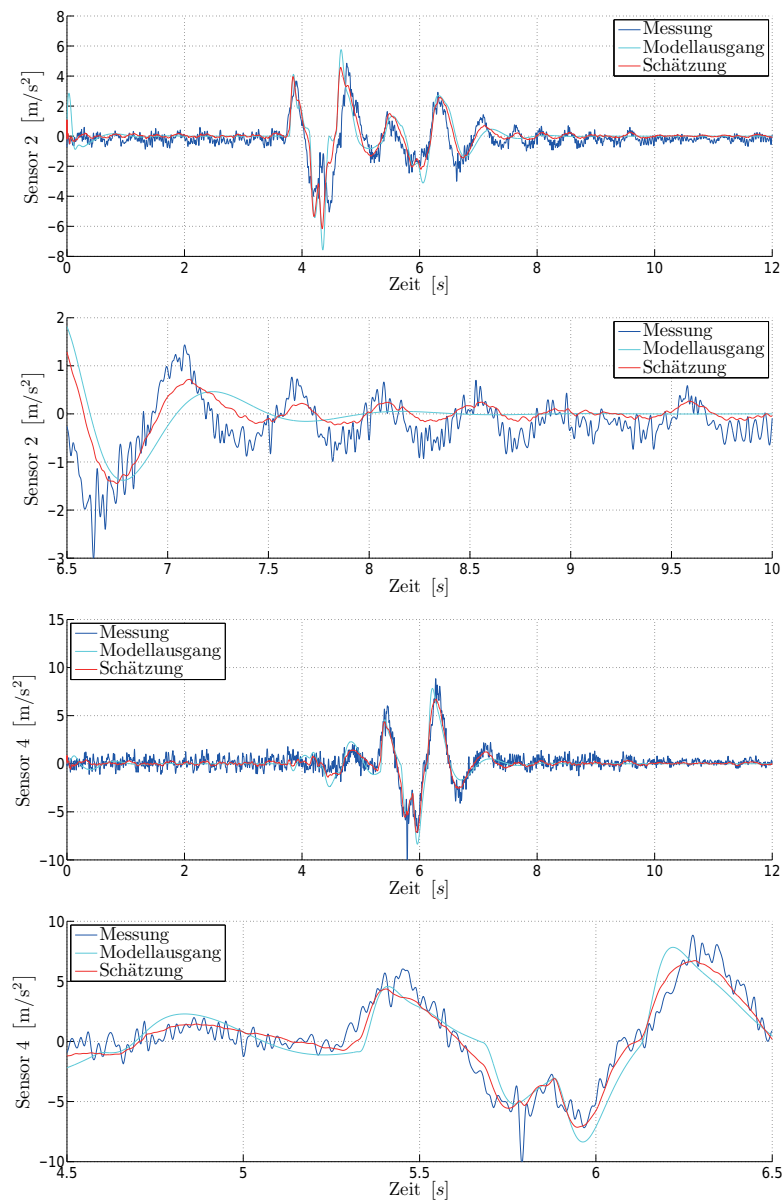


Bild 1-15: Rekonstruktion der realen Messungen eines Dingo 2 bei Überfahren von Trapezhindernissen

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, Mobilität, Vernetzung: Eine neue Schule des Entwurfs der 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 neun Professoren mit insgesamt 150 Mitarbeiterinnen und Mitarbeitern. Pro Jahr promovieren hier etwa 20 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, Mobility, Integration: Enroute to the 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 nine Professors and 150 researchers work at the Heinz Nixdorf Institute. Per year approximately 20 young researchers receive a doctorate.

Zuletzt erschienene Bände der Verlagsschriftenreihe des Heinz Nixdorf Instituts

- Bd. 339 MAO, Y.: 245 GHz Subharmonic Receivers For Gas Spectroscopy in SiGe BiCMOS Technology. Dissertation, Fakultät für Elektrotechnik, Informatik und Mathematik, Universität Paderborn, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 339, Paderborn, 2015 – ISBN 978-3-942647-58-8
- Bd. 340 DOROCIĄK, R.: Systematik zur frühzeitigen Absicherung der Sicherheit und Zuverlässigkeit fortschrittlicher mechatronischer Systeme. Dissertation, Fakultät für Maschinenbau, Universität Paderborn, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 340, Paderborn, 2015 – ISBN 978-3-942647-59-5
- Bd. 341 BAUER, F.: Planungswerkzeug zur wissensbasierten Produktionssystemkonzipierung. Dissertation, Fakultät für Maschinenbau, Universität Paderborn, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 341, Paderborn, 2015 – ISBN 978-3-942647-60-1
- Bd. 342 GAUSEMEIER, J.; GRAFE, M.; MEYER AUF DER HEIDE, F. (Hrsg.): 12. Paderborner Workshop Augmented & Virtual Reality in der Produktentstehung. Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 342, Paderborn, 2015 – ISBN 978-3-942647-61-8
- Bd. 343 GAUSEMEIER, J.; DUMITRESCU, R.; RAMMIG, F.; SCHÄFER, W.; TRÄCHTLER, A. (Hrsg.): 10. Paderborner Workshop Entwurf mechatronischer Systeme. Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 343, Paderborn, 2015 – ISBN 978-3-942647-62-5
- Bd. 344 BRÖKELMANN, J.: Systematik der virtuellen Inbetriebnahme von automatisierten Produktionssystemen. Dissertation, Fakultät für Maschinenbau, Universität Paderborn, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 344, Paderborn, 2015 – ISBN 978-3-942647-63-2
- Bd. 345 SHAREEF, Z.: Path Planning and Trajectory Optimization of Delta Parallel Robot. Dissertation, Fakultät für Maschinenbau, Universität Paderborn, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 345, Paderborn, 2015 – ISBN 978-3-942647-64-9
- Bd. 346 VASSHOLZ, M.: Systematik zur wirtschaftlichkeitsorientierten Konzipierung Intelligenter Technischer Systeme. Dissertation, Fakultät für Maschinenbau, Universität Paderborn, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 346, Paderborn, 2015 – ISBN 978-3-942647-65-6
- Bd. 347 GAUSEMEIER, J. (Hrsg.): Vorausschau und Technologieplanung. 11. Symposium für Vorausschau und Technologieplanung, Heinz Nixdorf Institut, 29. und 30. Oktober 2015, Berlin-Brandenburgische Akademie der Wissenschaften, Berlin, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 347, Paderborn, 2015 – ISBN 978-3-942647-66-3
- Bd. 348 HEINZEMANN, C.: Verification and Simulation of Self-Adaptive Mechatronic Systems. Dissertation, Fakultät für Elektrotechnik, Informatik und Mathematik, Universität Paderborn, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 348, Paderborn, 2015 – ISBN 978-3-942647-67-0
- Bd. 349 MARKWART, P.: Analytische Herleitung der Reihenfolgeregeln zur Entzerrung hochauslastender Auftragsmerkmale. Dissertation, Fakultät für Wirtschaftswissenschaften, Universität Paderborn, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 349, Paderborn, 2015 – ISBN 978-3-942647-68-7
- Bd. 350 RÜBBELKE, R.: Systematik zur innovationsorientierten Kompetenzplanung. Dissertation, Fakultät für Maschinenbau, Universität Paderborn, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 350, Paderborn, 2016 – ISBN 978-3-942647-69-4
- Bd. 351 BRENNER, C.: Szenariobasierte Synthese verteilter mechatronischer Systeme. Dissertation, Fakultät für Elektrotechnik, Informatik und Mathematik, Universität Paderborn, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 351, Paderborn, 2016 – ISBN 978-3-942647-70-0
- Bd. 352 WALL, M.: Systematik zur technologieinduzierten Produkt- und Technologieplanung. Dissertation, Fakultät für Maschinenbau, Universität Paderborn, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 352, Paderborn, 2016 – ISBN 978-3-942647-71-7

Zuletzt erschienene Bände der Verlagsschriftenreihe des Heinz Nixdorf Instituts

- Bd. 353 CORD-LANDWEHR, A.: Selfish Network Creation - On Variants of Network Creation Games. Dissertation, Fakultät für Elektrotechnik, Informatik und Mathematik, Universität Paderborn, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 353, Paderborn, 2016 – ISBN 978-3-942647-72-4
- Bd. 354 ANACKER, H.: Instrumentarium für einen lösungsmusterbasierten Entwurf fortschrittlicher mechatronischer Systeme. Dissertation, Fakultät für Maschinenbau, Universität Paderborn, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 354, Paderborn, 2016 – ISBN 978-3-942647-73-1
- Bd. 355 RUDTSCH, V.: Methodik zur Bewertung von Produktionssystemen in der frühen Entwicklungsphase. Dissertation, Fakultät für Maschinenbau, Universität Paderborn, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 355, Paderborn, 2016 – ISBN 978-3-942647-74-8
- Bd. 356 SÖLLNER, C.: Methode zur Planung eines zukunftsfähigen Produktportfolios. Dissertation, Fakultät für Maschinenbau, Universität Paderborn, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 356, Paderborn, 2016 – ISBN 978-3-942647-75-5
- Bd. 357 AMSHOFF, B.: Systematik zur musterbasierten Entwicklung technologie-induzierter Geschäftsmodelle. Dissertation, Fakultät für Maschinenbau, Universität Paderborn, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 357, Paderborn, 2016 – ISBN 978-3-942647-76-2
- Bd. 358 LÖFFLER, A.: Entwicklung einer modellbasierten In-the-Loop-Testumgebung für Waschautomaten. Dissertation, Fakultät für Maschinenbau, Universität Paderborn, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 358, Paderborn, 2016 – ISBN 978-3-942647-77-9
- Bd. 359 LEHNER, A.: Systematik zur lösungsmusterbasierten Entwicklung von Frugal Innovations. Dissertation, Fakultät für Maschinenbau, Universität Paderborn, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 359, Paderborn, 2016 – ISBN 978-3-942647-78-6
- Bd. 360 GAUSEMEIER, J. (Hrsg.): Vorausschau und Technologieplanung. 12. Symposium für Vorausschau und Technologieplanung, Heinz Nixdorf Institut, 8. und 9. Dezember 2016, Berlin-Brandenburgische Akademie der Wissenschaften, Berlin, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 360, Paderborn, 2016 – ISBN 978-3-942647-79-3
- Bd. 361 PETER, S.: Systematik zur Antizipation von Stakeholder-Reaktionen. Dissertation, Fakultät für Maschinenbau, Universität Paderborn, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 361, Paderborn, 2016 – ISBN 978-3-942647-80-9
- Bd. 362 ECHTERHOFF, O.: Systematik zur Erarbeitung modellbasierter Entwicklungsaufträge. Dissertation, Fakultät für Maschinenbau, Universität Paderborn, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 362, Paderborn, 2016 – ISBN 978-3-942647-81-6
- Bd. 363 TSCHIRNER, C.: Rahmenwerk zur Integration des modellbasierten Systems Engineering in die Produktentstehung mechatronischer Systeme. Dissertation, Fakultät für Maschinenbau, Universität Paderborn, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 363, Paderborn, 2016 – ISBN 978-3-942647-82-3
- Bd. 364 KNOOP, S.: Flachheitsbasierte Positionsregelungen für Parallelkinematiken am Beispiel eines hochdynamischen hydraulischen Hexapoden. Dissertation, Fakultät für Maschinenbau, Universität Paderborn, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 364, Paderborn, 2016 – ISBN 978-3-942647-83-0
- Bd. 365 KLIEWE, D.: Entwurfssystematik für den präventiven Schutz Intelligenter Technischer Systeme vor Produktpiraterie. Dissertation, Fakultät für Maschinenbau, Universität Paderborn, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 365, Paderborn, 2017 – ISBN 978-3-942647-84-7
- Bd. 366 IWANEK, P.: Systematik zur Steigerung der Intelligenz mechatronischer Systeme im Maschinen- und Anlagenbau. Dissertation, Fakultät für Maschinenbau, Universität Paderborn, Verlagsschriftenreihe des Heinz Nixdorf Instituts, Band 366, Paderborn, 2017 – ISBN 978-3-942647-85-4