

Methodische Evaluierung von wahrscheinlichkeitstheoretischen Inferenzverfahren für dynamische Systeme

Dissertation

zur

Erlangung des Grades

Doktor-Ingenieur

der

Fakultät für Maschinenbau
der Ruhr-Universität Bochum

von

Marc-André Beyer

aus Essen

Bochum 2008

Dissertation eingereicht am:	26. November 2007
Tag der mündlichen Prüfung:	06. März 2008
Erster Referent:	Prof. Dr.-Ing. G. Reinig
Zweiter Referent:	Prof. Dr.-Ing. W. Eifler

Kurzfassung zur Dissertation

Methodische Evaluierung von wahrscheinlichkeitstheoretischen Inferenzverfahren für dynamische Systeme

Die wahrscheinlichkeitstheoretische Inferenz (WI) beschreibt das Problem, Informationen über die internen Variablen eines Systems, i.d.R. Zustände und/oder Parameter, aus unvollständigen und verrauschten Messdaten zu gewinnen. Anwendungsbeispiele von WI sind u.a. Filter für nichtlineare Regelungskonzepte, modellbasierte Fehlerdiagnose, Trägheitsnavigation u.v.m.. Wählt man einen statistischen, rekursiven Ansatz zur Lösung des WI-Problems, so ist diese durch die Rekursive BAYES'sche Schätztheorie gegeben, welche für die meisten praxisrelevanten Anwendungen nicht exakt lösbar ist. Das Spektrum an Näherungslösungen (z.B. Extended Kalman Filter, Unscented Kalman Filter, Partikel-Filter) ist in den letzten Jahren sehr umfangreich und unübersichtlich geworden. Die hohe Anzahl an Algorithmen und der Mangel an einer praxisorientierten Bewertungsmethodologie erschweren die Auswahl eines geeigneten Filters und inspirierten dieses Projekt.

Der wesentliche Grund für eine bisher vergleichsweise geringe Anwendungsbreite moderner Filterverfahren ist der Mangel an *aussagekräftigen, praxisrelevanten* und *glaubwürdigen* Informationen über die Möglichkeiten und Eigenheiten der Verfahren sowie einer benutzerfreundlichen Testumgebung zur Generierung quantifizierbarer Bewertungskriterien. In vorliegender Dissertation wird ein Zugang vorgestellt, der dazu beitragen soll, den Entscheidungsprozess zu objektivieren. Der vorgeschlagene Ansatz sieht vor, praxisrelevante Merkmale gehobener Filterverfahren zu erheben, wobei sowohl quantitative als auch qualitative Merkmale bewertet werden. Basierend auf einer Literaturrecherche, intensiven Befragungen von Automatisierungstechnikern und den während dieser Arbeit gesammelten Erfahrungen, wurde ein Kriterienkatalog zur Bewertung rekursiver Filter erarbeitet. Dazu war es notwendig das breite Spektrum an anspruchsvollen WI-Verfahren zu studieren, die Verfahren zu vereinheitlichen und für komplexe Zustands- und Parameterschätzaufgaben aufzubereiten. Die Evaluierungskriterien berücksichtigen die maßgeblichen Gesichtspunkte, die während des Selektionsprozesses zu beachten sind, u.a. *Filtergüte, Filterrobustheit, Filterrechenzeit, Anforderungen an den Anwender*. Die anwendungsfreundliche Umsetzung der entwickelten Methodik erfolgte innerhalb einer neu entwickelten MATLAB[®] basierten Toolbox ReBEL-IoN.

Das vorgestellte Konzept wurde an mehreren analytischen Experimenten demonstriert und die Praxiseignung am Beispiel der modellbasierten Diagnose eines Dieselpartikelfilters nachgewiesen. Letzteres Applikationsbeispiel wurde in engem Industriekontakt mit der Daimler AG bearbeitet. Hierbei kommen so genannte "graue Modelle" zum Einsatz, welche in weiten Teilen physikalisch fundiert sind, jedoch vereinfachende Annahmen beinhalten, wobei unsichere Modellparameter an Messdaten angepasst wurden. Die Motivation für die Diagnose mit grauen Modellen ist, ein im Fahrzeug On-Board rechenbares Modell mit physikalisch interpretierbaren System- und Fehlerparametern zu gewinnen, welche bei On-Line Schätzung eine direkte, detaillierte Aussage über Fehlerort und -ausmaß zulassen. Durch die zu Grunde liegenden physikalischen

Beziehungen sind solche Modelle in einem weiten Arbeitsbereich gültig und durch die Adaption weniger Parameter an baugleiche Systeme anpassbar.

Die während der Arbeit gewonnenen Ergebnisse und Erfahrungen zeigen, dass der entwickelte Zugang eine effiziente, praxisnahe und kostengünstige Methode darstellt, den Anwender beim Selektionsprozess zu unterstützen. Dabei ist besonders hervorzuheben, dass die innerhalb der Testumgebung gewonnenen Evaluierungsergebnisse sich direkt in der realen Anwendung widerspiegeln. Durch vorliegende Arbeit wurde ein Beitrag zur verbesserten Akzeptanz und zur verbreiterten industriellen Anwendung moderner Filterverfahren geleistet.

Zusätzlich zur Evaluierungsmethodik wurden für die Sequentielle-Monte-Carlo-Methoden (Partikel-Filter) Algorithmen entwickelt, um Informationen über harte Zustandsbeschränkungen in den Schätzprozess zu integrieren. Die Simulationsstudien zeigen, dass die gewonnenen Schätzverfahren bei beschränkten Schätzproblemen über bessere Schätzgüte und höhere Robustheit verfügen. Darüber hinaus wurde eine neue rekursive Bestimmungsgleichung für die Filter-Informationsmatrix zur Berechnung des Posterior Cramér-Rao Bounds für zeitinvariante Parameter hergeleitet.

Im Anhang dieser Dissertation ist eine Sammlung an Pseudo-Codes dreißig ausgewählter WI-Verfahren enthalten. Zweck dieser ist es, als Nachschlagewerk für WI-Verfahren mit einheitlicher Notation zu dienen. Dieser Abschnitt ist in englischer Sprache verfasst, um ihn auch internationalen Anwendern zur Verfügung stellen zu können.

Abstract of the thesis

Methodical evaluation of probabilistic inference algorithms for dynamic systems

Probabilistic inference (PI) is the problem of estimating the internal variables of a system (states and/or parameters) given only incomplete and or noisy measurements. Application examples of PI are filters for nonlinear control concepts, model-based fault detection, inertial navigation, etc.. By choosing a statistical recursive approach to solve the PI-problem, its solution is given by the Recursive Bayesian Estimation Theory, whose exact solution is intractable for most "real world" systems. The large number of approximate but tractable solutions, e.g. Extended Kalman Filter, Unscented Kalman Filter or particle filters, and the lack of a practical selection methodology makes the choice of an appropriate filter for a given system a challenging task and inspired this project.

The main reason for the small amount of (industrial) applications of modern filter algorithms is the lack of *significant, practice-oriented* and *authentic* information about the possibilities and the properties of the algorithms as well as a user friendly test environment for generation of quantifiable benchmark criteria. In present dissertation an approach is proposed that serves to objectify the selection progress. The proposed method allots the evaluation of quantitative and qualitative benchmark criteria which have been gathered by literature research, intensive interrogation of control engineers and the gained experience during this research project. This necessitated to study the broad spectrum of complex probabilistic inference algorithms, to unify the algorithms and to apply them to complex state and parameter estimation tasks. The benchmark criteria cover the most important issues that need to be studied during the selection progress, among others *filter performance, filter robustness, computational expensiveness or qualification requirements on the operator*. The user friendly implementation of the developed methodology is carried out with a new developed MATLAB[®] based Toolbox ReBEL-IoN.

The usage and the benefit of the concept is shown by means of several illustrative examples and the practical usefulness is demonstrated by means of the model-based fault detection of a diesel particulate filter. The latter application example has been processed in cooperation with the Daimler AG. For fault detection, grey-box models are used which are mostly based on physical relations and properties but also contain simplifying assumptions, where uncertain model parameters are estimated from measured data. The motivation for fault detection with grey-box models is to gain an on-board applicable model with physically interpretable system- and fault parameters which allow a determined statement about fault position, type and location if they are estimated on-line. Due to the underlying physical relations, such models are valid in a broad operating range and allow adjustment to systems identical in construction by adaption of few parameters.

The obtained results and gained experience during this research project allows to state that the developed approach is an efficient, practice-oriented, and inexpensive method to support the operator during the

selection process. The author would like to point out, that the simulatively obtained evaluation results are directly connected with the "real world" application performance of the algorithms. Present dissertation contributes to a better understanding and acceptance as well as to a broadened industrial application of modern probabilistic inference algorithms.

Aside from the comparative evaluation method, algorithms to deal with hard state constraints in particle filtering are proposed. Simulation studies show, that the developed algorithms consistently outperform their unconstrained counterparts in estimation performance and robustness if the state space is constrained. Furthermore a new recursive equation for the filter information matrix for Posterior Cramér-Rao Bound computation for time-invariant parameter estimation problems is proposed.

The appendix of this dissertation contains a collection of thirty pseudo-codes of selected inference algorithms. Its purpose is to serve as a reference book for probabilistic inference algorithms with a consistent notation.

FÜR FREUNDSCHAFT, UNTERSTÜTZUNG UND LIEBE. FÜR DICH, SARAH.

Danksagung

Die vorliegende Dissertation entstand während meiner Beschäftigung als wissenschaftlicher Mitarbeiter am Lehrstuhl für Regelungssysteme und Steuerungstechnik an der Ruhr-Universität Bochum. Das Forschungsprojekt wurde von der Daimler AG, Abteilung Advanced Engineering - Vehicle Diagnostics, gefördert.

Mein ganz besonderer Dank gilt Herrn Prof. Dr.-Ing. G. Reinig, der mir stets bei der Ausgestaltung der Arbeit ein hohes Maß an Eigenverantwortung einräumte und durch kritische Diskussionen den Fortschritt der Arbeit förderte.

Ebenso danke ich Herrn Prof. Dr.-Ing. W. Eifler für das Interesse an der Arbeit und die Übernahme des Koreferates.

Besonders geschätzt habe ich das stets freundschaftliche und konstruktive Arbeitsklima am Lehrstuhl. Besonders intensiv war die Zusammenarbeit mit Herrn Dipl.-Ing. Wolfgang Grote und Herrn Dipl.-Ing. Rainer Baese.

Der Daimler AG danke ich für die finanzielle und fachliche Unterstützung, wobei besonderer Dank den Herren Dipl.-Ing. Stephan Steinhauer und Dipl.-Ing. Werner Preuschoff für das mir entgegengebrachte Vertrauen sowie den Herren Dipl.-Ing. Bernd Rehfus und Dipl.-Ing. Joachim Konrad für die fachliche Unterstützung und konstruktive Zusammenarbeit zu Teil kommen soll.

Weiterhin danke ich allen Studenten, die im Rahmen von Studien- und Diplomarbeiten sowie HiWi-Tätigkeiten den Fortschritt der Arbeit gefördert haben.

Meinen Eltern und Freunden danke ich dafür, dass sie mich in diesem fordernden Lebensabschnitt stets unterstützt und so gesund gehalten haben.

Inhaltsverzeichnis

1	Einleitung, Motivation, Zielsetzung	1
1.1	Einleitung	1
1.2	Motivation	3
1.3	Zielsetzung	5
1.4	Gliederung	6
1.5	Beiträge für Zeitschriften und Kongresse	7
2	Das WI-Problem und seine Lösungen	8
2.1	Rekursive Bayes'sche Schätztheorie (RBS)	11
2.2	Näherungslösungen der RBS	12
2.2.1	Methoden mit Gauß'scher Approximation	13
2.2.2	Methoden für Zufallsvariablen beliebiger Verteilung	23
2.3	Kapitelzusammenfassung	40
3	Verschiedene Schätzprobleme, Variationen der Algorithmen-Implementierung und spezielle Systeme	41
3.1	Verschiedene Schätzprobleme	41
3.2	Variationen der Algorithmen-Implementierung	43
3.3	Systeme mit finiten Zustandsräumen	47
3.4	Systeme mit Totzeiten	55
3.5	Kapitelzusammenfassung	56
4	Evaluierungsmethodik und Matlab[®] basierte Toolbox ReBEL-IoN	57
4.1	Modellbildung Referenz-/Filtermodell	57
4.1.1	Zeitliche Diskretisierung kontinuierlicher Systeme	57
4.1.2	Analyse von System- und Messrauschen und der Initialverteilung	62
4.2	Beobachtbarkeit nichtlinearer, diskreter Systeme	65
4.3	Testen der Inferenzalgorithmen	66
4.3.1	Experimentelles Testen	66
4.3.2	Simulatives Testen	67
4.4	Wahl geeigneter Eingangsdaten (Wahl des Experiments)	70

4.5	Vorauswahl und qualitative Benchmarkkriterien	71
4.5.1	Qualifikationsanforderungen an den Anwender während der Einsatzvorbereitung . .	72
4.5.2	Aufwand und Unterstützung beim Entwurf des Filters	72
4.5.3	Installations- und Implementierungsaufwand	73
4.6	Numerische Bewertung über quantitative Benchmarkkriterien	74
4.6.1	Quantitativer Benchmark: Filtergüte über Posterior Cramér-Rao Bound	74
4.6.2	Quantitativer Benchmark: Filterrobustheit (Beschränkungssicherheit)	80
4.6.3	Quantitativer Benchmark: Rechenzeit	81
4.7	MATLAB [®] basierte Toolbox ReBEL-IoN	81
4.8	Kapitelzusammenfassung	83
5	Anwendungsbeispiele	87
5.1	Zwei anschauliche akademische Inferenzaufgaben	87
5.2	Reversible Gasgleichgewichtsreaktion (Demonstration beschränkte SMC-Methoden)	93
5.3	Terrain-Aided-Navigation	97
5.4	Nichtlineare Trajektorienregelung des Klatt-Engell-Reaktors durch E/A-Linearisierung . . .	102
5.5	Studie zur modellbasierten Diagnose eines Dieselpartikelfilters (DPF)	108
5.5.1	Versuchsträger	110
5.5.2	"Blitzdiagnose" über DPF-Kennlinienschätzung mit dem linearen Kalman Filter . .	111
5.5.3	Modellbildung, Fehlermodellierung und Nominalparameteridentifikation	113
5.5.4	Diagnose mittels WI-Verfahren	120
5.5.5	Umgang mit Part-to-Part-Streuung mittels WI-Verfahren	126
5.5.6	Bewertung der Methodik zur modellbasierten Diagnose mittels WI-Verfahren . . .	129
5.6	Kapitelzusammenfassung	131
6	Zusammenfassung und Ausblick	133
	Literaturverzeichnis	137
	Studentische Arbeiten	144
	A Pseudo-Codes of selected inference algorithms	145
	Verkürzter Lebenslauf	201
	Schriftenreihe des Lehrstuhls	203

Abbildungsverzeichnis

1.1	Anwendungsbeispiele wahrscheinlichkeitstheoretischer Inferenz	2
1.2	Umgang mit Part-to-Part-Streuung in der MBD	4
2.1	Wahrscheinlichkeitstheoretische Inferenz	8
2.2	Übersicht: Näherungslösungen der RBS	12
2.3	Sigma-Point Kalman Filter: Lokalisierung der Sigma-Vektoren einer 2D-Zufallsvariable	17
2.4	Beispiel: Multimodale Verteilung der Zustandsdichte $p(x_k)$	23
2.5	Gitternetzfilter: Integration über ein äquidistantes Gitternetz	24
2.6	SMC-Methoden: Integration über ein adaptives Gitternetz	26
2.7	SMC-Methoden: Verschiebung der Partikel der priori Dichte in Regionen hohen Likelihoods	30
2.8	SMC-Methoden: Der Generic Particle Filter	32
2.9	SMC-Methoden: Kernel Density Estimation mit GAUSS-Kern	35
2.10	SMC-Methoden: Der Gaussian Mixture Sigma-Point Particle Filter	39
3.1	1D-Gamma-Verteilung mit Zustandsschranken	53
3.2	2D-multimodale Verteilung mit Zustandsschranken	54
4.1	Stabilitätsbereich des expliziten EULER- und HEUN-Verfahrens	62
4.2	Konzept: Experimentelles Testen	67
4.3	Konzept: Simulatives Testen	68
4.4	Einfluss der POD auf den RMSE einer Schätzung	69
4.5	Geschwindigkeitsprofile NEDC und FTP	71
4.6	RMSE vs. PCRB für ein lineares und nichtlineares Radar-Verfolgungs-Problem	76
4.7	RMSE vs. Parameter-PCRB Kennlinienschätzung	78
4.8	Filterwirkungsgrad für ein nichtlineares Radar-Verfolgungs-Problem	79
4.9	Graphische Benutzeroberfläche ReBEL-IoN	86
5.1	RMSE vs. PCRB und Filterwirkungsgrad für das nicht normalverteilte Filterproblem	88
5.2	Skizze des nichtlinearen Radar-Verfolgungs-Problems	89
5.3	Beispieltrajektorien eines fallenden Flugobjektes	90
5.4	RMSE vs. PCRB und Filterwirkungsgrad für das Radar-Verfolgungs-Problem	92
5.5	RMSE des GPF und ASIR für das beschränkte Problem	95

5.6	RMSE des RPF und GMSPPF für das beschränkte Problem	96
5.7	Konzept Terrain-Aided-Navigation	97
5.8	Blockschaltbild Terrain-Aided-Navigation	98
5.9	Simulationsstudie Terrain-Aided-Navigation	100
5.10	RMSE vs. PCRB und Filterwirkungsgrad für das Terrain-Aided-Navigation-Problem	101
5.11	Skizze KLATT-ENGELL-Reaktor-Benchmark	102
5.12	Schema Exakte E/A-Linearisierung	103
5.13	RMSE vs. PCRB für das KLATT-ENGELL-Reaktor-Benchmark-Beispiel	107
5.14	Folgeregelung des KLATT-ENGELL-Reaktor-Benchmark-Beispiels	107
5.15	MBD DPF: Funktionsweise eines Honeycomb-DPF	109
5.16	MBD DPF: Datenlogger im Versuchsfahrzeug	110
5.17	MBD DPF: Rekursive DPF-Kennlinien-Schätzung mit dem linearen Kalman Filter	111
5.18	MBD DPF: 3D-Kennfeld zur Diagnose	112
5.19	MBD DPF: Längsschnitt durch einen Honeycomb-DPF	113
5.20	MBD DPF: Längsschnitt durch einen Honeycomb-DPF bei Verschlussausbruch	118
5.21	MBD DPF: Modellverifikation und Gütegebirge Honeycomb-DPF	120
5.22	MBD DPF: RMSE vs. PCRB und Filterwirkungsgrad Einzelschätzung Rußmasse	121
5.23	MBD DPF: RMSE vs. PCRB und Filterwirkungsgrad Einzeldiagnose Verschlussausbrüche	123
5.24	MBD DPF: Verifikation Einzeldiagnose Verschlussausbrüche an UDC-Fahrzyklusdaten	124
5.25	MBD DPF: RMSE vs. PCRB und Filterwirkungsgrad Schätzung Ruß+Verschlussausbrüche	126
5.26	MBD DPF: RMSE vs. PCRB und Filterwirkungsgrad On-Line Parameteridentifikation	128
5.27	MBD DPF: Verifikation Parameteridentifikation an UDC-Fahrzyklusdaten	128

Tabellenverzeichnis

3.1	Vergleich RS und PS am 1D-Beispiel	53
3.2	Vergleich RS und PS am 2D-Beispiel	54
4.1	Einstufung der Qualifikationsanforderungen an den Anwender	72
4.2	Anzahl der einzustellenden Parameter beim Filterentwurf	73
4.3	Einstufung der Unterstützung beim Filterentwurf	73
5.1	Studienparameter für das nicht normalverteilte Filterproblem	88
5.2	Studienergebnisse für das nicht normalverteilte Filterproblem	88
5.3	Studienparameter für das Radar-Verfolgungs-Problem	91
5.4	Studienergebnisse für das Radar-Verfolgungs-Problem	91
5.5	Studienparameter für das beschränkte Problem	94
5.6	Studienergebnisse für das beschränkte Problem	95
5.7	Studienparameter für das Terrain-Aided-Navigation-Problem	100
5.8	Studienergebnisse für das Terrain-Aided-Navigation-Problem	101
5.9	Studienparameter für das KLATT-ENGELL-Reaktor-Benchmark-Beispiel	106
5.10	MBD DPF: Technische Daten des Versuchsträgers	110
5.11	MBD DPF: Diagnoseergebnisse KF+KNN	112
5.12	MBD DPF: Konfidenzintervalle identifizierter Parameter	119
5.13	MBD DPF: Studienparameter Schätzung Rußmasse	121
5.14	MBD DPF: Studienparameter Diagnose Verschlussausrüche	122
5.15	MBD DPF: Studienergebnisse Diagnose Verschlussausrüche	122
5.16	MBD DPF: Studienparameter Schätzung Ruß+Verschlussausrüche	125
5.17	MBD DPF: Studienparameter Part-to-Part-Streuung	127
5.18	MBD DPF: Studienergebnisse Part-to-Part-Streuung	127
5.19	MBD DPF: Vergleich CDKF - Levenberg-Marquardt - Gauß-Newton	129

Symbolverzeichnis

Lateinische Buchstaben

$BIAS(\star)$	Verzerrung eines Schätzers von \star
\mathcal{BM}	Beobachtbarkeitsmatrix
d	Binäre Variable des Detektions-/Miss-Ereignisses
d_{Eq}	Vektor der Gleichheitsbeschränkung
d_{Ineq}	Vektor der Ungleichheitsbeschränkung
\mathbf{D}_{Eq}	Matrix der Gleichheitsbeschränkung
\mathbf{D}_{Ineq}	Matrix der Ungleichheitsbeschränkung
e	Einheitsvektor
f	$\mathbb{R}^n \times \mathbb{R}^p \rightarrow \mathbb{R}^n$ dimensionale deterministische Systemgleichung
\mathcal{F}	$\mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^v \rightarrow \mathbb{R}^n$ dimensionale stochastische Systemgleichung
G_r	Übertragungsfunktion Regler
G_s	Übertragungsfunktion Strecke
G_w	Übertragungsfunktion geschlossener Regelkreis
G, M, J	Anzahl Komponenten GMM-Modelle
h	$\mathbb{R}^n \times \mathbb{R}^s \rightarrow \mathbb{R}^m$ dimensionale deterministische Messgleichung Schrittweite der Finiten-Differenzen-Interpolation
\mathcal{H}	$\mathbb{R}^n \times \mathbb{R}^s \times \mathbb{R}^z \rightarrow \mathbb{R}^m$ dimensionale stochastische Messgleichung
\mathbf{I}	Einheitsmatrix
J	Filter-Informationsmatrix
$\mathcal{J}_\star^\bullet$	Differentialmatrix (Jacobi-Matrix) von \star , differenziert nach \bullet
K	Anzahl Observationen
K_p	Verstärkung Regler
K_λ	Kernel mit Bandbreite λ
\mathcal{K}	$n \times m$ -dimensionale Kalman-Verstärkung
\mathcal{L}	LIE-Ableitung
m	Dimension Messvektor
n	Dimension Zustandsvektor
N	Partikelanzahl
N_{eff}	Effektive Partikelanzahl

N_{thr}	Grenzwert der effektiven Partikelanzahl für Resampling
p	Dimension exogener Eingang
p_D	Detektionswahrscheinlichkeit
\mathbf{P}^*	Kovarianzmatrix von $*$
\mathbf{P}_k^*	Kovarianzmatrix von $*$ zum (Zeit)punkt k
$\mathbf{P}_{k k}^*$	Gefilterte Kovarianzmatrix von $*$ zum (Zeit)punkt k
$\mathbf{P}_{k k-1}^*$	Prädizierte Kovarianzmatrix von $*$ zum (Zeit)punkt k
r	Messkanal
\mathcal{R}	METROPOLIS-HASTINGS-Verhältnis
\mathcal{RG}	Relativer Grad
s	LAPLACE-Operator
\mathbf{S}_k^*	Cholesky-Faktorisierung der Kovarianzmatrix von $*$ zum (Zeit)punkt k
\mathcal{S}_k^*	Detektions-/Miss-Sequenz $*$ bis zum Zeitpunkt k
u	$\in \mathbb{R}^p$ Vektor exogener Eingänge in die System- und Messgleichung
U_k	$\in \mathbb{R}^p \times k$ Sequenz exogener Eingangsvektoren bis zum (Zeit)punkt k
v	Dimension Vektor Systemrauschen
w	Gewicht
\tilde{w}	Normalisiertes Gewicht
x	$\in \mathbb{R}^n$ Zustandsvektor
X_k	$\in \mathbb{R}^n \times k$ Sequenz sämtlicher Zustände bis zum (Zeit)punkt k
\mathfrak{X}	Wertebereich des Zustandes x
y	$\in \mathbb{R}^m$ Messvektor
Y_k	$\in \mathbb{R}^n \times k$ Sequenz sämtlicher Messungen bis zum (Zeit)punkt k
z	Dimension Vektor Messrauschen

Griechische Buchstaben

α	Skalierungsparameter SUT Komponentengewicht GMM-Modelle Akzeptanzwahrscheinlichkeit MCMC Faktor Varianz Messrauschen
β	Skalierungsparameter SUT Komponentengewicht GMM-Modelle
ε	Maximale Sensorfehler
γ	$\in \mathbb{R}^z$ Vektor des Messrauschens
Γ	$\mathbb{R}^n \times \mathbb{R}^p \rightarrow \mathbb{R}^{m \times z}$ Verteilungsmatrix der Messgleichung
η	Wirkungsgrad

θ	Parameter
κ	$\in \mathbb{R}^v$ Vektor des Systemrauschens (S)UT Skalierungsparameter
$\tilde{\kappa}$	$\in \mathbb{R}^v$ Vektor des künstlichen Systemrauschens
K	$\mathbb{R}^n \times \mathbb{R}^p \rightarrow \mathbb{R}^{n \times v}$ Verteilungsmatrix der Systemgleichung
λ	Eigenwert Kernel-Bandbreite
ξ	Punkte zur Auswertung der FDI
Ξ	Zustandsbeschränkungen
\mathcal{X}	Matrix mit Sigma-Vektoren als Einträge in den Spalten
ρ_B	BHATTACHARYYA-Koeffizient
σ	Einheitssprung Standardabweichung
τ	Integrationschrittweite/Schrittweite der Diskretisierung
φ	Komponentengewicht GMM-Modelle
Φ	Transitionsmatrix
$\Psi_*(\bullet)$	Schätzer von \star aus Messung \bullet

Wahrscheinlichkeitsfunktionen

$B(\star; A, B)_{GW_1, GW_2}^{GW_1}$	Betaverteilung von \star mit Parametern A und B innerhalb der Grenzen GW_1 und GW_2
$Bi(\star; A, B)$	Binomialverteilung von \star mit Parametern A und B
$\mathcal{N}(\star; \mu, \Sigma)$	Normalverteilung von \star mit Mittelwert μ und Kovarianzmatrix Σ
$\mathcal{G}a(\star; A, B)$	Gammaverteilung von \star mit Parametern A und B
$p(\star)$	Wahrscheinlichkeitsfunktion von \star
$p(\star \bullet)$	Wahrscheinlichkeitsfunktion von \star bedingt durch \bullet
$p(\star, \bullet)$	Gemeinsame Wahrscheinlichkeitsfunktion von \star und \bullet
$\mathcal{P}(\star)$	Verteilungsfunktion von \star
$Pr(\star)$	Wahrscheinlichkeit von/für \star
π	Ansatzverteilung
$\mathcal{U}(\star; GW_1, GW_2)$	Gleichverteilung von \star innerhalb der Grenzen GW_1 und GW_2

Mathematische Operatoren

$\delta(\star)$	Dirac-Delta-Funktion von \star
$\text{diag}(\star)$	Diagonalmatrix mit \star als Einträge, falls \star Vektor Vektor der diagonalen Elemente von \star , falls \star Matrix

$ \star $	Determinante von \star
$\ \star\ $	Euklidische Norm von \star
$\mathbb{E}[\star]$	Erwartungswert von \star
\star^T	Transponierte von \star
∇_\star	Nabla-Operator für \star
Δ_\star^\bullet	Delta-Operator für \star und \bullet
$(\star)^2$	Äußeres Vektorprodukt, falls \star ein Spaltenvektor ist Inneres Vektorprodukt, falls \star ein Zeilenvektor ist
$\inf(\star)$	Infimum (größte untere Schranke) von \star
$\lfloor\star\rfloor$	Ganzzahl-Funktion/Abrundungsfunktion von \star

Indizes

<i>dis</i>	diskret
<i>k</i>	(Zeit)index
<i>kon</i>	kontinuierlich

Hochgestellte Größen

$\hat{\star}$	Schätzung von \star
$\bar{\star}$	Mittelwert von \star
$\dot{\star}$	Zeitliche Ableitung von \star
\star^{-1}	Umkehrfunktion von \star Inverse von \star
(\star)	Zähler \star

DPF-Modell

a	Kanalbreite	[m]
A	Kontraktionsfläche	[m ²]
A_R	Durchströmte Rußfläche	[m ²]
A_{Filter}	Durchströmte Filterfläche	[m ²]
D	Effektiver Filterdurchmesser	[m]
F	Konstante	[-]
k_R	Permeabilität Ruß	[m ²]
k_w	Permeabilität Filterwand	[m ²]
L	Effektive Filterlänge	[m]
$m_{Ru\beta}$	Rußmasse	[kg]
$\tilde{m}_{Ru\beta}$	Relative Rußmasse	$[\frac{kg}{m^3}]$

\dot{m}_{Defekt}	Abgasmassenstrom durch Ausbruchkanäle	$[\frac{\text{kg}}{\text{s}}]$
\dot{m}_{Gesamt}	Gesamter Abgasmassenstrom	$[\frac{\text{kg}}{\text{s}}]$
\dot{m}_{Intakt}	Abgasmassenstrom durch intakte Kanäle	$[\frac{\text{kg}}{\text{s}}]$
N_{Cells}	Anzahl Zellen	$[-]$
N_{Defekt}	Anzahl defekter Zellen	$[-]$
N_{Intakt}	Anzahl intakter Zellen	$[-]$
p	Druck	$[\text{Pa}]$
R	Spezifische Gaskonstante	$[\frac{\text{J}}{\text{kgK}}]$
T	Temperatur	$[\text{K}]$
U	Strömungsgeschwindigkeit	$[\frac{\text{m}}{\text{s}}]$
V	Effektives Filtervolumen	$[\text{m}^3]$
\dot{V}	Volumenstrom	$[\frac{\text{m}^3}{\text{s}}]$
w_R	Rußschichtdicke	$[\text{m}]$
w_w	Wandstärke Filter	$[\text{m}]$
X_{Defekt}	Druckverlustterm defekter Zellen	$[\text{Pa}]$
X_{Intakt}	Druckverlustterm intakter Zellen	$[\text{Pa}]$
β	FORCHHEIMER-Koeffizient	$[\frac{1}{\text{m}}]$
Δp	Druckverlust/Druckdifferenz	$[\text{Pa}]$
$\zeta_{Kon,Exp}$	Kontraktions-/Expansionskoeffizient	$[-]$
η	Dynamische Viskosität	$[\text{Pa s}]$
ξ	Korrekturfaktor Saugeffekte	$[-]$
ρ	Dichte	$[\frac{\text{kg}}{\text{m}^3}]$
σ	Filterdichte	$[\frac{1}{\text{m}^2}]$

Abkürzungen

AEKF	Adaptive Extended Kalman Filter	Seite 16
AKPF	Auxiliary Kernel Particle Filter	Seite 36
APF	Adaptive Particle Filter	Seite 36
ASIR	Auxiliary Sequential Importance Resampling	Seite 33
BK	BHATTACHARYYA-Koeffizient	Seite 49
CDKF	Central Difference Kalman Filter	Seite 19
DPF	Dieselpartikelfilter	Seite 108
DSSM	Dynamic State Space Model	Seite 10
DTED	Digital Terrain Elevation Data	Seite 97
EKF	Extended Kalman Filter	Seite 15
EKPF	Extended Kalman Particle Filter	Seite 37

EUDC	Extra Urban Driving Cycle	Seite 70
FB-2-DOF	Flatness Based Two Degree of Freedom Control	Seite 70
FDI	Finite-Differenzen-Interpolation	Seite 19
GBF	Grid Based Filter	Seite 23
GF	Gauß-Filter	Seite 13
GHF	Gauß Hermite Filter	Seite 21
GMM	Gaussian Mixture Model	Seite 24
GMSPPF	Gaussian Mixture Sigma-Point Particle Filter	Seite 38
GPF	Generic Particle Filter	Seite 32
GSA	Gaussian Sum Approximation	Seite 24
GSKF	Gaussian Sum Kalman Filter	Seite 25
GSPF	Gaussian Sum Particle Filter	Seite 37
ICDKF	Iterated Central Difference Kalman Filter	Seite 23
IEKF ₁	Iterated Extended Kalman Filter (measurement equation)	Seite 15
IEKF ₂	Iterated Extended Kalman Filter (state and measurement equation)	Seite 16
IS	Importance Sampling	Seite 26
IUKF	Iterated Unscented Kalman Filter	Seite 22
KDE	Kernel Density Estimation	Seite 34
KI	Konfidenzintervall	Seite 119
KLD	Kullback-Leibler-Distanz	Seite 36
KNN	Künstlich Neuronales Netz	Seite 41
KPF	Kernel Particle Filter	Seite 36
MAP	Maximum A Posteriori	Seite 10
MBD	Modellbasierte Diagnose	Seite 3
MC	Monte Carlo	Seite 33
MCMC	Markov Chain Monte Carlo	Seite 34
MHE	Moving Horizon Estimation	Seite 55
MISE	Mean Integrated Squared Error	Seite 34
MMSE	Minimum Mean Square Estimation	Seite 10
MVS	Minimum Variance Sampling	Seite 32
NEDC	New European Driving Cycle	Seite 70
NMPC	Nonlinear Model Predictive Control	Seite 70
OBD	On-Board Diagnose	Seite 3
P2P	Part-to-Part	Seite 4
PCRB	Posterior Cramér-Rao Bound	Seite 74
PF	Partikel-Filter	Seite 25

PS	Projection Sampling	Seite 50
ReBEL-IoN	Recursive Bayesian Estimation Library - Investigation of Necessity	Seite 81
RI	Relative Improvement	Seite 79
RMSE	Root Mean Squared Error	Seite 74
RPF	Regularized Particle Filter	Seite 34
RRS	Regularization Resampling Step	Seite 34
RR	Residual Resampling	Seite 31
RS	Rejection Sampling	Seite 50
RTAMSE	Root Time Averaged Mean Squared Error	Seite 79
SIR	Sequential Importance Resampling	Seite 31
SIS	Sequential Importance Sampling	Seite 28
SMC	Sequential Monte Carlo Methods	Seite 25
SPKF	Sigma-Point Kalman Filter	Seite 22
SPPF	Sigma-Point Particle Filter	Seite 37
SRAEKF	Square Root Adaptive Extended Kalman Filter	Seite 45
SRCDKF	Square Root Central Difference Kalman Filter	Seite 45
SREKF	Square Root Extended Kalman Filter	Seite 45
SRIEKF ₁	Square Root Iterated Extended Kalman Filter	Seite 45
SRUKF	Square Root Unscented Kalman Filter	Seite 45
SUT	Scaled Unscented Transformation	Seite 18
TAN	Terrain-Aided-Navigation	Seite 97
TNS	Trägheitsnavigationssystem	Seite 97
UDC	Urban Driving Cycle	Seite 70
UDDS	Urban Dynamometer Driving Schedule	Seite 70
UKF	Unscented Kalman Filter	Seite 16
UT	Unscented Transformation	Seite 17
WI	Wahrscheinlichkeitstheoretische Inferenz	Seite 8

1 Einleitung, Motivation, Zielsetzung

1.1 Einleitung

"At the heart of the natural sciences and engineering there are the general notions of **observation** and **measurement**. ... it is legitimate to say that in the fields of natural science and engineering, experiments and observations (measurements) are fundamental."

PIETER EYKHOFF - "System Identifikation" - [1]

Die Informationsgewinnung aus Messungen und Beobachtungen ist ein zentraler Bestandteil der Natur- und Ingenieurwissenschaften. Basierend auf Messungen und Beobachtungen werden Theorie(n) über ein untersuchtes Problem aufgestellt, verifiziert oder falsifiziert, was die Korrektur der ursprünglichen Theorie(n) nach sich zieht. Dies erfolgt so lange, bis ein ausreichender Kompromiss zwischen Theorie(n) und Beobachtungen erzielt werden kann.

Die Formulierung einer Theorie wird in den Natur- und Ingenieurwissenschaften oftmals als Modellbildung bezeichnet, wobei ein Modell nach [1] definiert ist als "Abbild der grundlegenden Aspekte eines Systems, welches Kenntnis über das System in einer handhabbaren Form repräsentiert". Dabei ist ein System definiert als "Verbund von Objekten, angeordnet in einer strukturierten Form, welcher in irgendeiner Weise, zweck- oder zielgebunden ist und charakterisiert wird durch die Beziehung zwischen Eingangs- und Ausgangssignal". Ein System besitzt Zustände und Parameter. Der Zustand eines Systems ist eine Variable, welche zusammen mit den folgenden Systemeingangssignalen das Systemverhalten komplett bestimmt. Parameter eines Systems bestimmen zusätzlich das statische/dynamische und/oder (nicht)lineare Verhalten eines Systems. Daraus folgt, dass bei Kenntnis des Systemmodells, der Systemparameter und -zustände sowie der Systemeingänge das Systemverhalten präzisiert werden kann.

Schwerpunkt dieser Arbeit ist das Problem der "Schätzung" der Zustände und Parameter von dynamischen Systemen aus Messungen. Der Begriff Schätzung wird verwendet, da i.d.R. in allen realistischen Situationen die Observationen des untersuchten Systems durch Zufallsprozesse, welche sich sowohl auf das eigentliche untersuchte System (Systemrauschen) als auch auf den Messprozess (Messrauschen) auswirken

können, "kontaminiert" werden.

Die Parameterschätzung ist nach [1] definiert als die "experimentelle Bestimmung von Parametern, welche das dynamische, statische und/oder (nicht)lineare Verhalten bestimmen unter der Annahme, dass die Struktur des Systemmodells bekannt ist". Fälschlicherweise wird die Parameterschätzung oft auch als Systemidentifikation bezeichnet. Parameterschätzung erfordert aber a priori Kenntnis über das Systemmodell, das Problem der Systemidentifikation beschreibt den kompletten Modellbildungsprozess ohne jegliche a priori Kenntnis über Struktur und/oder Parameter des Modells. In vielen Anwendungen ist es notwendig, auch Kenntnis über die Systemzustände zu erlangen. Dies beschreibt das Problem der Zustandsschätzung. Betrachtet man dazu die Beispiele dargestellt in Abbildung 1.1: Wie kann aus der mit Messunsicherheiten behafteten Temperaturmessung eines chemischen Reaktors auf interne Reaktionsgrößen geschlossen werden? Oder wie kann die Position eines Flugzeuges alleinig aus der absoluten Höhenmessung, der Höhe über Grund und einer im Bordcomputer hinterlegten 3D-Karte der Topographie geschätzt werden? Des Weiteren, wie vertrauenswürdig sind diese Schätzungen?

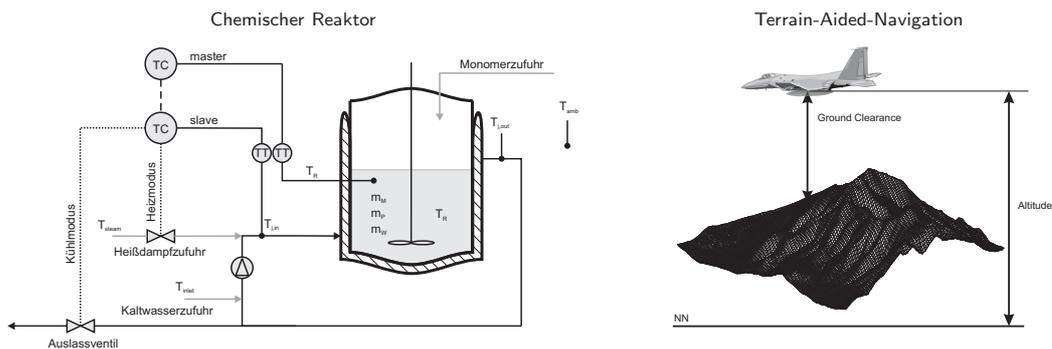


Abbildung 1.1: Beispiele zur wahrscheinlichkeitstheoretischen Inferenz.

Die Wahrscheinlichkeitstheorie bietet die Möglichkeit, solche Unsicherheiten in einer formalen Sprache abzubilden und dient als Rechenmethode dazu, mit Unsicherheiten umzugehen. Im Wesentlichen dient die Wahrscheinlichkeitstheorie dazu, aus Unsicherheiten schlussfolgern zu können. Dieser Prozess wird wahrscheinlichkeitstheoretische Inferenz (Rückschluss) (WI) genannt (Probabilistic Inference, PI).

Die zwei fundamentalen Ziele von wahrscheinlichkeitstheoretischer Inferenz sind also:

- Erlangung von Informationen über systeminterne Zustände und Parameter aus gemessenen Daten.
- Verringerung des Einflusses von System- und Messrauschen auf die Zustands- und Parameterschätzung.

In dem Fall, wo keine stochastischen Effekte auf System und Messung wirken, d.h. sowohl System als auch Messung deterministisch sind (bzw. Zufallsprozesse vernachlässigt werden können), werden Schätzverfahren als "Beobachter" gekennzeichnet. Treten Zufallsprozesse auf und werden a priori Informationen darüber im

Schätzalgorithmus verarbeitet, spricht man von "Filtern".

1.2 Motivation

In den letzten Jahren und Jahrzehnten ist das Spektrum an Verfahren zur Lösung komplexer (nichtlinearer) Inferenzaufgaben sehr umfangreich geworden, wobei als berühmteste Vertreter der Filterverfahren der Extended Kalman Filter, der Unscented Kalman Filter und die Partikel-Filter vorab genannt seien. Neben den wohlbekanntesten Verfahren sind eine Vielzahl von Variationen entwickelt worden, die mehr oder weniger starke Vereinfachungen bzw. Verkomplizierungen der Ausgangsverfahren darstellen. All diese Filteralgorithmen und ihre Variationen haben typische Vor- und Nachteile bezüglich ihrer Filtereigenschaften, wie z.B. Filtergüte, Implementierungs- und Rechenaufwand und Anforderungen an den Anwender.

Motivation allgemein

Aufgrund der unterschiedlichen Eigenschaften aller Filter stellt sich bei der Lösung einer konkreten Inferenzaufgabe immer die Frage, welcher Algorithmus Verwendung finden kann bzw. sollte. Grundsätzlich sollte immer eine möglichst hohe Filtergüte erzielt werden. Verlangt dies jedoch nach Verfahren, die verfügbare Ressourcen überschreiten, so muss unter den verbliebenen Algorithmen der bestmögliche gewählt werden. Ebenso gilt die Regel "Keep it simple": Liefern die aufwendigeren¹ Filteralgorithmen nur marginale Verbesserung, so sollte der einfachste¹ Algorithmus ausgewählt werden.

Der wesentliche Grund für eine bisher vergleichsweise geringe Anwendungsbreite moderner Filterverfahren ist der Mangel an *aussagekräftigen, praxisrelevanten* und *glaubwürdigen* Informationen über die Möglichkeiten und Eigenheiten der Verfahren sowie einer benutzerfreundlichen Testumgebung zur Generierung quantifizierbarer Bewertungskriterien.

Motivation im Hinblick auf die modellbasierte Diagnose (MBD)

Die Überwachung technischer Systeme gewinnt immer mehr an Bedeutung. Speziell im automobilen Sektor stellt der Gesetzgeber stetig neue Anforderungen speziell an die Diagnose emissionsrelevanter Komponenten. Zur Überwachung stehen viele Methoden zur Verfügung, siehe z.B. [2], wobei die modellbasierte Diagnose ein vielversprechendes Verfahren mit hoher Diagnosetiefe darstellt.

Basierend auf so genannten "grauen" Modellen, welche eine Kombination aus analytischen und empirischen Modellen darstellen und in einer möglichst optimalen Art und Weise das zu überwachende System nachbilden, können gute Diagnoseaussagen abgeleitet werden. Die physikalischen Streckenparameter geben detaillierten Aufschluss über die inneren Zustände eines Systems und bestimmen zudem dessen Dynamik.

¹Was im Zusammenhang mit Filteralgorithmen "aufwendig" bzw. "einfach" bedeutet, soll später im Text diskutiert werden.

Aus ihrer Kenntnis kann folglich eine sehr gute Diagnoseaussage abgeleitet werden. Nominale Systemkoeffizienten oder zusätzlich bestimmte Fehler- und/oder Verschleißparameter können somit zur Fehlerdiagnose verwendet werden [2]. Voraussetzung dafür ist, dass die physikalischen Streckenparameter und deren Änderungen in den Modell- und Verschleißparametern abgebildet werden.

Die Herausforderung bei der modellbasierten Diagnose ist somit die parallele/gleichzeitige On-Line Schätzung der Systemzustände und der System- bzw. Fehlerparameter.

Neben der grundlegenden dualen Inferenzaufgabe (parallele/gleichzeitige Schätzung von Zuständen und Parametern) bei der Diagnose spielt das Problem der Part-to-Part-Streuung (P2P-Streuung) eine zentrale Rolle in der MBD. Fertigungstoleranzen führen dazu, dass im Serienbetrieb eingebaute Komponenten in ihren nominalen Systemparametern schwanken. Da eine Off-Line Adaption dieser Nominalparameter nicht für jedes Serienbauteil erfolgen kann, muss die Adaption in den ersten Betriebsstunden On-Line durchgeführt werden. Abbildung 1.2 zeigt, wie durch die Verwendung von Inferenzalgorithmen mit dem Problem der Part-to-Part-Streuung (am Beispiel eines DPF) umgegangen werden kann.

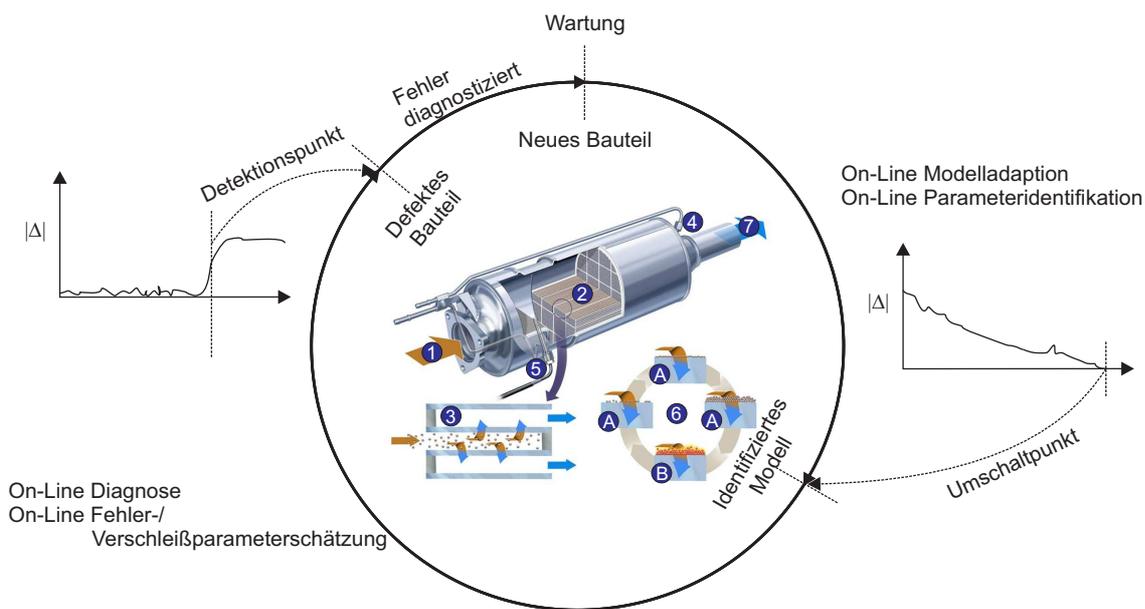


Abbildung 1.2: Umgang mit Part-to-Part-Streuung.

Als industrielles Anwendungsbeispiel wird in dieser Arbeit die modellbasierte Diagnose eines Dieselpartikelfilters betrachtet. Daher wird im Verlaufe der Dissertation immer wieder der Nutzen der vorgestellten Methodik für die MBD verdeutlicht.

Schlussfolgerung

Die bestehende Divergenz zwischen Forschung und praktischer Anwendung von nichtlinearen Filterverfahren im Allgemeinen sowie das Verlangen seitens der Industrie (Daimler AG) nach einem leicht anwendbaren Bewertungsinstrument im Speziellen motivieren zur Entwicklung einer Methode zum Testen und Evaluieren nichtlinearer Filter.

1.3 Zielsetzung

Systematisierung und Aufbereitung von wahrscheinlichkeitstheoretischen Inferenzverfahren für anspruchsvolle Zustands- und Parameterschätzaufgaben

Erstes Ziel der Arbeit ist es, einen Überblick über moderne wahrscheinlichkeitstheoretische Inferenzverfahren zu schaffen und deren Applikation als Zustands- und Parameterschätzer zu zeigen. Dabei soll besonders auf den mathematischen Ursprung und die Herleitung der Algorithmen eingegangen werden.

Entwicklung einer Methode zum Testen und zur Evaluierung von wahrscheinlichkeitstheoretischen Inferenzverfahren

Kernaufgabe dieser Arbeit ist es, eine Methode zu entwickeln, die es dem Anwender ermöglicht, nichtlineare Filterverfahren für komplexe (nichtlineare) Systeme kostengünstig, flexibel und ohne großen Zeitaufwand zu testen und zu bewerten. Diese Methode soll eine praxisrelevante Bewertung von Filterverfahren durch ein qualitatives und quantitatives Benchmarking ermöglichen.

Nachfolgend sind die wichtigsten Anforderungen an die Evaluierungsmethode beschrieben:

- Die Basis für die Evaluierungsmethode soll ein breites Spektrum an Filteralgorithmen bilden, eine **Filterbibliothek**, wobei die zu bewertenden Algorithmen nach Wünschen des Anwenders selektiert werden können.
- Die Bewertung der selektierten Filteralgorithmen soll anhand eines umfangreichen **Bewertungskriterien**-Katalogs erfolgen. Dabei sollen sowohl quantitative Kriterien, die eine numerische Aussage über die Eignung eines Filters ermöglichen, als auch qualitative Kriterien, die generelle Vor- und Nachteile der Filter beleuchten, Verwendung finden.
- Eine spezielle **Testumgebung** ist notwendig, um das Spektrum an Bewertungskriterien anwenden und auswerten zu können. Prinzipiell eignen sich labortechnische Versuchsanlagen und Simulationsumgebungen. Damit eine Reproduzierbarkeit der Versuchsergebnisse erreicht werden kann und identische Rahmenbedingungen für die Versuche garantiert werden können, soll die Testumgebung als Simulationsumgebung entwickelt werden. Dabei soll speziell auf die einfache Handhabung und

hohe Flexibilität der Simulationsumgebung geachtet werden, um die Evaluierungsmethode auch externen Anwendern anbieten zu können. Die Filteralgorithmen sollen aus der Simulationsumgebung exportierbar sein, um ein Post Processing anhand realer Messdaten zu ermöglichen und somit eine Verifikation der simulativ gewonnenen Ergebnisse zu gewährleisten.

Entwicklung einer modellbasierten Diagnosemethode unter Verwendung von wahrscheinlichkeitstheoretischen Inferenzverfahren

Am Beispiel eines Dieselpartikelfilters soll gezeigt werden, wie die nichtlinearen Filteralgorithmen zur modellbasierten Diagnose technischer Systeme eingesetzt werden können. Dabei sind besonders vom Gesetzgeber geforderte Randbedingungen zu berücksichtigen. Des Weiteren soll eine Methodik entwickelt werden, wie mit Nominalparameterschwankungen der zu überwachenden Komponenten, resultierend aus Serienfertigungstoleranzen, umgegangen werden kann. Ziel ist es, die Serientauglichkeit und die Akzeptanz modellbasierter Diagnosestrategien zu steigern.

Der entwickelte Diagnosealgorithmus soll anhand realer Fahrzeugdaten verifiziert und somit die Praxistauglichkeit gezeigt werden.

1.4 Gliederung

Die vorliegende Arbeit besteht, abgesehen von Einleitung, Zusammenfassung und Anhang, aus vier Kapiteln, welche wie nachfolgend beschrieben aufeinander aufbauen:

Kapitel 2 beschreibt die grundlegende Problematik der wahrscheinlichkeitstheoretischen Inferenz und wie dieses Problem gelöst werden kann. Es wird ein Einblick in die fundamentalen Grundlagen der in dieser Arbeit betrachteten Inferenzverfahren gegeben. Dieses Kapitel kann somit als Stand der Technik angesehen werden.

Kapitel 3 zeigt Arten von Schätzproblemen, sowie Variationen der in Kapitel 2 vorgestellten Algorithmen und vervollständigt den Stand der Technik. Es wird auf spezielle Systeme mit Gleichheits- und Ungleichheitsbeschränkungen sowie Systeme mit Totzeiten eingegangen. Darüber hinaus werden für die Sequentielle Monte-Carlo-Methoden Algorithmen zur Integration von harten Zustandsschranken in den Schätzprozess hergeleitet.

Kapitel 4 beinhaltet die Vorstellung der Evaluierungsmethodik sowie der Testumgebung. Es werden die quantitativen und qualitativen Bewertungskriterien aufgestellt und eine "Schritt für Schritt" Anleitung zur Auswahl eines Schätzverfahrens gegeben.

Kapitel 5 zeigt den Nutzen und die Nutzung der hier vorgestellten Methodik anhand mehrerer illustrativer Beispiele, u.a. wird die modellbasierte Diagnose eines Dieselpartikelfilters mit Hilfe von Inferenzverfahren demonstriert.

1.5 Beiträge für Zeitschriften und Kongresse

Die Veröffentlichung von Teilen dieser Arbeit ist in nachfolgenden Beiträgen zu wissenschaftlichen Fachzeitschriften und Kongressen bisher erfolgt:

- M.-A. Beyer and G. Reinig. *Evaluierung von wahrscheinlichkeitstheoretischen Inferenzverfahren für dynamische Systeme*. 42. Regelungstechnisches Kolloquium, Boppard, Germany, February 2008.
- M.-A. Beyer, W. Grote and G. Reinig. *Adaptive Exact Linearization Control of batch polymerization reactors using a Sigma-Point Kalman Filter*. Journal of Process Control, 2008, in press.

2 Das WI-Problem und seine Lösungen

Das Problem der wahrscheinlichkeitstheoretischen Inferenz (WI-Problem) beschreibt die Aufgabe, die internen Variablen eines Systems (Zustände und/oder Parameter) durch eine möglichst optimale Technik aus unvollständigen und/oder verrauschten Beobachtungen zu schätzen.

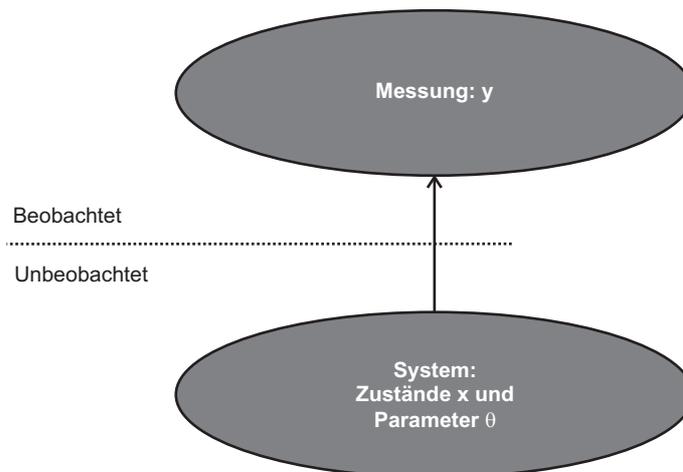


Abbildung 2.1: Wahrscheinlichkeitstheoretische Inferenz.

Welche Rückschlüsse über x und θ können aus der Kenntnis von y gezogen werden?

In der Filtertheorie wird das WI-Problem für dynamische Systeme in kontinuierlicher, kontinuierlich-diskreter und diskreter Form behandelt. Die kontinuierlichen, stochastischen Systeme sind durch die gewöhnliche nichtlineare stochastische LANGEVIN'sche Differentialgleichung mit zeitkontinuierlicher Messgleichung definiert:

$$\dot{x}(t) = f_{kon}(x, u; t) + K(x, u; t)\kappa(t) \quad (2.1)$$

$$y(t) = h_{kon}(x, u; t) + \Gamma(x, u; t)\gamma(t). \quad (2.2)$$

Dabei ist f_{kon} die $\mathbb{R}^n \times \mathbb{R}^p \rightarrow \mathbb{R}^n$ dimensionale kontinuierliche deterministische Systemgleichung, K die $\mathbb{R}^n \times \mathbb{R}^p \rightarrow \mathbb{R}^{n \times v}$ Verteilungsmatrix der Systemgleichung, κ der $\in \mathbb{R}^v$ Vektor des Systemrauschens, h_{kon} die $\mathbb{R}^n \times \mathbb{R}^p \rightarrow \mathbb{R}^m$ dimensionale kontinuierliche deterministische Messgleichung, Γ die $\mathbb{R}^n \times \mathbb{R}^p \rightarrow \mathbb{R}^{m \times z}$ Verteilungsmatrix der Messgleichung und γ der $\in \mathbb{R}^z$ Vektor des Messrauschens. Dabei müssen für die Verteilungsmatrizen die Bedingungen $v \geq n$ sowie $z \geq m$ erfüllt sein.

Die kontinuierlich-diskreten Systeme werden, wie schon der Name aussagt, durch eine zeitkontinuierliche gewöhnliche nichtlineare stochastische LANGEVIN'sche Differentialgleichung mit zeitdiskreter Messgleichung definiert:

$$\dot{x}(t) = f_{kon}(x, u; t) + K(x, u; t)\kappa(t) \quad (2.3)$$

$$y_k = h(x_k, u_k; t_k) + \Gamma(x_k, u_k; t_k)\gamma(t_k). \quad (2.4)$$

Die diskreten Systeme werden durch eine zeitdiskrete nichtlineare stochastische Differenzgleichung mit zeitdiskreter Messgleichung definiert:

$$x_k = f(x_{k-1}, u_{k-1}) + K(x_{k-1}, u_{k-1})\kappa_{k-1} \quad (2.5)$$

$$y_k = h(x_k, u_k) + \Gamma(x_k, u_k)\gamma_k. \quad (2.6)$$

Wie oben schon erwähnt, existieren in der Filtertheorie Lösungsansätze für alle oben aufgeführten stochastischen Systeme. In dieser Arbeit soll lediglich das WI-Problem für dynamische Systeme in diskreter Form aus nachfolgenden Gründen beleuchtet werden:

1. Die Anzahl an Lösungsalgorithmen/Inferenzverfahren für die zeitdiskreten Systeme ist deutlich umfangreicher als für die kontinuierlichen und kontinuierlich-diskreten Systeme. Somit macht ein sukzessives Testen und Bewerten für zeitdiskrete Systeme mehr Sinn.
2. Um ein sukzessives Testen zu ermöglichen, müssen Referenzzustände aus obigen stochastischen Systemgleichungen generiert werden, siehe Kapitel 4. Bei den kontinuierlichen und kontinuierlich-diskreten Systemen bedeutet das, dass die stochastischen Differentialgleichungen numerisch integriert werden müssen. Dies ist aus zweierlei Hinsicht problematisch: Formt man die stochastische Differentialgleichung in die äquivalente Integralgleichung um, so erhält man ein stochastisches Integral, welches der Form nach ein RIEMANN-STIELTJES-Integral [3] darstellt und definitionsgemäß durch RIEMANN-STIELTJES-Summen approximiert werden kann. Der Grenzwert der RIEMANN-STIELTJES-Summen hängt für dieses stochastische Integral wesentlich von der Wahl der Zwischenpunkte ab. Je nach Wahl der Zwischenpunkte erhält man unterschiedliche Lösungen des Integrals, die meistverbreiteten sind das ITÔ- und das STRATONOVICH-Integral. Die Frage, welche dieser Lösungen die "richtige" ist, ist ungeklärt [4]. Ebenso problematisch ist die numerische Integration (falls man sich für einen dieser Lösungsansätze entscheiden kann). Es existieren nur einfachste Algorithmen für die numerische Integration der stochastischen Differentialgleichungen (EULER, HEUN, RUNGE-KUTTA in stochastischer Form). Für steife stochastische Systeme existieren keine Lösungsalgorithmen.
3. Sind die kontinuierlichen Systeme nicht steif, so können sie ohne großen Aufwand stabil und konsistent diskretisiert werden und man erhält stochastische Differenzgleichung.
4. Sind die kontinuierlichen Systeme steif, so kann durch implizite Diskretisierung und Analyse der Stochastiken des Systems und der Messung eine stochastische Differenzgleichung gewonnen werden.

Die diskreten Systeme lassen sich also durch dynamische Zustandsraummodelle (Dynamic State Space Model - DSSM) beschreiben. Der verborgene Zustand x_k mit einer Initialverteilung $p(x_0)$ entwickelt sich mit der Zeit (k ist der diskrete Zeitindex) als indirekter oder partiell beobachteter MARKOV-Prozess erster Ordnung gemäß der bedingten Propagationsdichte/Transitionsdichte $p(x_k|x_{k-1})$. Die Beobachtungen y_k sind voneinander unabhängig und werden durch die bedingte Dichtefunktion $p(y_k|x_k)$ bestimmt (Likelihood).

Folgende Systeme seien somit für die Arbeit definiert:

Diskrete Systeme seien definiert durch ein nichtlineares Signalprozessmodell in zeitdiskreter Zustandsraumdarstellung, welches durch die Differenzengleichung

$$x_k = f(x_{k-1}, u_{k-1}) + K(x_{k-1}, u_{k-1})\kappa_{k-1} = \mathcal{F}(x_{k-1}, u_{k-1}, \kappa_{k-1}) \quad (2.7)$$

beschrieben ist. Dabei ist \mathcal{F} die $\mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^v \rightarrow \mathbb{R}^n$ dimensionale stochastische Systemgleichung. Die entsprechende (nichtlineare) Messgleichung sei definiert durch

$$y_k = h(x_k, u_k) + \Gamma(x_k, u_k)\gamma_k = \mathcal{H}(x_k, u_k, \gamma_k), \quad (2.8)$$

wobei \mathcal{H} die $\mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^z \rightarrow \mathbb{R}^m$ dimensionale stochastische Messgleichung ist.

Die vollständige Lösung des WI-Problems wird durch die bedingte posteriori Wahrscheinlichkeitsdichte (Filterdichte) des Zustandes x_k

$$p(x_k|Y_k) = ?$$

beschrieben, wobei

$$Y_k = \{y_1, y_2, \dots, y_k\} \quad (2.9)$$

die Sequenz sämtlicher Beobachtungs-Vektoren bis zu Zeitpunkt k repräsentiert. Ist die posteriori Wahrscheinlichkeitsdichte bekannt, so lassen sich die Zustände des Systems schätzen, z.B. Minimum-Mean-Square-Schätzung (MMSE) oder Maximum-A-Posteriori-Schätzung (MAP)

$$\hat{x}_k \stackrel{MMSE}{=} \mathbb{E}[p(x_k|Y_k)] = \int_{\mathbb{R}^n} x_k \cdot p(x_k|Y_k) dx_k \quad (2.10)$$

$$\hat{x}_k \stackrel{MAP}{=} \operatorname{argmax}_{x_k} p(x_k|Y_k). \quad (2.11)$$

Wie kann die posteriori Wahrscheinlichkeitsdichte bestimmt werden?

2.1 Rekursive Bayes'sche Schätztheorie (RBS)

Die meisten Methoden zur Parameter- und Zustandsschätzung, rekursiv und nichtrekursiv, linear und nicht-linear, haben ihren mathematischen Ursprung in BAYES Theorem, siehe z.B. [1], unterscheiden sich nur in der a priori Kenntnis über das System und die Stochastik. Die (Näherungs)lösung von BAYES Theorem hängt genau von diesem Wissen ab. Auch die "klassischen" Methoden, wie Least-Square oder Maximum-Likelihood sind Approximationslösungen von BAYES Theorem.

In dieser Arbeit werden lediglich rekursive Verfahren zur Bestimmung der posteriori Filterdichte untersucht. Dies liegt darin begründet, dass nichtrekursive Verfahren sich nicht zur On-Line Zustands- und Parameterschätzung eignen, da sie die Bereitstellung der gesamten Sequenz sämtlicher Beobachtungen Y_k erfordern und somit nicht parallel zum System On-Line einsetzbar sind.

Die posteriori Filterdichte kann rekursiv durch die BAYES'sche Schätztheorie bestimmt werden. Ausgehend von BAYES Theorem kann die posteriori Filterdichte zunächst expandiert und anschließend faktorisiert werden, um eine rekursive Bestimmungsgleichung zu erhalten:

$$\begin{aligned}
 p(x_k|Y_k) &= \frac{p(Y_k|x_k)p(x_k)}{p(Y_k)} \\
 &= \frac{p(y_k, Y_{k-1}|x_k)p(x_k)}{p(y_k, Y_{k-1})} \\
 &= \frac{p(y_k|Y_{k-1}, x_k)p(Y_{k-1}|x_k)p(x_k)}{p(y_k|Y_{k-1})p(Y_{k-1})} \\
 &= \frac{p(y_k|Y_{k-1}, x_k)p(x_k|Y_{k-1})p(Y_{k-1})p(x_k)}{p(y_k|Y_{k-1})p(Y_{k-1})p(x_k)} \\
 &= \frac{p(y_k|Y_{k-1}, x_k)p(x_k|Y_{k-1})}{p(y_k|Y_{k-1})} \\
 &= \frac{p(y_k|x_k)p(x_k|Y_{k-1})}{p(y_k|Y_{k-1})}. \tag{2.12}
 \end{aligned}$$

Um einen tieferen Einblick in die RBS zu erhalten, kann man die Zerlegung von (2.12) betrachten. Der Nenner von (2.12) ist eine Konstante, die als Normierungskonstante angesehen werden kann und dafür sorgt, dass die Filterdichte die Eigenschaft

$$\int_{\mathbb{R}^n} p(x_k|Y_k) dx_k = 1 \tag{2.13}$$

für sämtliche Dichtefunktionen aufweist. Somit müssen lediglich die Prädiktions-Verteilungsdichte (a priori Dichte) $p(x_k|Y_{k-1})$ und die Likelihood-Verteilungsdichte $p(y_k|x_k)$ in (2.12) bestimmt werden. Die posteriori Filterdichte zum Zeitpunkt $k - 1$ wird zunächst durch das wahrscheinlichkeitstheoretische Systemmodell (die Transitionsleichung) in der Zeit vorwärts zur a priori Filterdichte propagiert:

$$p(x_k|Y_{k-1}) = \int_{\mathbb{R}^n} p(x_k|x_{k-1}) \cdot p(x_{k-1}|Y_{k-1}) dx_{k-1}. \tag{2.14}$$

Als nächstes wird die aktuellste Messung verwendet, um durch die Likelihood-Verteilungsdichte die posteriori Filterdichte zum Zeitpunkt k zu bestimmen:

$$p(x_k|Y_k) = C \cdot p(y_k|x_k) \cdot p(x_k|Y_{k-1}). \quad (2.15)$$

Die normalisierende Konstante C ist dabei durch

$$C = \left(\int_{\mathbb{R}^n} p(y_k|x_k) \cdot p(x_k|Y_{k-1}) dx_k \right)^{-1} \quad (2.16)$$

gegeben und die Transitionsdichte und die Likelihood-Verteilungsdichte sind durch

$$p(x_k|x_{k-1}) = \int_{\mathbb{R}^v} \delta(x_k - \mathcal{F}(x_{k-1}, u_{k-1}, \kappa_{k-1})) \cdot p(\kappa_{k-1}) d\kappa_{k-1} \quad (2.17)$$

und

$$p(y_k|x_k) = \int_{\mathbb{R}^z} \delta(y_k - \mathcal{H}(x_k, u_k, \gamma_k)) \cdot p(\gamma_k) d\gamma_k \quad (2.18)$$

gegeben, wobei $\delta(\cdot)$ die Dirac-Delta-Funktion ist.

Die RBS ist die optimale Lösung des WI-Problems. Allerdings lassen sich die multidimensionalen Integrale nur für lineare, GAUSS'sche Systeme exakt lösen. In diesem Fall ist die geschlossene Lösung der RBS durch den berühmten Kalman Filter [5] gegeben. Für den allgemeinen Fall der nichtlinearen, nicht GAUSS'schen Systeme müssen praktische Näherungslösungen der RBS herangezogen werden.

2.2 Näherungslösungen der RBS

Eine Vielzahl an numerischen Näherungslösungen des RBS-Problems sind in den letzten Jahrzehnten veröffentlicht worden. Die nachfolgend aufgeführten Grundlagen zu den Näherungslösungen der RBS sind teilweise in Anlehnung an [3], [6] und [7] verfasst. Abbildung 2.2 ordnet die Näherungslösungen des WI-Problems in entsprechende Kategorien ein.

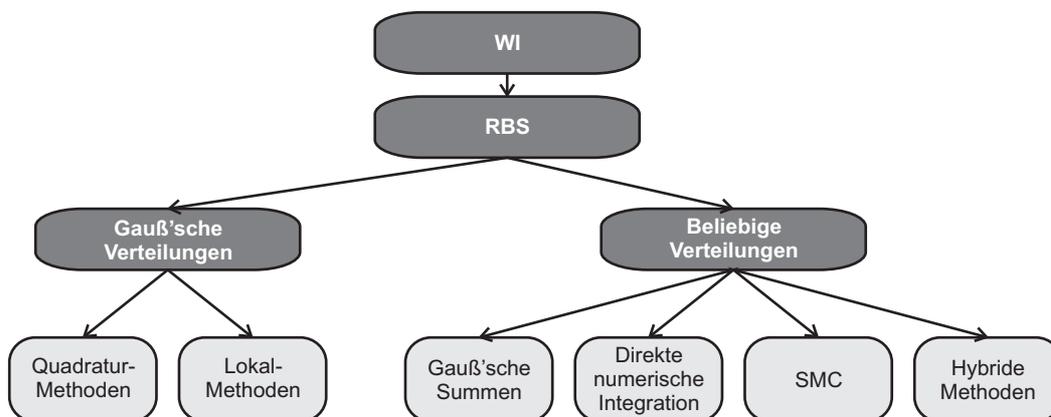


Abbildung 2.2: Näherungslösungen der RBS.

2.2.1 Methoden mit Gauß'scher Approximation

Methoden mit GAUSS'scher Approximation, auch Gauß-Filter (GF) genannt, gehen von GAUSS'schen Verteilungen sämtlicher Dichten in der BAYES'schen Rekursion aus, unter der Annahme, dass ein Minimum-Varianz-Schätzer (der posteriori Wahrscheinlichkeitsdichte) durch rekursives Propagieren und Updates des ersten und zweiten Moments der wahren Verteilung realisiert werden kann.

Unter der Annahme, dass alle Wahrscheinlichkeitsdichten der RBS normalverteilt sind, vereinfachen sich die BAYES'schen Rekursionsgleichungen erheblich. Die Filteraufgabe beschränkt sich somit auf die Bestimmung des bedingten Erwartungswertes und der Kovarianz (erstes und zweites Moment der Verteilung):

$$\begin{aligned}\hat{x}_{k|k} &= \mathbb{E}[x_k|Y_k] \\ \mathbf{P}_{k|k}^x &= \mathbb{E}\left[(x_k - \hat{x}_{k|k})(x_k - \hat{x}_{k|k})^T | Y_k\right].\end{aligned}$$

Gegeben sei das bereits oben vorgestellte diskrete Systemmodell mit diskreter System- und Messgleichung:

$$x_k = f(x_{k-1}, u_{k-1}) + K(x_{k-1}, u_{k-1})\kappa_{k-1} = \mathcal{F}(x_{k-1}, u_{k-1}, \kappa_{k-1}) \quad (2.19)$$

$$y_k = h(x_k, u_k) + \Gamma(x_k, u_k)\gamma_k = \mathcal{H}(x_k, u_k, \gamma_k). \quad (2.20)$$

Dabei sind κ und γ nun normalverteilte Zufallsvariablen

$$\kappa_{k-1} = \mathcal{N}(\kappa; \bar{\kappa}_{k-1}, \mathbf{P}_{k-1}^\kappa) \stackrel{\text{kurz}}{=} \mathcal{N}(\kappa; \bar{\kappa}, \mathbf{P}^\kappa) \quad (2.21)$$

$$\gamma_k = \mathcal{N}(\gamma; \bar{\gamma}_k, \mathbf{P}_k^\gamma) \stackrel{\text{kurz}}{=} \mathcal{N}(\gamma; \bar{\gamma}, \mathbf{P}^\gamma), \quad (2.22)$$

d.h. mit Mittelwertvektoren $\bar{\kappa}$ bzw. $\bar{\gamma}$ sowie Kovarianzmatrizen \mathbf{P}^κ bzw. \mathbf{P}^γ . An dieser Stelle sei angemerkt, dass für diskrete Systemdynamik in der Literatur [3], [6] nur Lösungen für mittelwertfreies Rauschen zu finden sind. Um jedoch noch allgemeinere stochastische Systeme behandeln zu können, wurden in dieser Arbeit die Gleichungen für mittelwertbehaftetes Rauschen erweitert.

Die Filterdichte ergibt sich aus der BAYES-Gleichung:

$$p(x_k|Y_k) = \frac{p(y_k|x_k)p(x_k|Y_{k-1})}{p(y_k|Y_{k-1})} = \frac{p(y_k|x_k)p(x_k|Y_{k-1})}{\int_{\mathbb{R}^n} p(y_k|x_k)p(x_k|Y_{k-1}) dx_k}. \quad (2.23)$$

Zur Lösung der BAYES-Gleichung müssen Likelihood $p(y_k|x_k)$ und die Transitionsdichte $p(x_k|x_{k-1})$ bestimmt werden. Sie ergeben sich in analoger Weise aus der Dichtefunktion mehrdimensionaler Normalverteilungen:

$$p(y_k|x_k) = \frac{1}{C} \cdot \exp\left[-\frac{1}{2}(y_k - \langle h + \Gamma\bar{\gamma} \rangle)^T \{\Gamma\mathbf{P}^\gamma\Gamma^T\}^{-1}(y_k - \langle h + \Gamma\bar{\gamma} \rangle)\right] \quad (2.24)$$

$$p(x_k|x_{k-1}) = \frac{1}{C} \cdot \exp\left[-\frac{1}{2}(x_k - \langle f + K\bar{\kappa} \rangle)^T \{K\mathbf{P}^\kappa K^T\}^{-1}(x_k - \langle f + K\bar{\kappa} \rangle)\right], \quad (2.25)$$

wobei die Konstanten \check{C} und \tilde{C} dafür sorgen, dass $\int_{\mathbb{R}^m} p(y_k|x_k)dy_k = \int_{\mathbb{R}^n} p(x_k|x_{k-1})dx_k = 1$ gilt, d.h.

$$\check{C} = \int_{\mathbb{R}^m} \exp \left[-\frac{1}{2} (y_k - \langle h + \Gamma \bar{\gamma} \rangle)^T \{ \Gamma \mathbf{P}^\gamma \Gamma^T \}^{-1} (y_k - \langle h + \Gamma \bar{\gamma} \rangle) \right] dy_k \quad (2.26)$$

$$\tilde{C} = \int_{\mathbb{R}^n} \exp \left[-\frac{1}{2} (x_k - \langle f + K \bar{\kappa} \rangle)^T \{ K \mathbf{P}^\kappa K^T \}^{-1} (x_k - \langle f + K \bar{\kappa} \rangle) \right] dx_k. \quad (2.27)$$

Aus den beiden Dichten lässt sich nun die BAYES'sche Rekursion in zwei Schritten lösen (bereits hier wird die Zwei-Schritt-Architektur sämtlicher Gauß-Filter offensichtlich):

1. Prädiktion

$$\begin{aligned} p(x_k|Y_{k-1}) &= \int_{\mathbb{R}^n} p(x_k|x_{k-1})p(x_{k-1}|Y_{k-1}) dx_{k-1} \\ &= \int_{\mathbb{R}^n} \frac{1}{\check{C}} \cdot \exp \left[-\frac{1}{2} (x_k - \langle f + K \bar{\kappa} \rangle)^T \{ K \mathbf{P}^\kappa K^T \}^{-1} (x_k - \langle f + K \bar{\kappa} \rangle) \right] \\ &\quad \cdot p(x_{k-1}|Y_{k-1}) dx_{k-1}. \end{aligned} \quad (2.28)$$

2. Filterung

$$\begin{aligned} p(x_k|Y_k) &= C \cdot p(y_k|x_k)p(x_k|Y_{k-1}) \\ &= C \cdot \frac{1}{\check{C}} \cdot \exp \left[-\frac{1}{2} (y_k - \langle h + \Gamma \bar{\gamma} \rangle)^T \{ \Gamma \mathbf{P}^\gamma \Gamma^T \}^{-1} (y_k - \langle h + \Gamma \bar{\gamma} \rangle) \right] \\ &\quad \cdot p(x_k|Y_{k-1}). \end{aligned} \quad (2.29)$$

Die Gleichungen (2.28) und (2.29) beschreiben den rekursiven Filter mit Prädiktions- (2.28) und Filterungsschritt (2.29). Dabei gilt die Anfangsbedingung:

$$p(x_0|Y_0) = p(x_0). \quad (2.30)$$

Zur theoretischen Lösung des Minimal-Varianz-Schätzproblems können die BAYES'schen Rekursionsformeln (2.28) und (2.29) verwendet werden, um die exakten Gleichungen für den bedingten Erwartungswert und die bedingte Kovarianzmatrix zu erhalten. Für die Prädiktion gilt:

$$\hat{x}_{k|k-1} = \mathbb{E}[x_k|Y_{k-1}] = \int_{\mathbb{R}^n} x_k \int_{\mathbb{R}^n} p(x_k|x_{k-1})p(x_{k-1}|Y_{k-1}) dx_{k-1} dx_k \quad (2.31)$$

$$\begin{aligned} \mathbf{P}_{k|k-1}^x &= \mathbb{E} \left[(x_k - \hat{x}_{k|k-1}) (x_k - \hat{x}_{k|k-1})^T | Y_{k-1} \right] \stackrel{[3]}{=} \mathbb{E} [x_k x_k^T | Y_{k-1}] - \hat{x}_{k|k-1} \hat{x}_{k|k-1}^T \\ &= \int_{\mathbb{R}^n} x_k x_k^T \int_{\mathbb{R}^n} p(x_k|x_{k-1})p(x_{k-1}|Y_{k-1}) dx_{k-1} dx_k - \hat{x}_{k|k-1} \hat{x}_{k|k-1}^T. \end{aligned} \quad (2.32)$$

Dabei ist zu beachten, dass $p(x_{k-1}|Y_{k-1})$ normalverteilt ist mit Mittelwert $\hat{x}_{k-1|k-1}$ und Kovarianzmatrix $\mathbf{P}_{k-1|k-1}^x$. Für die Filterung gilt in analoger Weise:

$$\hat{x}_{k|k} = \mathbb{E}[x_k|Y_k] = \int_{\mathbb{R}^n} x_k \cdot C \cdot p(y_k|x_k)p(x_k|Y_{k-1}) dx_k \quad (2.33)$$

$$\begin{aligned} \mathbf{P}_{k|k}^x &= \mathbb{E} \left[(x_k - \hat{x}_{k|k}) (x_k - \hat{x}_{k|k})^T | Y_k \right] \stackrel{[3]}{=} \mathbb{E} [x_k x_k^T | Y_k] - \hat{x}_{k|k} \hat{x}_{k|k}^T \\ &= \int_{\mathbb{R}^n} x_k x_k^T \cdot C \cdot p(y_k|x_k)p(x_k|Y_{k-1}) dx_k - \hat{x}_{k|k} \hat{x}_{k|k}^T. \end{aligned} \quad (2.34)$$

Es kann gezeigt werden [8], dass die aufgeführten Integrale auf die rekursive Schätzung

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + \mathcal{K}_k \cdot (y_k - \hat{y}_{k|k-1}) = \hat{x}_{k|k-1} + \mathcal{K}_k \cdot \tilde{y}_k \quad (2.35)$$

$$\mathbf{P}_{k|k}^x = \mathbf{P}_{k|k-1}^x - \mathcal{K}_k \cdot \mathbf{P}_k^{\tilde{y}} \cdot \mathcal{K}_k^T \quad (2.36)$$

führen, wobei die Kalman-Verstärkung \mathcal{K}_k gegeben ist durch

$$\mathcal{K}_k = \mathbf{P}_k^{xy} \cdot (\mathbf{P}_k^{\tilde{y}})^{-1} \quad (2.37)$$

mit der Kreuzkovarianz \mathbf{P}^{xy} und der Messfehlerkovarianz $\mathbf{P}^{\tilde{y}}$. Die oben aufgeführten Lösungen der RBS für GAUSS'sche Systeme können, wie bereits erwähnt, nur im linearen Fall exakt gelöst werden. Die Lösung entspricht dem Kalman Filter [5]. Der Kalman Filter ist somit Optimalfilter für lineare GAUSS'sche Systeme, es gibt keine bessere Lösung. Für den allgemeineren Fall der nichtlinearen Systeme müssen andere Ansätze zur Lösung verwendet werden.

Methoden mit Arbeitspunktlinearisierung (Lokal-Methoden)

Um die Lösung der BAYES-Integrale für lineare Systeme (Kalman Filter) auf nichtlineare Systeme anwenden zu können, werden die nichtlinearen System- und Messgleichungen in Taylor-Reihen entwickelt und nach dem ersten Glied abgebrochen. Die Systemgleichung wird um den bedingten posteriori Erwartungswert zum Zeitpunkt $k - 1$, $\hat{x}_{k-1|k-1}$, und die Messgleichung um den bedingten prädizierten (priori) Erwartungswert, $\hat{x}_{k|k-1}$, linearisiert. Man erhält so linearisierte Systeme, für die die Gleichungen des Kalman Filters verwendbar sind. Dieser Algorithmus ist als Extended Kalman Filter (EKF) [9] bekannt.

Je nach Ausmaß der Nichtlinearität von System- und Messgleichung kann die lineare Taylor-Entwicklung die Gleichungen mehr oder minder genau approximieren. Dies kann bei starken Nichtlinearitäten zu unerlaubt großen Abweichungen führen, die u.U. den Filter divergieren¹ lassen können. Aufgrund dieser Schwäche des klassischen Extended Kalman Filters sind einige Variationen entstanden, die die Problematik der örtlichen Approximation durch Linearisierung verringern:

- **Iterated Extended Kalman Filter**

Das Verfahren zur Iteration der Filtergleichungen (IEKF₁) wurde von JAZWINSKY [9] zur Verbesserung der Konvergenzeigenschaften des Filters, insbesondere bei schlechten Startwerten, durch Minimierung der Linearisierungsfehler der Messgleichung hergeleitet. Es ist auch ausführlich in [3] beschrieben. Die Prädiktion erfolgt analog zum EKF. Ausgehend von einem ersten Filterwert, der im statistischen Mittel näher am tatsächlichen Zustandsvektor liegt als der Prädiktionwert, welcher lediglich die Informationen aus den Messungen zu Zeitpunkt $k - 1$ verwertet hat, wird die nichtlineare Messgleichung um diesen ersten Filterwert nochmals linearisiert. Diese Iteration wird solange durchgeführt, bis die absolute Differenz zweier aufeinander folgender Iterationsergebnisse eine

¹Ein Filter wird als divergiert erklärt, wenn Kovarianzmatrizen ihre positive Definitheit verlieren. Dies resultiert in Kalman-Verstärkungen, die divergierte Schätzungen (inf, nan) liefern.

Abbruchschranke unterschreitet. Alternativ kann eine feste Anzahl von Iterationsschritten vorgegeben werden (bewährt sind 2-3). Die so iterativ gewonnenen Schätzungen für den Zustandsvektor und die Kalman-Verstärkung werden den übrigen EKF-Gleichungen zur Verfügung gestellt, die alle nur einmalig pro Tastschritt berechnet werden müssen. Das Verfahren zur Iteration von Filter- und Prädiktionsgleichungen (IEKF₂) setzt auf das IEKF₁-Verfahren auf. Um den Linearisierungsfehler der Systemnichtlinearität zu minimieren, werden die Prädiktionsgleichungen ebenfalls iteriert [3], [9]. Beim Eintreffen neuer Messwerte werden sowohl der prädizierte Zustand $\hat{x}_{k|k-1}$ als auch die zugehörige Kovarianzmatrix $\mathbf{P}_{k|k-1}^x$ nachträglich iterativ korrigiert. Der Abbruch der Iteration erfolgt analog zum IEKF₁.

- Es existieren noch weitere Verfahren, bei denen versucht wird, den Linearisierungsfehler zu minimieren, z.B. den Nonlinearity Compensating Extended Kalman Filter, publiziert von [10] oder den Adaptive Extended Kalman Filter (AEKF), publiziert erstmals von [11] für lineare Systeme und erweitert auf nichtlineare Systeme in [12]. Diese Verfahren sind den nachfolgenden Quadratur-Methoden meistens unterlegen und sind somit obsolet. Daher sei an dieser Stelle auf sie hingewiesen, auf eine detaillierte Beschreibung wird allerdings verzichtet. Der AEKF ist in der später vorgestellten Simulationsumgebung implementiert.

Methoden ohne Arbeitspunktlinearisierung (Quadratur-Methoden)

Um die Schwächen und den Aufwand der Arbeitspunktlinearisierung zu umgehen, wurden seit 1995 mehrere Algorithmen zur nichtlinearen Inferenz entworfen, die ohne eine Arbeitspunktlinearisierung auskommen. Das Integral der Form

$$I = \int_{\mathbb{R}^n} F(x) \cdot \frac{1}{\sqrt{(2\pi)^n |\mathbf{P}^x|}} \cdot \exp \left[-\frac{1}{2} (x - \bar{x})^T (\mathbf{P}^x)^{-1} (x - \bar{x}) \right] dx, \quad (2.38)$$

welches der Form nach den RBS-Integralen entspricht, substituiert durch

$$I = \int_{\mathbb{R}^n} \tilde{F}(b) \cdot \frac{1}{(2\pi)^{n/2}} \cdot \exp \left(-\frac{1}{2} b^T b \right) db, \quad (2.39)$$

mit $\mathbf{P}^x = \mathbf{S}^x (\mathbf{S}^x)^T$ und $x = \mathbf{S}^x b + \bar{x}$ wird durch verschiedene Quadratur-Regeln diskret approximiert.

1. JULIER&UHLMANN-Quadratur-Regel und der Unscented Kalman Filter (UKF)

In JULIER&UHLMANN [13] wird die Standard-Normalverteilung der Zufallsvariablen x durch die diskrete Verteilung

$$x^{(i)} = \begin{cases} \sqrt{n + \kappa} e^{(i)} & \text{für } 1 \leq i \leq n \\ -x^{(i)-n} & \text{für } n + 1 \leq i \leq 2n \\ 0 & \text{für } i = 2n + 1 \end{cases} \quad (2.40)$$

$$p(x^{(i)}) = \begin{cases} \frac{2\kappa}{2(n+\kappa)} & \text{für } i = 2n + 1 \\ \frac{1}{2(n+\kappa)} & \text{für } 1 \leq i \leq 2n \end{cases} \quad (2.41)$$

approximiert, wobei $\kappa > 0$ eine Konstante (nicht zu verwechseln mit dem Systemrauschen) und $e^{(i)}$ der i_{te} Einheitsvektor im \mathbb{R}^n ist. Das Integral (2.39) wird durch die Quadratur-Regel

$$I \approx I_{JU} = \sum_{i=1}^{2n+1} \tilde{F}(x^{(i)}) p(x^{(i)}) \quad (2.42)$$

approximiert. Basierend auf dieser Quadratur-Regel wurde die sog. Unscented Transformation (UT) und der Unscented Kalman Filter von JULIER&UHLMANN entwickelt. Nachfolgend sei die Unscented Transformation kurz beschrieben:

Die mehrdimensionale Zufallsvariable x sei durch eine Anzahl deterministisch gewählter Beispielpunkte (Sigma-Points) in ihrer Statistik näherungsweise beschrieben. \mathcal{X} sei eine Matrix mit $2n + 1$ Sigma-Points (Vektoren im Mehrdimensionalen) als Einträge der Spalten:

$$\mathcal{X}^{(i)} = \begin{cases} \bar{x} & \text{für } i = 1 \\ \bar{x} + \left(\sqrt{(n + \kappa) \mathbf{P}^x} \right)^{(i)} & \text{für } i = 2, \dots, n + 1 \\ \bar{x} - \left(\sqrt{(n + \kappa) \mathbf{P}^x} \right)^{(i)} & \text{für } i = n + 2, \dots, 2n + 1 \end{cases} \quad (2.43)$$

mit den Gewichten

$$\begin{aligned} w^{(1)} &= \frac{\kappa}{n + \kappa} \\ w^{(i)} &= \frac{1}{2(n + \kappa)} \quad \text{für } i = 2, \dots, 2n + 1, \end{aligned}$$

wobei $w^{(i)}$ das mit dem i_{ten} Sigma-Vektor assoziierte Gewicht darstellt und $\sum_{i=1}^{2n+1} w^{(i)} = 1$ ist. κ ist ein Skalierungsparameter und $\left(\sqrt{(n + \kappa) \mathbf{P}^x} \right)^{(i)}$ kennzeichnet die i_{te} Spalte (oder Reihe) der Cholesky-Faktorisierung der gewichteten Kovarianzmatrix $(n + \kappa) \mathbf{P}^x$. Abbildung 2.3 zeigt die Lokalisierung und das Gewicht (gekennzeichnet durch die Höhe) der Sigma-Vektoren eines typischen Satzes an Sigma-Vektoren für eine zweidimensionale GAUSS'sche Zufallsvariable.

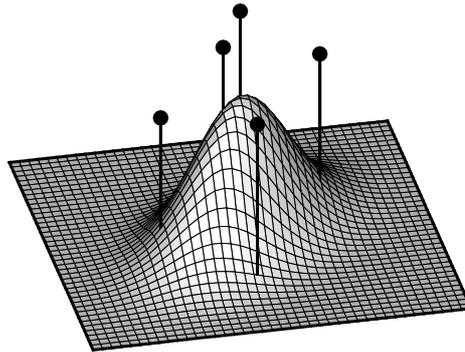


Abbildung 2.3: Typische Lokalisierung und Gewichtung (indiziert durch Höhe) von Sigma-Vektoren einer zweidimensionalen GAUSS'schen Zufallsvariable.

Um die Statistik (Mittelwert und Kovarianz) von $y = g(x)$ zu bestimmen, wird jeder der Sigma-Vektoren durch die nichtlineare Funktion propagiert

$$\Upsilon^{(i)} = g(\mathcal{X}^{(i)}) \quad (2.44)$$

und der approximierte Mittelwert sowie die Kovarianz und Kreuzkovarianz von y sind durch

$$\bar{y} \approx \sum_{i=1}^{2n+1} w^{(i)} \cdot \Upsilon^{(i)} \quad (2.45)$$

$$\mathbf{P}^y \approx \sum_{i=1}^{2n+1} w^{(i)} \cdot \left(\Upsilon^{(i)} - \bar{y} \right) \left(\Upsilon^{(i)} - \bar{y} \right)^T \quad (2.46)$$

$$\mathbf{P}^{xy} \approx \sum_{i=1}^{2n+1} w^{(i)} \cdot \left(\mathcal{X}^{(i)} - \bar{x} \right) \left(\Upsilon^{(i)} - \bar{y} \right)^T \quad (2.47)$$

gegeben. Dieser Vorgang ist als Unscented Transformation² bekannt. Bei der skalierten UT (Scaled Unscented Transformation, SUT) wird der originale Satz an Sigma-Vektoren durch

$$\tilde{\mathcal{X}}^{(i)} = \mathcal{X}^{(0)} + \alpha \left(\mathcal{X}^{(i)} - \mathcal{X}^{(0)} \right) \quad (2.48)$$

ersetzt, wobei α ein Skalierungsparameter ist, der beliebig innerhalb des Intervalls $0 \leq \alpha \leq 1$ gewählt werden kann. Diese Formulierung liefert einen zusätzlichen Freiheitsgrad in der Verteilung der Sigma-Vektoren und ermöglicht es, mögliche höhere Momente der Verteilung abzudecken. Setzt man

$$\lambda = \alpha^2 (n + \kappa) - n, \quad (2.49)$$

so erhält man einen neuen Satz an Sigma-Vektoren

$$\mathcal{X}^{(i)} = \begin{cases} \bar{x} & \text{für } i = 1 \\ \bar{x} + \left(\sqrt{(n + \lambda) \mathbf{P}^x} \right)^{(i)} & \text{für } i = 2, \dots, n + 1 \\ \bar{x} - \left(\sqrt{(n + \lambda) \mathbf{P}^x} \right)^{(i)} & \text{für } i = n + 2, \dots, 2n + 1 \end{cases} \quad (2.50)$$

mit den Gewichten

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

Die Gewichtung des ersten Sigma-Vektors beeinflusst direkt den Fehler im vierten und höheren Momenten symmetrischer Verteilungen [7]. Daher wird ein zusätzlicher Skalierungsparameter β eingeführt, der die Gewichtung des ersten Sigma-Vektors beeinflusst. Um die Statistik (Mittelwert und Kovarianz) von $y = g(x)$ zu bestimmen, wird wiederum jeder der Sigma-Vektoren durch die nichtlineare Funktion propagiert

$$\Upsilon^{(i)} = g(\mathcal{X}^{(i)}) \quad (2.51)$$

²Der Begriff "Unscented" → "Unparfümiert" wurde von JULIER&UHLMANN eingeführt. In ihren Arbeiten zeigten sie, dass das Konzept deterministisch gewählter Punkte zur Approximation von Statistiken auch außerhalb des Kalman Filter Kontextes Gültigkeit besitzt. Daher wurde die UT als der Transformationsprozess beschrieben, der "für sich selbst stehen kann" → "unparfümiert". Dabei legen sie Wert darauf, dass die Behauptung, der Begriff solle implizieren, dass der EKF "stinkt", reine Spekulation sei.

und der approximierte Mittelwert sowie die Kovarianz und Kreuzkovarianz von y sind durch

$$\bar{y} \approx \sum_{i=1}^{2n+1} w^{(i),m} \cdot \Upsilon^{(i)} \quad (2.52)$$

$$\mathbf{P}^y \approx \sum_{i=1}^{2n+1} w^{(i),c} \cdot \left(\Upsilon^{(i)} - \bar{y} \right) \left(\Upsilon^{(i)} - \bar{y} \right)^T \quad (2.53)$$

$$\mathbf{P}^{xy} \approx \sum_{i=1}^{2n+1} w^{(i),c} \cdot \left(\mathcal{X}^{(i)} - \bar{x} \right) \left(\Upsilon^{(i)} - \bar{y} \right)^T \quad (2.54)$$

gegeben. Die Entwicklung des Unscented Kalman Filters kann nun aus der UT/SUT und (2.35 – 2.37) direkt erfolgen [7].

II. Finite-Differenzen-Interpolation und der Central Difference Kalman Filter (CDKF)

Die Verfahren zur Finiten-Differenzen-Interpolation (FDI) wurden zeitgleich von [14] und [15] publiziert. Nachfolgende Erläuterungen sind in Anlehnung an diese Veröffentlichungen verfasst. Die Funktion $\tilde{F}(b)$ in (2.39) sei approximiert durch das Polynom Pol_2 zweiter Ordnung, welches

$$\tilde{F}(\xi^{(i)}) = Pol_2(\xi^{(i)}) \quad \text{für } 1 \leq i \leq \frac{(n+1)(n+2)}{2} \quad (2.55)$$

in den Punkten $\xi^{(i)}$ befriedigt. Diese sind gegeben durch

$$\xi^{(i)} = \begin{cases} \langle 0, he^{(i)}, -he^{(i)} \rangle & \text{für } 1 \leq i \leq n \\ he^{(i)} + he^{(j)} & \text{für } 1 \leq i < j \leq n, \end{cases} \quad (2.56)$$

wobei $h > 0$ die Schrittweite darstellt. Unter diesen Annahmen ist das Polynom Pol_2 gegeben durch

$$Pol_2(s) = \tilde{F}(0) + \sum_{i=1}^n a^{(i)} s^{(i)} + \frac{1}{2} s^T \mathbf{H} s, \quad (2.57)$$

wobei $s^{(i)}$ die i_{te} Koordinate eines Punktes $s \in \mathbb{R}^n$ ist. Der Vektor $a = (a^{(i)})$ und die symmetrische Matrix $\mathbf{H} = (\mathbf{H}^{(i),(j)})_{n \times n}$ sind definiert durch

$$a^{(i)} = \frac{\tilde{F}(he^{(i)}) - \tilde{F}(-he^{(i)})}{2h}, \quad \text{für } 1 \leq i \leq n \quad (2.58)$$

$$\mathbf{H}^{(i),(i)} = \frac{\tilde{F}(he^{(i)}) - 2\tilde{F}(0) + \tilde{F}(-he^{(i)})}{h^2}, \quad \text{für } 1 \leq i \leq n \quad (2.59)$$

$$\mathbf{H}^{(i),(j)} = \frac{\tilde{F}(he^{(i)} + he^{(j)}) - \tilde{F}(-he^{(i)}) - \tilde{F}(-he^{(j)}) + \tilde{F}(0)}{h^2}, \quad \text{für } 1 \leq i < j \leq n. \quad (2.60)$$

Somit wird (2.39) approximiert durch

$$I \approx I_{FDI} = \tilde{F}(0) + \sum_{i=1}^n \frac{1}{2} \mathbf{H}^{(i),(i)} \quad (2.61)$$

und das Integral

$$J = \int_{\mathbb{R}^n} \tilde{F}_1(b) \tilde{F}_2(b) \cdot \frac{1}{(2\pi)^{n/2}} \cdot \exp\left(-\frac{1}{2} b^T b\right) db \quad (2.62)$$

wird durch

$$\begin{aligned}
J &\approx \int_{\mathbb{R}^n} Pol_2^{\{1\}}(b) Pol_2^{\{2\}}(b) \cdot \frac{1}{(2\pi)^{n/2}} \cdot \exp\left(-\frac{1}{2}b^T b\right) db \\
&= J_{FDI} \\
&= I_{FDI}^{\{1\}} I_{FDI}^{\{2\}} + \sum_{i=1}^n \left(a^{(i),\{1\}} a^{(i),\{2\}} + \frac{1}{2} \mathbf{H}^{(i),(i),\{1\}} \mathbf{H}^{(i),(i),\{2\}} \right) \\
&\quad + \sum_{i \neq j} \mathbf{H}^{(i),(j),\{1\}} \mathbf{H}^{(i),(j),\{2\}}
\end{aligned} \tag{2.63}$$

approximiert, wobei $Pol_2^{\{i\}}$ die quadratische Polynomapproximation von \tilde{F}_i für $i = 1, 2$ ist. Alternativ kann \tilde{F} durch Pol_1 approximiert werden, wobei

$$Pol_1(s) = \tilde{F}(0) + \sum_{i=1}^n \frac{\tilde{F}(he^{(i)}) - \tilde{F}(-he^{(i)})}{2h} s^{(i)} + \sum_{i=1}^n \frac{1}{2} \mathbf{H}^{(i),(i)} (s^{(i)})^2 \tag{2.64}$$

ist. Diese Approximation basiert nur auf den Werten $\tilde{F}(\pm he^{(i)}) = Pol_1(\pm he^{(i)})$, $1 \leq i \leq n$ und $\tilde{F}(0) = Pol_1(0)$ und benutzt die diagonale Korrektur zweiter Ordnung $\sum_{i=1}^n (1/2) \mathbf{H}^{(i),(i)} (s^{(i)})^2$ für die zentrale Differenzenapproximation von \tilde{F} in (2.57). Somit werden die Integrale I und J approximiert durch

$$\begin{aligned}
I &\approx \int_{\mathbb{R}^n} Pol_1^{\{1\}}(b) \cdot \frac{1}{(2\pi)^{n/2}} \cdot \exp\left(-\frac{1}{2}b^T b\right) db \\
&= \tilde{I}_{FDI} = \tilde{F}(0) + \sum_{i=1}^n \frac{1}{2} \mathbf{H}^{(i),(i)}
\end{aligned} \tag{2.65}$$

und

$$\begin{aligned}
J &\approx \int_{\mathbb{R}^n} Pol_1^{\{1\}}(b) Pol_1^{\{2\}}(b) \cdot \frac{1}{(2\pi)^{n/2}} \cdot \exp\left(-\frac{1}{2}b^T b\right) db \\
&= \tilde{J}_{FDI} = \tilde{F}_1(0) \tilde{F}_2(0) + \sum_{i=1}^n \frac{\tilde{F}_1(he^{(i)}) - \tilde{F}_1(-he^{(i)})}{2h} \frac{\tilde{F}_2(he^{(i)}) - \tilde{F}_2(-he^{(i)})}{2h} \\
&\quad + \sum_{i=1}^n \frac{1}{2} \mathbf{H}^{(i),(i),\{1\}} \mathbf{H}^{(i),(i),\{2\}}.
\end{aligned} \tag{2.66}$$

Die Entwicklung des Central Difference Kalman Filters ist nun in analoger Form zum UKF möglich. Es kann gezeigt werden [7], dass die Funktionsauswertung durch die Approximation durch zentrale Differenzen in Analogie zur UT auf einen Datensatz an Sigma-Vektoren der Form

$$\mathcal{X}^{(i)} = \begin{cases} \bar{x} & \text{für } i = 1 \\ \bar{x} + \left(h\sqrt{\mathbf{P}^x}\right)^{(i)} & \text{für } i = 2, \dots, n+1 \\ \bar{x} - \left(h\sqrt{\mathbf{P}^x}\right)^{(i)} & \text{für } i = n+2, \dots, 2n+1 \end{cases} \tag{2.67}$$

mit den Gewichten

$$\begin{aligned}
w^{(1),m} &= \frac{h^2 - n}{h^2} \\
w^{(i),m} &= \frac{1}{2h^2} \quad \text{für } i = 2, \dots, 2n+1 \\
w^{(i),c1} &= \frac{1}{4h^2} \quad \text{für } i = 2, \dots, 2n+1 \\
w^{(i),c2} &= \frac{h^2 - 1}{4h^4} \quad \text{für } i = 2, \dots, 2n+1
\end{aligned}$$

führt. Der approximierte Mittelwert sowie die (Kreuz-)Kovarianz von $y = g(x)$ ergeben sich wiederum durch Propagation der Sigma-Vektoren durch g , d.h. $\Upsilon^{(i)} = g(\mathcal{X}^{(i)})$ und Berechnung durch

$$\bar{y} \approx \sum_{i=1}^{2n+1} w^{(i),m} \cdot \Upsilon^{(i)} \quad (2.68)$$

$$\mathbf{P}^y \approx \sum_{i=2}^{n+1} \left[w^{(i),c1} \cdot \left(\Upsilon^{(i)} - \Upsilon^{(n+i)} \right)^2 + w^{(i),c2} \left(\Upsilon^{(i)} + \Upsilon^{(n+i)} - 2\Upsilon^{(1)} \right)^2 \right] \quad (2.69)$$

$$\mathbf{P}^{xy} \approx \sqrt{w^{(2),c1} \mathbf{P}^x} \left[\Upsilon^{(2:n+1)} - \Upsilon^{(n+2:2n+1)} \right]^T. \quad (2.70)$$

Dabei beschreibt $(\cdot)^2$ das äußere Vektorprodukt, d.h. $(s)^2 = ss^T$. Somit sind der UKF und der CDKF in ihrer Struktur identisch, lediglich die Verteilungen werden durch verschiedene Sigma-Vektoren und Gewichte repräsentiert.

III. Gauß-Hermite-Quadratur und der Gauß Hermite Filter (GHF)

Das Verfahren zur Gauß-Hermite-Quadratur und der Gauß Hermite Filter wurden ebenfalls in [15] publiziert. Nachfolgende Erläuterungen sind teilweise in Anlehnung an diese Veröffentlichung verfasst. Im eindimensionalen Fall ist die Gauß-Hermite-Quadratur-Regel q_{ter} Ordnung durch (siehe, z.B. [16])

$$\int_{\mathbb{R}^n} g(x) \exp(-x^2) dx = \sum_{i=1}^q w^{(i)} g(x^{(i)}), \quad (2.71)$$

wobei die Gleichung für alle Polynome $g(x)$ bis zur Ordnung $2q-1$ hält. Die Quadratur-Punkte $x^{(i)}$ und die Gewichte $w^{(i)}$ können wie folgt berechnet werden (siehe, z.B. [17]): \mathbf{J} sei eine symmetrische, tridiagonale Matrix mit Nulleinträgen in der Diagonalen und $\mathbf{J}^{(i),(i+1)} = \sqrt{i/2}$, $1 \leq i \leq q-1$:

$$\mathbf{J} = \begin{bmatrix} 0 & \sqrt{1/2} & & & \\ \sqrt{1/2} & \ddots & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \sqrt{(q-1)/2} & \\ & & & \sqrt{(q-1)/2} & 0 \end{bmatrix}. \quad (2.72)$$

Dann sind die Quadratur-Punkte gleich den Eigenwerten $\{\lambda^{(i)}\}$ von \mathbf{J} und $w^{(i)} = |(v^{(i)})_1|^2$, wobei $(v^{(i)})_1$ das erste Element des i_{ten} normierten Eigenvektors von \mathbf{J} kennzeichnet. Die Normierung ist:

$$(v^{(i)})^T v^{(i)} = \int_{-\infty}^{\infty} \exp(-x^2) dx = \sqrt{\pi}. \quad (2.73)$$

Somit kann im eindimensionalen Fall

$$\begin{aligned} I &= \int_{\mathbb{R}^n} f(x) \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left(\frac{x-\bar{x}}{\sigma}\right)^2\right] dx = \int_{\mathbb{R}^n} \tilde{f}(b) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}b^2\right) db \\ &= \int_{\mathbb{R}^n} g(b) \exp(-b^2) db, \end{aligned} \quad (2.74)$$

wobei $g(b) = \tilde{f}(b) \frac{1}{\sqrt{2\pi}} \exp\left(\frac{1}{2}b^2\right)$ und $x = \sigma b + \bar{x}$ ($dx = \sigma db$) ist, durch

$$I = \sum_{i=1}^q w^{(i)} g(\lambda^{(i)}) \quad (2.75)$$

und im mehrdimensionalen Fall (2.39) durch

$$I \approx I_{GH} = \sum_{i_1=1}^q \cdots \sum_{i_n=1}^q \tilde{G}(\lambda^{(i_1)}, \lambda^{(i_2)}, \dots, \lambda^{(i_n)}) w^{(i_1)} w^{(i_2)} \dots w^{(i_n)} \quad (2.76)$$

approximiert werden, wobei $\tilde{G}(b) = \tilde{F}(b) \frac{1}{(2\pi)^{n/2}} \exp(\frac{1}{2}b^T b)$. Für die Berechnung von I_{GH} werden q^n Funktionsauswertungen benötigt.

Aus der Gauß-Hermite-Quadratur-Regel kann nun ein Filteralgorithmus für nichtlineare GAUSS'sche Systeme entworfen werden. Die Funktionsauswertung durch die Gauß-Hermite-Quadratur-Regel führt in Analogie zum UKF und zum CDKF auch auf einen Datensatz an Sigma-Vektoren.

Somit sind der UKF, der CDKF und die GHF in ihrer Struktur identisch, lediglich die Verteilungen werden durch verschiedene Sigma-Vektoren und Gewichte repräsentiert.

Sigma-Point Kalman Filter

Wie zuvor gezeigt, führen alle Quadratur-Regeln und hier vorgestellten Filter auf einen Satz an Sigma-Vektoren. Aufgrund dieser Analogie können UKF, CDKF und GHF in der Gruppe der Sigma-Point Kalman Filter (SPKF) zusammengefasst werden.

Der benötigte Rechenaufwand der drei Sigma-Point Kalman Filter wird durch

- $(n + m)(2n + 1)$ Funktionsauswertungen beim UKF
- $(n + m)(2n + 1)$ Funktionsauswertungen beim CDKF
- $(n + m)q^n$ Funktionsauswertungen beim GHF, wobei eine Quadratur-Regel q_{ter} Ordnung angewendet wird,

bestimmt. Hier zeigt sich auch schon der gravierende Vor- und Nachteil der Verfahren. UKF und CDKF approximieren die BAYES-Integrale, wohingegen der GHF die BAYES-Integrale nahezu exakt berechnet. Dies resultiert in einem starken Anstieg an Rechenaufwand, was bei höherdimensionalen Systemen schnell dazu führt, dass der GHF nicht mehr applizierbar ist.

Iterated Sigma-Point Kalman Filter

In [18] und [19] wurden in Analogie zum IEKF₁ iterierende Sigma-Point Kalman Filter entwickelt, welche die Konvergenzeigenschaft der Filter weiter verbessern sollen. In [19] wurde lediglich der Prädiktionsschritt des UKF beibehalten und mit dem Filterungsschritt des IEKF₁ verknüpft, also ein hybrider Filter aus beiden Verfahren. Der von [18] vorgeschlagene iterierende UKF (IUKF) hingegen verknüpft die Vorteile der statistischen Linearisierung (Sigma-Points) mit der Iteration und beschreibt den eigentlichen IUKF, welcher im Nachfolgenden auch mit IUKF abgekürzt sein soll. In Analogie zum IUKF kann auch der iterierende

CDKF (ICDKF) hergeleitet werden, was im Rahmen dieser Arbeit umgesetzt wurde, hier aber nicht näher beschrieben werden soll.

2.2.2 Methoden für Zufallsvariablen beliebiger Verteilung

Für den Fall, dass die Zufallsvariablen eines Systems nicht als normalverteilt angenommen werden können, liefern die zuvor vorgestellten Algorithmen u.U. unbefriedigende Ergebnisse. Daher sind Inferenzverfahren entwickelt worden, die die Gleichungen der RBS auch für nichtlineare, nicht GAUSS'sche Systeme approximieren.

Diese Verfahren sind notwendig, da die nichtlineare Transformation GAUSS'scher Verteilungen in nicht GAUSS'schen Verteilungen resultiert. Siehe dazu Abbildung 2.4.

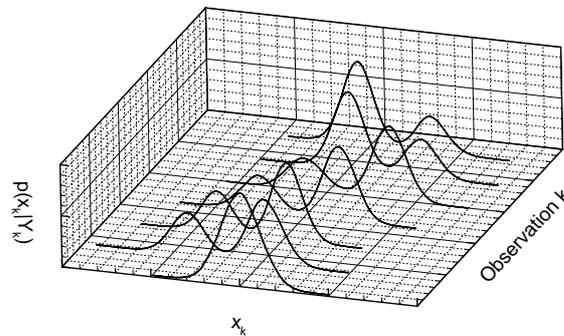


Abbildung 2.4: Verteilung von 10000 Partikeln durch das nichtlineare Systemmodell: $x_k = 0.5x_{k-1} + 25 \frac{x_{k-1}}{1+x_{k-1}^2} + 8 \cos(1.2k) + \kappa_{k-1}$, wobei $\kappa_{k-1} \sim \mathcal{N}(\kappa; 0, 10)$.

Hier wird besonders deutlich, dass das erste und zweite Moment die Verteilung nicht immer ausreichend genau beschreiben.

Methoden direkter numerischer Integration

Diese Methoden sind allgemein bekannt als Gitternetzfilter (Grid-based filters) (GBF). Sie approximieren die Integrale der RBS durch große, aber finite Summen über ein äquidistantes n -dimensionales Gitternetz über den gesamten Zustandsraum, siehe Abbildung 2.5. Die Methoden sind somit beschränkt auf Systeme mit finiten Zustandsräumen, und für höherdimensionale Systeme (höherdimensional im Sinne von Systemen mit mehreren Zuständen) wird der Rechenaufwand schnell extrem groß, weshalb diese Methoden für die meisten Systeme der "realen" Welt nicht praktikabel sind. Für weitere Informationen sei auf [20] und [21] verwiesen.

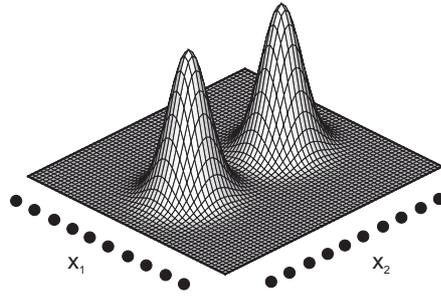


Abbildung 2.5: Integration über ein äquidistantes Gitternetz, aufgespannt über einen zweidimensionalen Zustandsraum.

Methoden mit Gauß'scher Summenapproximation (GSA)

Es kann gezeigt werden [22], dass jedwede Wahrscheinlichkeitsverteilung $p(x)$ beliebig genau durch eine Summe GAUSS'scher Normalverteilungen (GMM, Gaussian mixture model) approximiert werden kann:

$$p(x) \approx p_{GMM}(x) = \sum_{g=1}^G \alpha^{(g)} \cdot \mathcal{N}\left(x; \bar{x}^{(g)}, \mathbf{P}^{x,(g)}\right), \quad (2.77)$$

wobei G die Anzahl der Summanden darstellt, α die Komponentengewichte beschreibt und $\mathcal{N}(x; \bar{x}, \mathbf{P}^x)$ Normalverteilungen mit Mittelwertvektoren \bar{x} und Kovarianzmatrizen \mathbf{P}^x beschreiben. Basierend auf diesen GMM-Modellen wurden rekursive Filteralgorithmen entwickelt (wichtigste Literaturstellen sind [23, 24, 25, 26]), welche es erlauben, nicht GAUSS'sche Verteilungen abzudecken.

Legt man das dynamische Zustandsraummodell zu Grunde und nimmt ferner an, dass die Filterdichte $p(x_{k-1}|Y_{k-1})$ zum Zeitpunkt $k-1$ bekannt ist und auch die Dichtefunktionen des System- und Messrauschens durch ein GMM-Modell approximiert werden können, mit

$$p_{GMM}(x_{k-1}|Y_{k-1}) = \sum_{g=1}^G \alpha_{k-1}^{(g)} \cdot \mathcal{N}\left(x_{k-1}; \bar{x}_{k-1}^{(g)}, \mathbf{P}_{k-1}^{x,(g)}\right) \quad (2.78)$$

$$p_{GMM}(\kappa_{k-1}) = \sum_{m=1}^M \beta_{k-1}^{(m)} \cdot \mathcal{N}\left(\kappa_{k-1}; \bar{\kappa}_{k-1}^{(m)}, \mathbf{P}_{k-1}^{\kappa,(m)}\right) \quad (2.79)$$

$$p_{GMM}(\gamma_k) = \sum_{j=1}^J \varphi_k^{(j)} \cdot \mathcal{N}\left(\gamma_k; \bar{\gamma}_k^{(j)}, \mathbf{P}_k^{\gamma,(j)}\right), \quad (2.80)$$

so können nachfolgende Wahrscheinlichkeitsverteilungen ebenso als GMM approximiert werden:

1. Die prädierte priori Wahrscheinlichkeitsverteilung zum Zeitpunkt k :

$$p(x_k|Y_{k-1}) \approx p_{GMM}(x_k|Y_{k-1}) = \sum_{g'=1}^{G'} \alpha_{k|k-1}^{(g')} \cdot \mathcal{N}\left(x_{k|k-1}; \bar{x}_{k|k-1}^{(g')}, \mathbf{P}_{k|k-1}^{x,(g')}\right). \quad (2.81)$$

2. Die gefilterte posteriori Wahrscheinlichkeitsverteilung zum Zeitpunkt k :

$$p(x_k|Y_k) \approx p_{GMM}(x_k|Y_k) = \sum_{g''=1}^{G''} \alpha_{k|k}^{(g'')} \cdot \mathcal{N}\left(x_{k|k}; \bar{x}_{k|k}^{(g'')}, \mathbf{P}_{k|k}^{x,(g'')}\right). \quad (2.82)$$

Dabei sind $G' = G \cdot M$ und $G'' = G' \cdot J = G \cdot M \cdot J$ (G , M und J sind die Anzahl der Komponenten der GMM-Modelle für Verteilungen des Zustandes, des System- und Messrauschens). Die prädizierten und gefilterten Mittelwerte und Kovarianzen der GMM-Modell-Komponenten von $p_{GMM}(x_k|Y_{k-1})$ und $p_{GMM}(x_k|Y_k)$,

$$\text{Prädiktion : } \left\{ \bar{x}_{k|k-1}^{(g')}, \mathbf{P}_{k|k-1}^{x,(g')} \right\} \quad (2.83)$$

$$\text{Filterung : } \left\{ \bar{x}_{k|k}^{(g'')}, \mathbf{P}_{k|k}^{x,(g'')} \right\}, \quad (2.84)$$

werden durch eine Bank paralleler Filteralgorithmen (z.B. EKF) berechnet. Die resultierenden Algorithmen sind als Gaussian Sum Kalman Filter (GSKF) bekannt. Die Gewichte der GMM-Komponenten werden durch die von SØRENSEN&ALSPACH [23] vorgeschlagene Methode berechnet, d.h. die Gewichte für die Prädiktion sind gegeben durch

$$\alpha_{k|k-1}^{(g')} = \frac{\alpha_{k-1|k-1}^{(g)} \cdot \beta_{k-1}^{(m)}}{\sum_{g=1}^G \sum_{m=1}^M \alpha_{k-1|k-1}^{(g)} \cdot \beta_{k-1}^{(m)}} \quad (2.85)$$

und die Gewichte für die Filterung sind gegeben durch

$$\alpha_{k|k}^{(g'')} = \frac{\alpha_{k|k-1}^{(g')} \cdot \varphi_k^{(j)} \cdot z_k^{(j)}}{\sum_{g'=1}^{G'} \sum_{j=1}^J \alpha_{k|k-1}^{(g')} \cdot \varphi_k^{(j)} \cdot z_k^{(j)}}, \quad (2.86)$$

wobei

$$z_k^{(j)} = p^{(j)}(y_k|x_k) \quad (2.87)$$

die Likelihood-Funktion der Komponente j von $p_{GMM}(\gamma_k)$ ausgewertet mit

$$x_k = \bar{x}_k^{(g')}$$

ist.

Aus den hier aufgeführten Gleichungen ist ersichtlich, dass die Anzahl der GMM-Komponenten von G auf G' während des Prädiktionsschritts und von G' auf G'' während des Filterungsschritts anwächst. Dies hat zur Folge, dass mit fortschreitender Zeit die Anzahl der GMM-Komponenten exponentiell anwächst, was einen praktischen Nutzen des Filteralgorithmus in dieser Form verhindert. Daher muss ein Reduktionsschritt eingeführt werden, der die GMM-Komponenten so reduziert, dass die Anzahl der Filter pro Bank nicht exponentiell wächst und der durch die Reduzierung entstehende Fehler gering bleibt. In [23] wird ein Reduktionsschema für Verteilungsdichten identischer Varianz mit verschiedenen Gewichten und Mittelwerten vorgeschlagen. Alternativ können die Komponenten mit den geringsten Gewichten derart vernachlässigt werden, dass die Gesamtanzahl der GMM-Komponenten konstant gehalten wird.

Methoden mit sequentieller Monte-Carlo-Simulation (SMC-Methoden, Partikel-Filter)

Die Partikel-Filter basieren auf Monte-Carlo-Simulation mit Sequentiellem Importance Sampling (SIS) [27], [28]. Das verfolgte Ziel ist, die BAYES-Gleichungen direkt zu lösen. In Analogie zu den GBF approximieren

die SMC-Methoden die RBS-Integrale durch finite Summen. Im Gegensatz zu den GBF allerdings wird über ein stochastisch adaptives Gitternetz summiert. Dieses adaptive Netz, konzentriert an Orten hohen Likelihoods des Zustandsraumes, wird durch eine Menge gewichteter Beispielpunkte (Partikel) definiert, siehe Abbildung 2.6.

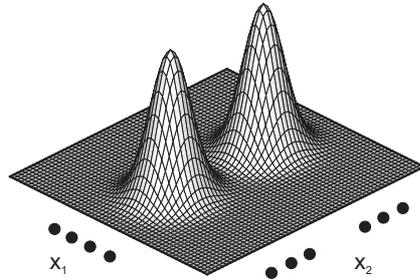


Abbildung 2.6: Integration über ein adaptives Gitternetz, aufgespannt über einen zweidimensionalen Zustandsraum.

Perfekte Monte-Carlo-Simulation

Die Filterdichte kann durch die empirische Näherung

$$p(x_k|Y_k) \approx \hat{p}(x_k|Y_k) = \frac{1}{N} \sum_{i=1}^N \delta(x_k - x_k^{(i)}) \quad (2.88)$$

approximiert werden, wobei die Partikel $\{x^{(i)}, i = 1, 2, \dots, N\}$ aus $p(x_k|Y_k)$ generiert werden und $\delta(\cdot)$ die Dirac-Delta-Funktion beschreibt. Als Konsequenz daraus kann ebenso jeder Erwartungswert der Form

$$\mathbb{E}[g(x_k)] = \int g(x_k) p(x_k|Y_k) dx_k \quad (2.89)$$

durch eine diskrete Summe approximiert werden:

$$\mathbb{E}[g(x_k)] \approx \tilde{\mathbb{E}}[g(x_k)] = \frac{1}{N} \sum_{i=1}^N g(x_k^{(i)}) \quad (2.90)$$

So resultiert daraus zum Beispiel für $g(x_k) = x_k$ die MMSE-Schätzung $\hat{x}_k \approx \mathbb{E}[x_k|Y_k]$. Die Partikel $x_k^{(i)}$ sind voneinander unabhängig und identisch verteilt (u.i.v.), damit die Approximation hält. Geht N gegen unendlich, so konvergiert die geschätzte Erwartung nahezu sicher (n.s.) gegen die wahre Erwartung:

$$\tilde{\mathbb{E}}[g(x_k)] \xrightarrow{n.s.} \mathbb{E}[g(x_k)]. \quad (2.91)$$

Bayesian Importance Sampling

Im Allgemeinen ist die Filterdichte nicht bekannt. Diese Problematik wird durch das Importance Sampling (IS) umgangen. Im Nachfolgenden sei $X_k = \{x_0, x_1, x_2, \dots, x_k\}$ die Sequenz der Zustände bis zum Zeitpunkt k . Die Partikel werden von einer einfach handzuhabenden Ansatzverteilung $\pi(X_k|Y_k)$ gewonnen.

Daraus folgt die Substitution:

$$\begin{aligned}
\mathbb{E}[g(X_k)] &= \int g(X_k) \frac{p(X_k|Y_k)}{\pi(X_k|Y_k)} \pi(X_k|Y_k) dX_k \\
&= \int g(X_k) \frac{p(Y_k|X_k)p(X_k)}{p(Y_k)\pi(X_k|Y_k)} \pi(X_k|Y_k) dX_k \\
&= \int g(X_k) \frac{w_k(X_k)}{p(Y_k)} \pi(X_k|Y_k) dX_k,
\end{aligned} \tag{2.92}$$

wobei die $w_k(X_k)$ als die unnormierten Gewichte bekannt sind, gegeben mit:

$$w_k(X_k) = \frac{p(Y_k|X_k)p(X_k)}{\pi(X_k|Y_k)}. \tag{2.93}$$

Die normalisierende Dichte $p(Y_k)$ kann wie folgt eliminiert werden:

$$\begin{aligned}
\mathbb{E}[g(X_k)] &= \frac{1}{p(Y_k)} \int g(X_k) w_k(X_k) \pi(X_k|Y_k) dX_k \\
&= \frac{\int g(X_k) w_k(X_k) \pi(X_k|Y_k) dX_k}{\int p(Y_k|X_k)p(X_k) \frac{\pi(X_k|Y_k)}{\pi(X_k|Y_k)} dX_k} \\
&= \frac{\int g(X_k) w_k(X_k) \pi(X_k|Y_k) dX_k}{\int w_k(X_k) \pi(X_k|Y_k) dX_k} \\
&= \frac{\mathbb{E}_\pi[g(X_k) w_k(X_k)]}{\mathbb{E}_\pi[w_k(X_k)]}.
\end{aligned} \tag{2.94}$$

Dabei deutet $\mathbb{E}_\pi[\cdot]$ an, dass die Erwartungswerte aus der Ansatzverteilung $\pi(X_k|Y_k)$ berechnet werden. Generiert man also Partikel von der Ansatzverteilung, so kann jedweder Erwartungswert wie folgt approximiert werden:

$$\begin{aligned}
\mathbb{E}[g(X_k)] \approx \tilde{\mathbb{E}}[g(X_k)] &= \frac{\frac{1}{N} \sum_{i=1}^N g(X_k^{(i)}) w_k(X_k^{(i)})}{\frac{1}{N} \sum_{i=1}^N w_k(X_k^{(i)})} \\
&= \sum_{i=1}^N \tilde{w}_k^{(i)} g(X_k^{(i)}),
\end{aligned} \tag{2.95}$$

wobei die normalisierten Gewichte gegeben sind durch:

$$\begin{aligned}
\tilde{w}_k^{(i)} &= \frac{w_k(X_k^{(i)})}{\sum_{j=1}^N w_k(X_k^{(j)})} \\
&\equiv \frac{w_k^{(i)}}{\sum_{j=1}^N w_k^{(j)}}.
\end{aligned} \tag{2.96}$$

Die Approximation in Gleichung (2.95) ist nicht erwartungstreu. Nichtsdestotrotz ist es möglich, asymptotische Konvergenz und einen zentralen Grenzwertsatz für $\tilde{\mathbb{E}}[g(X_k)]$ zu erlangen, wenn folgende Annahmen zutreffen [29]:

1. $X_k^{(i)}$ sind unabhängig und identisch verteilte Partikel, generiert aus der Ansatzverteilung $\pi(X_k|Y_k)$.
2. Der Gültigkeitsbereich der Ansatzverteilung schließt den Gültigkeitsbereich der posteriori Verteilung ein und $\mathbb{E}[g(X_k)]$ existiert und ist finit.

3. Die Erwartungswerte von \tilde{w}_k und $\tilde{w}_k g^2(X_k)$ der posteriori Verteilung existieren und sind finit.

Gelten diese Annahmen, so kann die posteriori Wahrscheinlichkeitsdichte beliebig genau durch folgende empirische Punktmassenapproximation beschrieben werden:

$$\hat{p}(X_k|Y_k) = \sum_{i=1}^N \tilde{w}_k^{(i)} \delta(X_k - X_k^{(i)}). \quad (2.97)$$

Sequential Importance Sampling (SIS)

Um eine sequentielle Schätzung der posteriori Verteilung zum Zeitpunkt k zu ermöglichen, ohne die zuvor simulierten Zustände $X_{k-1}^{(i)}$ modifizieren zu müssen, können Ansatzverteilungen der Form

$$\pi(X_k|Y_k) = \pi(X_{k-1}|Y_{k-1}) \pi(x_k|X_{k-1}, Y_k) \quad (2.98)$$

Verwendung finden. Hierbei wird angenommen, dass der aktuelle Zustand unabhängig von zukünftigen Messungen ist. Diese Annahme trifft auf die in dieser Arbeit betrachteten DSSM zu. Unter dieser Annahme, dass die Zustände einem MARKOV-Prozess gehorchen und dass die Messungen unabhängig sind, gilt:

$$p(X_k) = p(x_0) \prod_{j=1}^k p(x_j|x_{j-1}) \quad \text{und} \quad p(Y_k|X_k) = \prod_{j=1}^k p(y_j|x_j). \quad (2.99)$$

Setzt man die Gleichungen (2.98) und (2.99) in Gleichung (2.93) ein, so kann eine rekursive Gleichung für die Gewichte hergeleitet werden (siehe [7] für Details):

$$\begin{aligned} w_k &= \frac{p(Y_k|X_k) p(X_k)}{\pi(X_{k-1}|Y_{k-1}) \pi(x_k|X_{k-1}, Y_k)} \\ &= w_{k-1} \frac{p(Y_k|X_k) p(X_k)}{p(Y_{k-1}|X_{k-1}) p(X_{k-1})} \frac{1}{\pi(x_k|X_{k-1}, Y_k)} \\ &= w_{k-1} \frac{p(y_k|x_k) p(x_k|x_{k-1})}{\pi(x_k|X_{k-1}, Y_k)}. \end{aligned} \quad (2.100)$$

Diese Gleichung ermöglicht ein sequentielles Updaten der Gewichte, sofern eine geeignete Ansatzverteilung $\pi(x_k|X_{k-1}, Y_k)$ verfügbar ist. Die exakte Form dieser Verteilung ist ein kritisches Entwurfskriterium. Details dazu folgen im nächsten Abschnitt, siehe Seite 29. Da Partikel aus der Ansatzverteilung generiert werden können $x_k^{(i)} \sim \pi(x_k|X_{k-1}, Y_k)$, $i = 1, \dots, N$ und sowohl Likelihood $p(y_k|x_k^{(i)})$ als auch die Transitionsdichte $p(x_k^{(i)}|x_{k-1})$ auswertbar sind, müssen nur ein Initialdatensatz an Partikeln generiert und iterativ die Gewichte berechnet werden. Die Gewichte des Initialdatensatzes an Partikeln sind identisch $w_0^{(i)} = \frac{1}{N}$, $i = 1, \dots, N$. Diese Prozedur ist bekannt als Sequential Importance Sampling und erlaubt es, Schätzungen der Form

$$\mathbb{E}[g(x_k)] \approx \sum_{i=1}^N \tilde{w}_k^{(i)} g(x_k^{(i)}) \quad (2.101)$$

zu berechnen. Da in dieser Arbeit der Fokus auf den Filterungsprozess und nicht auf das Smoothing gerichtet ist, muss nicht die gesamte Historie der Partikel-Trajektorien berücksichtigt werden. Aus diesem

Grund können $g(X_k^{(i)})$ und $X_k^{(i)}$ durch $g(x_k^{(i)})$ und $x_k^{(i)}$ ersetzt werden. Somit kann jedwede posteriori Filterdichte beliebig genau durch die empirische Punktmassenapproximation

$$\hat{p}(x_k|Y_k) = \sum_{i=1}^N \tilde{w}_k^{(i)} \delta(x_k - x_k^{(i)}) \quad (2.102)$$

angenähert werden. Die Genauigkeit der Approximation ist lediglich durch die Anzahl der Partikel begrenzt. Im Gegensatz zu den GAUSS'schen Approximationsmethoden können so auch höhere Momente von Verteilungen wie Schiefe oder Wölbung sowie multimodale Verteilungen abgedeckt werden.

Wahl der Ansatzverteilungsdichte $\pi(x_k|X_{k-1}, Y_k)$

Wie bereits zuvor erwähnt, ist die Wahl der Ansatzverteilung ein wichtiges/kritisches Entwurfskriterium für die SIS-Algorithmen. DOUCET zeigt in [27], dass die optimale Ansatzverteilung die Varianz der Gewichte minimiert. Er beweist, dass die Ansatzverteilung³

$$\pi(x_k|X_{k-1}, Y_k) \stackrel{\circ}{=} p(x_k|x_{k-1}, y_k) \quad (2.103)$$

die Varianz der Gewichte minimiert. Diese Verteilung wurde auch von anderen Wissenschaftlern hergeleitet und ist als optimale Ansatzverteilung bekannt. Es kann gezeigt werden [7], dass die optimale Ansatzverteilung auf multidimensionale, nicht berechenbare Integrale führt. Somit müssen andere Ansätze gefunden werden. Die Transitionsdichte

$$\pi(x_k|X_{k-1}, Y_k) \stackrel{\circ}{=} p(x_k|x_{k-1}) \quad (2.104)$$

ist die meist verbreitete Ansatzverteilung, da sie aufgrund der Nichtberücksichtigung der aktuellsten Messung einfach zu implementieren ist. Diese vereinfachte Implementierung kann gezeigt werden, indem man die Ansatzverteilung in Gleichung (2.100) einsetzt:

$$\begin{aligned} w_k &= w_{k-1} \frac{p(y_k|x_k) p(x_k|x_{k-1})}{\pi(x_k|X_{k-1}, Y_k)} \\ &= w_{k-1} \frac{p(y_k|x_k) p(x_k|x_{k-1})}{p(x_k|x_{k-1})} \\ &= w_{k-1} p(y_k|x_k). \end{aligned} \quad (2.105)$$

Wählt man also die Transitionsdichte als Ansatzverteilung, reduziert sich das Update der Gewichte auf die Auswertung der Likelihood-Funktion der Partikel und Multiplikation mit den Gewichten w_{k-1} . Die Transitionsdichte wird durch das dynamische Zustandsraummodell (DSSM) definiert. Legt man z.B. ein additives, normalverteiltes Systemrauschen mit Mittelwert $\mathbb{E}[\kappa_{k-1}]$ und Kovarianzmatrix \mathbf{P}^{κ} zugrunde, so ist die Transitionsdichte durch

$$p(x_k|x_{k-1}) = \mathcal{N}(x_k; \mathcal{F}(x_{k-1}, u_{k-1}, \mathbb{E}[\kappa_{k-1}]), \mathbf{P}^{\kappa}) \quad (2.106)$$

für diskrete Systemdynamik gegeben.

³ $A \stackrel{\circ}{=} B$ impliziert, dass B gewählt wurde, um A zu approximieren.

Obwohl die Wahl der Transitionsdichte als Ansatzverteilung zwar zu einer vereinfachten Implementierung führt, kann sie erhebliche Probleme verursachen. Da sie im Gegensatz zur optimalen Ansatzverteilung nicht abhängig von der aktuellsten Messung y_k ist, kann dies dazu führen, dass bei Auswertung des Likelihoods nur eine geringe Anzahl an Partikeln signifikante Gewichte besitzt. Abbildung 2.7 veranschaulicht den Zusammenhang.

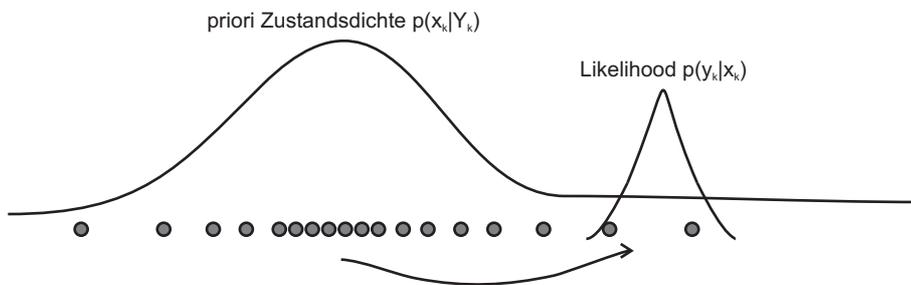


Abbildung 2.7: Verschiebung der Partikel der priori Zustandsdichte in Regionen hohen Likelihoods.

Es ist daher von enormer Wichtigkeit, die Partikel vorab in Regionen hohen Likelihoods zu verschieben. Dieses Problem tritt auch besonders dann auf, wenn die Likelihood-Verteilung sehr "gepeakt" /spitz im Vergleich zur Transitionsdichte ist. Dieser Effekt wird z.B. durch hochgenaue, gering rauschende Sensoren verursacht.

Degenerierung des SIS-Algorithmus

Der bis hierher vorgestellte SIS-Algorithmus hat einen gravierenden Nachteil: Die Varianz der Gewichte wächst stochastisch mit der Zeit [27]. Für Details und Beweise sei auf [7] verwiesen. Dieser Effekt hat zur Folge, dass in der Regel nach wenigen Iterationen eines der normalisierten Gewichte gegen eins konvergiert, während die restlichen Gewichte gegen null konvergieren. Das bedeutet, dass eine große Anzahl von Partikeln einfach vom gesamten Partikelsatz entfernt wird, da ihre Gewichtung gegen null konvergiert. Im nachfolgenden Abschnitt wird eine Strategie vorgestellt, wie dieser Effekt gemindert (aber nicht gänzlich beseitigt) werden kann.

Minderung der SIS-Degenerierung: Resampling

Der verfolgte Ansatz, den Degenerierungseffekt von Partikeln zu mindern, ist, Partikel mit geringen Gewichten auszusortieren und Partikel mit hohen Gewichten zu multiplizieren, wobei die Gesamtanzahl an Partikeln konstant gehalten wird. Ein brauchbarer Index für die Partikeldegenerierung ist die effektive Partikelanzahl N_{eff} , eingeführt in [30] und berechenbar durch:

$$N_{eff} = \frac{1}{\sum_{i=1}^N \left(\tilde{w}_k^{(i)}\right)^2}. \quad (2.107)$$

Es ist leicht zu zeigen, dass $1 \leq N_{eff} \leq N$, wobei $N_{eff} = N$, falls alle Gewichte uniform sind, d.h. $\tilde{w}_k^{(i)} = \frac{1}{N} \forall i$ und $N_{eff} = 1$, falls ein Gewicht gleich eins und alle anderen gleich null sind. Unterschreitet N_{eff} einen Grenzwert N_{thr} , dann wird ein Resample-Schritt nach folgenden Prinzipien durchgeführt: Ein Selektionsschema ordnet jedem Partikel $x_k^{(i)}$ eine Anzahl "Nachfahren" $N^{(i)}$ zu, wobei $\sum_{i=1}^N N^{(i)} = N$. Dieser Vorgang wird "Resampling" genannt. In der Literatur sind mehrere Selektionsalgorithmen zu finden, u.a. Sampling Importance Resampling (SIR) [31], Residual Resampling (RR) [28] und Minimum Variance Sampling (MVS) [27], welche nachfolgend kurz beschrieben werden sollen:

1. Sampling Importance Resampling

Beim SIR-Selektionsschema wird die diskrete Verteilung

$$\left\{ x_k^{(i)}, \tilde{w}_k^{(i)}; i = 1, \dots, N \right\} \quad (2.108)$$

in einer diskreten Verteilung mit Partikeln identischer Gewichtung

$$\left\{ x_k^{(j)}, \frac{1}{N}; j = 1, \dots, N \right\} \quad (2.109)$$

abgebildet. Dies geschieht, indem N neue Partikel identischer Gewichtung aus der diskreten Verteilung $\left\{ x_k^{(i)}; i = 1, \dots, N \right\}$ mit der Wahrscheinlichkeit $\left\{ \tilde{w}_k^{(i)}; i = 1, \dots, N \right\}$ generiert werden. Ein Verfahren zum Generieren von Partikeln aus diskreten Verteilungen kann z.B. [32] entnommen werden.

2. Residual Resampling

Residual Resampling ist ein Zwei-Schritt-Selektionsschema, wobei SIR integriert wird. Im ersten Schritt wird die Anzahl der "Nachfahren" $N^{(i)}$ durch Anwendung der Ganzzahl-Funktion/Abrundungsfunktion (floor function) deterministisch festgelegt:

$$N_a^{(i)} = \left\lfloor N \tilde{w}_k^{(i)} \right\rfloor. \quad (2.110)$$

Dies hat zur Folge, dass jedes Partikel $x_k^{(i)}$ $N_a^{(i)}$ -mal reproduziert wird. Im zweiten Schritt wird der SIR-Algorithmus verwendet, um die restlichen

$$\bar{N}_k = N - \sum_{i=1}^N N_a^{(i)} \quad (2.111)$$

Partikel zu generieren, wobei nun aus der diskreten Verteilung mit Partikeln $\left\{ x_k^{(i)}; i = 1, \dots, N \right\}$ und neuen Gewichten

$$\check{w}_k^{(i)} = \bar{N}_k^{-1} \left(\tilde{w}_k^{(i)} N - N_a^{(i)} \right) \quad (2.112)$$

die Partikel generiert werden. Diese Partikel beschreiben einen zweiten Satz an Partikeln $N_b^{(i)}$, wobei $\bar{N}_k = \sum_{i=1}^N N_b^{(i)}$ ist. Die Gesamtanzahl der Nachfahren eines Partikels ist dann durch $N^{(i)} = N_a^{(i)} + N_b^{(i)}$ gegeben. Residual Resampling ist weniger rechenaufwendig als SIR und resultiert in geringerer Varianz der Partikel, daher sollte RR in der Regel als Selektionsschema Verwendung finden.

3. Minimum Variance Sampling

Das MVS-Selektionsschema vereint das Stratified/Systematic Sampling, publiziert in [33] und den Tree Based Branching Algorithm, publiziert in [27]. Dabei werden zunächst N äquidistante Stützpunkte U im Intervall $[0, 1]$ mit Abstand $\frac{1}{N}$ gebildet. Die Anzahl an "Nachfahren" $N^{(i)}$ entspricht der Anzahl an Stützpunkten, die zwischen $\sum_{j=1}^{i-1} \tilde{w}_k^{(j)}$ und $\sum_{j=1}^i \tilde{w}_k^{(j)}$ liegen. Dieses Selektionsschema führt zu einer noch geringeren Varianz der "Nachfahren" als Residual Resampling, benötigt allerdings mehr Rechenkapazität.

Alle Selektionsschemata führen auf einen Satz von N Partikeln, ungefähr verteilt gemäß der posteriori Wahrscheinlichkeitsdichte. Es liegt in der Natur der Selektionsschemata, dass die Vervielfältigung der "fittesten" Partikel erfolgt, was u.U. dazu führen kann, dass einige Partikel keine "Nachfahren" besitzen ($N^{(i)} = 0$), wobei andere eine große Anzahl an "Nachfahren" haben. Im Extremfall wird nur ein Partikel N -mal reproduziert. In diesem Fall erfolgt trotz Resampling eine starke Degenerierung der Partikel.

Der gewöhnliche Partikel-Filter (Generic Particle Filter) (GPF)

Der Generic Particle Filter kann nun aus den vorgestellten Schritten Sequential Importance Sampling und Resampling zusammengesetzt werden. Ausgehend von einer Initialverteilung $p(x_0)$ wird ein Initialdatensatz an Partikeln generiert. Dieser wird sequentiell durch SIS und Resampling in der Zeit vorwärts propagiert. Schätzungen können über die empirische Näherung $\mathbb{E}[g(x_k)] \approx \frac{1}{N} \sum_{i=1}^N g(x_k^{(i)})$ gewonnen werden. Abbildung 2.8 veranschaulicht den Algorithmus des GPF.

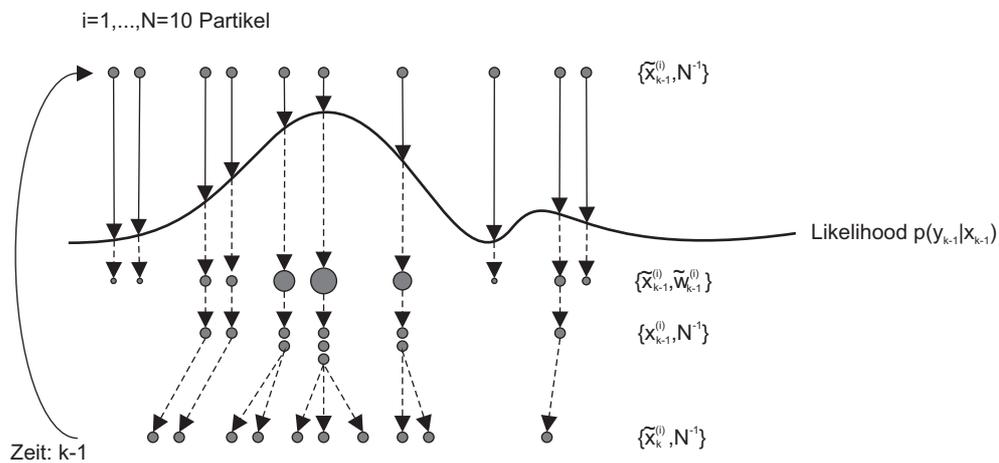


Abbildung 2.8: Veranschaulichung des GPF-Algorithmus. Ausgehend von einem Satz Partikel identischer Gewichtung nach der Generierung durch die Ansatzverteilung (Transitionsdichte) $\{\tilde{x}_{k-1}^{(i)}, N^{-1}\}$ wird durch Auswertung der Likelihood-Funktion $p(y_{k-1}|x_{k-1})$ der Partikelsatz gewichteter Partikel $\{\tilde{x}_{k-1}^{(i)}, \tilde{w}_{k-1}^{(i)}\}$ generiert. Durch den Resampling-Schritt entsteht der Partikelsatz $\{\tilde{x}_{k-1}^{(i)}, N^{-1}\}$. Dieser wird wiederum durch die Transitionsverteilung in der Zeit vorwärts propagiert zum Partikelsatz $\{\tilde{x}_k^{(i)}, N^{-1}\}$ und der Algorithmus beginnt von Neuem.

Die Güte des GPF hängt im Wesentlichen davon ab, ob die zwei Annahmen

- **Monte-Carlo-Annahme (MC-Annahme):** Die Dirac-Punktmassenapproximation liefert eine adäquate Repräsentation der wahren posteriori Verteilung.
- **Importance-Sampling-Annahme (IS-Annahme):** Es ist möglich, Partikel der posteriori Verteilung zu generieren, indem man von einer geeigneten Ansatzverteilung Partikel gewinnt und Gewichtungskorrekturen vornimmt.

gültig sind. Trifft eine dieser Annahmen nicht zu, liefert der GPF u.U. schlechte Schätzungen. Die approximierende Punktmassendiskretisierung führt dazu, dass die Auflösung zu gering ist. Durch den Resampling-Schritt wird u.U. ein Partikel viele Male reproduziert, was dazu führt, dass die repräsentative Wolke an Partikeln zu einem einzigen Partikel kollabiert. Als "brachiale" Methode, dieses Problem zu umgehen, kann die Anzahl der Partikel drastisch gesteigert werden, was allerdings in einem massiven und evtl. unnötigen Anstieg an Rechenzeit resultiert (und auch nicht immer Erfolg garantiert).

Nachfolgend werden Versionen des GPF vorgestellt, welche u.a. versuchen, die Problematik des Partikel-Kollaps zu beheben oder zumindest zu verringern. Die vorgestellten Methoden entsprechen dem aktuellen Stand der Technik auf dem Gebiet der Partikel-Filter. Dabei sei angemerkt, dass die nachfolgenden Versionen durchaus untereinander kombiniert werden können (und auch kombiniert werden).

Versionen von Partikel-Filtern: I. Auxiliary Particle Filter (ASIR)

Der Auxiliary Particle Filter (auch Auxiliary Sequential Importance Resampling genannt) wurde erstmals von PITT&SHEPHARD in [34] publiziert und von FEARNHEAD in [35] korrigiert. Für Details sei zusätzlich auf [36] verwiesen. Die grundlegende Idee des ASIR ist, den Resampling-Schritt zum Zeitpunkt $k - 1$ unter Berücksichtigung der aktuellsten Messung zum Zeitpunkt k durchzuführen. Ausgehend von einer Approximation $\mu_k^{(i)}$, welche den Zustand x_k bei gegebenen $x_{k-1}^{(i)}$ repräsentieren soll (z.B. $\mu_k^{(i)} = \mathbb{E}[x_k | x_{k-1}^{(i)}]$ oder $\mu_k^{(i)} \sim p(x_k | x_{k-1}^{(i)})$), werden die Partikel $x_{k-1}^{(i)}$ basierend auf der Likelihood-Funktion gewichtet und durch einen der vorgestellten Resample-Algorithmen "ge-resamplet". Die "ge-resampletten" Partikel zum Zeitpunkt $k - 1$ werden nun durch das stochastische Systemmodell in der Zeit vorwärts propagiert. Anschließend werden die generierten Partikel $x_k^{(i)}$ gemäß

$$w_k^{(i)} \propto \frac{p(y_k | x_k^{(i)})}{p(y_k | \mu_k^{(i)})} \quad (2.113)$$

neu gewichtet und der Algorithmus beginnt von Neuem.

Der ASIR hat gegenüber dem konventionellen GPF den Vorteil, dass Partikel aus der Partikelwolke zum Zeitpunkt $k - 1$ generiert werden, welche durch den Resampling-Schritt unter Verwendung der aktuellsten Messung zum Zeitpunkt k bereits in Regionen hohen Likelihoods verschoben sind. Bei geringem Systemrauschen ist der ASIR gegenüber dem GPF weniger anfällig gegenüber Ausreißen in der Partikelwolke. Ist das Systemrauschen hingegen groß, repräsentieren die prädierten Partikel $\mu_k^{(i)}$ die Verteilung

$p(x_k|x_{k-1}^{(i)})$ schlecht und ASIR führt den Resampling-Schritt folglich mit einer mangelhaften Repräsentation von $p(x_k|x_{k-1}^{(i)})$ durch. In diesem Fall kann die Verwendung des ASIR die Schätzungsgüte verschlechtern.

Versionen von Partikel-Filtern: II. Partikel-Filter mit verbesserter Partikelstreuung

Wie bereits zuvor erwähnt, kann der Resampling-Schritt im SIR zum Partikel-Kollaps führen. Nachfolgend werden zwei Methoden vorgestellt, welche das Problem des Partikel-Kollaps durch einen veränderten Resampling-Schritt adressieren. Die erste Methode ist der Regularization Resampling Step (RRS), die zweite der Markov Chain Monte Carlo Step (MCMC).

Die grundlegende Idee des RRS (RRS in Kombination mit dem GPF führt zum Regularized Particle Filter (RPF)) ist, die Filterdichte $p(x_k|Y_k)$, welche als diskrete Verteilung in Form einer gewichteten Partikelwolke $\{x_k^{(i)}, \tilde{w}_k^{(i)}\}_{i=1}^N$ vorliegt, durch Kernel-Schätzung durch eine kontinuierliche zu approximieren und von dieser kontinuierlichen Verteilung ein neues Partikelsystem zu generieren. Dies garantiert, dass die Partikelwolke nicht zu einem einzigen Partikel kollabiert.

Liegt das oben genannte Partikelsystem vor, so kann die Kernel Density Estimation (KDE) der Filterdichte durch

$$\hat{p}(x_k|Y_k) \approx \sum_{i=1}^N K_\lambda(x_k - x_k^{(i)}) \tilde{w}_k^{(i)} \quad (2.114)$$

formuliert werden, wobei K_λ der Kern (Kernel) skaliert durch die Kernel-Bandbreite λ gegeben ist durch

$$K_\lambda = \frac{1}{N\lambda^n} K\left(\frac{x_k - x_k^{(i)}}{\lambda}\right). \quad (2.115)$$

Der Kern K ist eine symmetrische Wahrscheinlichkeitsverteilung im \mathbb{R}^n mit den Eigenschaften:

- $K \geq 0$
- $\int K(x)dx = 1$
- $\int xK(x)dx = 0$
- $\int \|x\|^2 K(x)dx < \infty$

Der Kern und die Bandbreite werden so gewählt, dass der Mean Integrated Squared Error (MISE) $MISE = \int_{\mathbb{R}^n} \mathbb{E}[(\hat{p}(x_k|Y_k) - p(x_k|Y_k))^2] dx_k$ zwischen der Filterdichte und der KDE minimiert wird. Am einfachsten, speziell im mehrdimensionalen Raum, ist es, einen GAUSS-Kern mit Einheitskovarianzmatrix zu verwenden⁴. Dazu wird jedes Partikel $x_k^{(i)}$ transformiert zu $(\mathbf{S}^x)^{-1}x_k^{(i)}$, wobei $\mathbf{S}(\mathbf{S}^x)^T = \mathbf{P}_{dis}^x$ und \mathbf{P}_{dis}^x die empirische

⁴Es existieren auch andere Kerne, wie z.B. der optimale EPANECHNIKOV-Kern. Die Verwendung des EPANECHNIKOV-Kerns führt allerdings auf einen erhöhten Rechenaufwand mit nur geringfügiger Verbesserung der KDE. Für weitergehende Informationen sei auf [27] und [36] verwiesen.

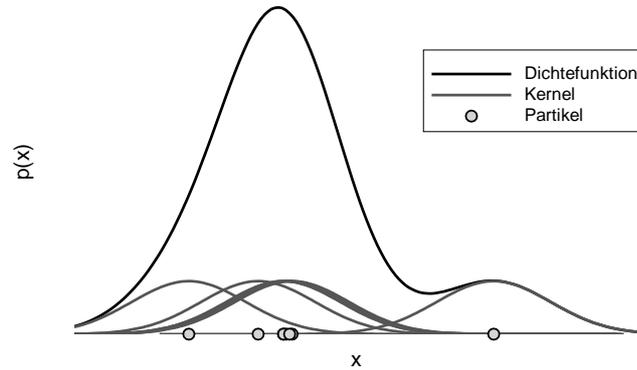


Abbildung 2.9: Kernel Density Estimation mit GAUSS-Kern.

Kovarianzmatrix des diskreten Partikelsatzes ist. Die optimale Bandbreite ist dann gegeben durch [27]:

$$\lambda_{Opt} = \left(\frac{4}{(n+2)N} \right)^{\frac{1}{n+4}}. \quad (2.116)$$

Beim RRS wird somit im Anschluss an den Resampling-Schritt aus der KDE ein neuer Partikelsatz gewonnen. Dazu wird lediglich die empirische Kovarianz des Partikelsatzes $\{x_k^{(i)}, \tilde{w}_k^{(i)}\}_{i=1}^N$ berechnet und jedem Index i ein neues Partikel $x_k^{(i)*} = x_k^{(i)} + \lambda_{Opt} \mathbf{S}^x \epsilon^{(i)}$ zugeordnet, wobei $\epsilon^{(i)} \sim \mathcal{N}(\epsilon; 0, \mathbf{I}_n)$. In der Praxis, wo die Verteilungen oft multimodal sind, wählt man $\lambda_{Opt} = \frac{\lambda_{Opt}}{2}$.

Die theoretischen Grundlagen des MCMC sind in [37] und [38] erläutert. Die Kernidee des MCMC ist, dass ein Partikel $x_k^{(i)}$ nur dann einem neuen Zustand $x_k^{(i)*}$ zugeordnet wird, wenn $\vartheta \leq \alpha$, wobei $\vartheta \sim \mathcal{U}(\vartheta; 0, 1)$. Dabei ist α die so genannte Akzeptanzwahrscheinlichkeit, wobei $\alpha = \min\{1, \mathcal{R}\}$. MCMC schließt sich immer an den Resampling-Schritt an, also ausgehend von der Situation, in der

- N Partikel $x_k^{(i)}$ aus dem Resampling-Schritt und $x_{k-1}^{(i)}$ die zu den "ge-resamplen" zugehörigen Partikel zum Zeitpunkt $k-1$ und
- N Partikel $x_k^{(i)*} \sim \pi(x_k | x_{k-1}^{(i)}, y_k)$ von einer Ansatzverteilung π

generiert wurden, dann ist das METROPOLIS-HASTINGS-Verhältnis \mathcal{R} gegeben durch

$$\mathcal{R} = \frac{p(y_k | x_k^{(i)*}) p(x_k^{(i)*} | x_{k-1}^{(i)}) \pi(x_k^{(i)} | x_{k-1}^{(i)}, y_k)}{p(y_k | x_k^{(i)}) p(x_k^{(i)} | x_{k-1}^{(i)}) \pi(x_k^{(i)*} | x_{k-1}^{(i)}, y_k)}. \quad (2.117)$$

Wählt man beispielsweise als Ansatzverteilung π die Transitionsdichte $p(x_k | x_{k-1})$, dann müssen lediglich die Likelihood-Funktionen zur Berechnung von $\mathcal{R} = \frac{p(y_k | x_k^{(i)*})}{p(y_k | x_k^{(i)})}$ ausgewertet werden.

Der RPF und der GPF mit MCMC-Schritt liefern i.d.R. eine verbesserte Schätzgüte als der konventionelle GPF, besonders dann, wenn das System- und Messrauschen klein und somit der Partikel-Kollaps schwerwiegend ist. Die zusätzlichen Schritte gehen naturgemäß auf Kosten der Rechenzeit.

Versionen von Partikel-Filtern: III. Kernel Particle Filter (KPF)

Der Kernel Particle Filter [39] sieht eine Verschiebung der Partikel in Regionen dominanter Moden der KDE durch Anwendung von Mean Shift Mode Seeking [40, 41, 42] vor. Bei gegebener KDE (2.114) wird der Gradient der Verteilung geschätzt und die Partikel entlang des Gradienten in Richtung der Moden der Verteilung verschoben.

Ausgehend von einer gewichteten Partikelwolke $\{x_k^{(i)}, \tilde{w}_k^{(i)}\}_{i=1}^N$ wird im Anschluss an den Resampling-Schritt zunächst ein RRS durchgeführt, d.h. jedem Partikel $x_k^{(i)}$ wird ein neues Partikel gemäß $x_k^{(i)*} = x_k^{(i)} + \lambda_{Opt} \mathbf{S}^{x_{\epsilon}^{(i)}}$ zugeordnet. Die Moden der KDE entsprechen den Nullstellen des KDE Gradienten $\nabla p(x_k | Y_k) = 0$. Mean Shift Mode Seeking ist eine elegante Methode, diese Moden zu lokalisieren. Jedes Partikel $x_k^{(i)}$ wird nun in Richtung der Moden verschoben durch

$$m(x_k^{(i)}) = \frac{\sum_{j=1}^N K_\lambda(x_k^{(i)} - x_k^{(j)}) \tilde{w}_k^{(j)} x_k^{(j)}}{\sum_{j=1}^N K_\lambda(x_k^{(i)} - x_k^{(j)}) \tilde{w}_k^{(j)}}. \quad (2.118)$$

Im Anschluss an die Verschiebung müssen die Partikel neu gewichtet werden [39]. $\{x_k^{(\cdot),(j)}\}$ sei die Partikelwolke nach der j_{ten} Mean-Shift-Prozedur zum Zeitpunkt k , dann werden die Partikel gemäß

$$w_k^{(i),(j)} = \frac{p(x_k^{(i),(j)} | Y_k)}{\sum_{g=1}^N K_\lambda(x_k^{(i),(j)} - x_k^{(g),(j)})} \quad (2.119)$$

mit

$$p(x_k^{(i),(j)} | Y_k) \propto p(y_k | x_k^{(i),(j)}) \sum_{g=1}^N p(x_k^{(i),(j)} | x_{k-1}^{(g),(j)}) w_{k-1}^{(g)} \quad (2.120)$$

neu gewichtet und die Mean-Shift-Iteration beginnt von Neuem. Von CHANG werden in [39] zwei bis fünf Iterationen für drei- bis neun-dimensionale Zustandsräume vorgeschlagen.

In [43] wird ein Partikel-Filter-Hybrid aus dem ASIR und dem KPF vorgestellt, der Auxiliary Kernel Particle Filter (AKPF), welcher die Vorteile beider Verfahren kombiniert und (in dem vorgestellten Beispiel) zu einer verbesserten Schätzgüte führt.

Versionen von Partikel-Filtern: IV. Adaptive Particle Filter (APF)

Typischerweise benötigt der GPF während der Initialphase, wo größte Unsicherheiten auftreten, mehr Partikel als in der folgenden Schätzphase. Die Initialverteilung ist in der Regel breit und multimodal, wobei im späteren Schätzprozess, nachdem eine Vielzahl von Observationen verwertet wurde, die posteriori Verteilung stark "gepeakt" um wenige Likelihood-Hypothesen wird. FOX [44, 45] hat in seinen Arbeiten dieses Problem adressiert und einen adaptiven Partikel-Filter entwickelt. Die Kernidee des APF besteht darin, den Approximationsfehler der wahren posteriori Wahrscheinlichkeitsdichte durch die Partikelwolke mit einem Grenzwert zu belegen. Als Maß für den Fehler wird die Kullback-Leibler-Distanz [46], [47] verwendet,

wodurch zu jedem Zeitschritt k die Anzahl der Partikel so adaptiert werden kann, dass ein spezifizierter Grenzwert der KLD eingehalten wird. Allerdings wird auch gezeigt (u.a. auch in [48]), dass aufgrund des Mehraufwandes der KLD-Berechnung nicht zwangsläufig die Rechenanforderung reduziert wird. Aus diesem Grund werden die adaptiven Partikel-Filter im weiteren Verlauf dieser Arbeit nicht weiter berücksichtigt.

Versionen von Partikel-Filtern: V. Gaussian Particle Filter (EKPF, SPPF)

Auf die Degenerierung der Partikel wirkt sich am stärksten die Verschiebung in Regionen hohen Likelihoods, also eine ungenaue Ansatzverteilung, aus. Obwohl die Wahl der Transitionsdichte als Ansatzverteilung zwar die IS-Annahme erfüllt, erfüllt sie diese allerdings auf Kosten der Nichtberücksichtigung der aktuellsten Observation, was im Gegenzug zu einer raschen Degenerierung der Partikel und somit auf die Verletzung der MC-Annahme führen kann. Um Partikel-Filter zu verbessern, müssen also bessere Ansatzverteilungen gefunden werden, die die aktuellste Messung berücksichtigen.

Ein effektiver Ansatz, dies zu ermöglichen, ist eine durch einen GAUSS'schen Inferenzalgorithmus (EKF, IEKF, UKF, CDKF, GHF, etc.) generierte Verteilung der posteriori Wahrscheinlichkeitsdichte eines jeden Partikels zu verwenden

$$\pi(x_k|X_{k-1}, Y_k) \stackrel{\circ}{=} p(x_k|x_{k-1}, y_k) \tag{2.121}$$

$$\approx q_{\mathcal{N}}(x_k|Y_k) \tag{2.122}$$

$$= \mathcal{N}\left(x_k; \bar{x}_k^{(i)}, \mathbf{P}_{k|k}^{x,(i)}\right), \quad i = 1, \dots, N \tag{2.123}$$

woraus aus jeder dieser posteriori Verteilungen die Ansatzverteilung generiert wird. So wird zum Zeitpunkt k aus jedem Partikel $x_{k-1}^{(i)}$ der Mittelwert und die Kovarianz durch einen EKF, IEKF, UKF, ... berechnet. Daraus werden jeweils die Partikel $x_k^{(i)}$ generiert, wobei die Gewichte nach der bereits vorgestellten Gleichung (2.100) bestimmt werden.

So entsteht eine Vielzahl von hybriden Algorithmen, die über eine bessere Schätzgüte als der gewöhnliche Partikel-Filter verfügen. Allerdings kann die Verwendung der GAUSS'schen Inferenzalgorithmen für jedes Partikel schnell extrem hohe Rechenanforderungen stellen. In der Regel wählt man einen SPKF als GAUSS'schen Inferenzalgorithmus und erhält so die Familie der Sigma-Point Particle Filter (SPPF).

Versionen von Partikel-Filtern: VI. Gaussian Sum Particle Filter (GSPF)

In [49] wird von KOTECHEA&DJURIC der Gaussian Sum Particle Filter vorgeschlagen. Ausgehend von einer GMM-Approximation des Zustandes und des Systemrauschens

$$p_{GMM}(x_{k-1}|Y_{k-1}) = \sum_{g=1}^G \alpha_{k-1}^{(g)} \mathcal{N}\left(x_{k-1}; \bar{x}_{k-1}^{(g)}, \mathbf{P}_{k-1}^{x,(g)}\right) \tag{2.124}$$

$$p_{GMM}(\kappa_{k-1}) = \sum_{m=1}^M \beta_{k-1}^{(m)} \mathcal{N}\left(\kappa_{k-1}; \bar{\kappa}_{k-1}^{(m)}, \mathbf{P}_{k-1}^{\kappa,(m)}\right) \tag{2.125}$$

wird durch Sampling und Propagation durch die nichtlineare Systemfunktion \mathcal{F}

$$x_{k-1}^{(g),(i)} \sim \mathcal{N}\left(x_{k-1}; \bar{x}_{k-1}^{(g)}, \mathbf{P}_{k-1}^{x,(g)}\right) \quad (2.126)$$

$$\kappa_{k-1}^{(m),(i)} \sim \mathcal{N}\left(\kappa_{k-1}; \bar{\kappa}_{k-1}^{(m)}, \mathbf{P}_{k-1}^{\kappa,(m)}\right) \quad (2.127)$$

$$x_{k|k-1}^{(gm),(i)} = \mathcal{F}\left(x_{k-1}^{(g),(i)}, u_{k-1}, \kappa_{k-1}^{(m),(i)}\right) \quad (2.128)$$

sowie Gewichtung durch $w^{(gm)} = \alpha_{k-1}^{(g)}\beta_{k-1}^{(m)} / \sum_{gm=1}^{GM} \alpha_{k-1}^{(g)}\beta_{k-1}^{(m)}$ ein GMM-Modell der priori Verteilung

$$p_{\mathcal{GMM}}(x_k | Y_{k-1}) = \sum_{gm=1}^{GM} w^{(gm)} \mathcal{N}\left(x_k; \bar{x}_{k|k-1}^{(gm)}, \mathbf{P}_{k|k-1}^{x,(gm)}\right) \quad (2.129)$$

mit

$$\bar{x}_{k|k-1}^{(gm)} = \frac{1}{N} \sum_{i=1}^N x_{k|k-1}^{(gm),(i)} \quad (2.130)$$

$$\mathbf{P}_{k|k-1}^{x,(gm)} = \frac{1}{N} \sum_{i=1}^N (x_{k|k-1}^{(gm),(i)} - \bar{x}_{k|k-1}^{(gm)})(x_{k|k-1}^{(gm),(i)} - \bar{x}_{k|k-1}^{(gm)})^T \quad (2.131)$$

gewonnen. Von dieser Ansatzverteilung werden wiederum Partikel generiert und durch die aktuellste Messung gewichtet, wodurch eine diskrete Punktmassenapproximation der posteriori Verteilung $p(x_k | Y_k)$ gewonnen wird, welche eine Zustandsschätzung (z.B. MMSE) zulässt. Die Rückgewinnung des GMM-Modells mit G Komponenten erfolgt durch die in [49] vorgeschlagene Resampling-Methode.

Versionen von Partikel-Filtern: VII. Gaussian Mixture Sigma-Point Particle Filter (GMSPPF)

Im Gaussian Mixture Sigma-Point Particle Filter wird die GMM-Approximation $p_{GMM}(x_k | Y_k)$ generiert durch GAUSS'sche Summenfilter (Sigma-Point Kalman Filter als Filterbank) als repräsentative Ansatzverteilung gewählt:

$$\pi(x_k | X_{k-1}, Y_k) \stackrel{\circ}{=} p_{GMM}(x_k | Y_k). \quad (2.132)$$

Die Partikel werden auch hier aus $p_{GMM}(x_k | Y_k)$ generiert, wobei die Gewichte wiederum nach Gleichung (2.100) bestimmt werden. Somit erhält man eine gewichtete Punktmassenapproximation der posteriori Verteilung. Die GMSPPF repräsentieren die posteriori Verteilung allerdings naturgemäß durch ein GMM-Modell. Somit muss aus der gewichteten Punktmassenapproximation wieder ein GMM-Modell erstellt werden. Dies kann auf zwei Wegen erfolgen:

1. *Resampling und Expectation Maximization Clustering*

Hierbei werden die gewichteten Partikel zunächst einem Resampling-Schritt unterzogen und anschließend wird ein Standard EM-Algorithmus verwendet, um das GMM-Modell zu generieren.

2. *Weighted Expectation Maximization Clustering*

Hierbei wird das GMM-Modell direkt aus der gewichteten Punktmassenapproximation der posteriori Verteilung gewonnen [50]. Da das Resampling auch zur Degenerierung von Partikeln beitragen kann, sollte das WEM-Verfahren bevorzugt werden. Allerdings ist grundlegend vom Inferenzproblem abhängig, welcher Algorithmus zu wählen ist.

Die Abbildung 2.10 zeigt schematisch, wie der Algorithmus der GMSPPF funktioniert. Nähere Details dazu

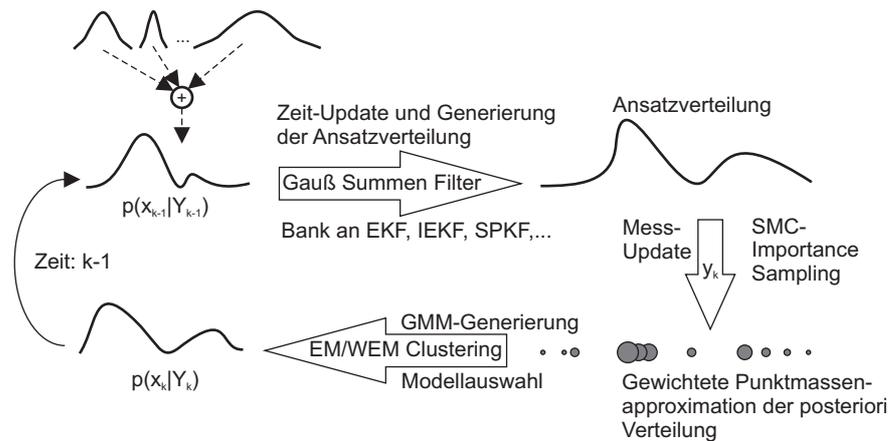


Abbildung 2.10: Gaussian Mixture Sigma-Point Particle Filter.

sind in [51] zu finden.

Sicherlich existieren noch weitere interessante Variationen von Partikel-Filtern, wie z.B. der Rejection Particle Filter (TANIZAKI) oder der Rao-Blackwellization Particle Filter [27]. Im Rahmen dieser Arbeit sollen diese allerdings nicht näher berücksichtigt werden und es wird sich auf die bisher vorgestellten Versionen beschränkt. Sie seien nur der Vollständigkeit halber genannt.

Schlussfolgerung

Alle hier aufgeführten approximierenden Methoden vereinfachen die RBS-Integrale in irgendeiner Form, z.B. bezüglich der Form der Wahrscheinlichkeitsdichten oder der Systemstruktur. Dies geschieht, um praktisch implementierbare Algorithmen zu erhalten. Die Genauigkeit und Gültigkeit der Vereinfachungen sind für jedes WI-Problem verschieden, wodurch gewisse Näherungslösungen immer attraktiver als andere sein werden. Wie bereits in der Einleitung erläutert, ist die Lösung dieses Filterauswahl-Problems oder zumindest eine Unterstützung Ziel dieser Arbeit.

Ein Teil der Algorithmen kann a priori vernachlässigt werden, da praktische Anwendungen nicht realisierbar sind und somit der Filter uninteressant für den Anwender und die Evaluierung wird. Aufgrund dessen werden in dieser Arbeit nur folgende viel versprechende Näherungslösungen untersucht:

- Methoden mit GAUSS'scher Approximation
- Methoden mit GAUSS'scher Summenapproximation
- Sequentielle-Monte-Carlo-Methoden
- Hybride Varianten aus den obigen Verfahren

2.3 Kapitelzusammenfassung

In diesem Kapitel wurde zunächst die theoretische Lösung des wahrscheinlichkeitstheoretischen Inferenzproblems gezeigt. Diese ist durch die rekursive BAYES'sche Schätztheorie gegeben. Allerdings ist die RBS nur für lineare GAUSS'sche Systeme exakt lösbar.

Daher wurde vorgestellt, wie die RBS näherungsweise auch für nichtlineare Systeme gelöst werden kann. Dabei wurden zunächst die Lösungen für GAUSS'sche Systeme (alle Dichten in der RBS sind normalverteilt) vorgestellt. Approximiert man die nichtlinearen Systemfunktionen durch Taylor-Reihen erster Ordnung und wendet auf die linearisierten Systeme die Gleichungen des Kalman Filters an, so erhält man die Lokal-Methoden mit dem Extended Kalman Filter als Basis und verschiedene Variationen des EKF, wie z.B. iterierende Verfahren IEKF. Approximiert man hingegen nicht die Systemgleichungen, sondern approximiert man die Zufallsverteilungen durch diskrete Summen, so erhält man die Quadratur-Methoden, bzw. die Sigma-Point Kalman Filter. Diese haben gegenüber den Lokal-Methoden den Vorteil, dass keine fortlaufende Taylor-Linearisierung durchgeführt werden muss.

Für den allgemeinsten Fall der nichtlinearen, nicht GAUSS'schen Systeme wurden ebenfalls Näherungslösungen vorgestellt. Nicht GAUSS'sche Verteilungen können durch eine Summe von gewichteten GAUSS'schen Verteilungen approximiert werden. Aus dieser Annahme entstehen die Methoden mit GAUSS'scher Summenapproximation. Als weit verbreitete Methode, nichtlineare, nicht GAUSS'sche Systeme zu behandeln, wurden die Sequentiellen-Monte-Carlo-Methoden eingeführt, die die RBS-Integrale durch diskrete Summen über ein adaptives Gitternetz, konzentriert an Orten hohen Likelihoods des Zustandsraums, approximieren. Des Weiteren wurden hybride Verfahren dargestellt, welche versuchen, die Vorteile der Hybrid-Komponenten zu vereinen.

In gewisser Weise stellen die Sigma-Point Kalman Filter auch SMC-Methoden dar, da sie die Integrale der RBS durch diskrete Summen approximieren. Allerdings besteht der gravierende Unterschied darin, dass die Auswahl der Beispielpunkte (Partikel) bei den SPKF deterministisch und nicht zufällig wie bei den SMC-Methoden erfolgt.

3 Verschiedene Schätzprobleme, Variationen der Algorithmen-Implementierung und spezielle Systeme

3.1 Verschiedene Schätzprobleme

Die Art des rekursiven Schätzproblems hängt vom untersuchten System und vom geforderten Inferenzziel ab. Grundsätzlich unterscheidet man zwischen der Zustands-, Parameter- und dualen Schätzung.

Zustandsschätzung

Die Zustandsschätzung ist die klassische Aufgabe nichtlinearer Inferenzverfahren. Das System wird beschrieben durch ein nichtlineares diskretes Zustandsraummodell, siehe (2.7-2.8). Dabei sind die x_k die unbeobachteten Zustände des Systems, u_k der bekannte Eingangsvektor und y_k der Vektor der bekannten Messungen. Das Systemrauschen κ treibt das dynamische System, und das Messrauschen γ korrumpiert den Messprozess. Die Systemdynamik und die Messgleichung seien beschrieben und bekannt durch \mathcal{F} und \mathcal{H} und seien parametrisiert durch einen unveränderlichen und bekannten Parametersatz.

Parameterschätzung zeit(in)varianter Parameter

Zeitinvariante Parameter

Die Parameterschätzung zeitinvarianter Parameter, oft auch bezeichnet als Parameteridentifikation, beschreibt die Aufgabe, ein nichtlineares, multidimensionales Kennfeld

$$y_k = \mathcal{KF}(\theta, [u_k \tilde{x}_k]) \quad (3.1)$$

zu bestimmen, wobei \tilde{x}_k den **bekanntem** Systemzustand, u_k den Systemeingangsvektor und y_k den Systemausgangsvektor beschreibt. Das nichtlineare Kennfeld $\mathcal{KF}(\cdot)$ wird durch den Vektor θ parametrisiert. Das nichtlineare Kennfeld kann z.B. ein künstlich neuronales Netz (KNN) (θ setzt sich aus den Gewichten und Bias zusammen) oder ein sonstiges statisches Systemmodell darstellen. Die Parameteridentifikation beschreibt die Aufgabe, den Parametervektor θ auf eine "optimale" Art und Weise zu schätzen. Typischerweise wird ein Trainings-Datensatz mit Betriebspunkten $\{\tilde{X}_k, U_k, Y_k\}$ bereitgestellt. Die zuvor vorgestellten Inferenzalgorithmen können zur Lösung dieses Schätzproblems verwendet werden, indem man eine

Zustandsraumdarstellung der Art

$$\theta_k = \theta_{k-1} = \mathcal{F}(x_{k-1} = \theta_{k-1}, u_{k-1} = 0, \kappa_{k-1} = 0) \quad (3.2)$$

$$y_k = \mathcal{KF}(\theta_k, [u_k \tilde{x}_k]) + \gamma_k = \mathcal{H}(x_k = \theta_k, u_k, \gamma_k) \quad (3.3)$$

wählt, wobei die Parameter θ_k durch einen stationären Prozess mit Einheitsmatrix \mathbf{I} als Zustandstransitionsmatrix Φ beschrieben werden. Dabei bleibt der prädizierte Erwartungswert konstant und der Parameter erhält lediglich während des Filterungsschritts eine Adaption. Die Ausgangswerte y_k entsprechen einer nichtlinearen Beobachtung parametrisiert durch θ_k . Die Systemzustände werden als bekannt vorausgesetzt und können den exogenen Eingängen zugeordnet werden.

Zeitvariante Parameter

In manchen Anwendungen können die Parameter nicht als zeitinvariant angesehen werden, da sie z.B. durch Alterungsprozesse mit der Zeit driften können. Um adaptionsfähige Algorithmen zu erhalten, muss eine Zustandsraumdarstellung der Form

$$\theta_k = \theta_{k-1} + \tilde{\kappa}_{k-1} = \mathcal{F}(x_{k-1} = \theta_{k-1}, u_{k-1} = 0, \kappa_{k-1} = \tilde{\kappa}_{k-1}) \quad (3.4)$$

$$y_k = \mathcal{KF}(\theta_k, [u_k \tilde{x}_k]) + \gamma_k = \mathcal{H}(x_k = \theta_k, u_k, \gamma_k) \quad (3.5)$$

gewählt werden. Die Zustandstransition des Parametervektors wird in diesem Fall durch einen stochastischen Prozess der Innovationsschrittweite $\tilde{\kappa}$ beschrieben. Dies hat zum Vorteil, dass der Parameter auch während der Prädiktionsphase einen Zuwachs erhält und der Filter ständig den Parameter updaten muss. Man erhält einen gegen Parameterschwankungen sensibleren Filter.

Die bestmögliche Wahl des "künstlichen" Systemrauschens $\tilde{\kappa}$ zur Parameterschätzung ist unterschiedlich für jede Inferenzaufgabe und ist immer noch ein Gebiet intensiver Forschung. In [7] werden neben der Wahl eines konstanten Systemrauschens $\tilde{\kappa}$ Methoden vorgestellt, welche das Systemrauschen On-Line adaptiv justieren. Als wichtigste Methoden seien λ -Decay-Method und Robbins-Monro genannt, wobei für Details auf [52, 53] (λ -Decay-Method) und [54, 55] (Robbins-Monro) hingewiesen sei. Die später vorgestellte MATLAB[®] Toolbox REBEL-IoN ermöglicht die komfortable Auswahl einer der Adaptionmethoden.

Das duale Schätzproblem

In vielen technischen Anwendungen, speziell in der modellbasierten Diagnose, ist es notwendig, Zustände und Parameter gleichzeitig zu schätzen. Die Modellparameter sind somit zeitlich veränderliche Größen mit (in der Regel) multiplikativem Eintrag auf die dynamischen Zustände:

$$x_k = \theta_{k-1} x_{k-1}. \quad (3.6)$$

Somit wird auch aus nahezu jedem linearen Schätzproblem ein nichtlineares Schätzproblem. Besteht die Schätzaufgabe darin, die Zustände und Parameter eines Systems gleichzeitig zu schätzen, so spricht man

von gemeinsamer Zustands- und Parameterschätzung (Duales Schätzproblem). Diese Aufgabe kann durch die in Kapitel 2 vorgestellten Näherungslösungen der RBS ebenso gelöst werden. Dazu muss lediglich der Zustandsvektor des Systems um den Parametervektor erweitert werden. Man erhält so den erweiterten Zustandsvektor (augmented state vector):

$$x_{aug} = [x \ \theta]^T. \quad (3.7)$$

Je nach Innovationsansatz für die Parameter (s.o.) erhält man so die erweiterte Zustandsraumdarstellung eines stochastischen DSSM. Der erweiterte Zustandsvektor wird für die Schätzung wie bei der einfachen Zustandsschätzung behandelt.

Die Näherungslösungen der RBS können somit auch für die gemeinsame Zustands- und Parameterschätzung verwendet werden. Der duale Ansatz verknüpft den Zustands- und Parametervektor zu dem erweiterten Zustandsvektor x_{aug} . Dadurch wird z.B. die Kreuzkovarianz zwischen Zustands- und Parameterschätzung effektiv modelliert:

$$\mathbb{E} \left[(\hat{x}_{aug,k} - \mathbb{E}[\hat{x}_{aug,k}]) (\hat{x}_{aug,k} - \mathbb{E}[\hat{x}_{aug,k}])^T \right] = \begin{bmatrix} \mathbf{P}_k^x & \mathbf{P}_k^{x\theta} \\ \mathbf{P}_k^{\theta x} & \mathbf{P}_k^\theta \end{bmatrix}. \quad (3.8)$$

Die Erweiterung des Zustandsvektors erlaubt es dem dualen Ansatz, nicht nur die Unsicherheiten von Zustands- und Parameterschätzung genau zu behandeln, sondern auch die Korrelation zwischen Zustands- und Parameterschätzung effektiv zu beschreiben.

3.2 Variationen der Algorithmen-Implementierung

Additives System- und Messrauschen

Für den speziellen, aber oft vorzufindenden Fall, dass das System- und Messrauschen rein additiv ist, reduziert sich die Rechen- und Implementierungskomplexität aller Algorithmen. Die vereinfachten Zustandsraummodelle sind nun durch

$$x_k = f(x_{k-1}, u_{k-1}) + \kappa_{k-1} = \mathcal{F}(x_{k-1}, u_{k-1}, \kappa_{k-1}) \quad (3.9)$$

$$y_k = h(x_k, u_k) + \gamma_k = \mathcal{H}(x_k, u_k, \gamma_k) \quad (3.10)$$

für diskrete Systeme gegeben.

Square-Root-Implementierung (SR-Implementierung) von Gauß-Filtern

Aufgrund von Rundungsfehlern (z.B. Festkommazahl/Gleitkommazahl-Prozessoren) kann es bei Gauß-Filtern zu dem Effekt kommen, dass die Kovarianzmatrix \mathbf{P}^x ihre positive Definitheit verliert, was zur sicheren Divergenz des Filters führt. Um diesen Effekt zu umgehen, wird bei der SR-Implementierung nicht

\mathbf{P}^x , sondern die Cholesky-Zerlegung \mathbf{S}^x propagiert, wobei

$$\mathbf{P}^x = \mathbf{S}^x (\mathbf{S}^x)^T \quad (3.11)$$

$$\mathbf{P}^\kappa = \mathbf{S}^\kappa (\mathbf{S}^\kappa)^T \quad (3.12)$$

sind. Dadurch wird die positive Definitheit der Kovarianzmatrix garantiert.

Lokal-Methoden

Nachfolgend werden die Beweise für die Gleichungen der SR-Implementierung von Lokal-Methoden am Beispiel des EKF für diskrete Systeme gezeigt:

Theorem 1. Die Cholesky-Faktorisierung der prädizierten (priori) Kovarianzmatrix $\mathbf{P}_{k|k-1}^x$ sei durch

$$\mathbf{S}_{k|k-1}^x = \begin{bmatrix} \mathcal{J}_{\mathcal{F}}^x \mathbf{S}_{k|k-1}^x & K \mathbf{S}^\kappa \end{bmatrix} \quad (3.13)$$

gegeben!

Beweis 1.

$$\begin{aligned} \mathbf{P}_{k|k-1}^x &= \mathbf{S}_{k|k-1}^x (\mathbf{S}_{k|k-1}^x)^T \\ &= \begin{bmatrix} \mathcal{J}_{\mathcal{F}}^x \mathbf{S}_{k|k-1}^x & K \mathbf{S}^\kappa \end{bmatrix} \cdot \begin{bmatrix} (\mathbf{S}_{k|k-1}^x)^T (\mathcal{J}_{\mathcal{F}}^x)^T \\ (\mathbf{S}^\kappa)^T K^T \end{bmatrix} \\ &= \mathcal{J}_{\mathcal{F}}^x \mathbf{S}_{k|k-1}^x (\mathbf{S}_{k|k-1}^x)^T (\mathcal{J}_{\mathcal{F}}^x)^T + K \mathbf{S}^\kappa (\mathbf{S}^\kappa)^T K^T \\ &= \mathcal{J}_{\mathcal{F}}^x \mathbf{P}_{k-1|k-1}^x (\mathcal{J}_{\mathcal{F}}^x)^T + K \mathbf{P}^\kappa K^T \end{aligned} \quad (3.14)$$

q.e.d.

Die o.a. Cholesky-Faktorisierung $\mathbf{S}_{k|k-1}^x$ ist nicht quadratisch. Für den Filterungsschritt muss diese in eine quadratische Form gebracht werden. Dies kann z.B. durch QR-Zerlegung (siehe Seite 45) erreicht werden. Dann kann ebenso die Kovarianzmatrix für den Filterungsschritt in SR-Form berechnet werden.

Theorem 2. Die Cholesky-Faktorisierung der gefilterten (posteriori) Kovarianzmatrix $\mathbf{P}_{k|k}^x$ sei durch

$$\mathbf{S}_{k|k}^x = \mathbf{S}_{k|k-1}^x \cdot \left(\mathbf{I} - \mathfrak{z} (\mathfrak{U}^T)^{-1} (\mathfrak{U} + \mathfrak{M})^{-1} \mathfrak{z}^T \right) \quad (3.15)$$

mit den Definitionen

$$\mathfrak{z} = (\mathbf{S}_{k|k-1}^x)^T (\mathcal{J}_{\mathcal{H}}^x)^T \quad (3.16)$$

$$\mathfrak{U} \mathfrak{U}^T = \Gamma \mathbf{P}^\gamma \Gamma^T + \mathfrak{z}^T \mathfrak{z} \quad (3.17)$$

$$\mathfrak{M} \mathfrak{M}^T = \Gamma \mathbf{P}^\gamma \Gamma^T \quad (3.18)$$

gegeben!

Beweis 2.

$$\begin{aligned}
\mathbf{P}_{k|k}^x &= \mathbf{S}_{k|k}^x (\mathbf{S}_{k|k}^x)^T \\
&= \mathbf{S}_{k|k-1}^x \left(I - \mathfrak{z} (\mathfrak{U}^T)^{-1} (\mathfrak{U} + \mathfrak{M})^{-1} \mathfrak{z}^T \right) \cdot \left(I - \mathfrak{z} (\mathfrak{U}^T)^{-1} (\mathfrak{U} + \mathfrak{M})^{-1} \mathfrak{z}^T \right)^T (\mathbf{S}_{k|k-1}^x)^T \\
&= \mathbf{S}_{k|k-1}^x \left(I - \mathfrak{z} (\mathfrak{U}^T)^{-1} (\mathfrak{U} + \mathfrak{M})^{-1} \mathfrak{z}^T \right) \cdot \left(I - \mathfrak{z} \left((\mathfrak{U} + \mathfrak{M})^T \right)^{-1} (\mathfrak{U})^{-1} \mathfrak{z}^T \right) (\mathbf{S}_{k|k-1}^x)^T \\
&= \mathbf{S}_{k|k-1}^x \left(I - \mathfrak{z} (\mathfrak{U}^T)^{-1} (\mathfrak{U} + \mathfrak{M})^{-1} \mathfrak{z}^T - \mathfrak{z} \left((\mathfrak{U} + \mathfrak{M})^T \right)^{-1} \mathfrak{U}^{-1} \mathfrak{z}^T \right. \\
&\quad \left. + \mathfrak{z} (\mathfrak{U}^T)^{-1} (\mathfrak{U} + \mathfrak{M})^{-1} \mathfrak{z}^T \mathfrak{z} \left((\mathfrak{U} + \mathfrak{M})^T \right)^{-1} \mathfrak{U}^{-1} \mathfrak{z}^T \right) (\mathbf{S}_{k|k-1}^x)^T \tag{3.19} \\
&= \mathbf{S}_{k|k-1}^x \left(I - \mathfrak{z} (\mathfrak{U}^T)^{-1} (\mathfrak{U} + \mathfrak{M})^{-1} \langle (\mathfrak{U} + \mathfrak{M}) (\mathfrak{U} + \mathfrak{M})^T \rangle \left((\mathfrak{U} + \mathfrak{M})^T \right)^{-1} \mathfrak{U}^{-1} \mathfrak{z}^T \right) (\mathbf{S}_{k|k-1}^x)^T \\
&= \mathbf{S}_{k|k-1}^x \left(I - \mathfrak{z} (\mathfrak{U}^T)^{-1} \mathfrak{U}^{-1} \mathfrak{z}^T \right) (\mathbf{S}_{k|k-1}^x)^T \\
&= \mathbf{S}_{k|k-1}^x \left(I - \mathfrak{z} (\mathfrak{U} \mathfrak{U}^T)^{-1} \mathfrak{z}^T \right) (\mathbf{S}_{k|k-1}^x)^T \\
&= \mathbf{S}_{k|k-1}^x (\mathbf{S}_{k|k-1}^x)^T \\
&\quad - \mathbf{S}_{k|k-1}^x (\mathbf{S}_{k|k-1}^x)^T (\mathcal{J}_{\mathcal{H}}^x)^T \left(\mathcal{J}_{\mathcal{H}}^x \mathbf{S}_{k|k-1}^x (\mathbf{S}_{k|k-1}^x)^T (\mathcal{J}_{\mathcal{H}}^x)^T + \Gamma \mathbf{P}^\gamma \Gamma^T \right)^{-1} \mathcal{J}_{\mathcal{H}}^x \mathbf{S}_{k|k-1}^x (\mathbf{S}_{k|k-1}^x)^T \\
&= \mathbf{P}_{k|k-1}^x - \mathbf{P}_{k|k-1}^x (\mathcal{J}_{\mathcal{H}}^x)^T \left(\mathcal{J}_{\mathcal{H}}^x \mathbf{P}_{k|k-1}^x (\mathcal{J}_{\mathcal{H}}^x)^T + \Gamma \mathbf{P}^\gamma \Gamma^T \right)^{-1} \mathcal{J}_{\mathcal{H}}^x \mathbf{P}_{k|k-1}^x \\
&\quad \text{q. e. d.}
\end{aligned}$$

Die SR-Formulierung lässt sich so für nahezu alle Variationen des EKF verfassen \rightarrow SREKF, SRIEKF₁, SRAEKF. Allerdings existiert keine SR-Formulierung für den IEKF₂.

Quadratur-Methoden

In [7] und [8] wurden ebenso für den UKF und den CDKF SR-Formulierungen (SRUKF, SRCDFK) verfasst, um die positive Definitheit der Kovarianzmatrix \mathbf{P}^x zu garantieren. In den Standardformulierungen aller Quadratur-Methoden, also auch im GHF, muss die Cholesky-Faktorisierung $\mathbf{S}^x (\mathbf{S}^x)^T$ der Kovarianzmatrix \mathbf{P}^x zu jedem Zeitschritt berechnet werden. Bei Rundungsfehlern kann dies dazu führen, dass \mathbf{P}^x seine positive Definitheit verliert und der Algorithmus somit versagt.

Die SR-Formen der SPKF verwenden drei Verfahren der linearen Algebra, *QR-Zerlegung*, *Cholesky-Faktor Updating* und *Effiziente Least-Square-Methoden*, welche nachfolgend kurz beschrieben werden:

1. QR-Zerlegung

Die QR-Zerlegung einer Matrix $\mathbf{A} \in \mathbb{R}^{L \times N}$ sei gegeben durch

$$\mathbf{A}^T = \mathbf{Q} \mathbf{R}, \tag{3.20}$$

wobei $\mathbf{Q} \in \mathbb{R}^{N \times L}$ orthogonal, $\mathbf{R} \in \mathbb{R}^{L \times L}$ obertriangulär und $N \geq L$ ist. Der obertrianguläre Teil von \mathbf{R} , $\tilde{\mathbf{R}}$ ist die Transponierte des Cholesky-Faktors von $\mathbf{P} = \mathbf{A} \mathbf{A}^T$, d.h. $\tilde{\mathbf{R}}^T \tilde{\mathbf{R}} = \mathbf{A} \mathbf{A}^T = \mathbf{P}$.

2. Cholesky-Faktor-Updating

Sei \mathbf{S} der untertrianguläre Cholesky-Faktor von $\mathbf{P} = \mathbf{A}\mathbf{A}^T$, dann ist der Cholesky-Faktor des Rang-1-Updates (oder Datedates)

$$\check{\mathbf{P}} = \mathbf{P} \pm \sqrt{\nu} \vartheta \vartheta^T \quad (3.21)$$

gekennzeichnet durch

$$\check{\mathbf{S}} = \text{cholupdate}\{\mathbf{S}, \vartheta, \pm\nu\}. \quad (3.22)$$

Falls ϑ eine Matrix und kein Vektor ist, resultieren aus dem Update \mathfrak{M} konsekutive Updates des Cholesky-Faktors, wobei die \mathfrak{M} Spalten von ϑ benutzt werden.

3. Effiziente Least-Square-Methoden

Die Lösungen der LS-Probleme

$$\mathbf{A}x = b \quad (3.23)$$

$$x\mathbf{A} = b \quad (3.24)$$

können effizient durch MATLABS "\"- bzw. "/"-Operatoren bestimmt werden (trianguläre QR-Zerlegung mit Pivotisierung bzw. GAUSS-Elimination mit teilweiser Pivotisierung).

Wie in den Standard-SPKF wird während der Initialisierung die Cholesky-Faktorisierung der Zustandskovarianzmatrix berechnet. Der propagierte und upgedatete Cholesky-Faktor wird anschließend verwendet, um die Sigma-Vektoren direkt zu berechnen. Die Prädiktion des Cholesky-Faktors $\mathbf{S}_{k|k-1}^x$ wird durch QR-Zerlegung der Verbundmatrix Λ , bestehend aus den propagierten, gewichteten Sigma-Vektoren $\mathcal{X}_{k|k-1}^x$ und dem Cholesky-Faktor der Kovarianz des Systemrauschens \mathbf{S}^c , berechnet. Das sich anschließende Cholesky-Faktor-Update/Downdate ist nur für den UKF notwendig, da das erste Gewicht $w^{(1),c}$ negativ sein kann. Diese zwei Schritte ersetzen das Zeit-Update der Zustandskovarianzmatrix $\mathbf{P}_{k|k-1}^x$. Der gleiche Zwei-Schritt-Ansatz wird zur Berechnung des Cholesky-Faktors der Messfehlerkovarianzmatrix $\mathbf{P}^{\tilde{y}}, \mathbf{S}^{\tilde{y}}$, verwendet. Um die Kalman-Verstärkung \mathcal{K}_k zu bestimmen, werden zwei geschachtelte LS-Lösungen verwendet:

$$\mathcal{K}_k \left(\mathbf{S}_k^{\tilde{y}} (\mathbf{S}_k^{\tilde{y}})^T \right) = \mathbf{P}_k^{xy} \quad (3.25)$$

$$\mathcal{K}_k = \left(\mathbf{P}_k^{xy} / (\mathbf{S}_k^{\tilde{y}})^T \right) / \mathbf{S}_k^{\tilde{y}}. \quad (3.26)$$

Da $\mathbf{S}_k^{\tilde{y}}$ quadratisch und triangulär ist, können effiziente Methoden zur Lösung nach der Kalman-Verstärkung verwendet werden, die ohne eine Matrixinversion auskommen. Zu guter Letzt muss der Cholesky-Faktor $\mathbf{S}_{k|k}^x$ der posteriori Zustandskovarianzmatrix $\mathbf{P}_{k|k}^x$ berechnet werden, indem m sequentielle Cholesky-Downdates von $\mathbf{S}_{k|k-1}^x$ berechnet werden. Die Downdate-Vektoren dabei sind die Spalten von $\mathfrak{U} \in \mathbb{R}^{n \times m} = \mathcal{K}_k \mathbf{S}_k^{\tilde{y}}$.

Entkoppelte Zustands- und Parameterschätzung (Gauß-Filter)

Die gemeinsame Zustands- und Parameterschätzung kann relativ schnell zu einem nichtlinearen Schätzproblem hoher Dimension führen. Rechenintensive Matrizenoperationen wie Inversionen können so schnell

dazu führen, dass die Algorithmen nicht mehr echtzeitfähig sind. Eine Methode, dieses Problem zu umgehen, ist die entkoppelte Zustands- und Parameterschätzung (Decoupled Estimation), siehe z.B. [8], [56] und [57]. Dabei erfolgt die Zustands- und Parameterschätzung durch zwei parallel laufende Schätzer für den Zustand und die Parameter.

Die zwei DSSM, welche durch die Entkopplung entstehen sind

$$x_k = \mathcal{F}([x_{k-1} \hat{\theta}_{k-1}], u_{k-1}, \kappa_{k-1}^x) \quad (3.27)$$

$$y_k = \mathcal{H}([x_k \hat{\theta}_{k-1}], u_k, \gamma_k) \quad (3.28)$$

für den Zustandsfilter und

$$\theta_k = \theta_{k-1} + \tilde{\kappa}_{k-1}^\theta \quad (3.29)$$

$$y_k = \mathcal{H}([\mathcal{F}([\hat{x}_{k-1} \theta_k], u_{k-1}, \kappa_{k-1}^x) \theta_k], u_k, \gamma_k) \quad (3.30)$$

für den Parameterfilter. Dies bedeutet, dass der geschätzte Zustandsvektor zum Zeitpunkt $k-1$ als bekannte Konstante im Parameterfilter und der geschätzte Parametervektor als bekannte Konstante im Zustandsfilter verwendet wird.

In [53] wird gezeigt, dass in der Theorie die gemeinsame Zustands- und Parameterschätzung eine höhere Güte als die entkoppelte Methode aufweist. Dies resultiert aus der Architektur der Algorithmen: Bei der Entkopplung geht die Information über die Korrelation zwischen Zustands- und Parameterschätzung verloren, d.h. die Kreuzkovarianzen verschwinden:

$$\mathbf{P}_k^{\theta x} = \mathbf{P}_k^{x\theta} = 0 \quad (3.31)$$

Neben dem Verlust der Kreuzkorrelation, kann die Beobachtbarkeit der Parameter durch die Entkopplung verloren gehen, wenn die Parameterdrift keine Auswirkungen auf die beobachteten Zustände in einem Zeitschritt hat. Die Entkopplung liefert dann einen Filter, der seine Aufgabe nicht mehr erfüllen kann.

3.3 Systeme mit finiten Zustandsräumen

In vielen technischen Anwendungen sind die Zustandsräume des zu beobachtenden Systems finit, z.B. ist die Summe von Molenbrüchen immer gleich eins oder Konzentrationen sind nicht negativ. Die bisher vorgestellten Methoden in ihrer allgemeinen Form können mit Zustands-/Parameterbeschränkungen nicht umgehen, was in speziellen Anwendungen zu schlechter Schätzgüte führen kann. In diesem Abschnitt wird gezeigt, wie Zustands-/Parameterbeschränkungen in die Filter integriert werden können.

Gauß-Filter

Unter der Annahme, dass man basierend auf der physikalischen Kenntnis des Systems, Bereichsbeschränkungen in den Formen

$$\mathbf{D}_{Eq}x = d_{Eq} \quad (3.32)$$

$$\mathbf{D}_{Ineq}x \leq d_{Ineq}, \quad (3.33)$$

d.h. lineare Gleichheits- und Ungleichheitsbeschränkungen, definieren kann, gibt es mehrere Ansätze, diese in GF einzubauen. Ein Ansatz ist die Modellreduktion [58], was i.d.R. zu Zustandsraumdarstellungen mit nicht interpretierbaren Zustandsvektoren führt. Ein weiterer Ansatz ist, Bereichsgrenzen als "perfekte" Messungen in die Filter einzubauen [59, 60]. Beide Ansätze ermöglichen lediglich die Behandlung von Gleichheitsbeschränkungen. Ein weit verbreiteter Ansatz, welcher auch die Behandlung von Ungleichheitsbeschränkungen erlaubt, ist der so genannte Projektionsansatz [61]. Ausgehend von einer ersten unbegrenzten Schätzung \hat{x} , welche der Maximierung $\hat{x}_k = \operatorname{argmax}_{x_k} p(x_k|Y_k)$ entspricht (bei GF entspricht die MMSE-Schätzung der MAP-Schätzung), kann ein neues Optimierungsproblem mit Bereichsgrenzen wie folgt definiert werden: Die Optimierung von $p(x_k|Y_k)$ entspricht bei Annahme GAUSS'scher Verteilung von $p(x_k|Y_k)$ der Minimierung $\hat{x}_k = \operatorname{argmin}_{\hat{x}_k} (\hat{x}_k - \bar{x}_k)^T (\mathbf{P}_k^x)^{-1} (\hat{x}_k - \bar{x}_k)$. Unter der Annahme, dass Gleichheits- und Ungleichheitsbeschränkungen bestehen, ergibt sich das neue Optimierungsproblem

$$\hat{x}_{k,constr} = \operatorname{argmin}_{\hat{x}_{k,constr}} (\hat{x}_{k,constr} - \hat{x}_k)^T (\mathbf{P}_k^x)^{-1} (\hat{x}_{k,constr} - \hat{x}_k), \quad (3.34)$$

so dass die Bedingungen

$$\mathbf{D}_{Eq} \hat{x}_{k,constr} = d_{Eq} \quad (3.35)$$

$$\mathbf{D}_{Ineq} \hat{x}_{k,constr} \leq d_{Ineq} \quad (3.36)$$

erfüllt sind. Der Form nach ist das o.a. Optimierungsproblem ein quadratisches Optimierungsproblem, welches in MATLAB[®] mit der Funktion *quadprog* gelöst werden kann. An dieser Stelle sei angemerkt, dass für nichtlineare Filterprobleme $p(x_k|Y_k)$ i.d.R. nicht durch das erste und zweite Moment hinreichend genau beschrieben werden kann, wodurch das o.a. Optimierungsproblem auch nur eine suboptimale Lösung liefert.

Für nichtlineare Gleichheits- und Ungleichheitsbeschränkungen entsteht ein nichtlineares quadratisches Optimierungsproblem, welches grundsätzlich durch von MATLAB[®] bereitgestellte Optimierer gelöst werden kann. Alternativ können auch die Nichtlinearitäten in den Modellgleichungen in zusätzlichen, künstlichen Zuständen "versteckt" werden, wodurch das Problem wieder auf ein lineares Optimierungsproblem zurückgeführt werden kann. Innerhalb dieser Arbeit soll aber nicht weiter darauf eingegangen werden.

SMC-Methoden

Nachfolgend wird ein Ansatz hergeleitet, wie harte Zustandsschranken der Form $\Xi_U \leq x \leq \Xi_O$, $\Xi_O > \Xi_U$ in SMC-Methoden integriert werden können. Die Entwicklung von Ansätzen zum Umgang mit linearen

Gleichheits- und Ungleichheitsbeschränkungen wie zuvor dargestellt, soll im Rahmen dieser Arbeit nicht erfolgen, sondern soll Ausblick auf weitere Arbeiten auf diesem Gebiet sein.

Der hier vorgestellte Ansatz sieht vor, die priori $p(x_k|Y_{k-1})$ und posteriori $p(x_k|Y_k)$ Verteilung durch GMM-Modelle zu approximieren und unter Berücksichtigung der Zustandsschranken neue (beschränkte) Partikelsätze zu generieren. Wie bereits zuvor erwähnt (Seite 24), kann jedwede Wahrscheinlichkeitsverteilung $p(x)$ beliebig genau durch ein GMM approximiert werden:

$$p(x) \approx p_{GMM}(x) = \sum_{g=1}^G \alpha^{(g)} \cdot \mathcal{N}\left(x; \bar{x}^{(g)}, \mathbf{P}^{x,(g)}\right). \quad (3.37)$$

Ferner ist ein GMM, dessen Komponenten lediglich diagonale Kovarianzmatrizen besitzen, d.h. es gilt

$$\mathbf{P}_{ij}^{x,(g)} = 0 \quad i \neq j, \quad (3.38)$$

ebenso in der Lage jedwede Verteilung genau anzunähern. Um die Güte approximierender Verteilungen zu bestimmen, kann der BHATTACHARYYA-Koeffizient (BK) ρ_B [62] herangezogen werden. Der BK zwischen einer Ansatzverteilung $\pi(x)$ und $p(x)$ ist definiert durch

$$\rho_B(p, \pi) = \int_{\mathbb{R}^n} \sqrt{p(x) \cdot \pi(x)} dx \times 100 \% \quad (3.39)$$

$$\rho_B(p, \pi) = \sum_{x \in \mathbb{R}^n} \sqrt{p(x) \cdot \pi(x)} \times 100 \% \quad (3.40)$$

für kontinuierliche und diskrete Verteilungen, wobei die Bedingungen

$$\int_{\mathbb{R}^n} p(x) dx = \int_{\mathbb{R}^n} \pi(x) dx \stackrel{!}{=} 1 \quad (3.41)$$

$$\sum_{x \in \mathbb{R}^n} p(x) = \sum_{x \in \mathbb{R}^n} \pi(x) \stackrel{!}{=} 1 \quad (3.42)$$

erfüllt sein müssen. Der BK ist nach oben beschränkt durch $\rho_B \leq 100 \%$, wobei $\rho_B = 100 \%$ eine optimale Ansatzverteilung, d.h. $\pi(x) = p(x)$, kennzeichnet. Um Partikel-Filter mit Zustandsschranken zu entwickeln, müssen Partikel aus GMM-Modellen generiert werden. Dazu kann folgender Algorithmus verwendet werden:

1. Bestimmung der pro Komponente g zu generierenden Partikelanzahl durch:
 - a) Generierung von N Zufallszahlen $u \sim \mathcal{U}(u; 0, 1)$.
 - b) Berechnung des Cumulative-Sum-of-Weights-Vektors $CSW^{(g)} = \sum_{i=1}^g \alpha^{(i)}$, $g = 1, \dots, G$, $CSW^{(0)} = 0$.
 - c) Die Anzahl an Zufallszahlen q , welche die Bedingung $CSW^{(g-1)} < u \leq CSW^{(g)}$ befriedigen, bestimmt die Anzahl der zu generierenden Partikel aus Komponente g .
2. Generierung von q ($\sum q = N$) Partikel aus jeder Komponente g , $x \sim \mathcal{N}(x; \bar{x}^{(g)}, \mathbf{P}^{x,(g)})$.
3. Verknüpfung der Partikel jeder Komponente zu einer Partikelwolke.

4. Bei Bedarf: Gewichtung jedes Partikels via $w^{(i)} = \frac{p_{GMM}(x^{(i)})}{\sum_i p_{GMM}(x^{(i)})}$.

Die vorgestellte Prozedur resultiert in einer Partikelwolke, welche gemäß $p_{GMM}(x) \approx p(x)$ verteilt ist.

Es sei nun angenommen, dass ein Satz an harten Zustandsschranken der Form $\Xi \in \mathbb{R}^{n \times 2}$ existiert. Die beschränkte Dichtefunktion $p(x, \Xi)$ der unbeschränkten Verteilung $p(x)$ ist dann gegeben durch

$$p(x, \Xi) = \begin{cases} \frac{p(x)}{\int_{x \in \Xi} p(x)}, & x \in \Xi; \\ 0, & \text{sonst.} \end{cases} \quad (3.43)$$

Im Nachfolgenden werden zwei Ansätze vorgestellt, wie aus der beschränkten Verteilung $p(x, \Xi)$ Partikelsätze generiert werden können.

I. Rejection Sampling (RS) aus einer beschränkten Verteilung

Wie zuvor vorgestellt, können aus der unbeschränkten GMM-Approximation $p_{GMM}(x)$ der Verteilung $p(x)$ leicht Partikel generiert werden. Über RS werden Partikel aus der unbeschränkten Verteilung gewonnen ($x \sim p_{GMM}(x)$), solange bis die Bedingung $x \in \Xi$ erfüllt ist. Der resultierende Partikelsatz ist gemäß der beschränkten Verteilung $p(x, \Xi)$ verteilt. RS ist eine simple und einfach zu implementierende Technik um Zustandsbeschränkungen in die Generierung von Partikeln einzubinden. Der gravierende Nachteil des eher naiven RS-Ansatzes ist, dass das Generieren von Partikeln aus beschränkten Verteilungen mit schmalen Definitions-/Gültigkeitsbereichen und/oder geringen Wahrscheinlichkeiten im gültigen Bereich in einer exponentiell anwachsenden Anzahl an abgewiesenen Partikeln und somit in exponentiell anwachsendem Rechenaufwand resultiert. Die später vorgestellten Beispiele demonstrieren diese Eigenschaft.

II. Projection Sampling (PS) aus einer beschränkten Verteilung

Die Verteilung $p(x)$ besitzt die Verteilungsfunktion

$$\mathcal{P}(x) = \int_{-\infty}^x p(\tilde{x}) d\tilde{x}. \quad (3.44)$$

Angenommen $p(x)$ sei beschränkt mit Grenzen $\Xi = [\Xi_U \ \Xi_O]$, $\Xi_O > \Xi_U$, dann ist $\mathcal{P}(x)$ ebenso beschränkt mit $\Lambda = [\Lambda_U \ \Lambda_O] = [\mathcal{P}(\Xi_U) \ \mathcal{P}(\Xi_O)]$, $\Lambda_O > \Lambda_U$. Existiert nun die Inverse der Verteilungsfunktion $\mathcal{P}^{-1}(x)$, d.h. $\mathcal{P}(x)$ ist bijektiv, so können aus $p(x, \Xi)$ wie folgt Partikel generiert werden:

1. Generierung einer Zufallszahl $u \sim \mathcal{U}(u; 0, 1)$.
2. Setze $\Omega = \Lambda_U + (\Lambda_O - \Lambda_U) \cdot u$.
3. $x = \mathcal{P}^{-1}(\Omega)$.

Dieser Algorithmus verwendet dabei das Theorem der wahrscheinlichkeitstheoretischen Integraltransformation:

Theorem 3. Sei x eine kontinuierliche Zufallsvariable mit streng monoton wachsender Verteilungsfunktion $\mathcal{P}(x)$ und $U = \mathcal{P}(x)$ sei ebenfalls eine Zufallsvariable, dann ist U gleichverteilt auf dem Intervall $(0, 1)$, d.h. $\Pr(U \leq u) = u$, $0 < u < 1$!

Beweis 3. [63] Die Aussage aus obiger Behauptung ist äquivalent mit der mathematischen Beschreibung

$$\mathcal{P}^{-1}(u) = \inf(x | \mathcal{P}(x) = u, 0 < u < 1), \quad (3.45)$$

d.h. die Inverse der Verteilungsfunktion $\mathcal{P}^{-1}(u)$ ist größte untere Schranke (Infimum) von x an der Stelle $\mathcal{P}(x) = u$. Daraus folgt

$$\Pr(\mathcal{P}^{-1}(U) \leq x) = \Pr(\inf(x | \mathcal{P}(x) = U) \leq x) \quad (3.46)$$

$$= \Pr(U \leq \mathcal{P}(x)) \quad (3.47)$$

$$= \mathcal{P}(x) \quad (3.48)$$

q.e.d.

PS garantiert, dass jedes resultierende Partikel x die Bedingung $x \in \Xi$ befriedigt. Dies erlaubt es Partikel aus beschränkten Zustandsräumen zu generieren, auch aus solchen die schmale Definitions-/Gültigkeitsbereiche und/oder geringe Wahrscheinlichkeiten im gültigen Bereich besitzen. Die Bedingung ist, dass die Inverse der Verteilungsfunktion existieren und bijektiv sein muss.

Die Restauflage ist also einen Ansatz zur Inversion beliebiger Verteilungsfunktionen $\mathcal{P}(x)$, $x \in \mathbb{R}^n$ herzuleiten. Wie bereits zuvor erwähnt, kann jedwede Wahrscheinlichkeitsdichte genau durch ein GMM approximiert werden, wobei die Komponenten lediglich diagonale Kovarianzmatrizen besitzen. Somit können nachfolgende Theoreme aufgestellt werden:

Theorem 4. Sind die Kovarianzmatrizen der Komponenten diagonal, so ist die n -dimensionale Dichtefunktion $p(x)^{(g)}$ jeder Komponente g gegeben durch

$$p(x)^{(g)} = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma_i^{(g)}} \cdot \exp\left[-\frac{1}{2}\left(\frac{x_i - \bar{x}_i^{(g)}}{\sigma_i^{(g)}}\right)^2\right] \quad (3.49)$$

$$= \prod_{i=1}^n p_i(x_i)^{(g)}, \quad (3.50)$$

wobei $\mathbf{P}^{x,(g)} = \text{diag}([\sigma_1^{(g)}]^2 \ (\sigma_2^{(g)})^2 \ \dots \ (\sigma_n^{(g)})^2]$!

Beweis 4. Die Dichtefunktion der n -variaten Normalverteilung mit Mittelwert \bar{x} und Kovarianzmatrix \mathbf{P}^x ist durch

$$p(x) = \frac{1}{\sqrt{2\pi}^n \sqrt{|\mathbf{P}^x|}} \cdot \exp\left[-\frac{1}{2}(x - \bar{x})^T \{\mathbf{P}^x\}^{-1} (x - \bar{x})\right] \quad (3.51)$$

gegeben. Die Kovarianzmatrix \mathbf{P}^x ist diagonal, d.h. $\mathbf{P}^x = \text{diag}([\sigma_1^2 \ \sigma_2^2 \ \dots \ \sigma_n^2])$. So gilt für die Determinante $|\mathbf{P}^x| = \prod_{i=1}^n \sigma_i^2$ und somit für die n -variate Dichtefunktion

$$p(x) = \frac{1}{\sqrt{2\pi}^n \sqrt{\prod_{i=1}^n \sigma_i^2}} \cdot \exp \left[-\frac{1}{2} \sum_{i=1}^n \left(\frac{x_i - \bar{x}_i}{\sigma_i} \right)^2 \right] \quad (3.52)$$

$$= \frac{1}{\sqrt{2\pi}\sigma_1} \dots \frac{1}{\sqrt{2\pi}\sigma_n} \cdot \exp \left[-\frac{1}{2} \left(\frac{x_1 - \bar{x}_1}{\sigma_1} \right)^2 \right] \dots \exp \left[-\frac{1}{2} \left(\frac{x_n - \bar{x}_n}{\sigma_n} \right)^2 \right] \quad (3.53)$$

$$= \prod_{i=1}^n p_i(x_i) \quad (3.54)$$

q. e. d.

Korollar 1. Da Theorem 4 bewiesen ist, gilt für die Verteilungsfunktion der Komponente g

$$\mathcal{P}(x)^{(g)} = \prod_{i=1}^n \int_{-\infty}^{x_i} \frac{1}{\sqrt{2\pi}\sigma_i^{(g)}} \cdot \exp \left[-\frac{1}{2} \left(\frac{\tilde{x}_i - \bar{x}_i^{(g)}}{\sigma_i^{(g)}} \right)^2 \right] d\tilde{x}_i \quad (3.55)$$

$$= \prod_{i=1}^n \mathcal{P}_i(x_i)^{(g)}. \quad (3.56)$$

Behauptung 1. (Ohne Beweis) Für eine univariate Normalverteilung mit Mittelwert \bar{x} und Varianz σ^2 gilt, dass die Verteilungsfunktion, gegeben durch

$$\mathcal{P}(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}\sigma} \cdot \exp \left[-\frac{1}{2} \left(\frac{\tilde{x} - \bar{x}}{\sigma} \right)^2 \right] d\tilde{x}, \quad (3.57)$$

bijektiv ist und somit die Inverse

$$x = \mathcal{P}^{-1}(\mathcal{P}(x)|\bar{x}, \sigma^2) \quad (3.58)$$

existiert!

Behauptung 2. (Ohne Beweis) Da Theoreme 3 und 4 sowie Behauptung 1 gültig sind, kann ein Partikel von Komponente g unter Berücksichtigung von Zustandsschranken über nachfolgenden Algorithmus gewonnen werden:

1. FOR $i = 1, \dots, n$

a) Generierung einer Zufallszahl $u \sim \mathcal{U}(u; 0, 1)$.

b) Setze $\Omega = \mathcal{P}_i^{(g)}(\Xi_{U,i}) + (\mathcal{P}_i^{(g)}(\Xi_{O,i}) - \mathcal{P}_i^{(g)}(\Xi_{U,i})) \cdot u$.

c) $x_i = (\mathcal{P}_i^{(g)})^{-1}(\Omega | \bar{x}_i^{(g)}, (\sigma_i^{(g)})^2)$

2. END FOR

3. Verknüpfung der einzelnen Zustände $x = [x_1 \ x_2 \ \dots \ x_n]^T$.

4. Bei Bedarf: Gewichtung jedes Partikels via $w^{(i)} = \frac{p_{GMM}(x^{(i)})}{\sum_i p_{GMM}(x^{(i)})}$.

Das resultierende Partikel befriedigt $x \in \Xi$. Die Kombination aus Verteilungsapproximation durch GMM und Partikelgenerierung mit Zustandsschranken erlaubt es Partikelwolken zu generieren, welche gemäß $p(x, \Xi)$ verteilt sind.

Nachfolgend werden der RS- und der PS-Ansatz an zwei illustrativen Dichtefunktionen mit verschiedenen Schranken demonstriert. Dazu wird zunächst eine eindimensionale Gamma-Verteilung mit Parametern \mathcal{A} und \mathcal{B} betrachtet. Die einseitige Wahrscheinlichkeitsdichte, auf dem semi-infiniten Intervall $x \in [0, \infty)$ definiert, hat die Dichtefunktion

$$p(x|\mathcal{A} = 3, \mathcal{B} = 1.25) = \frac{1}{\mathcal{B}^{\mathcal{A}}\Gamma(\mathcal{A})} x^{\mathcal{A}-1} e^{-\frac{x}{\mathcal{B}}}, \quad (3.59)$$

wobei Γ hier die Gamma-Funktion kennzeichnet (nicht zu verwechseln mit der Verteilungsmatrix $\Gamma(x_k, u_k)$ der Messgleichung). Abbildung 3.1 zeigt die Dichtefunktion und drei GMM-Approximationen mit verschiedener Komponentenanzahl G sowie beispielhaft Zustandsschranken und einen beschränkten Partikelsetz. Tabelle 3.1 zeigt die Rechenzeiten zur Generierung von 100 Partikeln aus dieser Verteilung unter

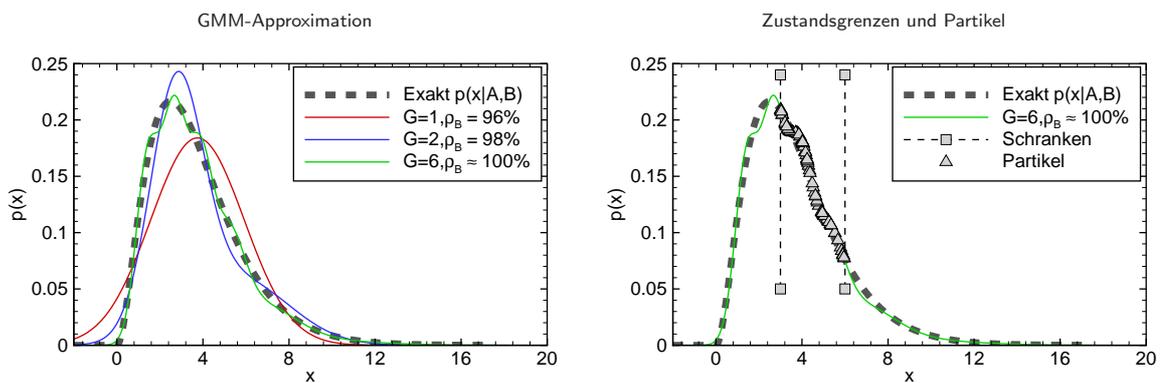


Abbildung 3.1: Eindimensionale Gamma-Verteilung mit GMM-Approximation, Zustandsschranken und beschränkten Partikeln.

Ξ	Rejection Sampling		Projection Sampling
	t^\ddagger	Abweisungen	t
[3 6]	0.100 s	162	0.0021 s
[8 10]	0.837 s	2446	0.0021 s
[12 14]	8.944 s	27560	0.0021 s
[14 14.5]	92.34 s	286600	0.0021 s

\ddagger : DELL PC, Intel CoreDuo CPU, 2x 1.86 GHz, 2GB RAM \rightarrow Alle Simulationsstudien

Tabelle 3.1: Vergleich des RS- und PS-Ansatzes am Beispiel einer eindimensionalen Gamma-Verteilung mit verschiedenen Zustandsschranken

Berücksichtigung verschiedener Zustandsschranken. Wie erwartet wachsen beim RS-Ansatz der Rechenaufwand sowie die Anzahl an Abweisungen mit schrumpfenden Definitions-/Gültigkeitsbereich und/oder geringen Wahrscheinlichkeiten im gültigen Bereich, wohingegen der PS-Ansatz für alle Ξ gleichen Rechenbedarf besitzt. Nun wird der Ansatz am Beispiel einer zweidimensionalen, multimodalen Verteilung gezeigt,

welche in Abbildung 3.2 dargestellt ist. Das approximierende GMM hat diagonale Kovarianzmatrizen und erreicht einen BK von $\rho_B = 99.97\%$. Tabelle 3.2 fasst wiederum den Rechenbedarf der beiden Ansätze zur Generierung von 100 Partikeln mit verschiedene Zustandsgrenzen Ξ zusammen. Auch hier zeigt der PS-Ansatz seine deutliche Überlegenheit, besonders bei schmalen Gültigkeitsbereichen. Wie in Abbildung 3.2 gezeigt, ist die resultierende Partikelwolke gemäß der beschränkten Verteilung verteilt. Somit ist der vorgestellte Ansatz für beliebige multidimensionale Verteilungen gültig und auf solche applizierbar.

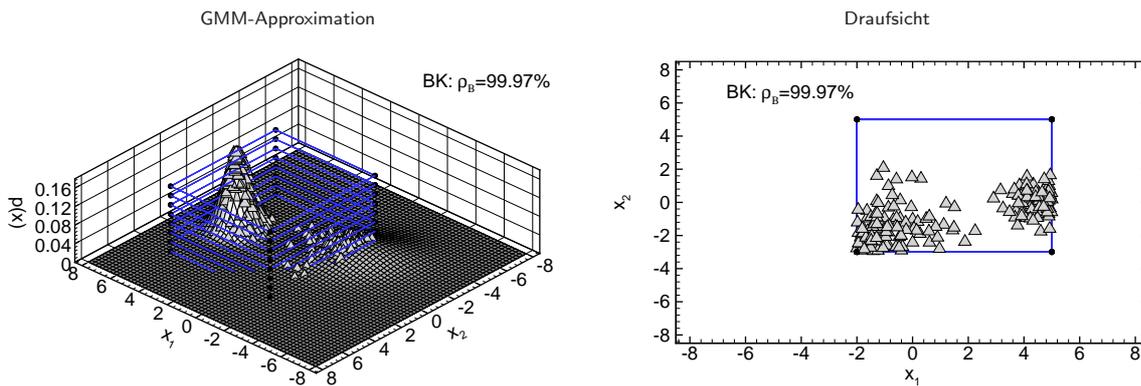


Abbildung 3.2: Zweidimensionale, multimodale Verteilung mit GMM-Approximation, Zustandsschranken und beschränkten Partikeln.

Ξ	Rejection Sampling		Projection Sampling
	t	Abweisungen	t
-2.0 5.0	0.060 s	69	0.005 s
-3.0 5.0			
-1.0 1.0	0.69 s	1349	0.005 s
-1.0 1.0			
-0.1 0.1	56.95 s	115328	0.005 s
-0.1 0.1			

Tabelle 3.2: Vergleich des RS- und PS-Ansatzes am Beispiel einer zweidimensionalen, multimodalen Verteilung mit verschiedenen Zustandsschranken.

Partikel-Filter mit Zustandsschranken

Die Herleitung von Partikel-Filtern mit Zustandsschranken ist nun offensichtlich. Die unbeschränkten priori und posteriori Verteilungen $p(x_k|Y_{k-1})$ und $p(x_k|Y_k)$, in Partikel-Filtern durch

$$p \approx \sum_{i=1}^N w_k^{(i)} \delta(x_k - x_k^{(i)}) \quad (3.60)$$

approximiert, werden durch ein GMM ersetzt und neue, beschränkte Partikelwolken $\{x_k^{(i)}, w_k^{(i)} | \Xi\}_{i=1, \dots, N}$, welche die beschränkten priori und posteriori Verteilungen

$$p(x_k | Y_{k-1}, \Xi) = \begin{cases} \frac{p(x_k | Y_{k-1})}{\int_{x_k \in \Xi} p(x_k | Y_{k-1})}, & x_k \in \Xi; \\ 0, & \text{sonst.} \end{cases} \quad (3.61)$$

und

$$p(x_k | Y_k, \Xi) = \begin{cases} \frac{p(x_k | Y_k)}{\int_{x_k \in \Xi} p(x_k | Y_k)}, & x_k \in \Xi; \\ 0, & \text{sonst.} \end{cases} \quad (3.62)$$

repräsentieren, können durch den vorgestellten Ansatz gewonnen werden. Diese Prozedur garantiert, dass jedes Partikel der Bedingung $x_k \in \Xi$ genügt. Dabei ist die Anzahl G an GMM-Komponenten ein kritisches Entwurfskriterium. Entweder kann eine feste Anzahl an Komponenten G gewählt werden oder G kann adaptiv bestimmt werden, so dass beispielsweise ein gewünschter BK erzielt wird. Letzterer Ansatz kann zu genaueren GMM führen, allerdings auf Kosten erhöhten Rechenaufwandes.

Moving Horizon Estimation

Moving Horizon Estimation (MHE) ist ein neues Konzept zur Lösung nichtlinearer Filterprobleme mit Bereichsbeschränkungen. Über einen bewegten Zeithorizont wird dabei zu jedem Tastschritt ein nichtlineares Least-Square-Problem gelöst. Der Ansatz basiert nicht wie bei den betrachteten Verfahren auf der RBS. Obwohl MHE eine vielversprechende Alternative darstellt, soll sie im weiteren Verlauf nicht näher beleuchtet werden. Wichtigste Literaturstellen sind [64, 65, 66, 67].

3.4 Systeme mit Totzeiten

Oftmals entstehen durch Modellbildung Systeme mit Totzeiten, welche sich in kontinuierlicher Form durch

$$\dot{x} = f(x(t), x(t - T_1), x(t - T_2), \dots, u) \quad (3.63)$$

beschreiben lassen. Durch Diskretisierung (siehe Kapitel 4.1.1, ab Seite 57), wobei $\tau_k = t_k - t_{k-1} = dt$ ist, entsteht daraus das diskretisierte Modell

$$x_k = f(x_{k-1}, x_{k-1-\frac{T_1}{\tau_k}}, x_{k-1-\frac{T_2}{\tau_k}}, \dots, u_{k-1}). \quad (3.64)$$

Sei $\mathfrak{M} = \frac{T_{\text{Max}}}{\tau_k}$, so wird das Zustandsraummodell wie folgt erweitert:

$$\tilde{x} = \begin{bmatrix} \tilde{x}_k \\ \tilde{x}_{k-1} \\ \vdots \\ \tilde{x}_{k-\mathfrak{M}} \end{bmatrix} = \begin{bmatrix} f(\tilde{x}_{k-1}, \tilde{x}_{k-2}, \dots, \tilde{x}_{k-1-\mathfrak{M}}, u_{k-1}) \\ \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \tilde{x}_{k-2} \\ \vdots \\ \tilde{x}_{k-1-\mathfrak{M}} \end{bmatrix} \\ \begin{bmatrix} 0 & \ddots & 0 & \vdots \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \tilde{x}_{k-2} \\ \vdots \\ \tilde{x}_{k-1-\mathfrak{M}} \end{bmatrix} \end{bmatrix}. \quad (3.65)$$

Hier wird bereits offensichtlich, dass Systeme mit großen Totzeiten in hochdimensionale Systeme übergehen.

3.5 Kapitelzusammenfassung

Das Kapitel 3 besteht aus vier Hauptteilen:

Im ersten Teil wurde gezeigt, welche Arten von Schätzaufgaben mit der RBS gelöst werden können. Neben der klassischen Aufgabe der Zustandsschätzung bieten die Inferenzverfahren die Möglichkeit, Systemparameter rekursiv, entweder einzeln als nichtlineares Kennfeld oder gemeinsam mit den Zuständen, zu schätzen. Die gemeinsame Zustands- und Parameterschätzung ist besonders für die modellbasierte Diagnose von Interesse, da aus variierenden Systemparametern oder festgelegten Fehlerparametern gute Diagnoseaussagen abgeleitet werden können. Die einfache Parameterschätzung kann ebenfalls zur MBD eingesetzt werden, wenn nur statische Modelle vorliegen. Dabei ist die Diagnosegeschwindigkeit von der Abtastfrequenz abhängig.

Im zweiten Teil wurden Variationen der Algorithmen-Implementierung aufgezeigt. Sind das System- und das Messrauschen rein additiv, so vereinfachen sich die Algorithmen erheblich. Für die GAUSS'schen Inferenzverfahren wurden Square-Root-Implementierungen angegeben. Diese propagieren die Cholesky-Faktorisierung der Zustandskovarianzmatrix anstelle der vollen Zustandskovarianzmatrix. Dies stabilisiert die Algorithmen numerisch bei gleichem Rechenaufwand auf Plattformen, wo Rundungsfehler auftreten können. Auch wurde die entkoppelte Zustands- und Parameterschätzung vorgestellt. Dabei wird die erweiterte Zustandsraumdarstellung aufgetrennt und es werden für die Zustands- und Parameterschätzung getrennte Schätzer verwendet. Dadurch reduziert sich der Rechenaufwand bei hochdimensionalen Systemen allerdings auf Kosten der Filtergüte, da die Korrelation zwischen Parametern und Zuständen nicht modelliert wird.

Im dritten Teil wurde gezeigt, wie mit Systemen mit finiten Zustandsräumen, Gleichheits- und Ungleichheitsbeschränkungen umgegangen werden kann. Es ergibt sich für GAUSS'sche Systeme ein quadratisches Optimierungsproblem, welches zusätzlich in die Algorithmen implementiert werden muss. Für SMC-Methoden können durch GMM-Approximation der unbeschränkten priori und posteriori Verteilungen neue beschränkte Punktmassenapproximationen über einen Projektionsansatz gewonnen werden.

Im vierten Teil wurde gezeigt, wie mit totzeitbehafteten Systemen umgegangen werden kann. Die Systemordnung erhöht sich dabei in Abhängigkeit von der Schrittweite der Diskretisierung und der Größe der Totzeiten.

4 Evaluierungsmethodik und Matlab[®] basierte Toolbox ReBEL-IoN

In diesem Kapitel wird gezeigt, wie ausgehend von einem technischen System, definiert in Kapitel 1 (Seite 1), Schritt für Schritt ein Inferenzverfahren ausgewählt werden kann. Dazu wird zunächst skizziert, wie ausgehend von Modellbildung und Diskretisierung ein stochastisches DSSM gewonnen werden kann. Im Anschluss daran wird gezeigt, wie die Beobachtbarkeit des stochastischen DSSM nachgewiesen wird. Abschließend folgt die Beschreibung der entwickelten Test- und Bewertungsmethodologie sowie eine Vorstellung der in MATLAB[®] entwickelten Simulations- und Testumgebung.

4.1 Modellbildung Referenz-/Filtermodell

Der Realisierung eines Filters geht eine Modellierung des Systems voraus, für das eine Inferenzaufgabe gelöst werden soll. Klassischerweise kann solch eine Modellbildung analytisch und/oder empirisch erfolgen. Bei der analytischen Modellbildung werden durch Analyse des Systems Erhaltungsgleichungen (Impulssatz, Energieerhaltung,...) gewonnen, welche i.d.R. in eine oder mehrere Differentialgleichungen münden. Physikalische Gesetze bilden die Grundlage solcher Systemgleichungen. Bei der empirischen Modellbildung wird durch Systemtests und Parameteridentifikation i.d.R. aus einem Pool an parametrierbaren, allgemeinen Modellen (DGL(n), Regressionsmodelle, KNN, ...) das bestmögliche ausgewählt.

4.1.1 Zeitliche Diskretisierung kontinuierlicher Systeme

Speziell im Ingenieurwesen, also bei technischen Systemen, resultiert aus einer Modellbildung eine (nicht)lineare, mehrdimensionale zeitkontinuierliche Differentialgleichung in Zustandsraumdarstellung der Form

$$\dot{x} = f_{kon}(x, u; t). \quad (4.1)$$

Um die in dieser Dissertation vorgestellten Inferenzverfahren und die Evaluierungsmethodik auch auf solche Systeme anwenden zu können, müssen diese zeitlich diskretisiert werden. Das heißt, die Differentialgleichung (4.1) wird approximiert durch die Differenzgleichung

$$x_k = f(x_{k-1}, u_{k-1}). \quad (4.2)$$

Wie schon in (4.2) indiziert, soll der Zustand x_k nur aus Informationen des vorherigen Zustands x_{k-1} , des externen Eingangs u_{k-1} und der nichtlinearen Systemgleichung f berechnet werden. Somit scheiden

Mehrschrittverfahren kategorisch aus, die durch Einbeziehen von mehreren vergangenen Zuständen den Folgezustand x_k berechnen. Die Anwendung von impliziten Diskretisierungsverfahren resultiert in nichtlinearen Gleichungssystemen für den Zustand x_k , die zu jedem Zeitschritt gelöst werden müssen. Eine solche Diskretisierung ist auf die Inferenzverfahren anwendbar, führt jedoch zu einem erhöhten Rechenaufwand durch benötigte Lösungsverfahren und zu einem erhöhten Implementierungsaufwand.

Explizite Einschrittverfahren

EULER-Verfahren

Die Basisversion der Diskretisierungsmethoden ist das explizite EULER-Verfahren. Dabei ist die Differenzgleichung/Verfahrensfunktion durch

$$\begin{aligned} x_k &= x_{k-1} + \tau_k \cdot f_{kon}(x, u; t = t_{k-1}) \\ &= x_{k-1} + \tau_k \cdot f_{kon}(x_{k-1}, u_{k-1}; t_{k-1}) \\ &= f(x_{k-1}, u_{k-1}) \end{aligned} \quad (4.3)$$

gegeben, wobei $\tau_k = t_k - t_{k-1}$ ist.

HEUN-Verfahren

Die HEUN-Diskretisierung verspricht verbesserte Genauigkeit im Vergleich zum EULER-Verfahren. Die Diskretisierung durch dieses Trapez-Verfahren ist gegeben durch:

$$x_k = x_{k-1} + \tau_k \cdot \frac{f_{kon}(x, u; t = t_{k-1}) + f_{kon}(x, u; t = t_k)}{2}. \quad (4.4)$$

Der rechte Term über dem Bruchstrich verlangt implizit den Zustand x sowie den exogenen Eingang zum Zeitpunkt t_k , daher ist das Verfahren ein implizites Verfahren. Da x_k nicht verfügbar ist, kann x_k durch die EULER-Diskretisierung ersetzt und ein neuer Eingang $\tilde{u}_{k-1} = [u_{k-1} \ u_k]^T$ definiert werden:

$$\begin{aligned} x_k &= x_{k-1} + \tau_k \cdot \frac{f_{kon}(x_{k-1}, u_{k-1}; t_{k-1}) + f_{kon}(x_{k-1} + \tau_k f_{kon}(x_{k-1}, u_{k-1}; t_{k-1}), u_k; t_k)}{2} \\ &= f(x_{k-1}, \tilde{u}_{k-1}). \end{aligned} \quad (4.5)$$

Ableitungsfreie Diskretisierungen höherer Ordnung

Das HEUN-Verfahren hat gezeigt, dass durch die verschachtelte Auswertung der Systemfunktion die Komplexität der diskretisierten Systemgleichungen rasch ansteigt. Prinzipiell ist die vorgestellte Synthese von diskretisierten Systemgleichungen auch auf höhere Diskretisierungsverfahren z.B. SIMPSON und RUNGE-KUTTA anwendbar, führt aber insbesondere bei höherdimensionalen Systemen schnell zu nicht handhabbaren (auf dem Papier) Gleichungen. Daher wird an dieser Stelle auf weitere Ausführungen verzichtet. Für Hinweise auf Diskretisierungen höherer Ordnung siehe z.B. [68], [69].

Implizite Einschrittverfahren

EULER-Verfahren

Die Differenzgleichung/Verfahrensfunktion des impliziten EULER-Verfahrens ist gegeben durch

$$\begin{aligned}x_k &= x_{k-1} + \tau_k \cdot f_{kon}(x, u; t = t_k) \\ &= x_{k-1} + \tau_k \cdot f_{kon}(x_k, u_k; t_k),\end{aligned}\quad (4.6)$$

wobei $\tau_k = t_k - t_{k-1}$ ist. Die Gleichung zeigt, dass zur Berechnung des Folgezustandes x_k der Folgezustand x_k gegeben sein muss, daher ist dieses Verfahren implizit. Es entsteht ein nichtlineares Gleichungssystem, welches zu jedem Zeitschritt gelöst werden muss. Stellt man um, erhält man

$$0 = x_k - x_{k-1} - \tau_k \cdot f_{kon}(x_k, u_k; t_k). \quad (4.7)$$

Die Nullstellen dieses Gleichungssystems bezüglich x_k entsprechen dem gesuchten Folgezustand x_k .

HEUN-Verfahren

Das implizite Trapez-Verfahren wurde bereits bei der expliziten HEUN-Diskretisierung vorgestellt:

$$x_k = x_{k-1} + \tau_k \cdot \frac{f_{kon}(x, u; t = t_{k-1}) + f_{kon}(x, u; t = t_k)}{2} \quad (4.8)$$

$$= x_{k-1} + \tau_k \cdot \frac{f_{kon}(x_{k-1}, u_{k-1}; t_{k-1}) + f_{kon}(x_k, u_k; t_k)}{2}. \quad (4.9)$$

Der rechte Term über dem Bruchstrich verlangt implizit den Zustand x zum Zeitpunkt t_k , daher ist das Verfahren ein implizites Verfahren. Auch hier muss zu jedem Zeitschritt ein nichtlineares Gleichungssystem gelöst werden. Stellt man um, so erhält man

$$0 = x_k - x_{k-1} - \tau_k \cdot \frac{f_{kon}(x_{k-1}, u_{k-1}; t_{k-1}) + f_{kon}(x_k, u_k; t_k)}{2}. \quad (4.10)$$

Die Nullstellen dieses Gleichungssystems bezüglich x_k entsprechen wiederum dem gesuchten Folgezustand x_k .

Jacobi-Matrizen für Lokal-Methoden

Die Lokal-Methoden benötigen die zeitvarianten Differentialmatrizen $\mathcal{J}_{\mathcal{F}}^x = \frac{\partial \mathcal{F}}{\partial x_{k-1}}$ sowie $K = \frac{\partial \mathcal{F}}{\partial \kappa_{k-1}}$ für die Kovarianzprädiktion. Wie diese aus den verschiedenen Diskretisierungen hergeleitet werden können, wird nachfolgend gezeigt.

Explizite Diskretisierung

Die expliziten Diskretisierungen (EULER, HEUN) liefern eine Differenzgleichung der Form

$$x_k = f(x_{k-1}, \tilde{u}_{k-1}). \quad (4.11)$$

Untersucht man für das gegebene System die Rauschgrößen (vgl. Kapitel 4.1.2, ab Seite 62), so entsteht eine Differenzgleichung der Form

$$x_k = \mathcal{F}(x_{k-1}, u_{k-1}, \kappa_{k-1}), \quad (4.12)$$

welche der Systemgleichung entspricht, für die die Inferenzverfahren in dieser Arbeit untersucht werden. Die Bestimmung der benötigten Differentialmatrizen kann direkt aus dieser Darstellung erfolgen.

Implizite Diskretisierung

Erfolgt die Diskretisierung durch ein implizites Verfahren, so ist die Bestimmung der Differentialmatrizen deutlich komplizierter und vor allem eingeschränkter. Um eine Berechnung zu ermöglichen, muss die Verfahrensfunktion zunächst derart umgeformt werden, dass eine Form

$$x_k = x_{k-1} + f_{Imp}(x_k, u_k; t_k, t_{k-1}) \quad (4.13)$$

resultiert. Die implizite EULER-Diskretisierung liefert sofort diese Darstellung. Bei der impliziten HEUN-Diskretisierung kann dies durch eine Verschachtelung analog zur expliziten HEUN-Diskretisierung erfolgen, indem der Zustand x_{k-1} durch eine implizite EULER-Diskretisierung ersetzt wird. Diese suboptimale Transformation resultiert in nur einer geringfügigen Verschlechterung der Genauigkeit. So resultiert die transformierte HEUN-Diskretisierung:

$$x_k = x_{k-1} + \tau_k \cdot \frac{f_{kon}(x_k - \tau \cdot f_{kon}(x_k, u_k; t_k), u_{k-1}; t_{k-1}) + f_{kon}(x_k, u_k; t_k)}{2}. \quad (4.14)$$

Sowohl die implizite EULER-Diskretisierung als auch die implizite HEUN-Diskretisierung führen nun auf die Systemgleichung

$$x_k = x_{k-1} + f_{Imp}(x_k, \tilde{u}_k; t_k, t_{k-1}) \quad (4.15)$$

bzw.

$$x_{k-1} = x_k - f_{Imp}(x_k, \tilde{u}_k; t_k, t_{k-1}) \approx f_{Exp}^{-1}(x_k, \tilde{u}_k), \quad (4.16)$$

wobei \tilde{u}_k wiederum den Vektor der exogenen Eingänge $[u_{k-1} \ u_k]^T$ darstellt. Da beide Eingänge ohnehin zur Verfügung gestellt werden müssen (Prädiktions- und Filterungsschritt), werden keine zusätzlichen Speicherressourcen benötigt. Die Berechnung der Jacobi-Matrizen $\frac{\partial \mathcal{F}}{\partial x_{k-1}}$ und $\frac{\partial \mathcal{F}}{\partial \kappa_{k-1}}$, welche zur Kovarianzprädiktion benötigt werden, kann nur suboptimal und im Fall additiven Systemrauschens erfolgen. Näherungen sind dann

$$\frac{\partial \mathcal{F}}{\partial x_{k-1}} \approx \left(\frac{\partial f_{Exp}^{-1}}{\partial x_k} \right)^{-1} \quad (4.17)$$

$$\frac{\partial \mathcal{F}}{\partial \kappa_{k-1}} = \mathbf{1}. \quad (4.18)$$

Kondition, Konsistenz und Stabilität

Drei in der Numerik verwendete Bewertungskriterien zur Beurteilung von Fehlern sind Kondition, Konsistenz und Stabilität [70]. Alle drei Kriterien bewerten die Entstehung von Fehlern, unterscheiden sich aber in der Definition der Fehlerquellen.

Kondition

Die Kondition beschreibt die Abhängigkeit der Störung einer genügend glatten mathematischen Operation $x = \mathcal{OP}(u)$ von der Störung der Eingangsdaten u , unabhängig von der Wahl des Lösungsverfahrens. Aufgrund dieser Unabhängigkeit ist die Kondition keine für dieses Kapitel interessante Größe.

Konsistenz

Die Konsistenz beschreibt, mit welcher Genauigkeit ein Verfahren die Differentialgleichung approximiert. Die Konsistenzordnung (Ordnung des globalen Fehlers) p eines Verfahrens mit Verfahrensfunktion $F(x, t, \tau)$ (τ : Schrittweite) entspricht der Anzahl der übereinstimmenden Terme in den Taylor-Entwicklungen nach τ von $x(t + \tau)$ und von $F(x, t, \tau)$. Der lokale Fehler des Verfahrens ist dann von der Ordnung $p + 1$.

- EULER explizit:
Konsistenzordnung: $\mathcal{O}(\tau)$
Lokale Fehlerordnung: $\mathcal{O}(\tau^2)$
- HEUN explizit:
Konsistenzordnung: $\mathcal{O}(\tau^2)$
Lokale Fehlerordnung: $\mathcal{O}(\tau^3)$
- EULER implizit:
Konsistenzordnung: $\mathcal{O}(\tau)$
Lokale Fehlerordnung: $\mathcal{O}(\tau^2)$
- HEUN implizit:
Konsistenzordnung: $\mathcal{O}(\tau^2)$
Lokale Fehlerordnung: $\mathcal{O}(\tau^3)$

Stabilität

Ein numerisches Verfahren gilt dann als stabil, wenn die numerische Lösung für große Zeitwerte gegen die exakte Lösung konvergiert. An dieser Stelle sei darauf hingewiesen, dass numerische Stabilität des Diskretisierungsverfahrens nichts mit der Stabilität der Differentialgleichung zu tun hat.

Für die hier vorgestellten Verfahren lässt sich analytisch ein Stabilitätsgebiet bestimmen, für die Herleitung sei z.B. auf [69] verwiesen, welches in Abbildung 4.1 gezeigt ist. Für lineare Systeme entspricht $\mu := \lambda\tau$

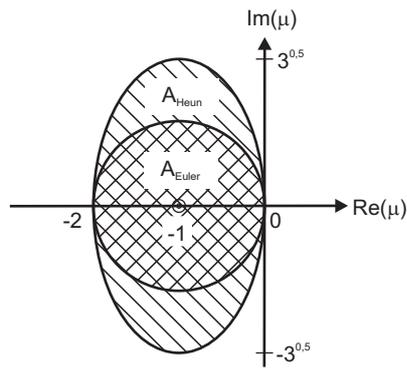


Abbildung 4.1: Stabilitätsbereich des expliziten EULER- und HEUN-Verfahrens [69] (global für lineare Systeme, lokal für nichtlineare Systeme).

dem Produkt aus Schrittweite und Eigenwerten des Systems und man erhält einen global gültigen Stabilitätsbereich. Bei nichtlinearen Systemen entspricht λ den Eigenwerten des linearisierten Systems und man erhält lediglich eine Aussage über lokale Stabilität. Besonders bei steifen Systemen mit großen Eigenwerten wird eine sehr kleine Schrittweite für alle Zeitschritte benötigt, was in sehr große Rechenzeiten münden kann. Hier bieten die impliziten Verfahren einen deutlichen Vorteil und ihr größerer Implementierungsaufwand macht sich bezahlt. Es kann gezeigt werden, dass für die implizite EULER- und HEUN-Methode die Stabilitätsgebiete

$$A_{\text{EULER}} = \{\mu \in \mathbb{C} \mid \text{Re}(\mu) < 0\} \quad (4.19)$$

$$A_{\text{HEUN}} = \{\mu \in \mathbb{C} \mid \text{Re}(\mu) < 0\} \quad (4.20)$$

gelten, d.h. die linke Halbebene von \mathbb{C} ist Stabilitätsbereich. Dies bedeutet, dass die implizite EULER- und HEUN-Methode absolut stabil (A-stabil) sind. Somit erhält man auch für steife Systeme stabile Diskretisierungen mit größeren Schrittweiten (wobei die Schrittweite durchaus die Genauigkeit beeinflusst).

4.1.2 Analyse von System- und Messrauschen und der Initialverteilung

Aus der Modellbildung und Diskretisierung erhält man ein nichtlineares, zeitdiskretes Zustandsraummodell der Form

$$x_k = f(x_{k-1}, u_{k-1}) \quad (4.21)$$

$$y_k = h(x_k, u_k), \quad (4.22)$$

welches das Systemverhalten hinreichend gut beschreiben sollte. Auch die impliziten Diskretisierungen resultieren prinzipiell in dieser Form, da der Zustand x_k lediglich aus Informationen des Zustandes x_{k-1} berechnet wird. Zwar muss hier ein Gleichungssystem gelöst werden, die Form der Gleichung ist aber analog zur expliziten Diskretisierung.

Neben Ungenauigkeiten in der Modellierung (Approximationen) treten in der Praxis Unsicherheiten auf, welche untersucht und in die Systemgleichungen integriert werden sollten. Dabei unterscheidet man zwischen Unsicherheiten, die auf das Systemmodell einwirken (Systemrauschen) und solchen, die sich auf das Messmodell auswirken (Messrauschen).

Systemrauschen

Die Analyse der Systemunsicherheiten gestaltet sich als relativ schwierig. Allgemeine Aussagen dazu konnten in der Literatur nicht gefunden werden. In [30] wird für die Radar-Verfolgung ballistischer Flugobjekte eine Einstellung für das Systemrauschen vorgeschlagen. Allgemein bietet sich die Möglichkeit, unsichere Größen in Gleichungen (z.B. Luftdichte oder optimierte Nominalparameter) als stochastische Größen zu interpretieren, was zwangsläufig auf ein stochastisches Systemmodell führt. Die Wahl der statistischen Größen sowie deren Verteilung ist stark abhängig vom untersuchten System, so können z.B. nominelle Parameter im Betrieb altern oder verschmutzen. Der zeitliche Verlauf des Alterungsgrades ist i.d.R. nur durch einen stochastischen Ansatz beschreibbar und verlangt somit tiefen Einblick in die System- und Umgebungseigenschaften.

Praxisorientierter ist es allerdings, die Verteilung des Modellierungsfehlers als stochastische Zufallsvariable zu interpretieren. Ausgehend von einer analytischen oder experimentellen Systemmodellierung und eventueller Nominalparameterschätzung kann die Verteilung des Modellfehlers $\Delta_{\text{Modellierung}} = x_{\text{Modell}} - x_{\text{Messung}}$ zur Bestimmung des Systemrauschens κ herangezogen werden. Werden die Systemzustände bei der Nominalmodellbildung nicht beobachtet, ist diese Vorgehensweise nicht anwendbar. In diesem Fall wird empfohlen, die Modellunsicherheiten dem Messrauschen zuzurechnen.

Messrauschen

Auf das Messsignal wirken sich im Allgemeinen folgende Einflussgrößen aus:

- Wandlerrauschen.
- Signalleitungseinfluss: Der Kabeltyp und die Peripherie können das Messsignal beeinflussen.
- Sensorrauschen: Widerstandsrauschen der Messbrücken, mangelnde Abschirmung und Temperatureinfluss können das Messsignal verfälschen.

Um die Charakteristik des Messrauschens zu bestimmen, werden für jeden Messkanal $r^{(i)}$, $i = 1, \dots, m$ Leerlaufmessungen (Nullmessungen) durchgeführt. Das aus dem o.a. Einzelrauschen kombinierte Messrauschen kann durch Klassifizierung und Identifikation durch eine approximierte Dichtefunktion beschrieben werden, aus der die Verteilungsparameter bestimmt werden können. Bei angenommener GAUSS'scher Verteilung können Mittelwert und Varianz auch direkt aus den Messdaten bestimmt werden.

Die Bestimmung des Messrauschens ist für den eingesetzten Filter neben einer ausreichend genauen Modellierung der Streckendynamik ein für die Schätzgüte maßgebendes Kriterium. Die angenommene Verteilung des Messrauschens beeinflusst die Filtergüte erheblich und kann u.U. zur Filterdivergenz führen. Eine ausreichend genaue Analyse und daraus folgend eine gute Einstellung des Filters kann Stabilität und Schätzgüte maßgeblich steigern.

Es wird angenommen, dass die verschiedenen Messungen unkorreliert sind. Im GAUSS'schen Fall folgt, dass die Messkovarianzmatrix \mathbf{P}^γ eine Diagonalmatrix mit den Varianzen der einzelnen Messkanäle als Einträge in der Diagonalen ist:

$$\mathbf{P}^\gamma = \begin{bmatrix} \mathbb{E}[(r^{(1)} - \bar{r}^{(1)})^2] & 0 & \dots & 0 \\ 0 & \mathbb{E}[(r^{(2)} - \bar{r}^{(2)})^2] & 0 & \vdots \\ \vdots & 0 & \ddots & \vdots \\ 0 & \dots & \dots & \mathbb{E}[(r^{(m)} - \bar{r}^{(m)})^2] \end{bmatrix}. \quad (4.23)$$

Besteht nicht die Möglichkeit, Nullmessungen durchzuführen, so können Herstellerangaben der Sensoren verwendet werden, um das Messrauschen näherungsweise normalverteilt zu charakterisieren. In der Regel wird der maximale absolute Fehler $|\varepsilon|$ vom Hersteller angegeben. Setzt man voraus, dass die Fehler mit einer x -prozentigen Wahrscheinlichkeit innerhalb eines Toleranzbandes liegen, so erhält man die Varianz aus:

$$\mathbf{P}^\gamma = \left(\frac{|\varepsilon|}{\alpha} \right)^2. \quad (4.24)$$

Dabei ist α ein von der x -prozentigen Wahrscheinlichkeit abhängiger Faktor, z.B.:

$$\alpha = \begin{cases} 1.96 & \text{für 95\% - Wahrscheinlichkeit} \\ 3.86 & \text{für 99\% - Wahrscheinlichkeit.} \end{cases} \quad (4.25)$$

Gewinn eines stochastischen Zustandsraummodells

Durch die Analyse von System- und Messrauschen erhält man das stochastische Zustandsraummodell

$$x_k = f(x_{k-1}, u_{k-1}) + K(x_{k-1}, u_{k-1})\kappa_{k-1} = \mathcal{F}(x_{k-1}, u_{k-1}, \kappa_{k-1}) \quad (4.26)$$

$$y_k = h(x_k, u_k) + \Gamma(x_k, u_k)\gamma_k = \mathcal{H}(x_k, u_k, \gamma_k) \quad (4.27)$$

des Systems. Die in Kapitel 2 beschriebenen und in dieser Arbeit untersuchten Inferenzverfahren sind auf diese System(modell)e anwendbar.

Initialverteilung

Die Bestimmung der Initialverteilung ist sowohl für ein ausgewogenes, realitätsnahes Testen von Wichtigkeit als auch für die spätere Initialisierung des Filters im Realbetrieb. Die Einstellung von $p(x_0)$ kann auf zwei Wegen erfolgen, die nachfolgend kurz diskutiert werden sollen:

1. Kenntnis über das System und seine Verteilung:

In diesem Fall muss der Anwender aus fundierter Kenntnis über das System, z.B. mögliche Verteilung von Parametern, die Initialverteilung intuitiv wählen. Dies erfordert tiefe Einsicht in das untersuchte System und muss von Anwendungsfall zu Anwendungsfall neu untersucht werden.

2. Initialisierung aus ersten Messungen:

Methodischer ist, die Filter über die ersten Messungen zu initialisieren. Dabei wird auf den aktuellen Zustand über die Mess- und Systemgleichung "zurückgerechnet". Da die Messungen naturgemäß stochastisch sind, entspricht dies einer nichtlinearen Transformation einer Zufallsvariablen. Die Verteilung der transformierten Variablen, in diesem Fall Initialzustände, kann über Monte-Carlo-Simulation oder approximiert durch das erste und zweite Moment durch die (S)UT berechnet werden.

4.2 Beobachtbarkeit nichtlinearer, diskreter Systeme

Grundsätzlich muss bei der Beobachtbarkeit von nichtlinearen Zustandsraummodellen zwischen Beobachtbarkeit des deterministischen Systems

$$x_k = f(x_{k-1}, u_{k-1}) \quad (4.28)$$

$$y_k = h(x_k, u_k) \quad (4.29)$$

und Beobachtbarkeit des stochastischen Systems

$$x_k = \mathcal{F}(x_{k-1}, u_{k-1}, \kappa_{k-1}) \quad (4.30)$$

$$y_k = \mathcal{H}(x_k, u_k, \gamma_k) \quad (4.31)$$

unterschieden werden. Notwendige Bedingung für die Beobachtbarkeit des stochastischen Systems ist die gegebene Beobachtbarkeit des deterministischen Systems. Aus diesem Grund empfiehlt es sich, die Beobachtbarkeit des deterministischen Systems zunächst zu überprüfen. Nachfolgende Ausführungen sind in Anlehnung an [71] und [72] verfasst.

Beobachtbarkeit des deterministischen Systems

Zwei Anfangszustände $x_{01}, x_{02} \in \mathcal{S} \subseteq \mathbb{R}^n$, $x_{01} \neq x_{02}$ (\mathcal{S} entspricht dem zu erwartenden Raum des Zustandes x), werden als ununterscheidbar bezeichnet $x_{01} \mathcal{I} x_{02}$ (indistinguishable), wenn die Ausgangsfunktionen $y(x_{01}, u)$ und $y(x_{02}, u)$ für die Anfangswerte $x(0) = x_{01}$ bzw. $x(0) = x_{02}$ und für jeden zulässigen Eingang u identisch sind. Das System heißt beobachtbar, wenn $x_{01} \mathcal{I} x_{02}$ bedeutet, dass $x_{01} = x_{02}$ für jedes $x \in \mathcal{S}$ gilt. Anders formuliert gilt ein System als beobachtbar, wenn jeder Anfangszustand aus Kenntnis des Ausgangssignals y und des Eingangssignal u eindeutig bestimmt werden kann. An dieser Stelle sei angemerkt, dass per Definition der Nachweis der Beobachtbarkeit an einem Punkt zu einem infiniten Test wird, da für jedes $x \in \mathcal{S}$ nachgewiesen werden muss, dass das Kriterium erfüllt ist.

Im Unterschied zu linearen Systemen hängt die Beobachtbarkeit nichtlinearer Systeme von den Eingangsgrößen u ab. In der Literatur konnten keine Beobachtbarkeitskriterien für diskrete nichtlineare, gesteuerte Systeme gefunden werden. Da die Beobachtbarkeit allerdings auch für $u = 0$ (zero-input-observability) gelten muss, wird im Folgenden das autonome System

$$x_k = f(x_{k-1}) \quad (4.32)$$

$$y_k = h(x_k) \quad (4.33)$$

betrachtet. Um den infiniten Test zu umgehen, führt man den Begriff der "lokalen Beobachtbarkeit" ein. In der Praxis kann es ausreichend sein, wenn man x_k von seinen Nachbarn unterscheiden kann. Ein System sei somit "lokal beobachtbar", wenn eine Umgebung $\mathcal{U} : |x_k - x_k^P|$ um x_k^P existiert, so dass die Beobachtbarkeitsmatrix

$$\mathcal{BM}(x_k^P \rightarrow x_k) = \begin{bmatrix} \frac{\partial h}{\partial x}(x_k) \\ \frac{\partial h}{\partial x}(x_{k+1}) \frac{\partial f}{\partial x}(x_k) \\ \vdots \\ \frac{\partial h}{\partial x}(x_{k+n-1}) \frac{\partial f}{\partial x}(x_{k+n-2}) \cdots \frac{\partial f}{\partial x}(x_k) \end{bmatrix} \quad (4.34)$$

vollen Rang besitzt. Das nichtlineare, autonome System ist "lokal beobachtbar", wenn dies für alle $x \in \mathcal{S}$ gilt. Allerdings ist diese Bedingung für die Beobachtbarkeit eines Systems nur notwendig: Erfüllt ein System das Rangkriterium nicht, so kann nicht daraus geschlossen werden, dass es nicht lokal beobachtbar ist. Beispielsweise erfüllt das skalare System $x_k = x_{k-1}(1 + \tau)$, $y_k = x_k^3$ das Rangkriterium bei $x_k = 0$ nicht, dennoch ist es beobachtbar für alle x_k , da $x_k = \sqrt[3]{y_k}$ injektiv ist.

Beobachtbarkeit des stochastischen Systems

In [73] wird am Beispiel des EKF gezeigt, dass die Beobachtbarkeit des stochastischen Systems stark vom Ausmaß der Stochastik (Rauschen) und der Initialverteilung $p(x_0)$ abhängt. Eine fundamentale Analyse der stochastischen Beobachtbarkeit für jedes Inferenzverfahren liegt nicht im Fokus dieser Arbeit, sie kann aber empirisch in der Simulationsumgebung ReBEL-IoN geprüft werden.

4.3 Testen der Inferenzalgorithmen

4.3.1 Experimentelles Testen

Die zentrale Aufgabenstellung besteht darin, wie bereits kurz in der Einleitung erwähnt, für eine Inferenzaufgabe den bestmöglichen Filteralgorithmus auszuwählen. Die Lösung dieses Problems erscheint zunächst trivial. Die Filteralgorithmen könnten am realen System ausgiebig getestet und nachfolgend anhand der Filtergüte bewertet werden, siehe Abbildung 4.2. So einfach diese Lösung ist, so unpraktikabel ist sie gleichermaßen. Zum einen ist der Aufwand, sämtliche Algorithmen zu implementieren, immens und zum

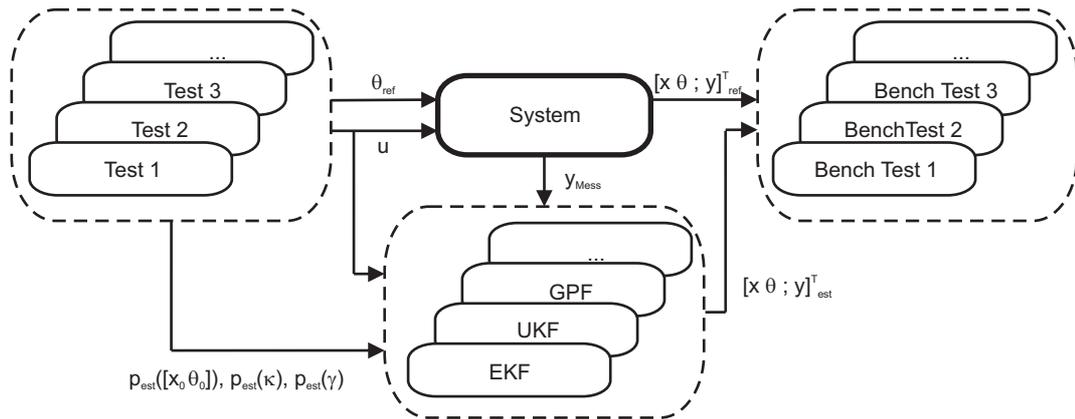


Abbildung 4.2: Testen von Inferenzverfahren am System.

anderen ist der zeitliche Aufwand, für jeden Algorithmus Testzyklen zu fahren, viel zu groß. Dadurch wird die in dieser Dissertation beschriebene Aufgabe um ein Vielfaches schwieriger. Das nachfolgend vorgestellte Konzept soll zeigen, wie eine Off-Line Simulation der verschiedenen Filteralgorithmen durchgeführt werden kann, deren Ergebnisse zur Auswahl herangezogen werden können.

4.3.2 Simulatives Testen

Abbildung 4.3 zeigt die schematische Darstellung des entwickelten Test- und Bewertungskonzepts. Das System wird durch ein hinreichend genaues Referenzmodell ersetzt, welches beschrieben ist durch ein nichtlineares diskretes stochastisches DSSM

$$x_k = \mathcal{F}_{Ref}(x_{k-1}, u_{k-1}, \kappa_{k-1}) \quad (4.35)$$

$$y_k = \mathcal{H}_{Ref}(x_k, u_k, \gamma_k) \quad (4.36)$$

mit Systemrauschen κ und Messrauschen γ .

Stochastische Initialisierung, Trajektorien- und Messdatengenerierung

Ein Test wird bestimmt durch einen vom Anwender vorzugebenden Satz an Eingangsdatenvektoren $U_K = \{u_1, u_2, \dots, u_K\}$ und den Wahrscheinlichkeitsverteilungen der Initialzustände $p([x_0 \ \theta_0])$ (im Folgenden $p(x_0)$), des Systemrauschens $p(\kappa)$, des Messrauschens $p(\gamma)$ sowie der Detektionswahrscheinlichkeit p_D ¹. Das Referenzmodell generiert Referenztrajektorien für die realen Zustände/Parameter sowie Referenzmessungen unter Verwendung von (4.35) und (4.36) und den o.a. Wahrscheinlichkeitsverteilungen. Die Referenzmesswerte $Y_K^{Ref} = \{y_1^{Ref}, y_2^{Ref}, \dots, y_K^{Ref}\}$, exogene Eingänge U sowie eingeschränkte Informationen

¹Zwei typische Sensorcharakteristika (speziell bei Radarsensoren) sind die Detektionswahrscheinlichkeit p_D POD (Probability Of Detection) und die Wahrscheinlichkeit für Fehldetektion p_{FA} POFA (Probability Of False Alert). Typische Werte für moderne Radarsysteme sind $p_D = 0.9$ und $p_{FA} = 10^{-6}$. Die POD beschreibt die Wahrscheinlichkeit dafür, dass eine Messung erfolgreich ist. Die POFA beschreibt die Wahrscheinlichkeit, dass ein Objekt falsch positiv detektiert wird. Die POFA ist i.d.R. viel kleiner als die POD und wird daher nicht näher berücksichtigt.

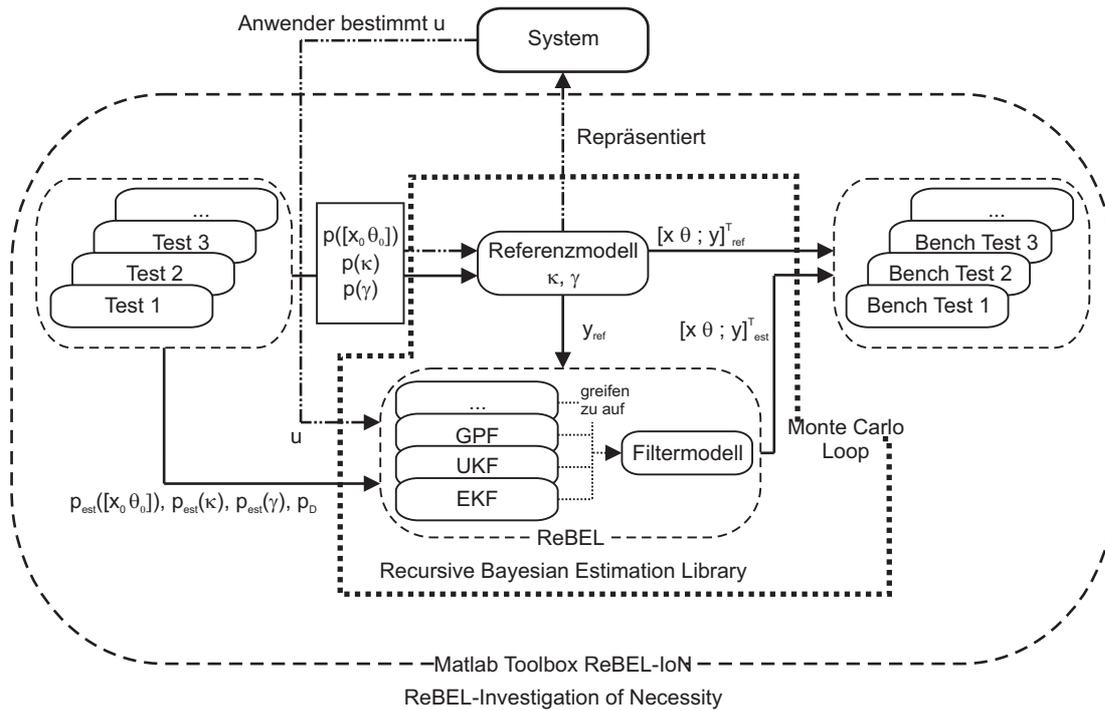


Abbildung 4.3: Testen von Inferenzverfahren innerhalb der Simulationsumgebung.

über die Statistiken des System- und Messrauschens und der Initialverteilung werden den Filtern, welche in einer Filterbibliothek hinterlegt sind, zur Verfügung gestellt. Unter Vorgabe dieser Informationen generieren die Filter unter Verwendung eines Filtermodells

$$x_k = \mathcal{F}_{Fil}(x_{k-1}, u_{k-1}, \kappa_{k-1}) \quad (4.37)$$

$$y_k = \mathcal{H}_{Fil}(x_k, u_k, \gamma_k) \quad (4.38)$$

geschätzte Zustands- und Parametertrajektorien.

Dieser Ansatz erlaubt es, einen Vergleich zu den Referenztrajektorien und eine quantitative Bewertung der Schätzung vorzunehmen. Um einen fairen Vergleich zwischen den Algorithmen und eine ausreichende Abdeckung der Statistiken zu gewährleisten, erfolgt die Trajektoriengenerierung und -schätzung innerhalb einer Monte-Carlo-Simulation.

Modellfehler

Sowohl die Referenztrajektorienengenerierung als auch die Trajektorien-schätzung basiert auf einem Modell des realen Systems. Die theoretische und empirische (und kombinierte) Modellbildung beschreibt i.d.R. das Systemverhalten nicht exakt, weshalb (fast) immer eine Divergenz zwischen Filtermodell und realem System existiert. Sinnvoll ist es daher zu testen, wie empfindlich die Filteralgorithmen auf Modellierungsfehler reagieren. Die MATLAB® Toolbox ReBEL-IoN bietet die Möglichkeit, unterschiedliche Modelle für die

Referenztrajektoriegenerierung und die Trajektorien-schätzung zu definieren.

Integrität

In der Regelungstechnik wird ein Regelungssystem dann als integer bezeichnet, wenn es trotz Ausfall eines oder mehrerer Elemente, z.B. Sensoren, (ggf. eingeschränkt) funktionsfähig und stabil bleibt [74]. Obwohl ein wichtiges praxisorientiertes Kriterium, wurde die Integrität sowohl für Regelungssysteme [75] als auch für Inferenzverfahren kaum untersucht.

Experimentelle Integritätsuntersuchung

In der entwickelten Simulationsumgebung kann die Integrität der Inferenzverfahren untersucht werden. Dazu definiert man die beobachtbare binäre Variable d_k , welche mit dem Ereignis korrespondiert, dass eine Messung zum Zeitpunkt k erfolgreich war:

$$d_k = \begin{cases} 1, & \text{Messung;} \\ 0, & \text{Fehl-messung.} \end{cases} \quad (4.39)$$

Die Wahrscheinlichkeit, dass $d_k = 1$ ist, wird über $0 \leq p_D \leq 1$ definiert, d.h. $d_k \sim \mathcal{Bi}(d_k; 1, p_D)$, wobei \mathcal{Bi} die Binomialverteilung kennzeichnet. Eine Simulation von Messaussetzern kann erreicht werden, indem im Filterungsschritt kein Update durch die aktuelle (nicht detektierte) Messung erfolgt. So definiert man bei Gauß-Filtern die Inverse der Messkovarianzmatrix \mathbf{P}^γ neu mit $(\tilde{\mathbf{P}}^\gamma)^{-1} = d_k(\mathbf{P}^\gamma)^{-1}$. Dies hat zur Folge, dass im Fall eines Messaussetzers $\tilde{\mathbf{P}}^\gamma \rightarrow \infty$ gilt und somit die Kalman-Verstärkung gegen null geht. So erfolgt kein Update durch die Messung. Bei den Partikel-Filtern wird im Fall eines Messaussetzers keine Gewichtung der prädierten Partikel vorgenommen, somit findet keine Verschiebung in Regionen höheren Likelihoods statt, was einem nicht durchgeführten Update durch die Messung entspricht. Die Abbildung 4.4 zeigt den Verlauf des Root Mean Squared Errors (RMSE) (Definition siehe Seite 74) eines Schätzers für verschiedene p_D an einem Beispiel.

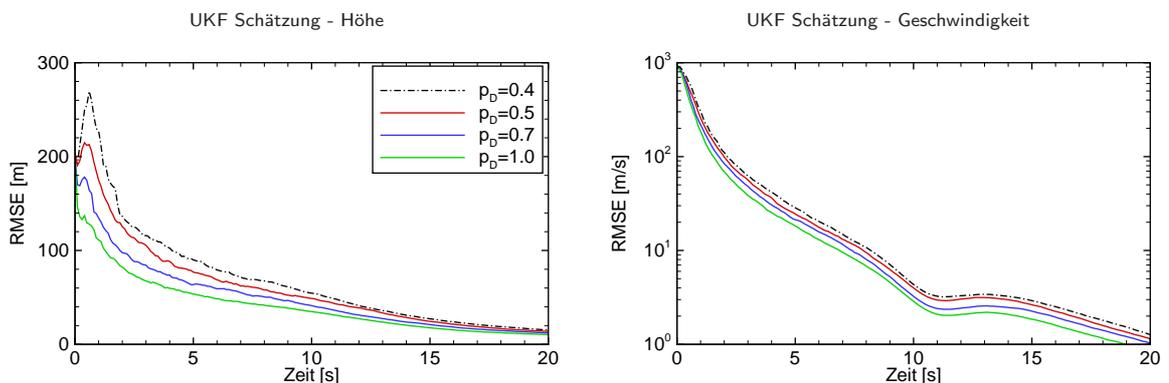


Abbildung 4.4: Root Mean Squared Error für die Schätzung der Höhe und Geschwindigkeit eines ballistischen Objekts beim Wiedereintritt in die Atmosphäre für verschiedene p_D

4.4 Wahl geeigneter Eingangsdaten (Wahl des Experiments)

Die Wahl der Eingangsdaten $U_K = \{u_1, u_2, \dots, u_K\}$ ist mit Sicherheit ein kritischer Gesichtspunkt für das Testen. Die Eingangsdaten sollten so gewählt werden, dass das System in Zustände versetzt wird, die repräsentativ sind. Sie sollten das System so anregen, dass sowohl alle Dynamiken zum Tragen kommen als auch ein möglichst repräsentatives Abbild des höchstwahrscheinlichen "Alltagsbetriebs" erzielt wird. Dies gestaltet sich als äußerst schwierig und muss von Anwendungsfall zu Anwendungsfall neu untersucht werden. Allgemeingültige Aussagen über die bestmögliche Wahl der Eingangsdaten können somit nicht formuliert werden, sondern die Wahl ist dem Ingenieur/Regelungstechniker überlassen. Für die in dieser Dissertation später vorgestellten Anwendungsbeispiele wurden die Eingangsdaten U aus folgenden Überlegungen heraus gewählt:

Regelungssysteme mit WI-Verfahren

Für geregelte Systeme ist es generell schwierig, geeignete Eingangsdaten für das System zu finden, da die Eingänge (Stellgrößen) aufgrund der Rückkopplung stark von der zu Grunde liegenden Regelung und den Schätzungen abhängen. Unter der Annahme der vollständigen Zustandsvektorrückführung (NMPC, FB-2DOF, Exakte E/A-Linearisierung,...) kann allerdings die bei einem erwartungstreuen Schätzer zu erwartende mittlere Stellgröße (entspricht U) gewonnen werden. Dies ermöglicht eine Off-Line Simulation und Evaluierung innerhalb der ReBEL-IoN Toolbox.

Systeme im Automobil, speziell modellbasierte Diagnose

Für Systeme im Automobil können geeignete Eingangsdaten U aus Standard-Fahrzyklen gewonnen werden. Diese sind darauf ausgelegt, das Automobil und somit auch das System in Zustände zu versetzen, die höchstwahrscheinlich im "Alltagsbetrieb" auftreten. Sie liefern somit reproduzierbare und repräsentative Informationen über die Systemzustände. Speziell in der Diagnose von Kfz-Komponenten (OBD, On-Board Diagnose) fordert der Gesetzgeber, dass Fehler innerhalb eines Fahrzyklusses detektiert werden müssen. Dies ist ein weiteres, nahezu zwingendes Argument für die Generierung von Eingangsdaten U aus Standard-Fahrzyklen. Weit verbreitete Fahrzyklen sind der NEDC (New European Driving Cycle), welcher aus einem Innerorts-Zyklus (UDC = Urban Driving Cycle) und einem Außerorts-Zyklus (EUDC = Extra Urban Driving Cycle) besteht, und der UDDS/FTP 72 (Urban Dynamometer Driving Schedule). Die Fahrzyklen geben ein abzufahrendes Geschwindigkeitsprofil mit spezifizierten Randbedingungen, wie Starttemperatur, Schaltpunkte, Beladung, usw. vor. Abbildung 4.5 zeigt die Geschwindigkeitsprofile des NEDC und des FTP. Der NEDC ist ein europäischer Fahrzyklus, der sowohl Stadt- als auch Überlandfahrten abdeckt. Der FTP ist der amerikanische Standard-Fahrzyklus.

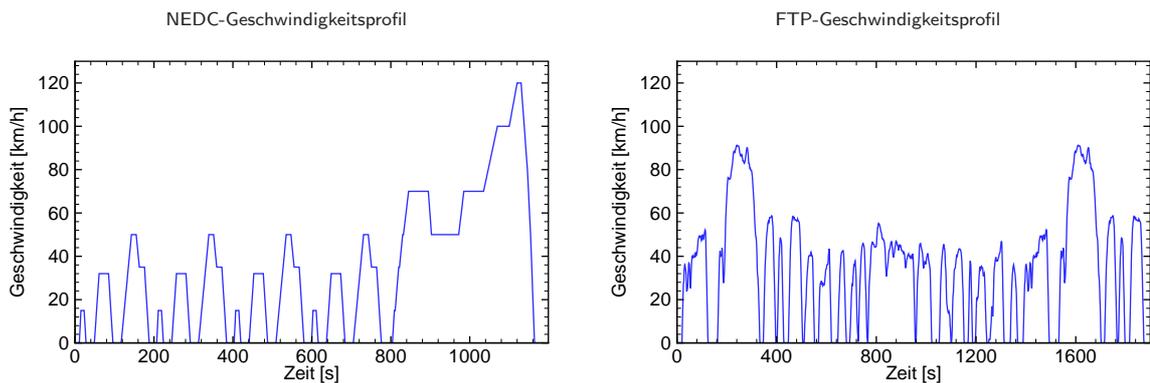


Abbildung 4.5: Geschwindigkeitsprofile NEDC und FTP.

4.5 Vorauswahl und qualitative Benchmarkkriterien

Neben quantitativen Bewertungskriterien (vorgestellt im Kapitel 4.6 ab Seite 74), die allein die Performance eines Filters bei einer Inferenzaufgabe beurteilen, existieren qualitative Bewertungskriterien, die wichtige Aspekte beleuchten, welche bei der Auswahl eines Filters zu beachten sind. Nachfolgend sind diese Kriterien aufgeführt. Einige sind in Anlehnung an [75] verfasst, wo qualitative Bewertungskriterien für gehobene Regelungsstrategien beschrieben werden.

Die vorgestellten Kriterien können sowohl als Knock-Out-Kriterien in der Vorauswahl als auch zur Ergänzung der quantitativen Bewertung verwendet werden, falls Algorithmen ähnliche Performance-Indizes liefern.

Verfügbarkeit und Differenzierbarkeit der Systemgleichungen - "Black-Box-Modelle"

Liegen die System- und/oder Messgleichung in einer Form vor, die keine analytische Differenzierung nach den Zuständen zulässt (Stichwort: "Black-Box-Modelle"), so können alle Verfahren, welche auf Taylor-Reihen-Linearisierung basieren (Lokal-Methoden), die benötigten Differentialmatrizen nur durch Differenzenmatrizen approximieren. Je nach System kann die Approximierung sehr sensitiv gegenüber Variationen in der Schrittweite der Differenzenapproximation sein. Daher sollten solche Verfahren nach Möglichkeit a priori ausgeschlossen werden.

Art des Schätzproblems (Zustands-, Parameter- oder Duale Schätzung)

Im Falle reiner Zustands- oder Parameterschätzung (kein duales Schätzproblem) entfallen alle Algorithmen, welche eine Entkopplung von Zustands- und Parametervektor vorsehen.

Rein additives System- und Messrauschen

Sind das System- und Messrauschen rein additiv, d.h. $K = \Gamma = \mathbf{I}$, so entfallen alle Algorithmen, welche eine Erweiterung des Zustandsvektors durch das System- und Messrauschen vorsehen ((SR)UKF, (SR)CDKF)

und deren vereinfachte Form kommt zum Einsatz (siehe dazu den Anhang A, ab Seite 145).

4.5.1 Qualifikationsanforderungen an den Anwender während der Einsatzvorbereitung

Die Lösung der nichtlinearen Schätz Aufgabe kann wie gezeigt durch eine Vielzahl von mehr oder weniger komplexen Algorithmen erfolgen. Dabei ist je nach Komplexitätsgrad des Algorithmus und Aufbereitung die Qualifikationsanforderung an den Benutzer unterschiedlich hoch. Es sei angemerkt, dass komplexe Algorithmen nicht unbedingt eine hohe Qualifikationsanforderung haben, sofern sie verständlich und ausführlich aufbereitet vorliegen. Die nachfolgend aufgeführten Bewertungen bieten einen Anhaltspunkt für die Anforderungen an den Anwender.

	Einordnung			
	Sehr Hoch	Hoch	Mittel	Niedrig
Eigenschaft	Bestimmung der Verteilungen des System- und Messrauschens $p(\kappa)$ und $p(\gamma)$ sowie Herleitung der Transitionsverteilung $p(x_k x_{k-1})$ und der Likelihoodverteilung $p(y_k x_k)$	Bestimmung der Verteilungen des System- und Messrauschens $p(\kappa)$ und $p(\gamma)$ sowie Herleitung der Likelihoodverteilung $p(y_k x_k)$	Bestimmung des Mittelwerts und der Kovarianz der Verteilungen des System- und Messrauschens κ und γ sowie Herleitung von Differentialmatrizen der System- und Messgleichung \mathcal{F} und \mathcal{H}	Bestimmung des Mittelwerts und der Kovarianz der Verteilungen des System- und Messrauschens κ und γ
Verfahren	SPPF, EKPF, KPF, RPF(mit MCMC)	Restlichen Partikel-Filter, GSKF, GSPF und GMSPPF	Alle Lokal-Methoden	Quadratur-Methoden

Tabelle 4.1: Einstufung der Qualifikationsanforderungen an den Anwender während der Einsatzvorbereitung.

4.5.2 Aufwand und Unterstützung beim Entwurf des Filters

Die Anzahl der Einstellparameter variiert von Filter zu Filter. Einige Algorithmen kommen ohne Einstellparameter aus, für einige gibt es Empfehlungen für die Einstellungen, allerdings gibt es auch Algorithmen, für die nicht einmal Empfehlungen für die Einstellung von Parametern existieren. Das Kriterium "Entwurfsaufwand" bewertet zum einen die Anzahl der einzustellenden Parameter und zum anderen, ob Einstellregeln für die Parameter existieren. Die Tabelle 4.2 listet für die ausgewählten Verfahren die Anzahl der Einstellparameter auf. Der EKF besitzt überhaupt keine Einstellparameter und dominiert dieses Kriterium. Beim Entwurf eines GMSPPF mit (SR)UKF als Hybridkomponente müssen die Anzahl der GMM-Komponenten für Zustand, System- und Messrauschen, die Anzahl der Partikel sowie die drei SUT-Parameter eingestellt werden. In der Summe sind das sieben Parameter, was den Algorithmus ans Ende des Feldes versetzt.

Anzahl Parameter	Inferenzverfahren
0	KF, (SR)EKF _(Entkoppelt)
1	(SR)IEKF ₁ , IEKF ₂ , (SR)AEKF, (SR)CDKF _(Entkoppelt) , GHF, ASIR
2	GPF, ICDKF, EKPF
3	(SR)UKF _(Entkoppelt) , RPF, GSPF
4	KPF, AKPF, IUKF
2+	SPPF
3+	GSKF
4+	GMSPPF

Das + deutet an, dass je nach Wahl des integrierten GAUSS'schen Filters die entsprechenden Parameter hinzukommen

Tabelle 4.2: Anzahl der einzustellenden Parameter beim Filterentwurf.

Neben der reinen Parameteranzahl muss unterschieden werden, ob die Parameter willkürlich einstellbar sind (trial and error) oder ob Empfehlungen existieren. Tabelle 4.3 ordnet die Algorithmen in entsprechende Kategorien ein.

	Einordnung		
	Schlecht	Mittel	Gut
Eigenschaft	Keine oder nur teilweise Empfehlungen vorhanden	Empfehlungen vorhanden	Keine Einstellparameter vorhanden
Verfahren	Alle Partikel-Filter, abgeleitete Hybride und Gauß-Summen-Filter	(SR)UKF, (SR)CDKF, (SR)IEKF ₁ , IEKF ₂ , (SR)AEKF, GHF, IUKF, ICDKF	KF, (SR)EKF

Tabelle 4.3: Einstufung der Unterstützung beim Filterentwurf.

4.5.3 Installations- und Implementierungsaufwand

Der Installations- und Implementierungsaufwand beinhaltet alle Arbeiten, die ausgehend von einer funktionierenden Systemperipherie erbracht werden müssen, um die Filteralgorithmen betreiben zu können. Dabei werden funktionierende Sensoren und der Zugriff auf die Messdaten vorausgesetzt.

Die Komplexität der Algebraoperationen variiert stark von Filter zu Filter, wodurch der Aufwand stark von der Unterstützung dieser Operationen durch die Zielplattform abhängig ist. Dies macht es schwer bzw. unmöglich, eine allgemeingültige Einordnung der Algorithmen zu gewinnen. Grundsätzlich gilt nach Meinung des Autors ungefähr, dass der Aufwand zur Implementierung wie folgt verläuft (ansteigend):

Lokal-Methoden → Partikel-Filter → Quadratur-Methoden → Gauß-Summen-Filter → Hybride

4.6 Numerische Bewertung über quantitative Benchmarkkriterien

Die Auswahl eines Filteralgorithmus hängt von einer Vielzahl von Faktoren ab. Die Basis für die Auswahl eines Algorithmus sollen quantitative und qualitative Bewertungskriterien bilden. Zur Unterstützung bei der Gewichtung der Bewertungskriterien kann vom Anwender nachfolgendes, hierarchisch geordnetes Lastenheft herangezogen werden:

1. Hohe Filtergüte und hohe Robustheit
2. Einhaltung verfügbarer "Ressourcen" (z.B. Anwenderqualifikation)
3. Geringer Implementierungsaufwand

Die wichtigste Aufgabe des Filters ist es, die verborgenen, unverfälschten (nicht durch System- und Messrauschen korrumpierten) Zustände des Systems möglichst genau, schnell und robust zu schätzen. Für die Bewertung der Qualität dieser Schätzung stehen einige quantitative Kriterien zur Verfügung, welche nachfolgend aufgeführt sind.

4.6.1 Quantitativer Benchmark: Filtergüte über Posterior Cramér-Rao Bound

Die Filtergüte ist zweifelsohne das wichtigste quantitative Kriterium zur Beurteilung von Filteralgorithmen. In den meisten Fällen wird der Root Mean Squared Error der Schätzung zum Zeitpunkt k , gemittelt über MC Monte-Carlo-Simulationen berechnet, wobei $RMSE = \sqrt{\frac{1}{MC} \sum_{i=1}^{MC} (\hat{x}_k^{(i)} - x_k^{(i)})^2}$ ist, und darüber werden die Verfahren miteinander verglichen. Der erhebliche Nachteil besteht darin, dass man so zwar eine Aussage darüber treffen kann, welcher der Algorithmus mit dem kleinsten Schätzfehler ist, jedoch ob die Schätzung generell gut ist, bleibt weiter ungewiss. Man erhält also nur eine relative Aussage über die Filtergüte. Wünschenswert ist aber auch, eine absolute Aussage über die Leistung eines Inferenzverfahrens treffen zu können. Der Posterior Cramér-Rao Bound (PCRB) liefert einen unteren Grenzwert für den RMSE eines erwartungstreuen Schätzers. Die Verzerrung ($BIAS$) eines Schätzers $\Psi_{X_k}(Y_k)$ der Zustandstrajektorie X_k mit Messungen Y_k sei nach [76] definiert durch

$$BIAS(X_k) = \mathbb{E}[\Psi_{X_k}(Y_k) - X_k | X_k], \quad (4.40)$$

d.h. ein Schätzer $\Psi_{X_k}(Y_k)$ der Trajektorie X_k ist erwartungstreu, falls $BIAS(X_k) \cong 0$ ist. Der PCRB liefert einen unteren Grenzwert für den RMSE, sofern die Annahme der asymptotischen Erwartungstreue, definiert durch

$$\lim_{X_j^{(i)} \rightarrow \mathfrak{X}^{(i),+}} BIAS(X_k)p(X_k) = \lim_{X_j^{(i)} \rightarrow \mathfrak{X}^{(i),-}} BIAS(X_k)p(X_k) \quad (4.41)$$

$$\forall j \in \{0, \dots, k\} \quad , \quad \forall i \in \{1, \dots, n\},$$

wobei $\mathfrak{X}^{(i)}$ der Wertebereich von $X_j^{(i)} \forall j \in \{0, \dots, k\}$ ist und $\{\mathfrak{X}^{(i),-}, \mathfrak{X}^{(i),+}\}$ die Bereichsgrenzen darstellen, erfüllt ist. Ist die Annahme erfüllt, kann der PCRB theoretisch nicht von den Inferenzverfahren

unterschriften werden. An dieser Stelle sei angemerkt, dass der PCRB lediglich einen Grenzwert für Fehler zweiter Ordnung liefert. Die posteriori Verteilungsdichte $p(x_k|Y_k)$ ist generell nicht normalverteilt und hat auch höhere Momente (als das zweite Moment). Eine vollständige statistische Bewertung kann somit nicht mit dem PCRB erzielt werden, falls $p(x_k|Y_k)$ nicht normalverteilt ist. Trotz der o.a. Einschränkungen ist der PCRB in vielen Anwendungen ein wirksames Verfahren zur Charakterisierung von Filtergüte.

Es gilt also [27]

$$\sqrt{\frac{1}{MC} \sum_{i=1}^{MC} (\hat{x}_k^{(i)} - x_k^{(i)})^2} \gtrsim \sqrt{\text{diag}(\text{PCRB}_k)}, \quad (4.42)$$

wobei \gtrsim andeutet, dass die Ungleichung für finite MC näherungsweise erfüllt ist. PCRB_k entspricht der Inversen der Filter-Informationsmatrix J_k , welche der Rekursion

$$J_k = D_{k-1}^{22} - D_{k-1}^{21} (J_{k-1} + D_{k-1}^{11})^{-1} D_{k-1}^{12} \quad (4.43)$$

mit

$$D_{k-1}^{11} = \mathbb{E} \left[-\Delta_{x_{k-1}}^{x_{k-1}} \log p(x_k | x_{k-1}) \right] \quad (4.44)$$

$$D_{k-1}^{12} = \mathbb{E} \left[-\Delta_{x_{k-1}}^{x_k} \log p(x_k | x_{k-1}) \right] \quad (4.45)$$

$$D_{k-1}^{21} = \mathbb{E} \left[-\Delta_{x_k}^{x_{k-1}} \log p(x_k | x_{k-1}) \right] = [D_{k-1}^{12}]^T \quad (4.46)$$

$$D_{k-1}^{22} = \mathbb{E} \left[-\Delta_{x_k}^{x_k} \log p(x_k | x_{k-1}) \right] \\ + \mathbb{E} \left[-\Delta_{x_k}^{x_k} \log p(y_k | x_k) \right] \quad (4.47)$$

genügt. Dabei wird die Rekursion initialisiert mit $J_0 = \mathbb{E} \left[-\Delta_{x_0}^{x_0} \log p(x_0) \right]$. Für eine Herleitung sei auf [77] verwiesen. ∇ und Δ sind Operatoren für die partiellen Ableitungen erster bzw. zweiter Ordnung einer Funktion, d.h.

$$\nabla_\psi = \left[\frac{\partial}{\partial \psi_1}, \dots, \frac{\partial}{\partial \psi_n} \right]^T \quad (4.48)$$

$$\Delta_\eta^\psi = \nabla_\eta \nabla_\psi^T \\ = \frac{\partial^2}{\partial \eta_i \partial \psi_j} \quad i = j = 1, \dots, n. \quad (4.49)$$

Die rekursive Gleichung zur Bestimmung des PCRB erfordert die Berechnung einiger sehr komplexer Erwartungswerte. Für GAUSS'sche Systeme mit additivem Rauschen können einige dieser Erwartungswerte analytisch bestimmt werden. Im Allgemeinen müssen sie aber durch Monte-Carlo-Simulation approximiert werden. In [36, 77] sind einige dieser Spezialfälle aufgeführt. Die Abbildung 4.6 zeigt für ein lineares und ein nichtlineares Beispiel den zeitlichen Verlauf des PCRB und des RMSE verschiedener Filteralgorithmen (im linearen nur Kalman Filter).

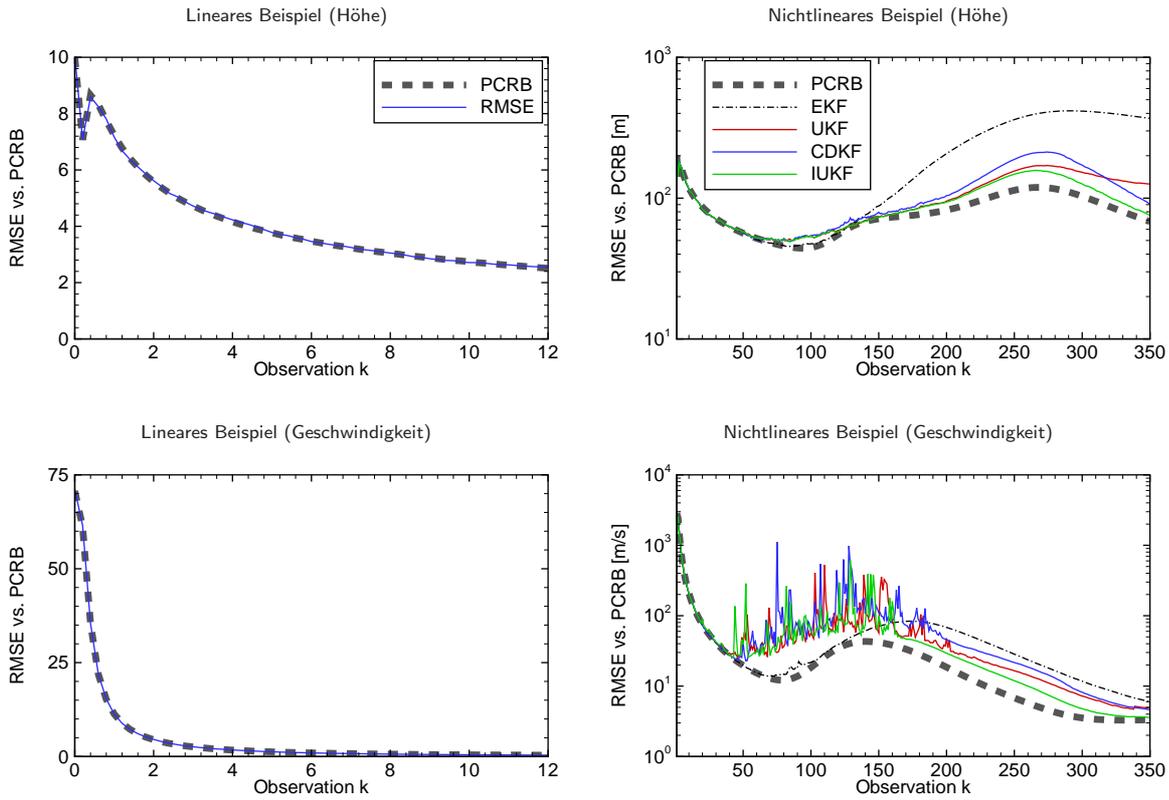


Abbildung 4.6: Links: RMSE und PCRB des Kalman Filters für die Schätzung der Höhe und Geschwindigkeit eines fallenden ballistischen Objekts mit konstanter Beschleunigung (Erdbeschleunigung) im konstanten Dichtefeld. Das Beispiel aus [9, Seite 287] wurde leicht abgewandelt. Rechts: RMSE und PCRB für die Schätzung der Höhe und Geschwindigkeit eines fallenden ballistischen Objekts mit konstanter Beschleunigung (Erdbeschleunigung) im variablen Dichtefeld mit nichtlinearer Messgleichung (Beispiel siehe Kapitel 5).

PCRB für die Parameterschätzung zeitinvarianter Parameter

Die rekursive Bestimmungsgleichung der Filter-Informationsmatrix J_k für Parameterschätzung zeitinvarianter Parameter wird nachfolgend hergeleitet:

Theorem 5. Für die Parameterschätzung zeitinvarianter Parameter ist die rekursive Bestimmungsgleichung der Filter-Informationsmatrix J_k durch

$$J_k = \mathbb{E} \left[-\Delta_{x_k}^{x_k} \log p(y_k | x_k) \right] + J_{k-1} \quad (4.50)$$

gegeben!

Beweis 5. Für die Parameter sei ein Variationsmodell der Form

$$\theta_k = \theta_{k-1} + \tilde{\kappa}_{k-1} \quad (4.51)$$

angenommen, wobei $\tilde{\kappa}_{k-1} \sim \mathcal{N}(\tilde{\kappa}; 0, \mathbf{P}^{\tilde{\kappa}}) \forall k$ sei. Unter dieser Annahme gilt für die Transitionsdichte

und deren logarithmierter Form ($\theta \rightarrow x$)

$$p(x_k|x_{k-1}) = \frac{1}{\sqrt{(2\pi)^n |\mathbf{P}^{\tilde{\kappa}}|}} \cdot \exp \left[-\frac{1}{2} (x_k - \mathbf{I}x_{k-1})^T (\mathbf{P}^{\tilde{\kappa}})^{-1} (x_k - \mathbf{I}x_{k-1}) \right] \quad (4.52)$$

$$-\log p(x_k|x_{k-1}) = \underbrace{-\frac{1}{\sqrt{(2\pi)^n |\mathbf{P}^{\tilde{\kappa}}|}}}_{\text{konst.}} + \frac{1}{2} (x_k - \mathbf{I}x_{k-1})^T (\mathbf{P}^{\tilde{\kappa}})^{-1} (x_k - \mathbf{I}x_{k-1}). \quad (4.53)$$

Wendet man nun die benötigten Delta-Operatoren an, so erhält man

$$-\Delta_{x_{k-1}}^{x_{k-1}} \log p(x_k|x_{k-1}) = (\mathbf{P}^{\tilde{\kappa}})^{-1} \quad (4.54)$$

$$-\Delta_{x_{k-1}}^{x_k} \log p(x_k|x_{k-1}) = -(\mathbf{P}^{\tilde{\kappa}})^{-1} \quad (4.55)$$

$$-\Delta_{x_k}^{x_k} \log p(x_k|x_{k-1}) = (\mathbf{P}^{\tilde{\kappa}})^{-1}. \quad (4.56)$$

Da die Kovarianzmatrix $\mathbf{P}^{\tilde{\kappa}}$ konstant ist, fallen die Erwartungswert-Operatoren $\mathbb{E}[\cdot]$ weg (außer für die Likelihood-Verteilung). Da die Kovarianzmatrix $\mathbf{P}^{\tilde{\kappa}}$ symmetrisch ist, gilt ferner $((\mathbf{P}^{\tilde{\kappa}})^{-1})^T = ((\mathbf{P}^{\tilde{\kappa}})^T)^{-1} = (\mathbf{P}^{\tilde{\kappa}})^{-1}$ und es resultiert folgende Rekursionsgleichung für die Filter-Informationsmatrix

$$J_k = (\mathbf{P}^{\tilde{\kappa}})^{-1} + \mathbb{E}[-\Delta_{x_k}^{x_k} \log p(y_k|x_k)] - (\mathbf{P}^{\tilde{\kappa}})^{-1} (J_{k-1} + (\mathbf{P}^{\tilde{\kappa}})^{-1})^{-1} (\mathbf{P}^{\tilde{\kappa}})^{-1}. \quad (4.57)$$

Gesucht ist jetzt der Grenzwert von J_k für $\mathbf{P}^{\tilde{\kappa}} \rightarrow 0$, da zufällige, aber zeitinvariante Parameter über kein Systemrauschen verfügen:

$$\begin{aligned} \lim_{\mathbf{P}^{\tilde{\kappa}} \rightarrow 0} J_k &= \lim_{\mathbf{P}^{\tilde{\kappa}} \rightarrow 0} (\mathbf{P}^{\tilde{\kappa}})^{-1} + \mathbb{E}[-\Delta_{x_k}^{x_k} \log p(y_k|x_k)] - (\mathbf{P}^{\tilde{\kappa}})^{-1} (J_{k-1} + (\mathbf{P}^{\tilde{\kappa}})^{-1})^{-1} (\mathbf{P}^{\tilde{\kappa}})^{-1} \\ &\stackrel{MIL}{=} \lim_{\mathbf{P}^{\tilde{\kappa}} \rightarrow 0} \mathbb{E}[-\Delta_{x_k}^{x_k} \log p(y_k|x_k)] + (J_{k-1}^{-1} + \mathbf{P}^{\tilde{\kappa}})^{-1} \\ &= \mathbb{E}[-\Delta_{x_k}^{x_k} \log p(y_k|x_k)] + (J_{k-1}^{-1})^{-1} \\ &= \mathbb{E}[-\Delta_{x_k}^{x_k} \log p(y_k|x_k)] + J_{k-1} \end{aligned} \quad (4.58)$$

$$\mathbf{q.e.d.} \quad (4.59)$$

Dabei deutet MIL die Verwendung des Matrix-Inversions-Lemma an, siehe z.B. [36, Seite 47].

Zur Berechnung der Filter-Informationsmatrix bzw. des PCRb muss somit lediglich $\mathbb{E}[-\Delta_{x_k}^{x_k} \log p(y_k|x_k)]$ berechnet werden. Die Initialisierung der Rekursion erfolgt analog zur Zustandsfilterung (s.o.). Die Abbildung 4.7 zeigt den Verlauf des Parameter-PCRb sowie die exogenen Eingänge u zur Schätzung der Parameter der quadratischen Kennlinie $y = \theta_1 u + \theta_2 u^2 + \gamma_k = [u \ u^2][\theta_1 \ \theta_2]^T + \gamma_k$, wobei $\gamma_k \sim \mathcal{N}(\gamma; 0, 6)$ und $\theta_0 \sim \mathcal{N}(\theta_0; [0.2 \ 0.1]^T, \mathbf{I})$. Das System (die "Messgleichung") ist linear in den Parametern, d.h. der lineare Kalman-Filter löst die RBS für dieses Beispiel exakt. Die Verläufe des RMSE des Kalman-Filters sowie des neuen PCRb sind identisch, womit noch einmal an einem einfachen Beispiel die Korrektheit demonstriert wird.

PCRb für $p_D < 1$

Wie bereits zuvor angedeutet (Abbildung 4.4), wird die Güte der Inferenzverfahren stark durch Messaussetzer beeinflusst. In [78] wird der theoretische PCRb für Systeme mit $p_D < 1$ hergeleitet. Wie dort gezeigt,

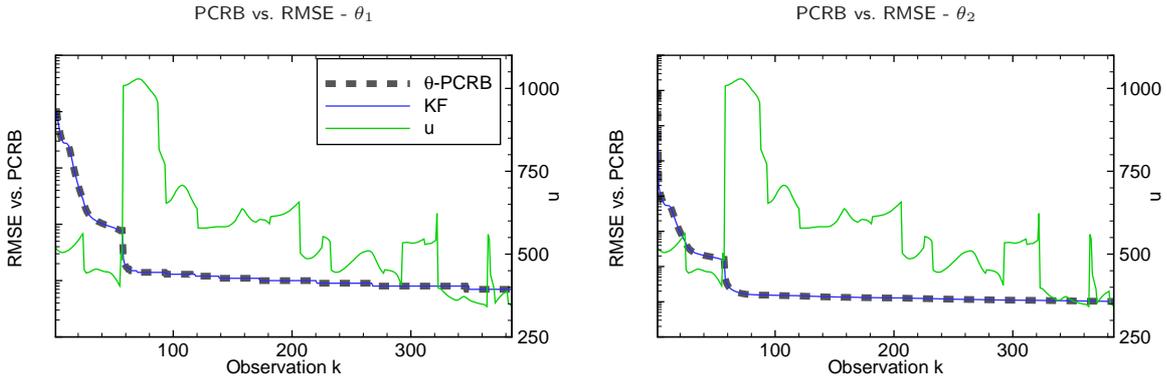


Abbildung 4.7: Vergleich des neuen Parameter-PCRB mit dem RMSE des KF am Beispiel der Parameterschätzung einer quadratischen Kennlinie.

wächst der Rechenaufwand zur Berechnung des theoretischen PCRB für $k \gg 1$ (k : Zeitindex) exponentiell an. Zum Zeitindex k können 2^k verschiedene Detektions-/Miss-Sequenzen nach

$$\mathcal{S}_k^l : d_{i_1}^l, d_{i_2}^l, \dots, d_{i_k}^l \quad (4.60)$$

definiert werden, wobei $i_{\tilde{k}}$ der Index des Detektions-/Miss-Ereignisses zum Zeitpunkt $\tilde{k} = 1, 2, \dots, k$ für eine Sequenz $l = 1, 2, \dots, 2^k$ ist. Die Anzahl an korrekten Detektionen in einer Sequenz \mathcal{S}_k^l ist gegeben durch

$$\Delta_k^l = \sum_{\tilde{k}=1}^k d_{i_{\tilde{k}}}^l. \quad (4.61)$$

Die Anzahl an Nicht-Detektionen ist somit $\bar{\Delta}_k^l = k - \Delta_k^l$. Für eine gegebene Detektionswahrscheinlichkeit p_D , konstant über k , ist die Wahrscheinlichkeit für das Auftreten einer Detektions/Miss-Sequenz somit

$$Pr\{\mathcal{S}_k^l\} = \binom{k}{\Delta_k^l} p_D^{\Delta_k^l} \cdot (1 - p_D)^{k - \Delta_k^l} = \binom{k}{\Delta_k^l} p_D^{\Delta_k^l} \cdot (1 - p_D)^{\bar{\Delta}_k^l}. \quad (4.62)$$

Bei einer gegebenen Detektions-/Miss-Sequenz \mathcal{S}_k^l ist die Kovarianz einer erwartungstreuen Schätzung nach unten beschränkt, gemäß

$$\mathbb{E}\left[(x_k - \hat{x}_k)(x_k - \hat{x}_k)^T | \mathcal{S}_k^l\right] \geq J_k^{-1}(\mathcal{S}_k^l). \quad (4.63)$$

Dabei ist nun $J_k(\mathcal{S}_k^l) = PC\tilde{R}B^{-1}$ die durch die Sequenz \mathcal{S}_k^l bedingte Filter-Informationsmatrix. Die Bestimmung des unbedingten PCRB erfordert die Berechnung des Erwartungswertes des bedingten $PC\tilde{R}B$ unter Berücksichtigung von \mathcal{S}_k^l , d.h.

$$PCRB = \sum_{l=1}^{2^k} PC\tilde{R}B \cdot Pr\{\mathcal{S}_k^l\}. \quad (4.64)$$

In der Praxis ist dieser Erwartungswert nicht mehr berechenbar (z.B.: $k = 50$, $\rightarrow 1.1259 \cdot 10^{15}$ Sequenzen). Approximationen werden in [78] vorgestellt, doch auch diese erfordern schon für kleine k die Berechnung

sehr vieler Sequenzen \mathcal{S}_k^l und der zugehörigen Filter-Informationsmatrix $J_k(\mathcal{S}_k^l)$ (z.B.: $k = 50, p_D = 0.7 \rightarrow 2.702 \cdot 10^{14}$ Sequenzen [78]).

Aus diesem Grund wird im Folgenden auf Berechnungen des exakten PCRB für $p_D < 1$ verzichtet. Der Standard-PCRb liefert dennoch weiterhin einen unteren Grenzwert, da für $p_D < 1$ der PCRb analog zum RMSE der Schätzungen wächst (vgl. Abbildung 4.4). Somit bleibt man mit dem PCRb für $p_D = 1$ immer auf der sicheren Seite, d.h. die Beschränkung des RMSE bleibt garantiert erhalten.

PCRb für verschiedene Modelle

Finden verschiedene Modelle zur Trajektoriengenerierung und -schätzung Verwendung, wird der PCRb mit dem Referenzmodell berechnet. So bleibt man auch auf der sicheren Seite.

Definition Filterwirkungsgrad

Wie zuvor hergeleitet und in Abbildung 4.6 zu sehen, liefert der PCRb einen unteren Grenzwert für den RMSE von \hat{x}_k . Dies ermöglicht die Definition des zeitabhängigen Filterwirkungsgrades η_k :

$$\eta_k = \frac{\sqrt{PCRb_k}}{RMSE_k} \times 100\%. \quad (4.65)$$

Der so definierte Wirkungsgrad ist beschränkt mit $0\% \leq \eta_k \lesssim 100\%$, wobei $\eta_k \cong 100\%$ einen effizienten (Güte-optimalen) Filter kennzeichnet. In der Abbildung 4.8 sind die Verläufe der Wirkungsgrade für die zwei oben beschriebenen (Abbildung 4.6) nichtlinearen Beispiele gezeigt. Im linearen Fall ist der Wirkungsgrad $\eta_k = 100\%$. Die durch die Rekursion berechneten Werte für J_k entsprechen der Inversen der Kovarianzmatrix $\mathbf{P}_{k|k}^x$ des KF. Somit ist nochmals an dieser Stelle gezeigt, dass der lineare KF der optimale Filter für lineare GAUSS'sche Systeme ist.

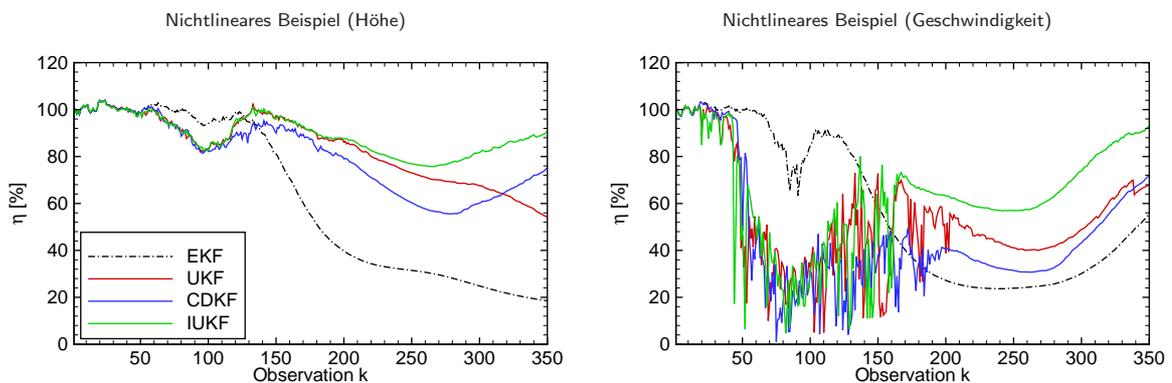


Abbildung 4.8: Wirkungsgrade der oben gezeigten nichtlinearen Beispiele.

Als weiteres interessantes Kriterium sei noch die relative Verbesserung RI (Relative Improvement) ein-

geführt:

$$RTAMSE = \sqrt{\frac{1}{K \cdot MC} \sum_{k=1}^K \sum_{i=1}^{MC} \left(\hat{x}_k^{(i)} - x_k^{(i)} \right)^2} \quad (4.66)$$

$$RI = \frac{RTAMSE_{Ref} - RTAMSE_{Filter}}{RTAMSE_{Ref}} \times 100\%. \quad (4.67)$$

Dabei ist $RTAMSE_{Ref}$ der Root Time Averaged Mean Squared Error eines Referenzfilters, z.B. EKF. K bezeichnet die Gesamtanzahl an Observationen.

Im Falle stückweit definierter Systemmodelle kann ein exakter PCRb nicht mehr berechnet werden, da die partiellen Ableitungen der logarithmierten Verteilungsdichten benötigt werden. Als Alternative können allgemeinere Grenzwerte wie der BHATTACHARYA-, der BOBROVSKY-ZAKAI- oder der WEISS-WEINSTEIN-Grenzwert (siehe [79]) berechnet werden. Dies ist allerdings mit einem stark erhöhten Rechenaufwand verbunden.

4.6.2 Quantitativer Benchmark: Filterrobustheit (Beschränkungssicherheit)

Die Sicherheit auf Beschränkung des Schätzfehlers, nachfolgend Filterrobustheit genannt, ist ein maßgebendes Kriterium für die Brauchbarkeit eines Filters. Die Modellfehler resultieren in der Regel aus nicht bzw. schlecht modellierten Systemdynamiken und -statiken sowie aus nicht erkannten Parameteränderungen im Betrieb, welche maßgebenden Einfluss auf das Systemverhalten haben.

Analytische Robustheitsuntersuchung

Die Robustheit und Konvergenz von Inferenzverfahren ist bisher nicht ausreichend untersucht worden. Lediglich in Spezialfällen wurden Robustheits- und Konvergenzkriterien untersucht, z.B. in [80] für Parameterschätzer linearer Systeme oder in [81] und [82] für Beobachter nichtlinearer Systeme. In [83] wurde das Robustheitsverhalten des zeitkontinuierlichen und in [73] des zeitdiskreten Extended Kalman Filters untersucht. Es konnte gezeigt werden, dass der Schätzfehler des zeitkontinuierlichen und zeitdiskreten EKF unter bestimmten Bedingungen beschränkt bleibt. Dies ist besonders bei hinreichend kleinen Anfangsschätzfehlern und hinreichend kleinem Rauschen gewährleistet. Prinzipiell ist eine analytische Robustheitsuntersuchung wünschenswert, diese ist allerdings sehr zeitaufwendig, wenn mehrere Algorithmen miteinander verglichen werden sollen. Allgemeine Kennwerte, die eine vergleichende Aussage über Robustheit der Verfahren liefern, existieren auch nicht.

Empirische Robustheitsuntersuchung

Um dennoch eine Aussage über die Robustheit eines Filters treffen zu können, kann man die Anzahl der divergierten Monte-Carlo-Simulationen pro Test als Anhaltspunkt für die Robustheit eines Filters nehmen. Eine Schätzung wird als divergiert erklärt, wenn die Algorithmen numerisch divergieren z.B. wenn

Kovarianzmatrizen ihre positive Definitheit verlieren oder wenn die Abweichung zwischen Referenz- und geschätzter Trajektorie einen benutzerdefinierten Grenzwert überschreitet. Somit wird definiert:

$$J_{Robustness} = \left(1 - \frac{\#Divergenz}{MC}\right) \times 100\%. \quad (4.68)$$

Man erhält so zwar keine Aussage über die Entfernung zur Robustheitsgrenze, allerdings einen empirisch gewonnenen Wert über die Robustheit und Beschränkung des Schätzfehlers, der beschränkt ist mit $0\% \leq J_{Robustness} \leq 100\%$, wobei $J_{Robustness} = 100\%$ einen robusten Filter kennzeichnet.

4.6.3 Quantitativer Benchmark: Rechenzeit

Die unterschiedliche Rechenkomplexität der Inferenzalgorithmen resultiert in unterschiedlicher Rechenzeit. Ein wichtiges Kriterium für die Tauglichkeit eines Filters ist, dass die Rechenanforderungen zur Verfügung stehende Kapazitäten nicht überschreiten und die Echtzeitfähigkeit des Algorithmus nicht gefährdet wird. Die hier durchgeführten Untersuchungen haben gezeigt, dass die benötigte Rechenzeit nicht nur von der Taktfrequenz der Rechner, sondern auch von der Rechnerarchitektur bestimmt wird. Die Rechengeschwindigkeit ist somit nur schwer von einer Plattform auf eine andere übertragbar, d.h. um eine absolute Aussage treffen zu können, ob ein Algorithmus echtzeitfähig ist oder nicht, muss dieser auf der Zielplattform implementiert und getestet werden. Untersuchungen haben allerdings auch gezeigt, dass die Ordnung der relativen Rechenzeit nahezu plattformunabhängig ist. Somit lässt sich ein relatives Gütekriterium für die Rechenanforderung festlegen mit

$$TI = \frac{\sum_{i=1}^{MC} t_{Sim,Ref}^{(i)} - \sum_{i=1}^{MC} t_{Sim,Filter}^{(i)}}{\sum_{i=1}^{MC} t_{Sim,Ref}^{(i)}} \times 100\%, \quad (4.69)$$

wobei $t_{Sim,Ref}$ die Simulationszeit eines Referenz-Algorithmus (i.d.R. EKF) bezeichnet. Somit erhält man zumindest eine Aussage, wie sich die Inferenzverfahren untereinander im Rechenaufwand unterscheiden.²

4.7 Matlab[®] basierte Toolbox ReBEL-IoN

Im Rahmen dieser Arbeit wurde eine MATLAB[®] basierte Toolbox ReBEL-IoN entwickelt. ReBEL-IoN ist eine Toolbox mit graphischer Benutzeroberfläche zum Testen, Evaluieren und Selektieren von wahrscheinlichkeitstheoretischen Inferenzverfahren. Der modulare Aufbau von ReBEL-IoN wurde inspiriert vom ReBEL Toolkit, welches von v. D. MERWE&WAN [84] entwickelt wurde. ReBEL-IoN verfügt über folgende besondere Eigenschaften:

1. Simulations- und Bewertungskonzept:

Das vorgestellte Simulations- und Bewertungskonzept ist in dieser Software umgesetzt worden. Dabei

²Aufgrund des modularen Aufbaus der MATLAB[®] Toolbox ReBEL-IoN wird Rechenkapazität eingebüßt, d.h. bei einer losgelösten Algorithmenimplementierung stimmen die von ReBEL-IoN gelieferten Kennwerte über den Rechenaufwand nicht zwingend mit den realen überein.

wurden die beschriebene Evaluierungsroutine (Monte-Carlo-Simulation) und die automatische Generierung der quantitativen Kennwerte *Filtergüte*, *Robustheit* und *Rechenanforderung* implementiert. Hauptarbeitspunkt war dabei die Entwicklung der kompletten Organisation des benötigten Datentransfers innerhalb der Monte-Carlo-Simulation. Es erfolgt ein Ranking der getesteten Verfahren bezüglich der o.a. Kriterien.

2. **Modularer Aufbau:**

ReBEL-IoN ist modular aufgebaut, so dass der Anwender sich vollständig auf Modellierung des stochastischen Systems (Referenz- und Filtermodell) konzentrieren kann, ohne Zeit zur Implementierung von Filteralgorithmen oder Evaluierungsroutinen zu "vergeuden". Dabei greifen alle Filter auf ein Filtermodell zu, welches zusammen mit dem Referenzmodell vom Anwender in einem einzigen m-File abgelegt werden muss.

3. **Pool an Verteilungen:**

Um ein möglichst breites Spektrum an stochastischen DSSM abdecken zu können, bietet ReBEL-IoN die Möglichkeit, unter einer Vielzahl von uni- und multivariaten Verteilungen für System- und Messrauschen sowie Initialverteilungen zu wählen.

4. **Simulation von Messaussetzern:**

Durch Vorgabe der POD werden Messaussetzer in den quantitativen Bewertungsprozess integriert.

5. **Gleichheits- und Ungleichheitsbeschränkungen:**

In vielen technischen Anwendungen treten Gleichheits- und Ungleichheitsbeschränkungen (nicht negative Konzentrationen, Summe von Molenbrüchen gleich eins) auf, die in der Software benutzerfreundlich vorgegeben werden können und in den Schätzalgorithmen (GF) umgesetzt werden.

6. **Graphische Benutzeroberfläche:**

Zur intuitiven Anwendung der Software wurde eine graphische Benutzeroberfläche in MATLAB[®] entwickelt, siehe Abbildung 4.9. Auf eine detaillierte Dokumentation des GUIs wird an dieser Stelle verzichtet.

7. **Post Processing:**

Die untersuchten Algorithmen können aus der Simulationsumgebung exportiert und anhand "realer" Messdaten getestet werden. Dies erlaubt eine Verifikation der simulativ gewonnenen Bewertungsergebnisse und eine Überprüfung der Brauchbarkeit in der realen Anwendung.

In der aktuellen Version³ beinhaltet ReBEL-IoN folgende Algorithmen:

³Diese Arbeit wurde im März 2008 publiziert. Die zu diesem Zeitpunkt aktuellste Version war ReBEL-IoN v1.0.1.

- (SR)AEKF
- (SR)EKF
- (SR)IEKF₁
- IEKF₂
- (SR)UKF
- (SR)CDKF
- IUKF, ICDKF
- GHF
- GPF
- ASIR
- RPF
- KPF
- SPPF
- GSKF
- GSPF
- GMSPPF

Für den EKF und die SPKF (ohne GHF und iterierende SPKF) wurden ebenfalls entkoppelte Algorithmen implementiert, welche bei der gemeinsamen Zustands- und Parameterschätzung evaluiert werden können. Ebenso ist ein optionaler MCMC-Schritt für den GPF, den RPF und die SPPF implementiert.

ReBEL-IoN mit all seinen Komponenten wurde unter Windows XP mit MATLAB[®] in Version R2007a codiert. Die Toolbox benötigt MATLAB[®] in Version R2007a oder eine neuere Version und die Statistics Toolbox.

Die Abkürzung ReBEL-IoN bedeutet 'Recursive Bayesian Estimation Library - Investigation of Necessity' und soll ausdrücken, dass die Toolbox dazu dient, den "notwendigsten" Filter zu finden, um bestmögliche Kennwerte in punkto Güte, Robustheit und Rechenzeit zu erlangen. Also die Frage, welcher Filter mit geringsten qualitativen Anforderungen **notwendig** ist, um z.B. einen Wirkungsgrad von $X\%$ mit einer Robustheit von $Y\%$ zu erreichen, kann mit ReBEL-IoN untersucht werden.

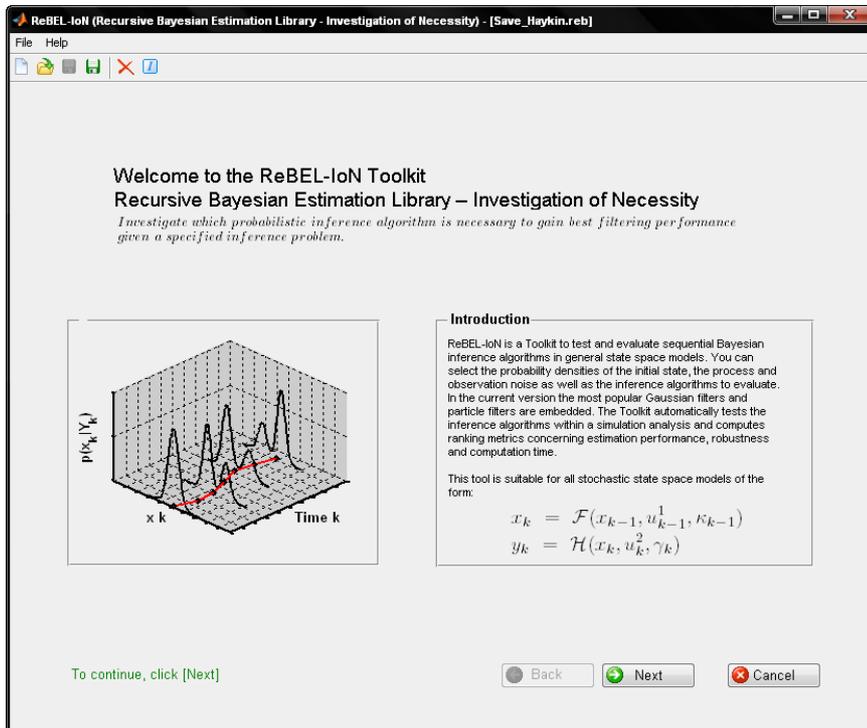
Weitergehende Information (Demos, Tutorials, Lizenzierung und Dokumentation) können im Internet unter <http://www.rus.rub.de> oder <http://www.rebel-ion.de> gefunden werden⁴.

4.8 Kapitelzusammenfassung

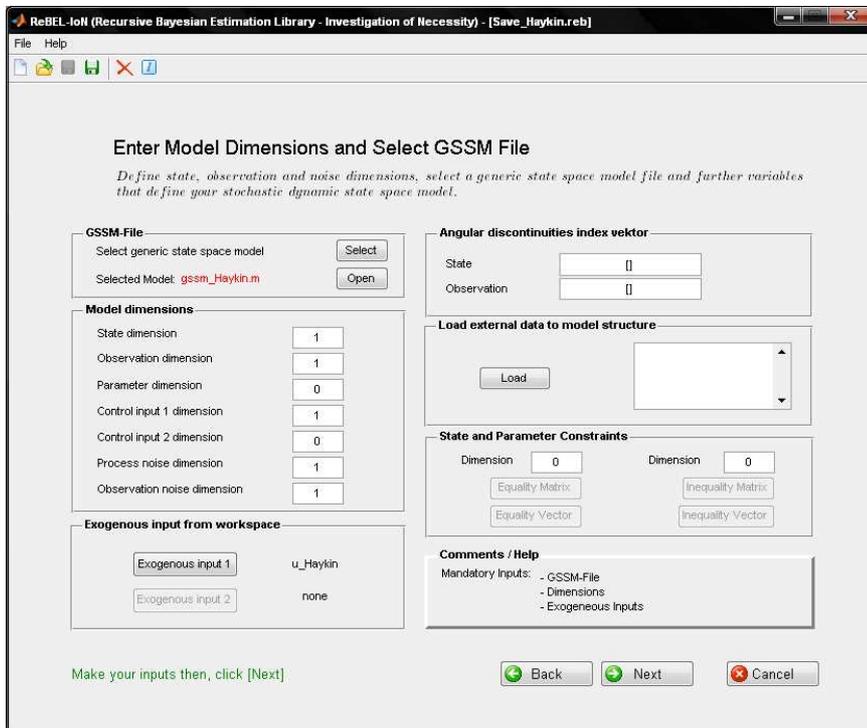
In diesem Kapitel wurde vorgestellt, wie bei einer vorgegebenen Inferenzaufgabe verschiedene Inferenzverfahren sukzessiv gegeneinander getestet und bewertet werden können. Grundsätzliche vom Anwender zu leistende Vorarbeit ist die Entwicklung des stochastischen DSSM, was Modellbildung, eventuell notwendige Diskretisierung und die Untersuchung der Stochastik des Systems und der Messung umfasst. Sind diese Arbeiten vollzogen, erfolgt die objektive Bewertung durch qualitative und quantitative Bewertungskriterien. Mit der in dieser Arbeit entwickelten MATLAB[®] Toolbox ReBEL-IoN wird dem Anwender eine Plattform zur Verfügung gestellt, die es ihm ermöglicht, schnell, flexibel und somit kostengünstig verschiedene Inferenzverfahren zu testen und letztendlich das bestmögliche bzw. "notwendigste" auszuwählen.

⁴Bei Publikation dieser Arbeit befanden sich die Webseiten noch im Aufbau.

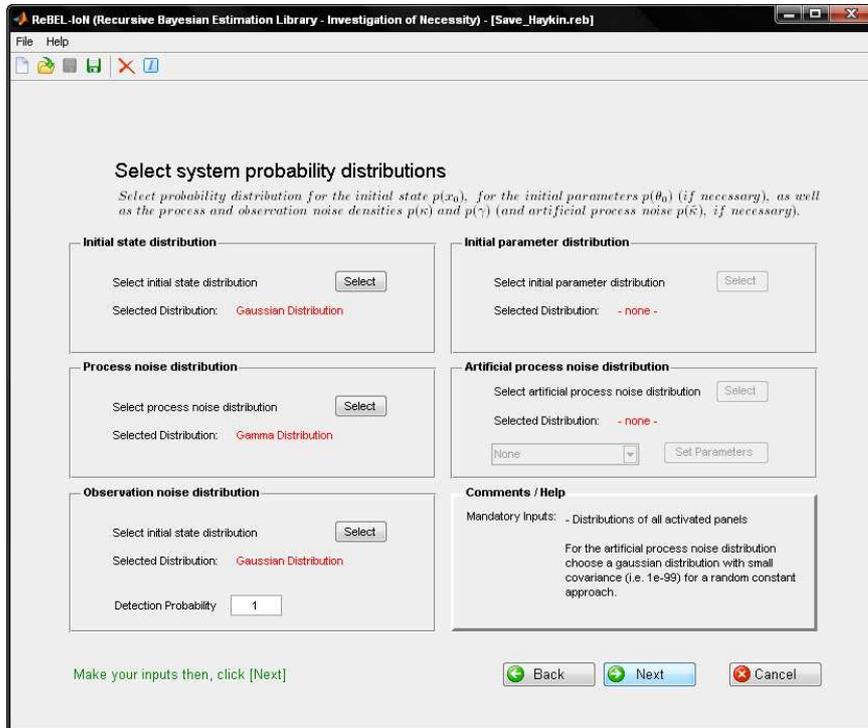
ReBEL-IoN GUI Opener



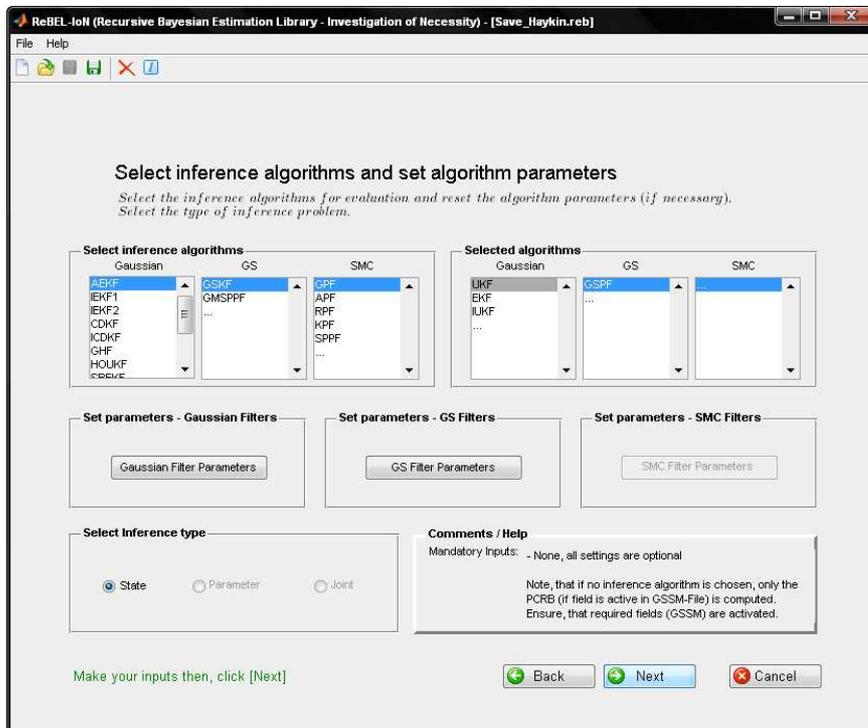
ReBEL-IoN GUI Fenster 1



ReBEL-IoN GUI Fenster 2



ReBEL-IoN GUI Fenster 3



ReBEL-IoN GUI Fenster 4

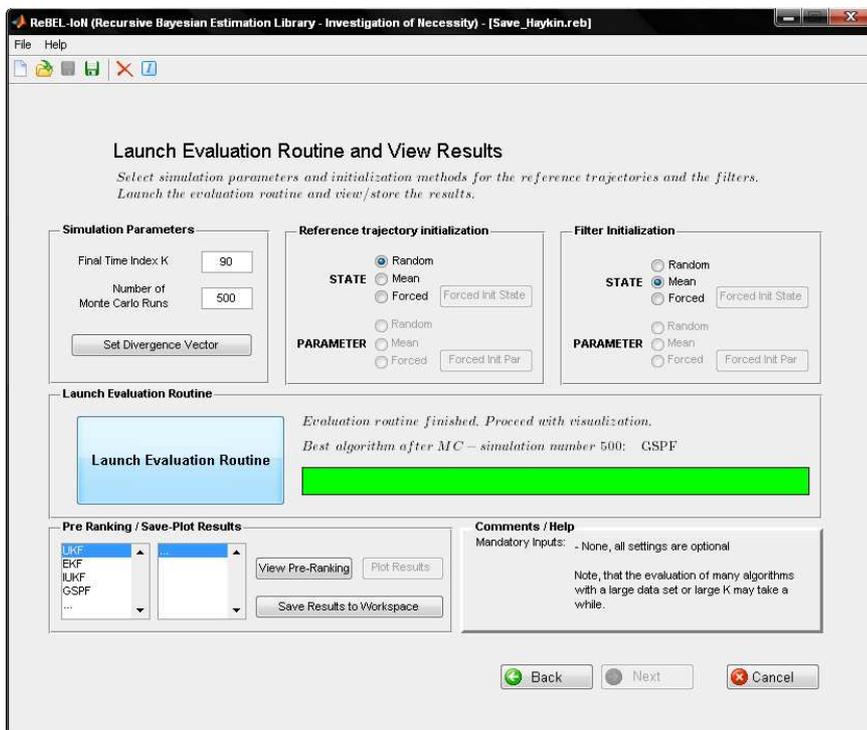


Abbildung 4.9: Bedienoberfläche ReBEL-IoN.

5 Anwendungsbeispiele

5.1 Zwei anschauliche akademische Inferenzaufgaben

In diesem Abschnitt soll der Nutzen und die Nutzung der in dieser Arbeit vorgestellten Methodik an zwei ausgewählten exemplarischen, aber eher akademischen Inferenzaufgaben gezeigt werden. Sie sind leicht nachzuvollziehen und erleichtern somit das Verständnis der zu Grunde liegenden Theorien und die Anwendung von ReBEL-IoN. Die Beispiele sind so aufbereitet, dass sie von den Lesern verifiziert werden können, sofern Zugriff auf ReBEL-IoN besteht.

Zustandsschätzung: Analytisches, skalares, nichtlineares, nicht Gauß'sches Experiment

In diesem Abschnitt soll der "notwendigste" Algorithmus für ein nichtlineares System bestimmt werden, dessen Systemrauschen nicht normalverteilt ist. Eine leicht abgeänderte Version wird auch in [8, S. 263] behandelt. Das System sei durch das diskrete Zustandsraummodell

$$x_k = 1 + \sin(\omega\pi k) + \phi_1 x_{k-1} + \kappa_{k-1} = \mathcal{F}_{Ref, Fil}(x_{k-1}, u_{k-1} = k, \kappa_{k-1}) \quad (5.1)$$

$$y_k = \phi_2 x_k^2 + \gamma_k = \mathcal{H}_{Ref, Fil}(x_k, u_k, \gamma_k) \quad (5.2)$$

beschrieben, wobei die Konstanten $\omega = 0.04$, $\phi_1 = 0.5$ und $\phi_2 = 0.2$ sind.

Beobachtbarkeit des deterministischen Systems

Die Beobachtbarkeitsmatrix (Skalar) ist für dieses Beispiel mit

$$\mathcal{BM}(x_k) = 2\phi_2 x_k \quad (5.3)$$

gegeben und besitzt vollen Rang $\forall x_k \in \mathbb{R}^n \setminus \{0\}$. Da im stochastischen Fall die Wahrscheinlichkeit dafür, dass der Zustand exakt den Wert null annimmt, gleich null ist, gilt das autonome System als lokal beobachtbar.

Simulative Evaluierung mit ReBEL-IoN

Für die Simulationsstudie wurden die in Tabelle 5.1 aufgeführten Rahmenbedingungen festgelegt. Die Studienergebnisse einiger ausgewählter Verfahren sind in Abbildung 5.1 zu sehen und für alle Verfahren in Tabelle 5.2 zusammengefasst.

MC-Simulationen	$MC = 2000$
Observationen	$K = 90$
Initialisierung	$x_0 [-] \sim \mathcal{N}(x_0; 0, 2)$
Systemrauschen	$\kappa_{k-1} [-] \sim \mathcal{Ga}(\kappa; 3, 1.25)$
Messrauschen	$\gamma_k [-] \sim \mathcal{N}(\gamma; 0, 2)$
POD	$p_D = 1$
Divergenzgrenze	$DIV_{Thres} [-] = 5$
Filter	EKF, UKF, RPF, ... (500 Partikel für PF+Hybride (SPPF: 200 Partikel))

Tabelle 5.1: Studienparameter für das analytische Experiment.

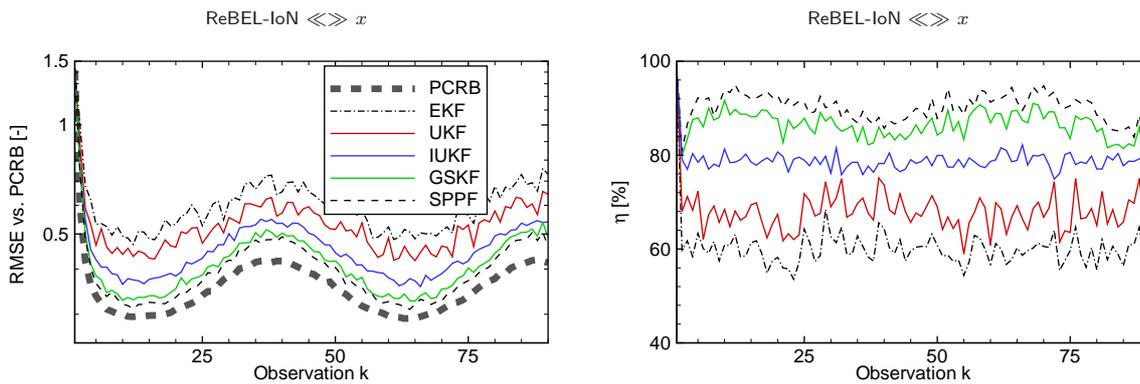


Abbildung 5.1: PCRB vs. RMSE und Filterwirkungsgrad für das analytische Experiment.

Algorithmus	$\bar{\eta}$	$RTAMSE$	RI	$J_{Robustness}$	\bar{t}_{sim}	TI
	[%]	[-]	[%]	[%]	[s]	[%]
EKF	60.17	0.61	0.0	89.2	0.0284	0
AEKF	65.50	0.56	8.17	96.3	0.0304	-7.04
UKF	67.98	0.54	11.47	93.2	0.0393	-38.35
IUKF	78.85	0.47	22.95	100	0.4483	-1.48e3
GSKF ₅₋₅₋₁	86.53	0.44	27.87	98.7	1.7815	-6.18e3
IEKF ₁	88.13	0.43	29.51	100	0.0676	-138.22
RPF	88.83	0.43	29.51	98.6	0.5506	-1.84e3
GMSPPF ₅₋₅₋₁	89.81	0.42	31.15	99.8	2.0065	-6.97e3
GSPF ₅₋₅	90.22	0.42	31.15	99.2	0.8446	-2.87e3
SPPF	90.23	0.42	31.15	99.9	20.5234	-7.22e4
PCRB	-	0.38	37.70	-	-	-

Tabelle 5.2: Quantitative Bewertungskriterien für das analytische Experiment.

Beurteilung der Ergebnisse

Man erkennt deutlich, dass die einfachen GAUSS'schen Verfahren das stark nicht GAUSS'sche Systemrauschen nicht ausreichend genau abbilden können, was in schlechter Filtergüte resultiert. Der GSKF mit einem 5-5-1 GMM besitzt eine bessere Filtergüte, wobei die Partikel-Filter (hier RPF) eine noch bessere Filtergüte erreichen. Der IEKF₁ und die Hybriden zeigen nahezu identische Filtergüte bei ähnlicher Robustheit, was eine Auswahl allein basierend auf den quantitativen Kriterien, besonders für ungeübte Anwender, schwierig

macht. Unter Berücksichtigung der vorgeschlagenen qualitativen Kriterien, dass die Anforderungen an den Anwender deutlich geringer und der Entwurf deutlich einfacher ist, ist der IEKF₁ die angemessene Wahl für dieses Beispiel.

Duale Schätzung: Verfolgung eines ballistischen Flugobjektes beim Wiedereintritt in die Atmosphäre [36]

Die Verfolgung ballistischer Objekte ist eine der am meisten untersuchten Anwendungen in der Filtertheorie und in der Raumfahrttechnologie. Die Aufgabe besteht darin, ballistische Objekte zu verfolgen, abzufangen und zu zerstören, bevor sie auf der Erdoberfläche aufschlagen und u.U. Schaden verursachen können. Ohne Zweifel ist das große Interesse militärisch begründet, da die Anzahl an taktischen Raketen und die Anzahl an besitzenden Ländern weltweit zunimmt.

In diesem Anwendungsbeispiel wird die Verfolgung eines ballistischen Objektes beim Wiedereintritt in die Erdatmosphäre behandelt. Während des Wiedereintritts wirken hauptsächlich die Erdbeschleunigung und der Luftwiderstand auf das Objekt, das System ist autonom. Mit abnehmender Höhe steigt der Luftwiderstand durch die exponentiell ansteigende Dichte nichtlinear an, was sich in extremen Verzögerungen bemerkbar macht. Die Radar-Verfolgung eines senkrecht fallenden Flugkörpers ist aufgrund dieses stark nichtlinearen Charakters in vielen Veröffentlichungen [3], [6], [14], [15], [36] als Benchmark für die Filtergüte von nichtlinearen Filtern verwendet worden. Abbildung 5.2 veranschaulicht den Zusammenhang.

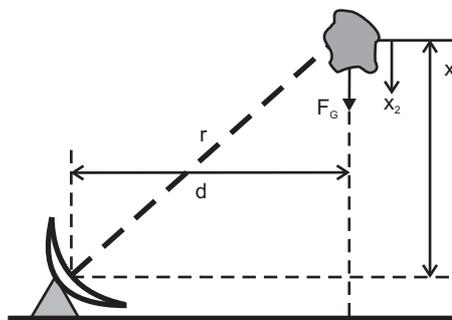


Abbildung 5.2: Skizze zur Radar-Verfolgung eines fallenden Flugobjektes.

Das dreidimensionale Zustandsraummodell ist durch

$$x_{1,k} = x_{1,k-1} + \tau(-x_{2,k-1}) + \kappa_{1,k-1} \quad (5.4)$$

$$x_{2,k} = x_{2,k-1} + \tau \left(g - \frac{\eta e^{-\lambda x_{1,k-1}} g x_{2,k-1}^2}{2\theta_{k-1}} \right) + \kappa_{2,k-1} \quad (5.5)$$

$$\theta_k = \theta_{k-1} \quad (5.6)$$

$$y_k = \sqrt{x_{1,k}^2 + d^2} + \gamma_k = \mathcal{H}_{Ref, Fil}(x_k, u_k, \gamma_k) \quad (5.7)$$

gegeben, wobei $[x_{1,k} \ x_{2,k} \ x_{3,k} = \theta_k]^T = \mathcal{F}_{Ref, Fil}(x_{k-1}, u_{k-1}, \kappa_{k-1})$ ist und θ die ballistische Konstante des fallenden Objekts beschreibt, welche simultan mit der Position und Geschwindigkeit des Objekts mitgeschätzt werden muss. Die Beschleunigung durch Gravitation ist $g = 9.81 \frac{m}{s^2}$ und die konstanten Parameter sind $\eta = 1.754$, $\lambda = 1.49 \cdot 10^{-4}$ und $d = 10$ km. Die Integrationsschrittweite ist $\tau = 0.1$ s.

Eine typische Objekt-Trajektorie ist in Abbildung 5.3 zu sehen. Folgende Parameter wurden benutzt:

- Anfangshöhe: 61000 m
- Anfangsgeschwindigkeit: $3048 \frac{m}{s}$
- Ballistische Konstante: $19161 \frac{kg}{ms^2}$

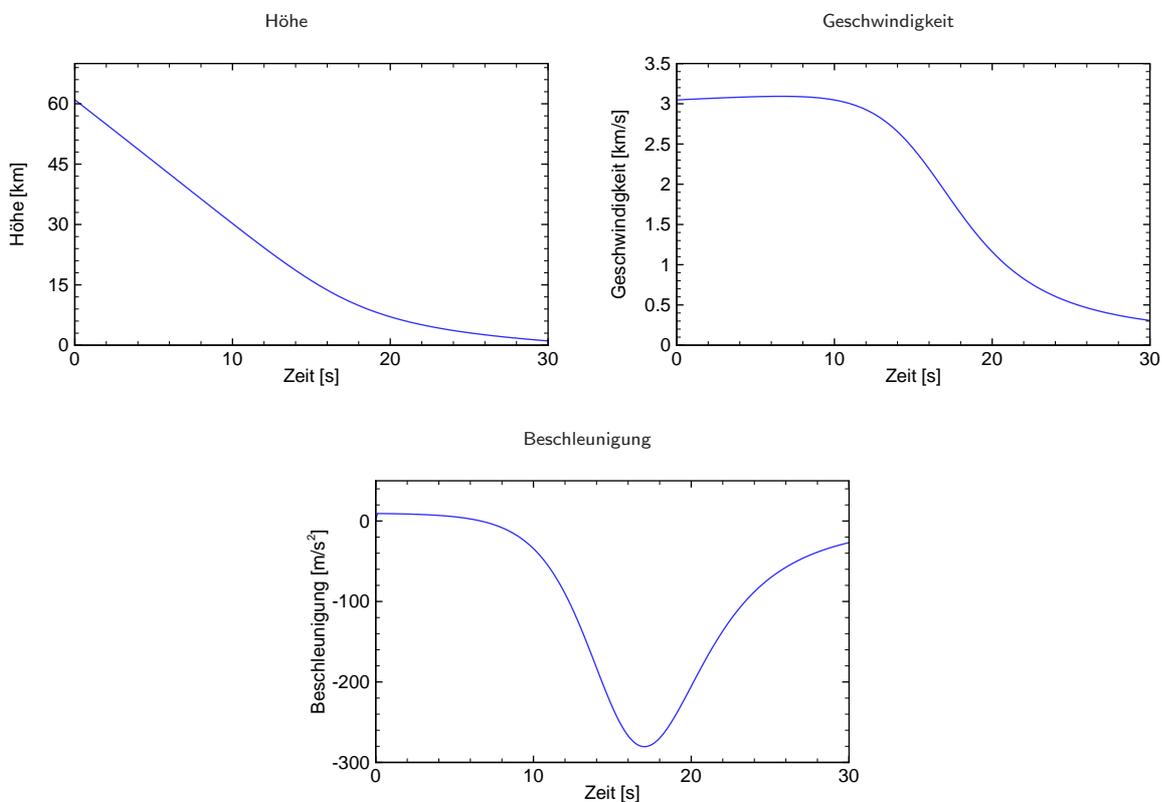


Abbildung 5.3: Referenztrajektorien.

Man erkennt deutlich, dass die Geschwindigkeit zunächst nahezu konstant ist, das Objekt aber beim Eintreten in dichtere Luftschichten durch die starke Verzögerung scharf abgebremst wird. Die Abbremsung und somit die Nichtlinearität der Systemgleichung wächst mit fallender ballistischer Konstante des Objekts.

Beobachtbarkeit des deterministischen Systems

Es wird die Beobachtbarkeit für das System mit horizontaler Entfernung $d = 0$ km untersucht, da sich die Beobachtbarkeitsmatrix deutlich vereinfacht. Der nachfolgende Beweis für die Beobachtbarkeit des

deterministischen Systems kann aber auch analog für andere horizontale Entfernungen geführt werden. Die Beobachtbarkeitsmatrix ist für dieses Beispiel mit

$$\mathcal{BM}(x_k) = \begin{bmatrix} 1 & 0 & 0 \\ 1 & -\tau & 0 \\ \frac{\partial \mathcal{F}_1}{\partial x_{1,k-1}} - \tau \frac{\partial \mathcal{F}_2}{\partial x_{1,k-1}} & \frac{\partial \mathcal{F}_1}{\partial x_{2,k-1}} - \tau \frac{\partial \mathcal{F}_2}{\partial x_{2,k-1}} & \frac{\partial \mathcal{F}_1}{\partial x_{3,k-1}} - \tau \frac{\partial \mathcal{F}_2}{\partial x_{3,k-1}} \end{bmatrix} \Big|_{x_{k-1}=x_k} \quad (5.8)$$

gegeben und besitzt vollen Rang $\forall x_k \in \mathbb{R}^n \setminus \{x_{2,k} = 0\}$. Da die Geschwindigkeit niemals den Wert null annimmt, ist das autonome System lokal beobachtbar.

Simulative Evaluierung mit ReBEL-IoN

Für die Simulationsstudie wurden die in Tabelle 5.3 gezeigten Rahmenbedingungen festgelegt. Die Studienergebnisse der untersuchten Filter sind in Abbildung 5.4 zu sehen und in Tabelle 5.4 zusammengefasst.

MC-Simulationen	$MC = 2000$
Observationen	$K = 350$
Initialisierung	$x_0 \text{ [m } \frac{\text{m}}{\text{s}}]^T \sim \mathcal{N}(x_0; [60960 \ 3048]^T, \begin{bmatrix} \mathbf{P}^\gamma & \mathbf{P}^\gamma/\tau \\ \mathbf{P}^\gamma/\tau & 2\mathbf{P}^\gamma/\tau^2 \end{bmatrix})$
	$\theta_0 \text{ [} \frac{\text{kg}}{\text{ms}^2}] \sim \mathcal{B}(\theta_0; 1.1, 1.1) \frac{10000}{63000}$
Systemrauschen	$\kappa_{k-1} \text{ [m } \frac{\text{m}}{\text{s}}]^T \sim \mathcal{N}(\kappa; 0, \begin{bmatrix} q_1 \frac{\tau^3}{3} & q_1 \frac{\tau^2}{2} \\ q_1 \frac{\tau^2}{2} & q_1 \tau \end{bmatrix}), \quad q_1 = 5$
Messrauschen	$\gamma_k \text{ [m]} \sim \mathcal{N}(\gamma; 0, 200^2)$
POD	$p_D = 1$
Divergenzgrenze	$DIV_{Thres} \text{ [m } \frac{\text{m}}{\text{s}} \frac{\text{kg}}{\text{ms}^2}]^T = [5000 \ \infty \ \infty]^T$
Filter	EKF, UKF, CDKF, IUKE

Tabelle 5.3: Studienparameter Radar-Verfolgung.

Algorithmus	$\bar{\eta}$	$RTAMSE$	RI	$J_{Robustness}$	\bar{t}_{sim}	TI
	[%]	[m m/s kg/ms ²]	[%]	[%]	[s]	[%]
EKF	[61.9 56.9 59.7]	[250 231 9137]	[0 0 0]	97.2	0.144	0
UKF	[83.9 52.5 73.7]	[109 239 8875]	[56.2 -3.2 2.9]	98.7	0.157	-8.99
CDKF	[80.3 45.7 65.6]	[121 258 9185]	[51.8 -11.4 -0.5]	97.8	0.189	-31.1
IUKE	[89.1 62.3 80.3]	[99.4 239 8851]	[60.3 -3.3 3.3]	99.0	0.887	-516.5
PCRB	-	[85.2 227 8867]	[65.9 1.9 2.9]	-	-	-

Tabelle 5.4: Quantitative Bewertungskriterien Radar-Verfolgung.

Beurteilung der Ergebnisse

Der Verlauf des PCRB zeigt, dass zunächst keine Schätzung der ballistischen Konstante möglich ist, solange bis das Objekt in Schichten höherer Luftdichte eindringt und abgebremst wird (Zeitpunkt ca. 12s). Bis zu dem Zeitpunkt (ca. 20s), in dem das Objekt die Radarstation passiert, ist eine Adaption möglich. Beim Passieren der Radarstation kann aufgrund der geringen Änderung in der gemessenen Schrägentfernung

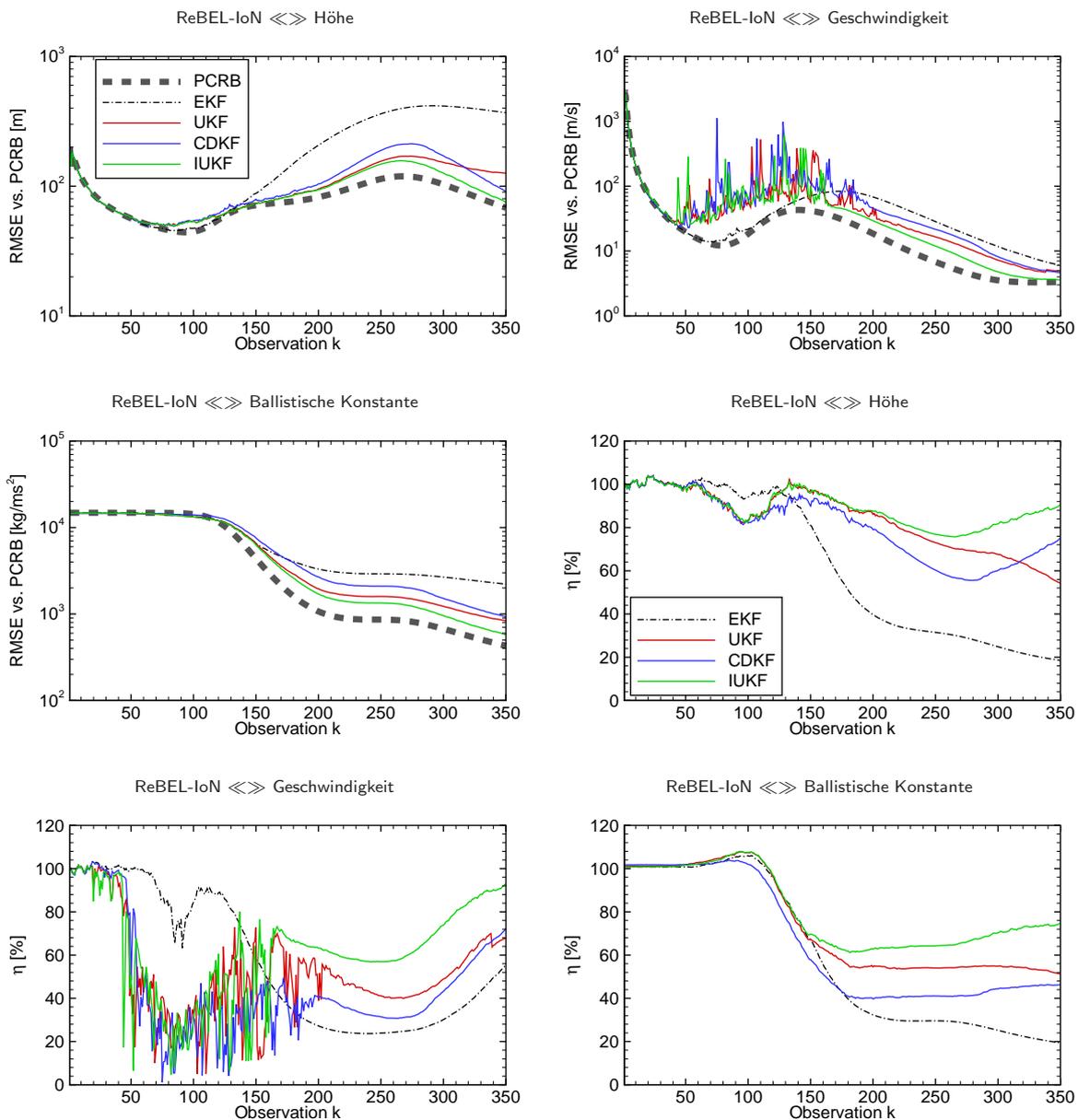


Abbildung 5.4: RMSE vs. PCRB und Filterwirkungsgrad Radar-Verfolgung.

zunächst keine weitere Adaption erfolgen. Mit fortschreitender (vertikaler) Entfernung von der Radarstation wird wieder eine Schätzung möglich.

Die numerische Bewertung zeigt, dass der EKF eine deutlich niedrigere Filtergüte und Robustheit als der UKF, der CDKF und der IUKF aufweist, was in der ungenauen Approximation der stark nichtlinearen Systemgleichung durch die lokale Linearisierung begründet liegt. Dies wird besonders deutlich, wenn das ballistische Objekt in Regionen hoher Luftdichte eintritt und folglich stark abgebremst wird. Der UKF, CDKF und IUKF sind in der Lage, das Objekt mit erhöhter Genauigkeit zu verfolgen, wobei der IUKF sowohl

bei der Filtergüte als auch bei der Robustheit beste Ergebnisse aufweist. Dies ist besonders bemerkenswert, da divergierte Schätzungen nicht in der quantitativen Evaluierung der Filtergüte berücksichtigt werden. Der IUKF besitzt allerdings eine deutlich erhöhte Rechenzeit durch die mehrmalige Iteration der Messgleichung. Unter der Berücksichtigung der qualitativen Kriterien ist der CDKF zwar der Algorithmus mit geringsten Anforderung, die deutlich bessere Leistung des IUKF rechtfertigt allerdings den Mehraufwand während des Filterentwurfs und den erhöhten Rechenaufwand. An dieser Stelle sei angemerkt, dass mit geringerer Breite des Intervalls der ballistischen Konstante sich die Filtergüte der Algorithmen annähert, besonders wenn das Intervall nach oben hin verkürzt wird. Dies liegt darin begründet, dass die Nichtlinearitäten besonders bei Objekten mit kleiner ballistischer Konstante zum Tragen kommen, da diese stärker als Objekte mit großer ballistischer Konstante abgebremst werden.

5.2 Reversible Gasgleichgewichtsreaktion (Demonstration beschränkte SMC-Methoden)

Betrachtet werden soll die reversible Gasgleichgewichtsreaktion [66]



mit Stöchiometriematrix $\nu = [-2 \ 1]$ und Reaktionsrate $r = \bar{k} \cdot \phi_A^2$. Die Zustände des Systems sind $x = [\phi_A \ \phi_B]^T$, wobei ϕ_j den Partialdruck von Komponente j kennzeichnet. Es sei angenommen, dass die Voraussetzungen für ein ideales Gas gültig sind und dass die Reaktion in einem gut gerührten (homogene Gasverteilung), isothermen Batch-Reaktor stattfindet. In diesem Fall ergibt sich das zeitdiskretisierte Zustandsraummodell

$$x_k = x_{k-1} + \tau \cdot \nu^T r_{k-1} = x_{k-1} + \tau \cdot \nu^T \bar{k} x_{1,k-1}^2 \quad (5.10)$$

mit Messgleichung

$$y_k = [1 \ 1]x_k \quad (5.11)$$

und Diskretisierungsschrittweite $\tau = 0.1$ s. Für die stochastische Zustandsfilterung werde nun folgendes Zustandsraummodell betrachtet:

$$\begin{aligned} x_k &= x_{k-1} + \tau \cdot \nu^T \bar{k} x_{1,k-1}^2 + \kappa_{k-1} \\ &= \mathcal{F}(x_{k-1}, u_{k-1}, \kappa_{k-1}) \end{aligned} \quad (5.12)$$

$$y_k = [1 \ 1]x_k + \gamma_k = \mathcal{H}(x_k, u_k, \gamma_k), \quad (5.13)$$

wobei κ und γ mittelwertfreie, normalverteilte Zufallsvariablen seien.

Beobachtbarkeit des deterministischen Systems

Die Beobachtbarkeitsmatrix für dieses Beispiel ist

$$\mathcal{BM}(x_k) = \begin{bmatrix} 1 & 1 \\ 1 - 2\tau\bar{k}x_{1,k} & 1 \end{bmatrix} \quad (5.14)$$

und besitzt vollen Rang $\forall x_k \in \mathbb{R}^n \wedge x_{1,k} > 0$. Da die Partialdrücke nichtnegativ und größer null sind ist das autonome System lokal beobachtbar.

Simulative Evaluierung mit ReBEL-IoN

Die angenommene Verteilung des initialen Systemzustands, welche zur Initialisierung der Filter verwendet wird, sei $\hat{x}_0 \sim \mathcal{N}(\hat{x}_0; [0.1 \ 4.5]^T, \text{diag}([6^2 \ 6^2]))$, wohingegen der wahre Initialzustand des Systems $x_0 = [3 \ 1]^T$ sei. Hier sei angemerkt, dass die initiale Schätzung (die Annahme des Anwenders) mangelhaft ist. Um den in Kapitel 3.3 vorgestellten Ansatz zu testen, wurden vier SMC-Methoden (GPF, ASIR, RPF und GMSPPF) mit dem Projection-Sampling-Ansatz für beschränkte Verteilungen ausgestattet ($G = 5$) und am obigen System getestet. Alle Partikel-Filter haben eine Partikelanzahl $N = 200$. Der GMSSPF verwendet ein 5-1-1 GMM, d.h. die priori und posteriori Verteilung des Zustandes werden als GMM mit fünf Komponenten, das System- und Messrauschen als GMM mit jeweils einer Komponente approximiert. Weitere Simulationsparameter sind der Tabelle 5.5 zu entnehmen.

MC-Simulationen	$MC = 2000$
Observationen	$K = 300$
Systemrauschen	$\kappa_{k-1} [\text{Pa Pa}] \sim \mathcal{N}(\kappa; 0, \text{diag}([0.001^2 \ 0.001^2]))$
Messrauschen	$\gamma_k [\text{Pa}] \sim \mathcal{N}(\gamma; 0, 0.1^2)$
POD	$p_D = 1$
Divergenzgrenze	$DIV_{Thres} [\text{Pa Pa}]^T = [10 \ 10]^T$
Filter	GPF, ASIR, RPF, GMSPPF

Tabelle 5.5: Studienparameter reversible Gasreaktion.

In zwei Szenarien wurde der Einfluss verschiedener Informationen über die Zustandsschranken auf den Schätzfehler und die Robustheit untersucht und verglichen mit den unbeschränkten Algorithmen:

1. **Szenario 1:** Die Partialdrücke des Systems sind nichtnegativ. In diesem Szenario (Ξ_1) wurde diese Information den Partikel-Filtern zur Verfügung gestellt.
2. **Szenario 2:** In diesem speziellen Beispiel überschreiten die Partialdrücke einen oberen Grenzwert von 3 Pa nicht. In diesem Szenario (Ξ_2) wurde die komplette Information über die Schranken den Filtern zur Verfügung gestellt.

In Abbildungen 5.5 und 5.6 sind die Verläufe der RMSE der vier Filter für die zwei Szenarien sowie den unbeschränkten Fall abgebildet. Die Ergebnisse der Simulationsstudie sind zusätzlich in Tabelle 5.6 zusammengefasst.

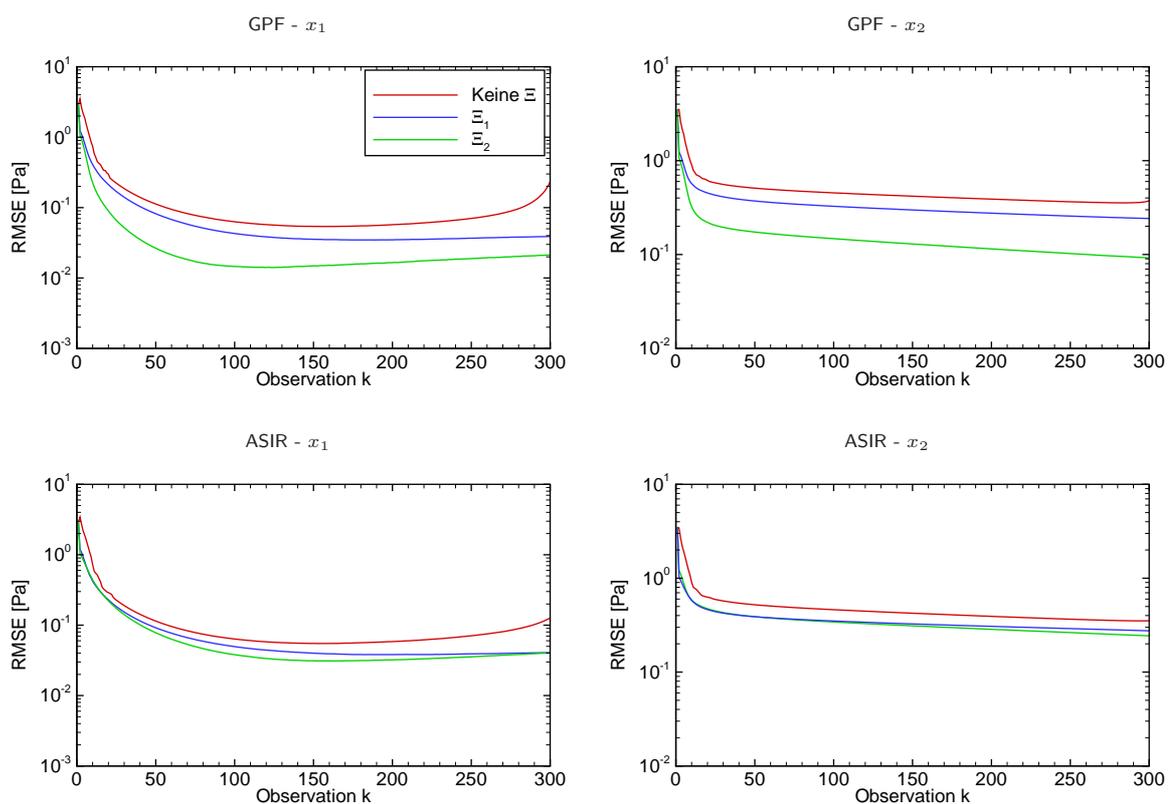


Abbildung 5.5: RMSE des GPF und ASIR mit und ohne Informationen über die Beschränkungen des Zustandsraumes für die reversible Gasreaktion.

Algorithmus	Unbeschränkte Algorithmen		Beschränkte Algorithmen		
	$RTAMSE$	$J_{Robustness}$	Ξ	$RTAMSE$	$J_{Robustness}$
GPF	0.39\0.59 Pa	64.40 %	Ξ_1 :	0.23\0.39 Pa	99.90 %
			Ξ_2 :	0.20\0.27 Pa	99.90 %
APF	0.39\0.59 Pa	65.70 %	Ξ_1 :	0.23\0.41 Pa	99.10 %
			Ξ_2 :	0.22\0.41 Pa	99.15 %
RPF	0.45\0.47 Pa	78.80 %	Ξ_1 :	0.27\0.44 Pa	100.0 %
			Ξ_2 :	0.21\0.27 Pa	100.0 %
GMSPPF	0.61\0.59 Pa	100.0 %	Ξ_1 :	0.21\0.24 Pa	100.0 %
			Ξ_2 :	0.20\0.24 Pa	100.0 %

$\Xi_1 : x > 0$
 $\Xi_2 : 0 < x \leq 3$

Tabelle 5.6: RTAMSE und Robustheit der untersuchten Partikel-Filter mit und ohne Informationen über die Beschränkungen des Zustandsraumes für die reversible Gasreaktion.

Beurteilung der Ergebnisse

Wie erwartet reduziert sich der Schätzfehler in Form des RMSE und RTAMSE je mehr Information über die Beschränkungen des Zustandsraumes den SMC-Methoden zur Verfügung gestellt wird. Darüber hinaus kann man den Abbildungen entnehmen, dass die Schätzgüte signifikant zwischen den einzelnen Partikel-

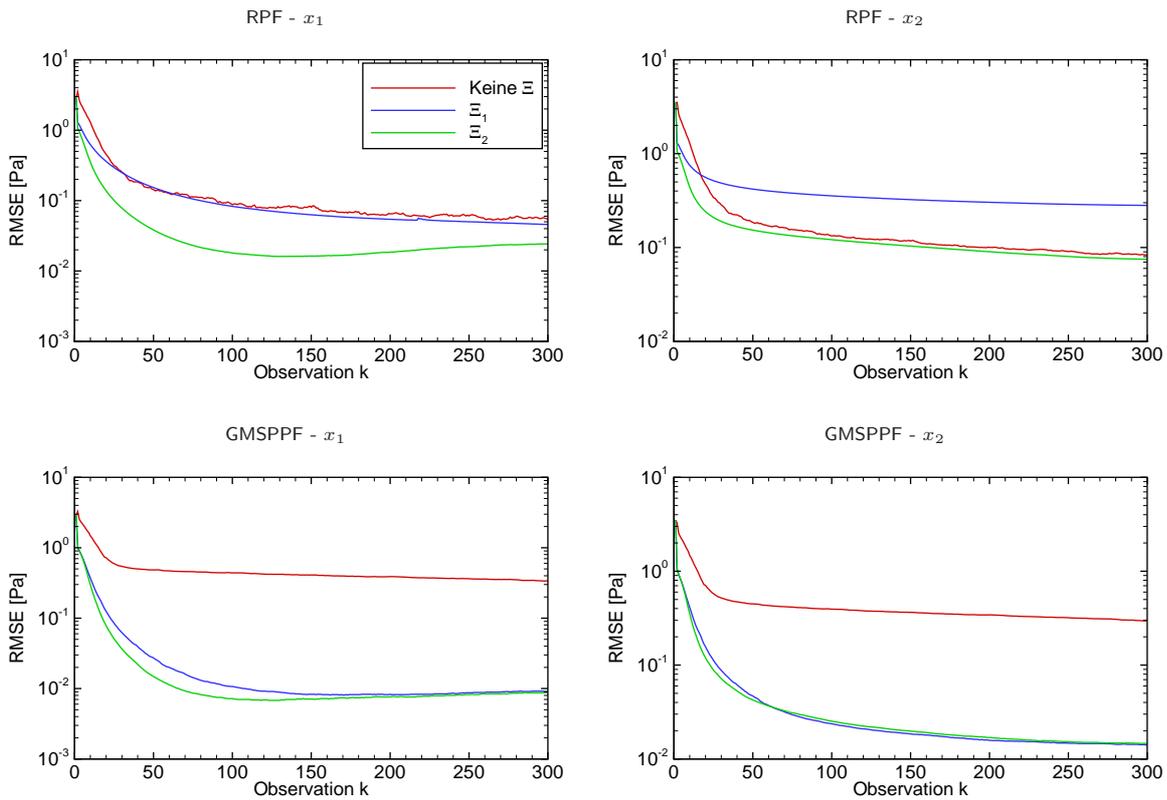


Abbildung 5.6: RMSE des RPF und GMSPPF mit und ohne Informationen über die Beschränkungen des Zustandsraumes für die reversible Gasreaktion.

Filtern variiert, was am allgemeinen Problem der Approximationslösungen der RBS für das nichtlineare Filterproblem liegt. Somit zeigt auch dieses Beispiel noch einmal, dass eine Überprüfung der Algorithmen für ein System und dafür zu lösende Schätzaufgabe unabdingbar ist. Tabelle 5.6 zeigt nochmals einen Vergleich der Algorithmen in Form des RTAMSE und Robustheit. Abermals ist offensichtlich, dass je mehr Information über die Zustandsschranken den Algorithmen zur Verfügung gestellt werden, desto bessere Zustandsschätzungen mit höherer Robustheit erzielt werden können. Die RTAMSE der unbeschränkten Algorithmen sind höher mit niedriger Robustheit. Die Einbindung der Information über die Nichtnegativität führt bereits zu robusten Schätzern (Robustheitsindizes nahe 100 %) mit verbesserter Schätzgüte. Dies ist besonders hervorzuheben, da divergierte Monte-Carlo-Simulationen nicht in der Berechnung des RTAMSE berücksichtigt werden. Wird die Obergrenze von 3 Pa zusätzlich den Algorithmen zur Verfügung gestellt, so kann die Schätzgüte der Algorithmen bei gleichbleibender Robustheit weiter gesteigert werden.

5.3 Terrain-Aided-Navigation

Die primäre Navigationseinheit für die Lagebestimmung und den rechtsweisenden¹ Kurs moderner Flugzeuge ist das Trägheitsnavigationssystem (TNS), z.B. Honeywell LaserNavII. Es besteht aus drei senkrecht zueinander angeordneten Beschleunigungssensoren und drei Lasergyroskopen. Im Ausrichtungsmodus vor dem Start berechnet das TNS die Lage seiner Referenzachsen bezüglich des lokalen Horizontsystems. Zusätzlich wird während der Ausrichtungsphase die Flugzeugausrichtung bezüglich geographisch Nord und die geographische Breite durch Messung der Erddrehraten bestimmt. Für eine Konsistenzüberprüfung und zur Festlegung der Startposition müssen die geographische Breite und Länge in das Gerät eingegeben werden. Im Navigationsmodus während des Fluges liefert das TNS Angaben über Flugzeugdrehraten, Flugzeuglage im Raum, rechtsweisenden Kurs, Beschleunigungen, Geschwindigkeit sowie Position.

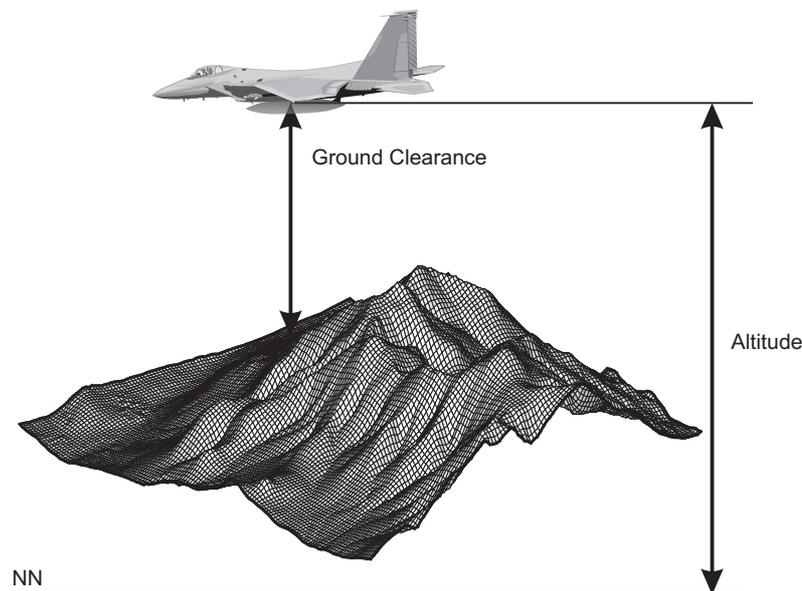


Abbildung 5.7: Konzept Terrain-Aided-Navigation.

Aufgrund von Fehlern in der Initialisierung und Messfehlern der Beschleunigungssensoren wird die wahre Position des Flugzeugs von der Positionsschätzung des TNS wegdriften. Das TNS liefert somit lediglich hochgenaue Positionsschätzungen über einen kurzen Zeitraum. Terrain-Aided-Navigation (TAN) ist ein neues Konzept zur Navigation, wo unter Verwendung einer 3D-Topographie-Karte (DTED: Digital Terrain Elevation Data) des überflogenen Terrains der Fehler des TNS durch WI-Verfahren korrigiert wird. So kann erreicht werden, dass die Navigationseinheit über sowohl über kurze, als auch über lange Zeiträume genaue Positionsangaben liefert. Oftmals findet eine weitere Unterstützung der Positionsbestimmung über

¹Ein Kurs in Bezug auf den geographischen Nordpol ist rechtsweisend, ein Kurs in Bezug auf den magnetischen Nordpol ist missweisend.

GPS statt. Eingesetzt wird TNS mit GPS und Mapmatching (TERPROM: Terrain-Profile-Matching) in modernen Kampfflugzeugen wie beispielsweise der F-16 und dem Eurofighter Typhoon sowie in der Storm Shadow Cruise Missile.

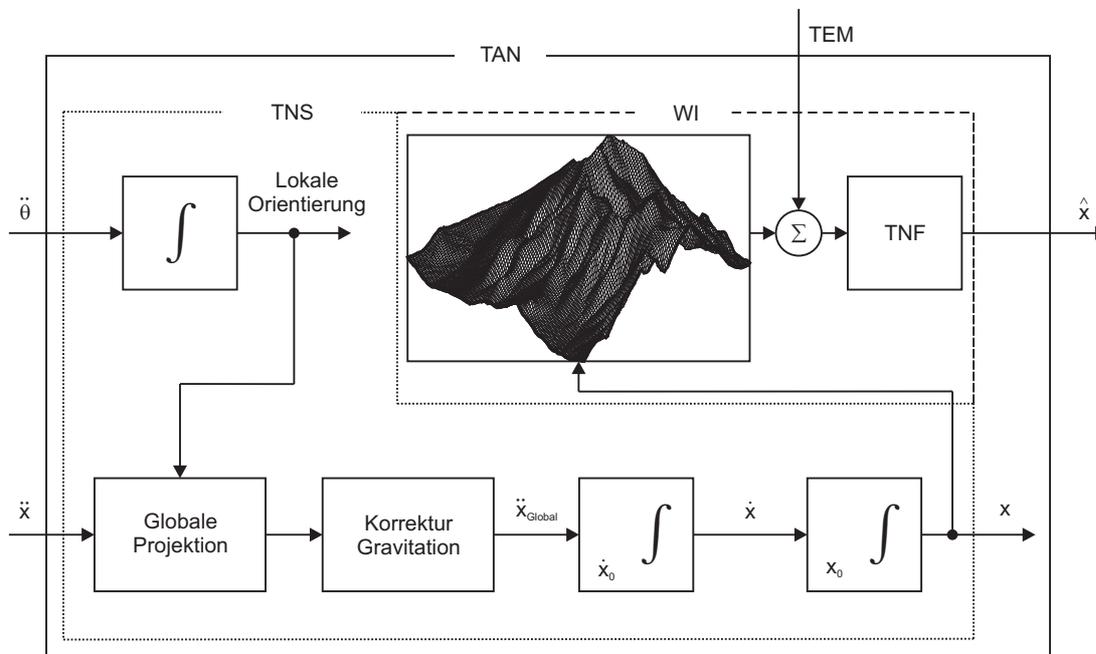


Abbildung 5.8: Blockschaltbild Terrain-Aided-Navigation - Strapdown TNS [85] mit WI.

Abbildungen 5.7 und 5.8 veranschaulichen das TAN-Konzept. An Bord des Flugzeugs ist eine 3D-Karte des überflogenen Terrains hinterlegt. Fliegt das Flugzeug nun innerhalb dieses Gebiets, wird aus der Differenz der absoluten Höhenmessung (über Null) und der Höhe über Grund eine "Messung" der Topographie (TEM: Terrain-Elevation-Measurement) erstellt. Diese Messung wird mit der hinterlegten 3D-Karte verglichen und erlaubt die Korrektur der über das TNS gelieferten Position (TNF: Terrain-Nav-Filter). TAN hat den Vorteil gegenüber anderen Navigationsmethoden (z.B. GPS), dass das Verfahren Positionsschätzungen autonom generiert und nicht auf unterstützende Informationen, von z.B. Satelliten, angewiesen ist. TAN ist nicht anwendbar über großen Gewässern (hier kommen dann magnetische Karten zum Einsatz), sehr flachen Terrains oder bei sehr großen Höhen über Grund. Andererseits ist es ohne Einbußen bei jeglichen Wetterbedingungen, tags und nachts einsetzbar und robust gegenüber elektronischen Maßnahmen zur Navigationsstörung. Letztere Eigenschaft ist besonders bei militärischen Applikationen nützlich.

Das resultierende Schätzproblem ist sehr herausfordernd aufgrund der hohen Anforderungen an die Genauigkeit sowie Robustheit und die variierende Charakteristik eines Terrains. Das System ist zudem hochgradig

nichtlinear. Das DSSM ist durch

$$x_k = x_{k-1} + u_{k-1} + \kappa_{k-1} = \mathcal{F}_{Ref, Fil}(x_{k-1}, u_{k-1}, \kappa_{k-1}) \quad (5.15)$$

$$y_k = h(x_k) + \gamma_k = \mathcal{H}_{Ref, Fil}(x_k, u_k, \gamma_k) \quad (5.16)$$

gegeben, wobei das TNS die relative Bewegung, u_{k-1} , zwischen zwei Messungen liefert. Die Messgleichung h ist gegeben durch die 3D-Karte der Topographie, die Zustände entsprechen hier der X - und Y -Position auf der Karte (Abbildung 5.9).

Beobachtbarkeit des deterministischen Systems

Auf den Nachweis der Beobachtbarkeit wird hier verzichtet, da die analytischen Jacobi-Matrizen der Messgleichung nicht herleitbar sind. Prinzipiell könnten die benötigten Jacobi-Matrizen über den Differenzenquotienten berechnet werden und der Rangnachweis simulativ erfolgen. Dies erfordert aber die Simulation des gesamten möglichen Zustandsraumes der Flugzeugtrajektorie und wird deshalb hier nicht weiter betrachtet.

Simulative Evaluierung mit ReBEL-IoN

Um die Filteralgorithmen für verschiedene Topographien zu testen, wurden zwei Simulationsstudien mit unterschiedlichem Terrain durchgeführt. In der ersten Studie (Szenario 1) wurde das Tracking-Verhalten über bergigem Terrain (Mount Washington), in der zweiten Studie (Szenario 2) über flachem Terrain (Anderson County, Texas) ausgewertet. Die DTED-Daten wurden vom US Geological Survey zur Verfügung gestellt und sind unter <http://data.geocomm.com> frei zugänglich abrufbar. Beide DTED-Datensätze haben eine Gitternetzauflösung von $\delta = 30$ m. Abbildung 5.9 zeigt die Terrains als 3D-Plots und Höhenlinienplots mit jeweils 50 Flugzeugtrajektorien.

Für beide Simulationsstudien wurden die in Tabelle 5.7 aufgeführten Rahmenbedingungen gewählt. Zur Berechnung des PCRB werden die partiellen Ableitungen zweiter Ordnung der logarithmierten Likelihood-Verteilung $\log p(y_k|x_k)$ benötigt. Für additives, normalverteiltes Messrauschen reduziert sich die Berechnung auf die partiellen Ableitungen erster Ordnung der Messgleichung. Diese sind hier nicht analytisch bestimmbar (DTED) und wurden in den Simulationsstudien durch die Differenzenquotienten ersetzt. Die Ergebnisse der Simulationsstudie sind in Abbildung 5.10 zu sehen und in Tabelle 5.8 zusammengefasst.

Beurteilung der Ergebnisse

Bereits am Verlauf des PCRB erkennt man, dass die Positionsschätzung mit abflachendem Terrain schwieriger wird, da weniger Informationsgehalt in den DTED-Daten steckt. Sowohl der Gradient als auch der finale Wert des PCRB indizieren, dass für bergiges Terrain genauere Positionsschätzungen möglich sind. Dies spiegelt sich auch in den Schätzungen der Filter wider. Für das bergige Terrain liefern alle untersuchten Algorithmen nahezu gleiche Filtergüte und konvergieren relativ schnell gegen den PCRB, wobei der GPF die höchste Konvergenzrate besitzt. Der Partikel-Filter ist ebenso der robusteste Algorithmus. Für

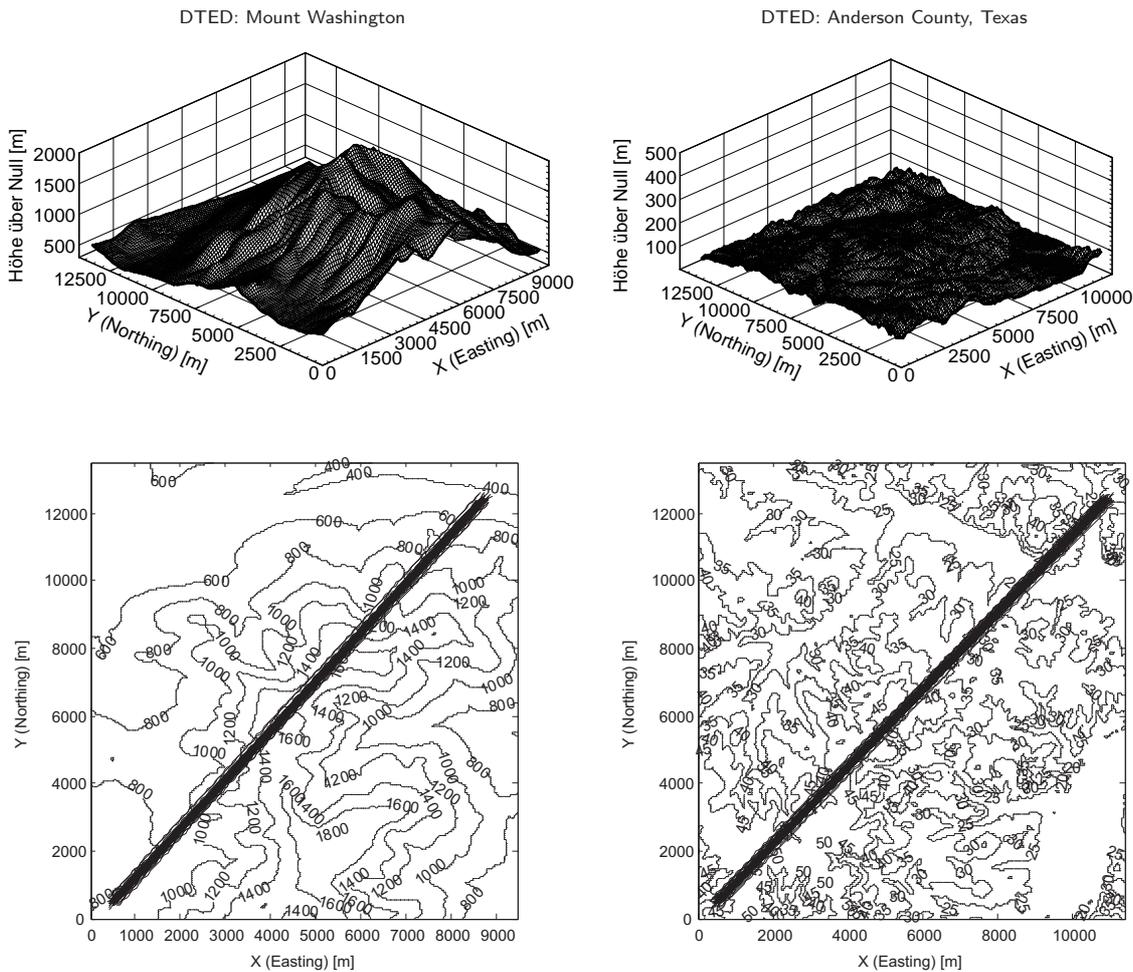


Abbildung 5.9: Oben Links: DTED Mt. Washington Oben Rechts: DTED Anderson County, Texas (Gitternetzabstand im Bild $\delta = 90$ m); Unten: Höhenlinien und Flugzeugtrajektorien Szenario 1 (Links) und Szenario 2 (Rechts).

MC-Simulationen	$MC = 500$
Observationen	$K = 150$
Initialisierung	x_0 [m m] ^T $\sim \mathcal{N}(x_0; [500 \ 500]^T, \mathbf{I} \cdot 80^2)$
Systemrauschen	κ_{k-1} [m m] ^T $\sim \mathcal{N}(\kappa; 0, \mathbf{I} \cdot 4)$
Messrauschen	γ_k [m] $\sim \mathcal{N}(\gamma; 0, 16)$
POD	$p_D = 1$
Divergenzgrenze	DIV_{Thres} [m m] ^T = $[300 \ 300]^T$
TNS _{Szenario 1}	u_{k-1} [m m] ^T = $[55 \ 80]^T$
TNS _{Szenario 2}	u_{k-1} [m m] ^T = $[70 \ 80]^T$
Filter	EKF, UKF, GPF (2000 Partikel)

Tabelle 5.7: Studienparameter TAN.

das flache Terrain liefern die Gauß-Filter im Gegensatz zum bergigen Terrain nur unbefriedigende Positionsschätzungen. Lediglich der Partikel-Filter konvergiert gegen den PCRB und wird optimal. Des Weiteren neigen die Gauß-Filter zu verschlechterter Robustheit, der Partikel-Filter hingegen bleibt robust. In der

Rechenzeit schneidet der EKF in beiden Szenarien am schlechtesten ab, da in jedem Filterungsschritt die numerische Jacobi-Matrix der Messgleichung (DTED) bestimmt werden muss. Aus diesem Grund scheiden auch iterierende Verfahren (IEKF_{1,2}) in dieser Evaluierung aus, da durch die mehrmalige Iteration des Filterungsschritts die Rechenzeit stark ansteigt.

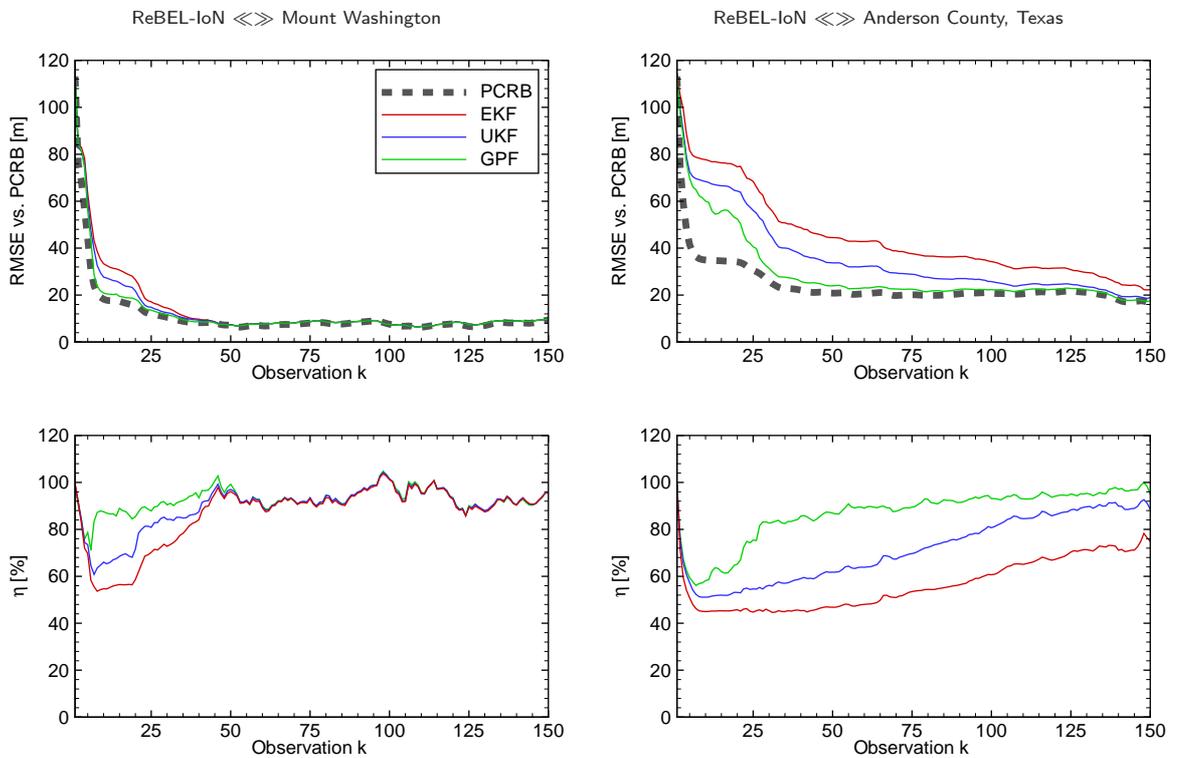


Abbildung 5.10: Oben: RMSE vs. PCRB und Filterwirkungsgrad für das bergige Terrain. Unten: RMSE vs. PCRB und Filterwirkungsgrad für das flache Terrain. Um Platz zu sparen, wurden die RMSE-Verläufe der X- und Y-Positionsschätzung über die Euklidische Norm zusammengefasst.

Algorithmus	Szenario 1 - Mt. Washington						Szenario 2 - Anderson County					
	$\bar{\eta}$	$RTAMSE$	RI	$J_{Rob.}$	\bar{t}_{sim}	TI	$\bar{\eta}$	$RTAMSE$	RI	$J_{Rob.}$	\bar{t}_{sim}	TI
	[%]	[m]	[%]	[%]	[s]	[%]	[%]	[m]	[%]	[%]	[s]	[%]
EKF	86.6	21.02	0.0	99.4	8.40	0.0	56.6	48.13	-18.8	99.4	8.56	-546.7
UKF	89.4	19.87	5.5	99.6	1.39	83.4	71.8	39.89	0.0	99.4	1.32	0
GPF	92.5	18.92	10.1	100	1.71	79.6	86.8	33.57	15.5	100	1.54	-16.7
PCRB	-	16.86	19.7	-	-	-	-	26.38	33.5	-	-	-

Tabelle 5.8: Simulationsergebnisse TAN.

Obwohl der GPF sowohl in der Qualifikationsanforderung als auch im Filterentwurf gemäß der vorgeschlagenen Kriterien schlechter abschneidet, lohnt sich der zusätzliche Mehraufwand in der Einsatzvorbereitung in diesem Beispiel.

Für weitere Information mit detaillierten Applikationsbeispielen und Vergleichen zu alternativen Navigationsmethoden sei auf [86] und [87] verwiesen. Für generelle Informationen über Avionik seien [88] und [89] empfohlen. Einen guten Einstieg in TNS liefert [85].

5.4 Nichtlineare Trajektorienregelung des Klatt-Engell-Reaktors durch E/A-Linearisierung

Moderne nichtlineare Regelungskonzepte basieren i.d.R. auf der Kenntnis des vollen Zustandsvektors. Werden diese nicht alle gemessen, müssen Beobachter bzw. Filter zum Einsatz kommen. In diesem Beispiel soll die entwickelte Methodik am Beispiel einer nichtlinearen Folgeregelung des KLATT-ENGELL-Reaktor-Benchmarks [90] durch exakte E/A-Linearisierung (benötigt volle Zustandsvektorrückführung) gezeigt werden. Der von KLATT&ENGELL untersuchte Prozess beschreibt die Herstellung von Cyclopentenol (Stoff A) aus Cyclopentadien (Stoff B) durch säurekatalysierte elektrophile Addition von Wasser. Die Umsetzung erfolgt bei 15 bar in einem mit Stickstoff inertisierten, kontinuierlich betriebenen Rührkesselreaktor, der in Abbildung 5.11 schematisch dargestellt ist. Der Reaktor ist mit einer Mantelkühlung versehen.

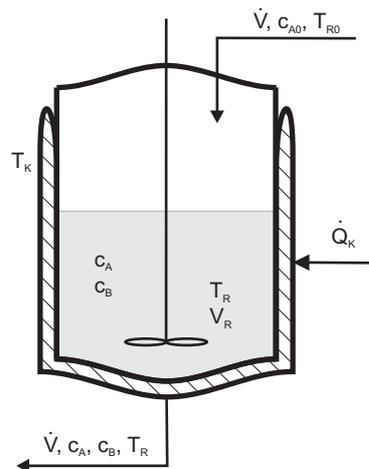


Abbildung 5.11: Schematische Darstellung des KLATT-ENGELL-Reaktor-Benchmarks.

Zusammenfassung exakte E/A-Linearisierung für SISO-Systeme

Die exakte Linearisierung ist eine Methode zur Regelung und Stabilisierung von steuerbaren Systemen, die stetig differenzierbare Nichtlinearitäten aufweisen. Wichtigste Literaturstellen sind [90] und [91].

Während herkömmliche Arbeitspunkt-Linearisierungen lediglich für eine gewisse Umgebung um den Linearisierungspunkt gültig sind, bietet die exakte Linearisierung eine Möglichkeit, Systemen durch eine statische

Zustandsvektorrückführung einen globalen linearen Charakter zuzuweisen. Es können dann für das gewonnene, über den gesamten Arbeitsbereich statisch und dynamisch lineare Verhalten, gewöhnliche lineare Regler mit einem einzigen bedarfs optimalen Parametersatz angewendet werden. Abbildung 5.12 zeigt das Schema einer Regelung durch exakte E/A-Linearisierung.

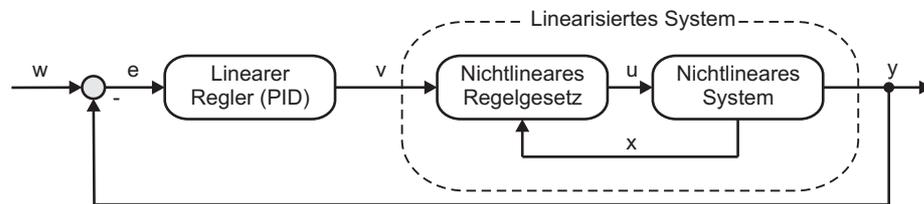


Abbildung 5.12: Schematische Darstellung der exakten E/A-Linearisierung.

Für diese Methoden werden spezielle differentialgeometrische Operatoren benötigt, die aus der Theorie der LIE-Algebra stammen. Für die Anwendung der Methode muss das zu regelnde System in einer definierten Normalform

$$\dot{x} = f(x, u) = f(x) + g(x)u \quad (5.17)$$

$$y = h(x) \quad (5.18)$$

vorliegen, d.h. das System muss eine lineare Steuerung (u multiplikativ verknüpft mit g) aufweisen, was oft der Fall ist. Dem Entwurf des nichtlinearen Regelgesetzes geht die Bestimmung des relativen Grades voraus. Für lineare Systeme entspricht der relative Grad der Differenz zwischen Polen und Nullstellen (Polüberschuss). Für nichtlineare Systeme wird mit Hilfe der LIE-Ableitungen der relative Grad $\mathcal{R}\mathcal{G}$ eingeführt. Für das System ((5.17),(5.18)) werden die LIE-Ableitungen

$$\mathcal{L}_g \mathcal{L}_f^{(0)} h(x) \stackrel{?}{=} 0 \dots \mathcal{L}_g \mathcal{L}_f^{(1)} h(x) \stackrel{?}{=} 0 \dots \mathcal{L}_g \mathcal{L}_f^{(2)} h(x) \stackrel{?}{=} 0 \quad (5.19)$$

solange gebildet, bis

$$\mathcal{L}_g \mathcal{L}_f^{(\mathcal{R}\mathcal{G}-1)} h(x) \neq 0 \quad \forall x \in \mathcal{S} \quad (5.20)$$

ist. Der relative Grad ist dann $\mathcal{R}\mathcal{G}$. Dabei bezeichnet $\mathcal{S} \in \mathbb{R}^n$ den Raum der erwarteten Zustandsgrößen x . Der relative Grad $\mathcal{R}\mathcal{G}$ des Systems muss, um einen vollständigen dynamischen Ausgleich erreichen zu können, gleich der Systemordnung n sein, im linearen Fall besäße das System also keine Nullstellen. Wenn das System einen relativen Grad besitzt, der kleiner ist als die Systemordnung, verbleiben auch nach der exakten Linearisierung noch $n - \mathcal{R}\mathcal{G}$ nicht kompensierbare Dynamiken im System (das wären im linearen Fall die Nullstellen, die auch durch einen Regler nicht zu manipulieren sind). Solange diese Nulldynamik jedoch stabil ist, ist deren Einfluss oft nicht groß, so dass die Methode trotzdem ohne merkbare Einbußen angewandt werden kann. Das einfache Regelgesetz der exakten E/A-Linearisierung lautet:

$$u(t) = \frac{v - \mathcal{L}_f^{(\mathcal{R}\mathcal{G})} h(x)}{\mathcal{L}_g \mathcal{L}_f^{(\mathcal{R}\mathcal{G}-1)} h(x)}. \quad (5.21)$$

Das durch (5.21) linearisierte System besitzt integralen Charakter, das Eingangssignal v wird \mathcal{RG} -mal integriert, d.h. im LAPLACE-Raum gilt $y(s) = \frac{1}{s^{\mathcal{RG}}}v(s)$. Durch geschickte Wahl des Eingangs v kann man bestimmte Systemeigenschaften gewinnen. Werden weitere LIE-Ableitungen, also zeitliche Ableitungen des Ausgangs, hinzugenommen, können Übertragungsfunktionen mit gewünschten dynamischen Eigenschaften gestaltet werden. Mit

$$v = \bar{v} - \sum_{m=0}^{\mathcal{RG}-1} \beta^{(m)} \cdot \mathcal{L}_f^{(m)} h(x) \quad (5.22)$$

erfolgt eine Polvorgabe ähnlich zur linearen Zustandsvektorrückführung. Das durch (5.21) linearisierte System mit gewähltem Eingang (5.22) besitzt die Übertragungsfunktion

$$G(s) = \frac{y(s)}{\bar{v}(s)} = \frac{1}{s^{\mathcal{RG}} + \beta^{(\mathcal{RG}-1)}s^{\mathcal{RG}-1} + \dots + \beta^{(1)}s + \beta^{(0)}}. \quad (5.23)$$

Die Koeffizienten $\beta^{(i)}$ werden nun so gewählt, dass ein gewünschtes dynamisches Verhalten entsteht. Der Reglerentwurf kann dann nach den bewährten Methoden der linearen Regelungstheorie erfolgen. Aus (5.23) erkennt man, dass im stabilen, konvergenten Fall die Übertragungsfunktion die statische Verstärkung $\frac{1}{\beta^{(0)}}$ besitzt. Um dies zu kompensieren, muss die Führungsgröße mit $\beta^{(0)}$ multipliziert werden, um eine korrekte Folgeregelung zu erhalten.

Prozessmodell, Regelungsaufgabe, Reglerentwurf und Evaluierung mit ReBEL-IoN

Prozessmodell

Die Herleitung des Prozessmodells ist ausführlich in [90] beschrieben, wo auch sämtliche Parameter sowie Toleranzen angegeben werden. Daher wird an dieser Stelle nur kurz das kontinuierliche Systemmodell vorgestellt. Die Zustände $x = [c_A \ c_B \ T_R \ T_K]^T$ entsprechen den Konzentrationen der Stoffe A und B sowie der Reaktor- und Kühlmanteltemperatur. Das nichtlineare Differentialgleichungssystem ist durch

$$\dot{c}_A = \frac{\dot{V}}{V_R}(c_{A0} - c_A) - k_1(T_R)c_A - k_3(T_R)c_A^2 \quad (5.24)$$

$$\dot{c}_B = -\frac{\dot{V}}{V_R}c_B + k_1(T_R)c_A - k_2(T_R)c_B \quad (5.25)$$

$$\begin{aligned} \dot{T}_R &= \frac{\dot{V}}{V_R}(T_{R0} - T_R) - \frac{1}{\rho C_p} [k_1(T_R)c_A \Delta H_{RAB} + k_2(T_R)c_B \Delta H_{RBC} + k_3(T_R)c_A^2 \Delta H_{RAD}] \\ &\quad + \frac{k_W A_R}{\rho C_p V_R} (T_K - T_R) \end{aligned} \quad (5.26)$$

$$\dot{T}_K = \frac{1}{m_K C_{PK}} [\dot{Q}_k + k_W A_R (T_R - T_K)] \quad (5.27)$$

gegeben, wobei als Stellgrößen der bezogene Volumenstrom $\frac{\dot{V}}{V_R}$ sowie die Leistung des Wärmetauschers \dot{Q}_k zur Verfügung stehen. k_1 , k_2 und k_3 beschreiben temperaturabhängige (Reaktortemperatur) Reaktionsgeschwindigkeitskoeffizienten. Für die nominellen Parameter und weitergehende Informationen sei nochmals auf [90] verwiesen. Gemessen wird die Produktkonzentration c_B .

Regelungsaufgabe

Es soll eine Folgeregelung für die Produktkonzentration c_B entworfen werden.

Reglerentwurf

Um eine Entkopplung der Gleichung für den Kühlkreislauf zu erreichen, wird das einfache Regelgesetz

$$\dot{Q}_k = u_2 = -k_W A_R (T_R - T_{K0}) = -k_W A_R (x_3 - T_{K0}) \quad (5.28)$$

gewählt und es gilt $\dot{x}_4 = -\frac{k_W A_R}{m_K C_{PK}} (x_4 - T_{K0})$ mit T_{K0} als Temperatur des Kühlmittels. Damit ist die obige Gleichung von den anderen Zustandsgrößen entkoppelt und die Kühlmitteltemperatur x_4 wird konstant gehalten. Da x_4 nicht gemessen werden kann, kommt der triviale Beobachter $\dot{\hat{x}}_4 = -\frac{k_W A_R}{m_K C_{PK}} (\hat{x}_4 - T_{K0})$ zum Einsatz.

Das System ist nun auf ein SISO-System reduziert und kann durch die vorgestellte E/A-Linearisierung geregelt werden. Dazu muss zunächst der relative Grad des Systems bestimmt werden:

$$\mathcal{L}_g \mathcal{L}_f^{(0)} h(x) = \mathcal{L}_g h(x) = \frac{\partial h}{\partial x} \cdot g(x) = -x_2 \neq 0 \quad \forall x_2 \in \mathbb{S}. \quad (5.29)$$

Der relative Grad ist somit $\mathcal{RG} = 1$. Zur Bestimmung des Regelgesetzes muss lediglich

$$\mathcal{L}_f^{(\mathcal{RG}=1)} h(x) = k_1(x_3)x_1 + k_2(x_3)x_2 \quad (5.30)$$

berechnet werden und das nichtlineare Regelgesetz ergibt sich zu:

$$u = \frac{v - (k_1(x_3)x_1 + k_2(x_3)x_2)}{-x_2}. \quad (5.31)$$

Die linearisierte Regelstrecke weist somit integrales Verhalten mit Übertragungsfunktion $G_s(s) = \frac{1}{s}$ auf. Eine strukturstabile Regelung kann mit einem einfachen P-Regler (Übertragungsfunktion $G_r = K_p$) erfolgen. Durch die Rückkopplung ergibt sich die Übertragungsfunktion des geschlossenen Systems $G_w(s) = \frac{1}{K_p s + 1}$. Die Verstärkung muss nun so eingestellt werden, dass Beschränkungen des Stellapparates, z.B. Stellwegbeschränkung oder maximale zeitliche Differentiale der Stellgröße u , eingehalten werden. Für nachfolgende Studien wird $K_p = 2$ festgelegt.

Das Regelgesetz erfordert die Kenntnis der ersten drei Zustände, wobei lediglich x_2 gemessen wird. Die restlichen Zustände müssen durch WI-Verfahren geschätzt werden. Auf den Nachweis der Beobachtbarkeit wird hier aus Gründen der Übersicht verzichtet.

Simulative Evaluierung mit ReBEL-IoN

Das kontinuierliche Prozessmodell wurde mit einer Schrittweite von $\tau = 0.01$ s durch explizite EULER-Diskretisierung diskretisiert. Der benötigte (statistisch mittlere) exogene Eingang (Stellsignal) wurde durch

die nominale Regelung (Zustandsvektor vollständig bekannt) ohne System- und Messrauschen generiert. Für die Simulationsstudie in ReBEL-IoN wurden die in Tabelle 5.9 aufgeführten Rahmenbedingungen gewählt. Die Ergebnisse der Studie sind in Abbildung 5.13 zu sehen.

MC-Simulationen	$MC = 500$
Observationen	$K = 2001$
Initialisierung	$x_0 [\frac{\text{mol}}{\text{L}} \quad \frac{\text{mol}}{\text{L}} \quad K \quad K]^T \sim \mathcal{N}(x_0; [3.22 \ 0.9 \ 386.5 \ 386.1]^T, \text{diag}([0.1 \ 0.1 \ 10 \ 10]))$
Systemrauschen	$\kappa_{k-1} [\frac{\text{mol}}{\text{L}} \quad \frac{\text{mol}}{\text{L}} \quad K \quad K]^T \sim \mathcal{N}(\kappa; 0, \mathbf{I} \cdot 10^{-6})$
Messrauschen	$\gamma_k [\frac{\text{mol}}{\text{L}}] \sim \mathcal{N}(\gamma; 0, 0.4)$
POD	$p_D = 1$
Divergenzgrenze	$DIV_{Thres} [\frac{\text{mol}}{\text{L}} \quad \frac{\text{mol}}{\text{L}} \quad K \quad K]^T = [\infty \ \infty \ \infty \ \infty]^T$
Führungssignal	$x_{2,soll} [\frac{\text{mol}}{\text{L}}] = 0.8 + 0.27\sigma(t - 5) - 0.27\sigma(t - 15) \rightarrow$ Zur Bestimmung von U , siehe Seite 70
Filter	EKF, UKF, CDKF

Tabelle 5.9: Studienparameter des KLATT-ENGELL-Reaktor-Benchmarks.

Beurteilung der Ergebnisse

Die Simulationsstudie zeigt, dass alle Filter für alle Zustände dem PCRB exakt folgen, effizient mit Wirkungsgrad $\eta = 100\%$ sind und nicht divergieren. Aus diesen Gründen wird sowohl auf die Darstellung der RMSE-Verläufe der einzelnen Filter als auch auf die Wirkungsgradverläufe verzichtet. Die bisher stets gezeigte tabellarische Auswertung wird hier ebenso nicht aufgeführt, da die Filter keine signifikanten Unterschiede in den quantitativen Kriterien liefern.

Aufgrund der in dieser Arbeit vorgeschlagenen qualitativen Kriterien wird für die weiteren Studien ein CDKF ausgewählt, da dieser im Filterentwurf gegenüber dem EKF geringere Anforderungen hat und gegenüber dem UKF weniger Parameter zum Einstellen besitzt. Der geringere Installations- und Implementierungsaufwand des EKF kann hier vernachlässigt werden, da die Implementierung in MATLAB® erfolgt und die erforderlichen Algebraoperationen komfortabel unterstützt werden. Bei einer anderen Zielplattform muss neu zwischen einem EKF und einem CDKF abgewägt werden.

Adaptive Regelung

Zur Demonstration des Reglerentwurfs mit Zustandsschätzung durch einen CDKF werden zwei Studien für die Führungssignale

$$x_{2,soll,1} = 0.8 - 0.3\sigma(t - 5) + 0.3\sigma(t - 15) \quad (5.32)$$

$$x_{2,soll,2} = 0.8 + 0.27\sigma(t - 5) - 0.27\sigma(t - 15) \quad (5.33)$$

durchgeführt. Dabei wird eine nominale Regelung (Zustandsvektor bekannt + aktuelle Messung) mit einer adaptiven Regelung (Zustandsvektor geschätzt + Messung gefiltert) verglichen. Der Verlauf der Regelgröße für die zwei Simulationen ist in Abbildung 5.14 zu sehen.

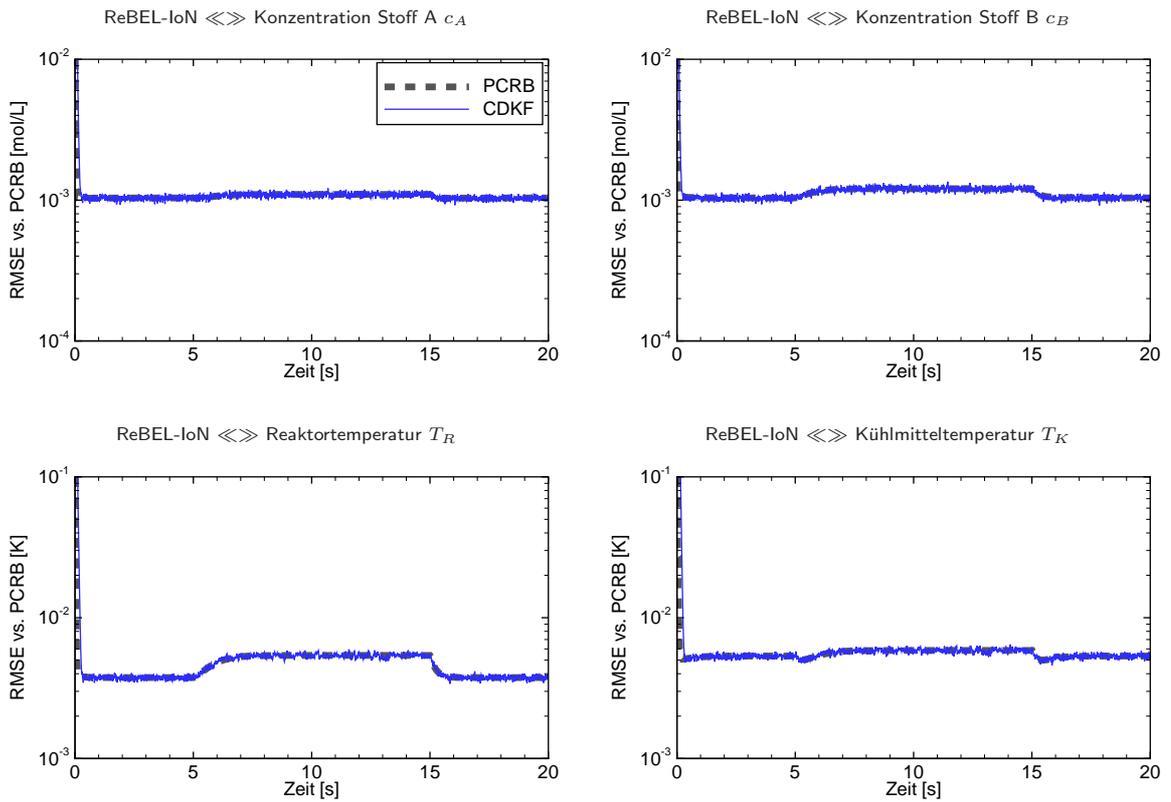


Abbildung 5.13: RMSE vs. PCRB für die Zustände des KLATT-ENGELL-Reaktors.

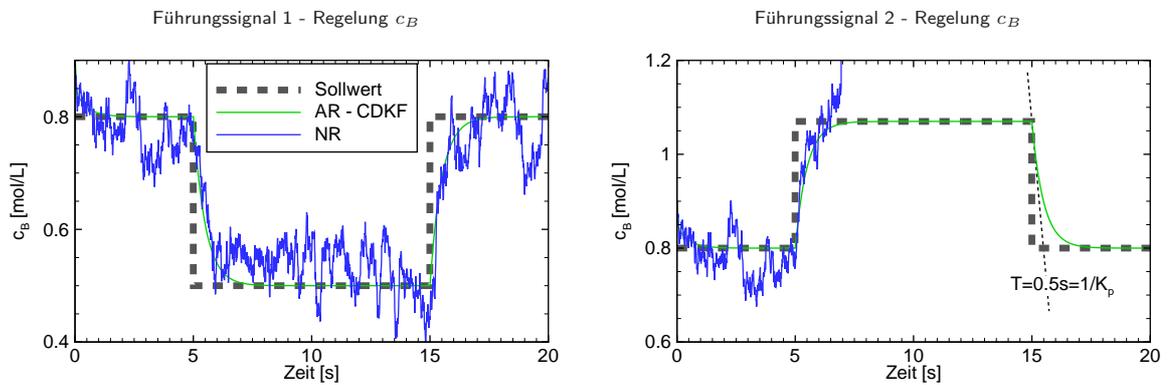


Abbildung 5.14: Regelung der Produktkonzentration c_B für zwei Führungssignale (AR: Adaptive Regelung - NR: Nominale Regelung).

Bereits in der ersten Studie wird deutlich, dass der ungefilterte Einfluss des Messrauschens auf die Regelgröße gravierend ist und zu erheblich schlechterer Regelgüte als beim gefilterten System führt. Die adaptive Regelung ermöglicht eine sehr genaue Regelung mit erwartetem linearem PT1-Verhalten bei jeglichem Arbeitspunktwechsel. Die konstanten Parameter des Reaktors wurden gemäß dem optimalen Betriebspunkt (siehe [90]) gewählt. In dieser Fahrweise besitzt die Nulldynamik des Systems in der Nähe von $c_B = 1.09 \frac{\text{mol}}{\text{L}}$

eine Zentrumsmannigfaltigkeit, d.h. in der Umgebung des optimalen Betriebspunktes gibt es Gebiete mit asymptotisch stabiler Nulldynamik (minimalphasiges E/A-Verhalten) und Gebiete mit instabiler Nulldynamik (nicht minimalphasiges E/A-Verhalten). Die Regelung wird außerordentlich erschwert, da Störungen (z.B. Messrauschen) zu einem Vorzeichenwechsel der stationären Verstärkung und damit zu einer Phasenunsicherheit von 180° führen können. In der zweiten Studie wird deutlich, dass eine stabile Regelung in der Nähe des optimalen Betriebspunktes nur mit der adaptiven Regelung möglich ist. Der eingesetzte CDKF kompensiert das Messrauschen vollständig. Die nominale Regelung wird instabil. An dieser Stelle sei angemerkt, dass der Einsatz eines einfachen Tiefpass-Filters zur Unterdrückung des Messrauschens evtl. eine stabile Regelung liefern kann, aber in der Anregelzeit der adaptiven Regelung mittels CDKF deutlich unterlegen ist. Der Tiefpass würde zwar das Rauschen unterdrücken, allerdings auch das eigentliche Nutzsignal zeitlich verzögern, was zu langsamerer Regelung führt.

5.5 Studie zur modellbasierten Diagnose eines Dieselpartikelfilters (DPF)

Die Überwachung sicherheits- und emissionsrelevanter Komponenten im modernen Kraftfahrzeug gewinnt immer mehr an Bedeutung. So müssen z.B. seit Januar 2001 alle Neufahrzeuge abgasrelevante Bauteile und Systeme während der Fahrt überwachen, um die Grenzwerte vorgeschriebener Abgasemissionen einzuhalten. Immer schärfere Vorgaben des Gesetzgebers erfordern so immer genauere Überwachung.

Der Oberflächenfilter

Bei Oberflächenfiltern erfolgt die Rußabscheidung überwiegend an der Oberfläche der Filterwand. Auf der Filteroberfläche entsteht dadurch der so genannte Filterkuchen, der die eigentliche, hochwirksame Filterschicht darstellt und somit die Rußabscheidung erheblich unterstützt. Durch diesen Effekt können Abscheidegrade von bis zu 99.9 % erreicht werden. Aufgrund des mit wachsendem Filterkuchen zunehmenden Strömungswiderstandes, was zu Leistungsverlust aufgrund erhöhter Ausschubarbeit des Motors führt, muss der Filter periodisch abgereinigt (regeneriert) werden. Dafür existieren verschiedene Abreinigungsmechanismen. Eine Möglichkeit ist es, einen elektronisch gesteuerten Brenner in das Filtersystem einzubauen, der bei einem bestimmten Druckverlust gezündet wird und die Rußpartikel durch ein Heraufsetzen der Abgastemperatur abbrennt. Eine spezielle Bauform des Oberflächenfilters ist der von Daimler verwendete Honeycomb-Filter, der in dieser Arbeit betrachtet wird.

Der Honeycomb-Filter ist ein wabenförmiger Wandstromfilter, bei dem das mit Rußpartikeln versetzte Abgas gefiltert wird, indem es durch eine poröse Filterwand strömt. Dieser Strömungsverlauf wird dadurch erzwungen, dass die Kanäle des Filters gegenseitig versetzt verschlossen sind und somit keinen anderen Weg für das Abgas zulassen. Die Abbildung 5.15 zeigt einen solchen DPF. Die Partikel bleiben durch Adhäsion

an der Oberfläche der porösen Filterwand haften und bilden den Filterkuchen.

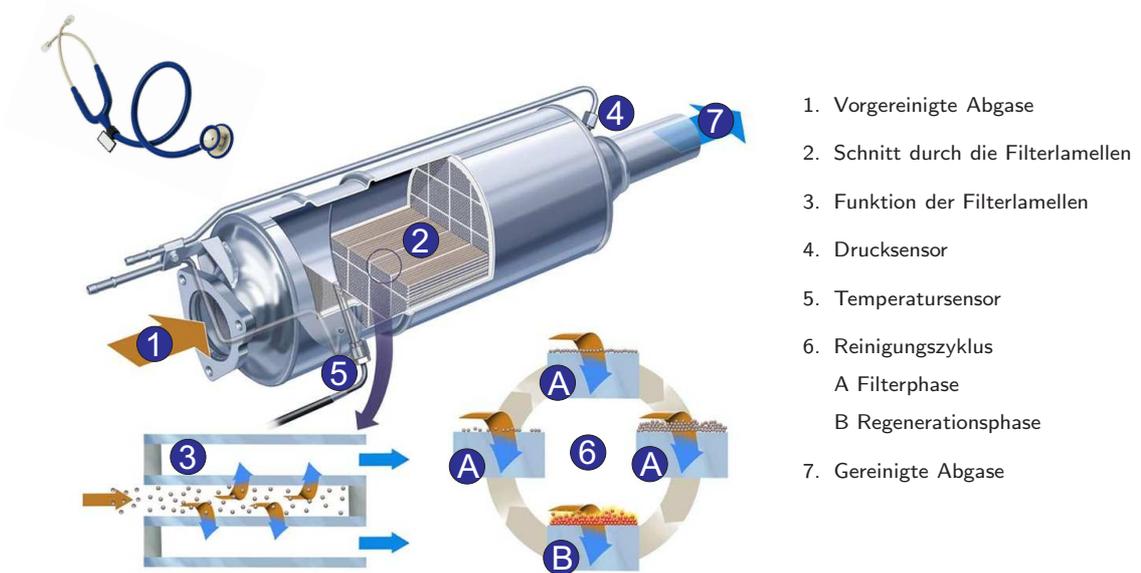


Abbildung 5.15: Honeycomb-DPF.

Diagnoseaufgabe

Wie bereits in der Einleitung erwähnt, ermöglicht es die modellbasierte Diagnose, detaillierte Diagnoseausagen zu treffen. Insbesondere die Kenntnis der Systemzustände und System- bzw. Fehlerparameter lässt eine hohe Diagnosetiefe erreichen. Im nachfolgenden Beispiel wird eine Studie durchgeführt, die aufzeigt, welche wahrscheinlichkeitstheoretischen Inferenzverfahren für die Diagnose eines Dieselpartikelfilters am besten geeignet sein könnten.

Die Diagnoseaufgabe für einen DPF besteht aus drei Teilaufgaben:

1. Der Gesetzgeber schreibt vor, innerhalb eines Fahrzyklusses abgas-/emissionsbeeinflussende Durchbrüche im Filter detektieren.
2. Nominalparameterschwankungen, bedingt durch Fertigungstoleranzen in der Serienproduktion, machen eine On-Line Adaption dieser Parameter innerhalb der ersten Betriebsstunden sinnvoll.
3. Die Schätzung der im Filter befindliche Rußmasse, um eine DPF-Regeneration (gezieltes Abbrennen von Ruß) aktiv einleiten zu können, soll wirkungsvoll unterstützt werden.

5.5.1 Versuchsträger

Als Versuchsträger diente eine Mercedes-Benz S-Klasse (S 320 CDI), Baujahr 2006, mit einem V6-Dieselmotor und einer Leistung von 173 kW. Einige technische Daten des Motors im Versuchsträger sind in Tabelle 5.10 aufgelistet. Das Versuchsfahrzeug wurde mit einem Datenlogger (MultiLOG) der Firma GiN ausgerüstet,

Motordaten	
Zylinderanordnung/-anzahl	V6
Hubraum [cm ³]	2987
Nenn Drehmoment [Nm bei 1/min]	540 / 1600-2400
Nennleistung [kW bei 1/min]	173 / 3600
Kompressionsverhältnis	17.7:1

Tabelle 5.10: Einige technische Daten des Motors im Versuchsträger.

der sowohl CAN-Größen mitloggt als auch durch CCP-Anfragen (CCP: CAN Calibration Protocol) über den CAN-Bus Signale aus Steuergeräten ausliest. Die für die Diagnose relevanten Größen können so dem CAN-Bus entnommen werden. Die Abbildung 5.16 zeigt schematisch den Aufbau der im Fahrzeug installierten Messtechnik. Die Erfassung der CAN-Bus-Signale erfolgt direkt über den Datenlogger. Die geloggten

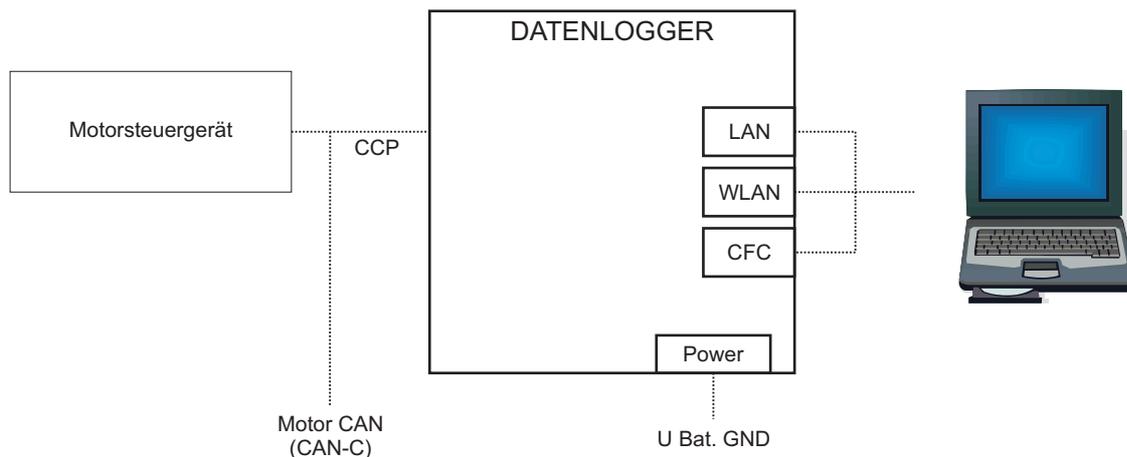


Abbildung 5.16: Datenlogger im Versuchsfahrzeug.

Daten werden auf einer Compact-Flash-Karte gespeichert und können über LAN, WLAN oder ein CF-Karten-Lesegerät aus dem Datenlogger exportiert werden.

Erfasste Messgrößen

Die CAN-Signale sowie die aus der Motorsteuerung abgerufenen Signale unterliegen der Vertraulichkeit und sind daher nicht in dieser Dissertation aufgeführt. Gleiches gilt für Informationen über die Abstraten. Für weitergehende Informationen muss die Absprache mit dem Autor und der Daimler AG erfolgen.

5.5.2 "Blitzdiagnose" über DPF-Kennlinienschtzung mit dem linearen Kalman Filter

Die Kennlinie eines DPF (Δp über \dot{V}) lässt sich durch eine quadratische Gleichung approximieren. Dies resultiert aus der Physik dissipativer Strömungen, welche vereinfacht durch eine quadratische Gleichung beschrieben werden können, siehe z.B. [StA 1]. Es folgt also:

$$\Delta p \approx \theta_1 \cdot \dot{V} + \theta_2 \cdot \dot{V}^2. \quad (5.34)$$

Die Parameter θ_1 und θ_2 lassen sich rekursiv durch einen linearen Kalman Filter schätzen, da o.a. Messgleichung linear in den Parametern ist, d.h.

$$y = [\dot{V} \ \dot{V}^2][\theta_1 \ \theta_2]^T + \gamma_k = H(u)\theta + \gamma_k. \quad (5.35)$$

Der lineare Kalman Filter ist optimal für lineare, GAUSS'sche Systeme. Daher muss keine Evaluierung erfolgen und der KF kann direkt appliziert werden. Die Abbildung 5.17 zeigt die Kennlinien-Schtzung des KF anhand realer UDC-Messdaten.

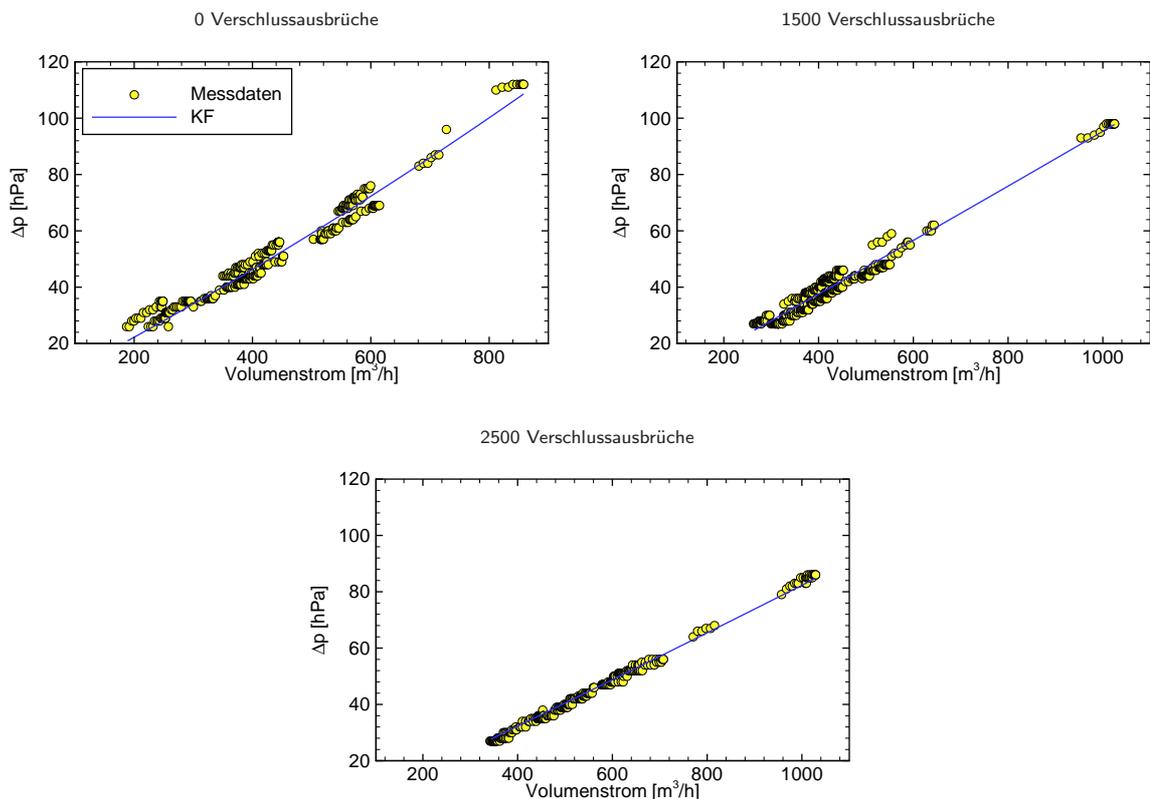


Abbildung 5.17: Rekursive DPF-Kennlinien Schätzung mit dem linearen Kalman Filter.

Aus dem Verlauf der deutlich unterschiedlichen Kennlinien für die drei verschiedenen Fehlerfälle lässt sich

eine Fehlerdiagnose ableiten. Mögliche Ansätze, den Rückschluss von der Kennlinie auf das Fehlerausmaß zu vollziehen, sind Expertensysteme oder kausale Netze (Probabilistisches oder Fuzzy-Logisches Schließen), siehe z.B. [2] oder [92].

Der nachfolgend vorgestellte Ansatz sieht vor, Merkmale der Kennlinie dem Fehlerausmaß zuzuordnen und diese Beziehung in einem künstlich neuronalen Netz (KNN) abzulegen. Die Verwendung von KNN in der MBD ist durchaus gängig, siehe z.B. [93], [94], [95] oder [96]. Die quadratische Kennlinie wird durch die zwei Parameter θ_1 und θ_2 vollständig beschrieben. Sie dienen als Eingangsdaten für das KNN, welches das Fehlerausmaß aus den Parametern berechnet. Als KNN wurde ein Feedforward-Netz mit zwei verdeckten Schichten, jeweils 50 Neuronen und einer Tan-Sigmoidal-Übertragungsfunktion gewählt. Die Trainingsdaten für das KNN wurden durch das im nachfolgenden Kapitel vorgestellte Modell simulativ generiert. Aus dem KNN kann dann für das Fehlerausmaß in Abhängigkeit von θ_1 und θ_2 ein Kennfeld berechnet werden, welches in Abbildung 5.18 zu sehen ist. Aus den UDC-Messdaten mit Kennlinienschätzung und anschließender KNN-Diagnose resultieren die in Tabelle 5.11 gezeigten Diagnoseergebnisse.

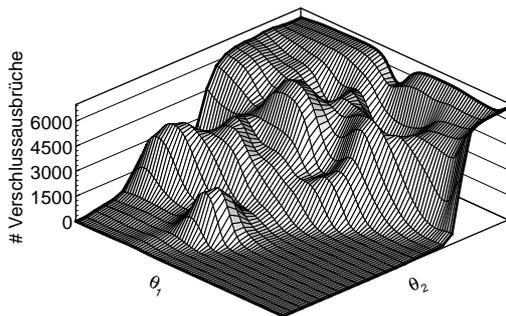


Abbildung 5.18: Kennfeld Diagnose (KNN).

UDC - # Verschlussausbrüche	0	1500	2500
Diagnose KF	77	1036	2378

Tabelle 5.11: Diagnoseergebnisse KF+KNN anhand realer UDC-Daten.

Die Bewertung der Ergebnisse der Blitzdiagnose sowie der nachfolgenden DPF-Modellierung mit Parameteridentifikation und Applikation zur MBD des DPF mittels WI-Verfahren erfolgt geschlossen im Kapitel 5.5.6, ab Seite 129.

5.5.3 Modellbildung, Fehlermodellierung und Nominalparameteridentifikation

Modellierung des statischen Druckverlustes

Von einer gasdynamischen Modellierung des Druckverlustes wird abgesehen, da diese in einem System partieller Differentialgleichungen resultieren würde. Zum einen ist der Aufwand, ein solches System zu lösen, immens (FEM) und zum anderen würde die Modellierung die Berücksichtigung der axialen und radialen Verteilung des Rußes im DPF erfordern. Diese ist jedoch nicht bekannt und unmöglich modellierbar. Des Weiteren ermöglicht die Messdatenaufnahme keine Erfassung der hochdynamischen Gasbewegungen. Das System kann und muss also als quasistatisch angenommen werden. Die statische Druckverlustmodellierung wurde bereits von KONSTANDOPOULOS in [97, 98, 99, 100] ausgiebig untersucht und erfolgreich an statischen Prüfstandsdaten verifiziert. Die nachfolgenden Ausführungen fassen die wichtigsten Ergebnisse zusammen, wobei an einigen Stellen zum Verständnis notwendige Erweiterungen angeführt sind.

Bei der Modellierung des statischen Druckverlustes über den DPF wird davon ausgegangen, dass es sich um eine kompressible Strömung handelt. Die Gaskompressibilität wird vorausgesetzt, da sie sich besonders bei hohen Durchsätzen auswirkt und damit eine Gültigkeit der Ergebnisse auch für hohe Durchsatzraten gegeben ist.

Geometrien des DPF und Druckverlusteffekte

Für die Filterwirkung des DPF sind verschiedene Geometriedaten, wie z.B. Wandstärke w_w , Zelldichte und Porengröße von Bedeutung. Charakteristische Größen sind der effektive Durchmesser D_{eff} , die effektive Kanallänge L_{eff} und die Filterdichte σ , die die Anzahl der Zellen pro Flächeneinheit beschreibt.

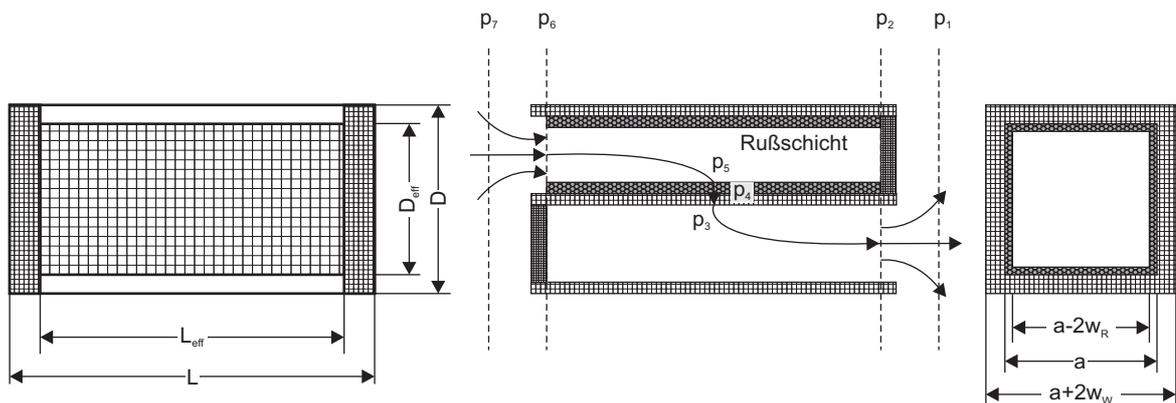


Abbildung 5.19: Längsschnitt durch einen DPF mit seinen charakteristischen Größen und Drücke entlang eines Filterkanals.

Aus diesen charakteristischen Abmaßen lassen sich weitere geometrische Daten errechnen, die für die Modellierung benötigt werden:

- Kanalbreite a : $a = \frac{1}{\sqrt{\sigma}} - w_w$
- Effektives Filtervolumen V_{eff} : $V_{eff} = \frac{\pi}{4} D_{eff}^2 L_{eff}$

Im Folgenden werden alle V_{eff}, L_{eff} durch V bzw. L ersetzt.

Zur Analyse des globalen Druckverlustes über den gesamten Rußpartikelfilter von Eintrittsseite bis Austrittsseite müssen die nachfolgenden Effekte berücksichtigt werden:

- Reibungsverluste entlang der Kanalwände.
- Reibungsverluste beim Durchströmen der porösen Filterwand.
- Reibungsverluste beim Durchströmen der Rußschicht.
- Druckverluste/-anstiege durch Kompression und Expansion am Ein- bzw. Austritt.

Da die Strömung von p_7 am Eingang durch den Filter bis hin zu p_1 am Ausgang verläuft, ergibt sich der gesamte Druckverlust zu $\Delta p = p_7 - p_1$, wobei sich dieser wie folgt auf die verschiedenen Einzeldruckverluste aufteilt:

$$\Delta p = \underbrace{(p_7 - p_6)}_{\text{Kontraktion}} + \underbrace{(p_6 - p_5)}_{\text{Reibung}} + \underbrace{(p_5 - p_4)}_{\text{Rußschicht}} + \underbrace{(p_4 - p_3)}_{\text{Filterwand}} + \underbrace{(p_3 - p_2)}_{\text{Reibung}} + \underbrace{(p_2 - p_1)}_{\text{Expansion}}. \quad (5.36)$$

Um eine geschlossene Gleichung für den gesamten DPF berechnen zu können, werden zunächst die Druckverlustgleichungen der einzelnen Abschnitte bestimmt.

Druckverlust durch Kontraktion

Um den Druckverlust am Eingang des Filters zu bestimmen, muss die Kontraktion der Strömung betrachtet werden. Diese lässt sich annähern durch:

$$\Delta p_{\text{Kontraktion}} = p_7 - p_6 = \zeta_{Kon} \cdot \frac{\rho U^2}{2}. \quad (5.37)$$

Dabei ist ζ_{Kon} der Kontraktionskoeffizient. Für die Strömungsgeschwindigkeit wird $U = \frac{\dot{V}}{A} = \frac{\dot{m}RT}{pA}$ eingeführt mit der spezifischen Gaskonstante R und der mittleren Gastemperatur T sowie der Dichte $\rho = \frac{p}{RT}$ und dem mittleren Druck $p = \frac{p_7 + p_6}{2}$. Durch Einführung des mittleren Drucks wird die Gaskompressibilität berücksichtigt. Setzt man nun noch die Kontraktionsfläche $A = \frac{a^2 V}{(a + w_w)^2 L}$ ein, erhält man am Eingang folgende Gleichung für den Druckverlust:

$$\Delta p_{\text{Kontraktion}} = p_7 - p_6 = \sqrt{p_6^2 + 2 \cdot \zeta_{Kon} \cdot \dot{m}^2 RT \cdot \frac{(a + w_w)^4}{V^2 a^2} \cdot \left(\frac{L}{a}\right)^2} - p_6. \quad (5.38)$$

Druckverlust durch Reibung im Einlasskanal

Der Druckverlust im Einlasskanal des DPF kommt durch die Reibung der Strömung an den Filterwänden zustande. Angenommen wird hierbei, dass es sich um eine völlig entwickelte laminare Strömung handelt, da für einen $5.66'' \times 6''$ 200 cpsi DPF die Reynoldszahl typischerweise unter 1000 liegt. Der Druckverlust

einer solchen Strömung bei einer Kanallänge L berechnet sich über

$$\Delta p_{Einlass} = p_6 - p_5 = \frac{\eta U L}{(a - 2w_R)^2} \cdot F \cdot \xi, \quad (5.39)$$

wobei w_R die Rußschichtdicke und η die dynamische Viskosität darstellt. Der Korrekturfaktor für starke Saugeffekte ξ kann hier gleich eins gesetzt werden, da die Strömungsgeschwindigkeit meist so gering ist, dass die Reynoldszahl unter zwei bleibt. Die Konstante F , welche Reynoldszahl und einen Reibungsfaktor zusammenfasst, wird durch Parameteridentifikation an Messdaten angepasst. Für die Strömungsgeschwindigkeit U und den mittleren Druck gelten die analogen Annahmen wie zuvor. Damit ergibt sich der Druckverlust für den Einlasskanal zu:

$$\Delta p_{Einlass} = \sqrt{p_5^2 + \eta \dot{m} R T \cdot \frac{(a + w_w)^2}{V} \cdot \frac{2FL^2}{(a - 2w_R)^4}} - p_5. \quad (5.40)$$

Druckverlust beim Durchströmen der Rußschicht

Der Druckverlust beim Durchströmen der Rußschicht folgt dem Gesetz von DARCY

$$\frac{dp}{dx} = -\frac{\eta}{k_R} \cdot \frac{\dot{V}}{A_R}, \quad (5.41)$$

mit der Rußpermeabilität k_R und der durchströmten Rußfläche A_R . Ersetzt man den Volumenstrom durch den Massenstrom und die Dichte durch die allgemeine Gasgleichung, ergibt sich:

$$\frac{dp}{dx} = -\frac{\eta}{k_R} \cdot \frac{\dot{m}}{A_R} \cdot \frac{RT}{p}. \quad (5.42)$$

Mit der durchströmten Rußfläche

$$A_R = 4 \cdot (a - 2w_R) \cdot L \cdot N_{Cells} = 4 \cdot (a - 2w_R) \cdot \sigma \cdot V = a \cdot \frac{(a - 2w_R)}{(a + w_w)^2} \cdot V \quad (5.43)$$

ergibt sich folgende Differentialgleichung für den Druckverlust:

$$p dp = -\frac{\eta}{k_R} \cdot \frac{\dot{m} R T}{4\sigma V} \cdot \frac{1}{a - 2x} dx. \quad (5.44)$$

Die Integration der Differentialgleichung über den Druck p und die variable Rußschicht w_R

$$\int_{p_5}^{p_4} = -\frac{\eta}{k_R} \cdot \frac{\dot{m} R T}{4\sigma V} \int_0^{w_R} \frac{1}{a - 2x} dx \quad (5.45)$$

liefert die Gleichung für den Druckverlust beim Durchströmen der Rußschicht:

$$\Delta p_{Ruß} = p_5 - p_4 = \sqrt{p_4^2 + \frac{\eta}{k_R} \cdot \frac{\dot{m} R T \cdot (a + w_w)^2}{4V} \cdot \ln\left(\frac{a}{a - 2w_R}\right)} - p_4. \quad (5.46)$$

Druckverlust beim Durchströmen der Filterwand

Der Druckverlust beim Durchströmen der Filterwand folgt ebenso dem Gesetz von DARCY, das sich auf Grund der Annahme, dass in der Filterwand keine Querströmungen auftreten, auf eine Dimension reduziert. Da das Gesetz nur für eine schleichende laminare Durchströmung gilt, müssen bei hohen

Strömungsgeschwindigkeiten zunehmende Trägheitseffekte, die FORCHHEIMER-Effekte, mit berücksichtigt werden. Dementsprechend gilt:

$$\frac{dp}{dx} = - \underbrace{\frac{\eta}{k_w} \cdot \frac{\dot{V}}{A_{Filter}}}_{\text{DARCY}} - \beta \cdot \rho \cdot \underbrace{\frac{\dot{V}^2}{A_{Filter}^2}}_{\text{FORCHHEIMER}}. \quad (5.47)$$

Hierbei ist k_w die Permeabilität der Filterwand und β der FORCHHEIMER-Koeffizient. Wie zuvor werden auch hier Volumenstrom durch Massenstrom und Dichte durch allgemeine Gasgleichung ersetzt. Es folgt:

$$\begin{aligned} \frac{dp}{dx} &= - \frac{\eta}{k_w} \cdot \frac{\dot{m}}{A_{Filter}} \cdot \frac{RT}{p} - \beta \cdot \frac{p}{RT} \cdot \frac{\dot{m}^2}{A_{Filter}^2} \cdot \left(\frac{RT}{p} \right)^2 \\ &= - \frac{\eta}{k_w} \cdot \frac{\dot{m}}{A_{Filter}} \cdot \frac{RT}{p} - \beta \cdot \frac{\dot{m}^2}{A_{Filter}^2} \cdot \frac{RT}{p}. \end{aligned} \quad (5.48)$$

Zusammen mit der durchströmten Filterfläche $A_{Filter} = a\sigma LN_{Cells} = \frac{4aV}{(a+w_w)^2}$ ergibt sich folgende Differentialgleichung:

$$pdp = \left(- \frac{\eta}{k_w} \cdot \frac{\dot{m}RT}{4a\sigma V} - \beta \cdot \frac{\dot{m}^2 RT}{(4a\sigma V)^2} \right) dx. \quad (5.49)$$

Nach Integration über p und der konstanten Filterwandstärke w_w

$$\int_{p_4}^{p_3} pdp = \left(- \frac{\eta}{k_w} \cdot \frac{\dot{m}RT}{4a\sigma V} - \beta \cdot \frac{\dot{m}^2 RT}{(4a\sigma V)^2} \right) \int_0^{w_w} 1 dx \quad (5.50)$$

ergibt sich folgende Gleichung für den Druckverlust:

$$\Delta p_{Wand} = p_4 - p_3 = \sqrt{p_3^2 + \frac{\eta}{k_w} \cdot \frac{\dot{m}RT(a+w_w)^2}{4aV} \cdot w_w + \beta \cdot \frac{\dot{m}^2 RT(a+w_w)^4}{(4aV)^2} \cdot w_w} - p_3. \quad (5.51)$$

Druckverlust durch Reibung im Auslasskanal

Die Druckverluste im Auslasskanal ergeben sich wie die Druckverluste im Einlasskanal durch die Reibung an den Filterwänden. Mit identischen Annahmen wie beim Einlasskanal ergibt sich folgende Gleichung für den Druckverlust:

$$\Delta p_{Auslass} = p_3 - p_2 = \sqrt{p_2^2 + \eta \dot{m} RT \cdot \frac{(a+w_w)^2}{V} \cdot \frac{2FL^2}{a^4}} - p_2. \quad (5.52)$$

Druckänderung bei der Expansion

Der Druckanstieg bei der Expansion am Filterausgang ist analog zur Druckminderung durch Kontraktion am Eingang. Dementsprechend gilt:

$$\Delta p_{Expansion} = \zeta_{Exp} \cdot \frac{\rho U^2}{2}. \quad (5.53)$$

Dabei ist ζ_{Exp} der Expansionskoeffizient. Daraus folgt durch Einsetzen von Strömungsgeschwindigkeit, allgemeiner Gasgleichung, mittlerem Druck und Expansionsfläche die Gleichung für den Druckanstieg am Ausgang des Filters:

$$\Delta p_{Expansion} = p_2 - p_1 = \sqrt{p_1^2 + 2 \cdot \zeta_{Exp} \cdot \dot{m}^2 RT \cdot \frac{(a+w_w)^4}{V^2 a^2} \cdot \left(\frac{L}{a} \right)^2} - p_1. \quad (5.54)$$

Dieselpartikelfilter - Statische Modellgleichung

Zur Herleitung einer Gesamtverlustgleichung für den kompletten DPF aus den abschnittswisen Druckverlustgleichungen gibt es zwei Möglichkeiten. Zum einen kann unter der Annahme, dass der statische Druck an der Austrittsseite des Filters bekannt ist, eine Lösung gefunden werden. Die Gesamtverlustgleichung lässt sich dementsprechend herleiten, indem von p_1 ausgehend, alle Gleichungen ineinander eingesetzt werden. Der Druckverlust ergibt sich somit zu:

$$\Delta p_{DPF} = \left\{ p_1^2 + \frac{\eta \dot{m} RT}{\sigma V} \left[2FL^2 \left(\frac{1}{a^4} + \frac{1}{(a-2w_R)^4} \right) + \frac{\ln \left(\frac{a}{a-2w_R} \right)}{4k_R} + \frac{w_w}{4ak_w} \right] + \frac{\dot{m}^2 RT}{\sigma^2 V^2 a^2} \left[2(\zeta_{Kon} + \zeta_{Exp}) \left(\frac{L}{a} \right)^2 + \frac{\beta w_w}{16} \right] \right\}^{\frac{1}{2}} - p_1. \quad (5.55)$$

Die zweite Möglichkeit, eine Gesamtlösung zu finden, besteht darin, dass der Druck an der Eintrittsseite des Filters bekannt ist. Für diesen Fall werden die Gleichungen von p_7 ausgehend ineinander eingesetzt und es ergibt sich die folgende Gleichung für den Gesamtdruckverlust des DPF:

$$\Delta p_{DPF} = p_7 - \left\{ p_7^2 - \frac{\eta \dot{m} RT}{\sigma V} \left[2FL^2 \left(\frac{1}{a^4} + \frac{1}{(a-2w_R)^4} \right) + \frac{\ln \left(\frac{a}{a-2w_R} \right)}{4k_R} + \frac{w_w}{4ak_w} \right] - \frac{\dot{m}^2 RT}{\sigma^2 V^2 a^2} \left[2(\zeta_{Kon} + \zeta_{Exp}) \left(\frac{L}{a} \right)^2 + \frac{\beta w_w}{16} \right] \right\}^{\frac{1}{2}}. \quad (5.56)$$

Fehlermodellierung - Einlasskanalverschlussausbruch

Wie bereits kurz zuvor in der Diagnoseaufgabe erwähnt, sollen über den Differenzdruck Verschlussausbrüche diagnostiziert werden. Der Verschluss, der den Einlasskanal verschließt und die Strömung zwingt, durch die Filterwand zu strömen, bricht hierbei aus², siehe dazu Abbildung 5.20. Dies hat den Effekt, dass die Strömung naturgemäß den Weg des geringsten Widerstands "wählt" und durch das offene Ende ungefiltert ausströmt. Dies führt zu einer erheblichen Reduktion des DPF-Wirkungsgrades und muss daher möglichst schnell und genau diagnostiziert werden. Um diesen Effekt erfassen zu können, wird ein Modell vorgestellt, das den veränderten Druckverlust beschreibt, wenn ein oder mehrere Verschlussausbrüche im DPF auftreten.

Bricht ein Verschluss aus einem Einlasskanal des DPF heraus, so strömen zwei Strömungen parallel durch den DPF. Solche parallele Strömungen wurden bereits in [StA 1] für einen Kühlerkreislauf untersucht. Für die Strömung durch zwei parallele Rohrleitungen gelten die Annahmen:

- Der Druckabfall ist in allen Kanälen gleich groß.

²Brüche in den Filterwänden sind ein weiterer auftretender Fehlerfall. Die detaillierte Modellierung dieser ist jedoch nicht möglich. Brüche in den Filterwänden werden somit als Einlasskanalverschlussausbrüche detektiert, da der Einfluss auf den Druckverlust identisch ist ($\Delta p \downarrow$ durch verringerten Strömungswiderstand).

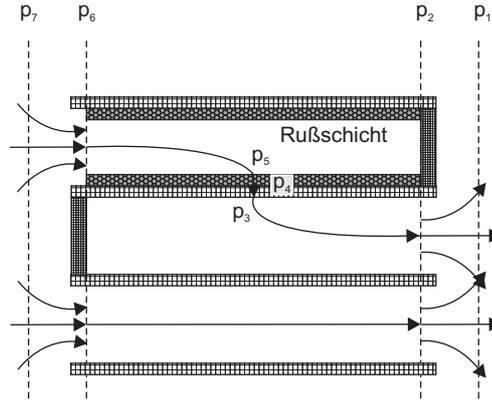


Abbildung 5.20: Verschlussausbruch.

- Der Massenstrom teilt sich je nach Widerstand der einzelnen Kanäle auf.

Diese Annahmen lassen sich in den mathematischen Beziehungen

$$\Delta p^{(1)} = \Delta p^{(2)} = \dots = \Delta p^{(N_{Cells})} = \Delta p \quad (5.57)$$

$$\dot{m}_{Gesamt} = \dot{m}^{(1)} + \dot{m}^{(2)} + \dots + \dot{m}^{(N_{Cells})} = \sum_i^{N_{Cells}} \dot{m}^{(i)} \quad (5.58)$$

erfassen, d.h.

$$\Delta p = \underbrace{(p_7 - p_6)}_{\text{Kontraktion}} + \underbrace{\left\{ \begin{array}{l} \underbrace{(p_6 - p_5)}_{\text{Reibung}} + \underbrace{(p_5 - p_4)}_{\text{Rußschicht}} + \underbrace{(p_4 - p_3)}_{\text{Filterwand}} + \underbrace{(p_3 - p_2)}_{\text{Reibung}} \\ \underbrace{(p_6 - p_2)}_{\text{Reibung}} \end{array} \right\}}_{\Delta p_{Intakt} = \Delta p_{Defekt}} + \underbrace{(p_2 - p_1)}_{\text{Expansion}}. \quad (5.59)$$

Im Fall des defekten DPF mit N_{Intakt} intakten Kanälen und N_{Defekt} defekten Kanälen teilt sich die Strömung auf sämtliche parallele Kanäle auf, wobei die oben aufgeführten Bedingungen erfüllt sein müssen. Des Weiteren gilt die Annahme, dass der in einem Ausbruchkanal befindliche Ruß beim Ausbruch durch das offene Ende ausgeblasen wird und somit die Rußmasse in allen Ausbruchkanälen gegen null geht. Für die parallelen Kanäle des Modells wird das folgende nichtlineare Gleichungssystem aufgestellt, um die Einzeldruckverluste Δp_{Intakt} und Δp_{Defekt} sowie die unterschiedliche Aufteilung der Massenströme \dot{m}_{Intakt} und \dot{m}_{Defekt} zu berechnen:

$$\dot{m}_{Gesamt} - \dot{m}_{Defekt} - \dot{m}_{Intakt} = 0 \quad (5.60)$$

$$\Delta p_{Defekt} - \Delta p_{Intakt} = 0 \quad (5.61)$$

$$X_{Defekt}(\dot{m}_{Defekt}) - \Delta p_{Defekt} = 0 \quad (5.62)$$

$$X_{Intakt}(\dot{m}_{Intakt}) - \Delta p_{Intakt} = 0. \quad (5.63)$$

Dabei sind die Verlustterme X_{Intakt} und X_{Defekt} durch

$$X_{Intakt} = p_7 - \left\{ p_7^2 - \frac{\eta \dot{m}_{Intakt} RT}{\sigma V \frac{N_{Intakt}}{N_{Cells}}} \left[2FL^2 \left(\frac{1}{a^4} + \frac{1}{(a-2w_R)^4} \right) + \frac{\ln \left(\frac{a}{a-2w_R} \right)}{4k_R} + \frac{w_w}{4ak_w} \right] - \frac{\dot{m}_{Intakt}^2 RT}{\sigma^2 \left(V \frac{N_{Intakt}}{N_{Cells}} \right)^2} \frac{\beta w_w}{16} - \frac{\dot{m}_{Gesamt}^2 RT}{\sigma^2 V^2 a^2} 2 (\zeta_{Kon} + \zeta_{Exp}) \left(\frac{L}{a} \right)^2 \right\}^{\frac{1}{2}} \quad (5.64)$$

und

$$X_{Defekt} = p_7 - \left\{ p_7^2 - \frac{\dot{m}_{Gesamt}^2 RT}{\sigma^2 V^2 a^2} 2 (\zeta_{Kon} + \zeta_{Exp}) \left(\frac{L}{a} \right)^2 - \frac{\eta \dot{m}_{Defekt} RT}{\sigma \left(V \frac{N_{Defekt}}{N_{Cells}} \right)} \frac{2FL^2}{a^4} \right\}^{\frac{1}{2}} \quad (5.65)$$

gegeben (ausgehend von p_7), wobei die Rußschichtdicke w_R mit

$$w_R = \frac{a - \sqrt{a^2 - \frac{\tilde{m}_{Ru\beta} \cdot V}{N_{Intakt} \cdot L \cdot \rho_R}}}{2} \quad (5.66)$$

gegeben ist. Dabei ist $\tilde{m}_{Ru\beta} = \frac{m_{Ru\beta}}{V}$ die relative Rußmasse. Das resultierende nichtlineare Gleichungssystem muss numerisch gelöst werden. Bei Implementierung in MATLAB[®] bietet sich die Funktion *fsolve* an.

Parameteridentifikation und Modellverifikation anhand von Messdaten

Die Identifikation der unsicheren Modellparameter ζ_{Kon} , ζ_{Exp} und F wurde in einer Reihe von Studienarbeiten [StA 2] untersucht. Unter Verwendung realer UDC-Fahrzyklus-Messdaten mit unterschiedlicher Anzahl an Verschlußausbrüchen (0, 1500 und 2500) wurde das Modell angepasst. Da keine verlässlichen Messdaten über die Rußentwicklung im DPF zur Verfügung standen, wurde für die Optimierung angenommen, dass $\tilde{m}_{Ru\beta} = 0 \frac{g}{L}$ ist. Die Optimierungsergebnisse sind in Abbildung 5.21 zu sehen. Für die drei Messdatensätze ergeben sich die RMSE zu

$$RMSE_0 = 3.0 \text{ hPa}$$

$$RMSE_{1500} = 2.6 \text{ hPa}$$

$$RMSE_{2500} = 0.9 \text{ hPa},$$

wobei der mittlere $RMSE_{Mittel} = 2.4 \text{ hPa}$ beträgt. Zur weiteren Beurteilung der Regressionsgüte wurden 95 %-ige Konfidenzintervalle (KI) für die Parameter berechnet (mit einer 95 %-igen Wahrscheinlichkeit liegt der wahre Parameter innerhalb des Intervalls), siehe Tabelle 5.12.

Parameter	ζ_{Exp}	ζ_{Kon}	F
KI	±30.8 %	±14.9 %	±0.8 %

Tabelle 5.12: Konfidenzintervalle.

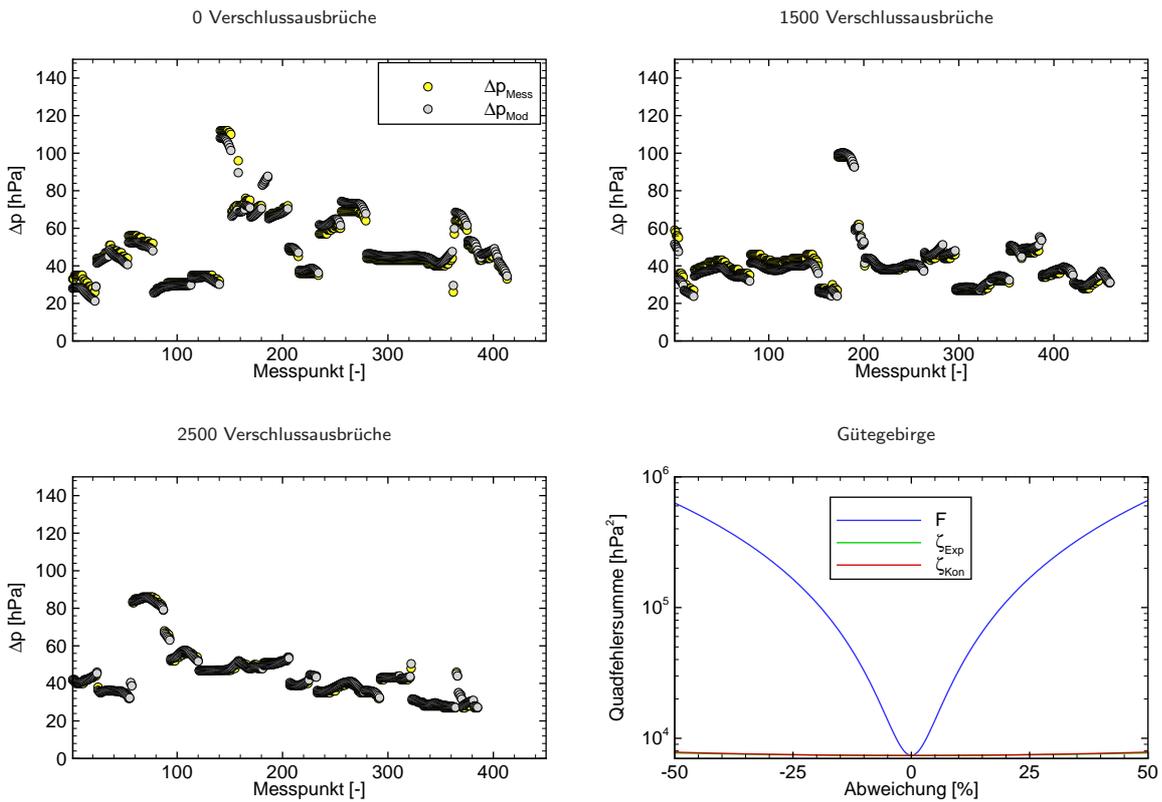


Abbildung 5.21: Modellverifikation und Gütegebirge der identifizierten Parameter.

5.5.4 Diagnose mittels WI-Verfahren

In den folgenden Abschnitten wird gezeigt, wie und mit welchen WI-Verfahren unter Verwendung des zuvor entwickelten Modells eine modellbasierte Diagnose für einen DPF realisiert werden kann. Auf einen Nachweis der Beobachtbarkeit für die folgenden Studien wird verzichtet, da in Analogie zum TAN-Beispiel die benötigten Jacobi-Matrizen durch den Differenzenquotienten (nichtlineares Gleichungssystem) ersetzt werden müssten und der Rangnachweis simulativ erfolgen müsste. Auch hier würde das die Simulation des gesamten möglichen Zustandsraumes erfordern und wird daher nicht näher betrachtet. Die exogenen Eingänge für sämtliche Simulationsstudien wurden aus UDC-Zyklendaten gewonnen.

Einzelschätzung Rußmasse

Simulative Evaluierung mit ReBEL-IoN

Die Schätzung der Rußmasse (Diagnose ist ein belegter Begriff für Fehlerdiagnose) ist für den laufenden Betrieb erforderlich, um den Zeitpunkt einer Regeneration zu bestimmen. Um zu überprüfen, welches WI-Verfahren dazu am besten geeignet und ob die Rußmasse aus der Druckdifferenz schätzbar ist, wurde mit der MATLAB[®] Toolbox ReBEL-IoN eine Simulationsstudie durchgeführt, wobei die in Tabelle 5.13 gezeigten

Rahmenbedingungen festgelegt wurden. Die Abbildung 5.22 zeigt die Ergebnisse der Simulationsstudie. Auf die tabellarische Darstellung der quantitativen Bewertungskriterien wird hier verzichtet, da die Filter keine signifikanten Unterschiede besitzen (Der GPF wird auch effizient, sobald die Partikelanzahl leicht erhöht wird).

Rußmassenmodell	$\theta_k = \theta_{k-1} + \tilde{\kappa}_{k-1}$
MC-Simulationen	$MC = 500$
Observationen	$K = 350$
Initialisierung	$\theta_0 [\frac{g}{L}] \sim \mathcal{U}(\theta_0; 0.5, 1.5)$
Systemrauschen	$\tilde{\kappa}_{k-1} [\frac{g}{L}] \sim \mathcal{N}(\tilde{\kappa}; 0.04, 0.01)$
Messrauschen	$\gamma_k [\text{hPa}] \sim \mathcal{N}(\gamma; 0, 6)$
POD	$p_D = 1$
Divergenzgrenze	$DIV_{Thres} [\frac{g}{L}] = \infty$
Filter	EKF, UKF, CDKF, GPF (25 Partikel)

Tabelle 5.13: Studienparameter Schätzung Rußmasse.

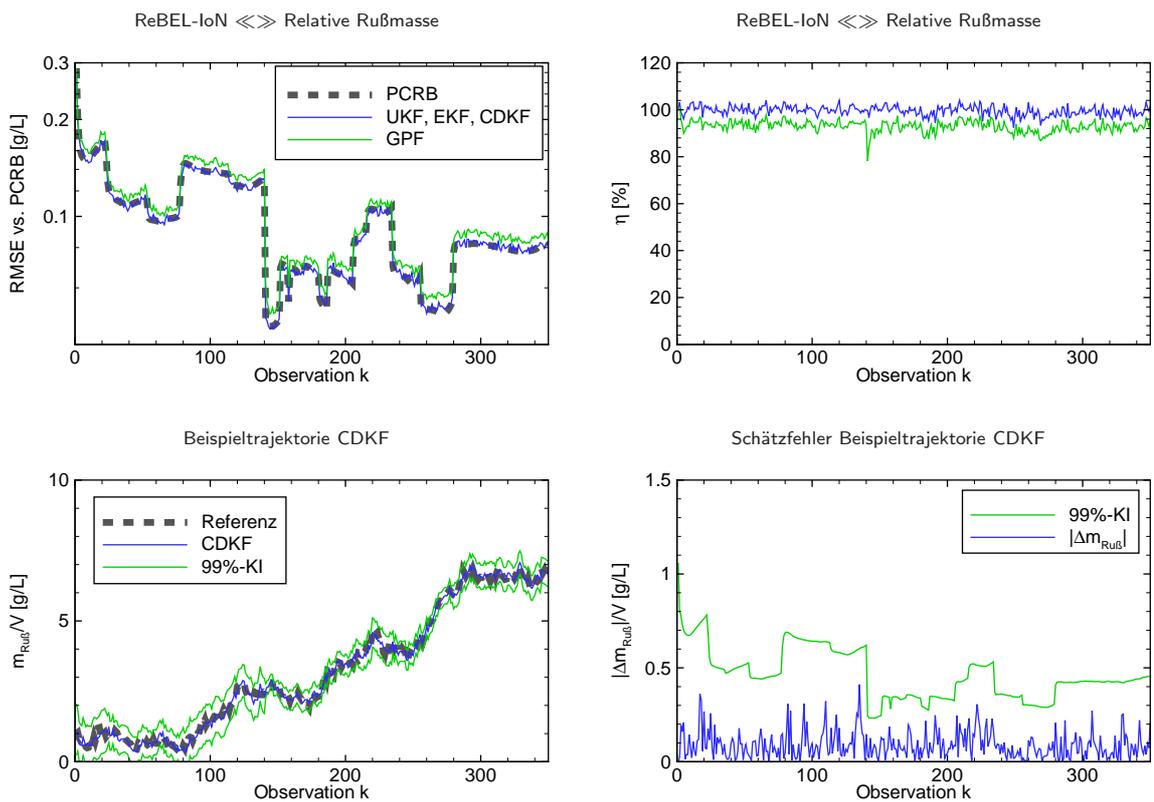


Abbildung 5.22: Einzelschätzung Rußmasse.

Einzel diagnose Ausbrüche Einlasskanalverschlüsse

Simulative Evaluierung mit der ReBEL-IoN

Analog zur Rußmassenschätzung wurde zunächst eine Simulationsstudie mit ReBEL-IoN durchgeführt, um einen Filter auszuwählen. Hierbei wurden zwei Szenarien mit verschiedener Initialisierung der Filter betrachtet. Durch die Notwendigkeit der Lösung des nichtlinearen Gleichungssystems bei einer eintreffenden Observation eignen sich die Partikel-Filter nicht für dieses Problem, obwohl sie durchaus bessere Filtergüte als die Gauß-Filter erzielen könnten. Sie sind rechentechnisch nicht handhabbar (zumindest bei heutiger Rechnertechnologie). Die Berechnung des PCRB erfolgt über den in dieser Arbeit hergeleiteten rekursiven Ansatz für zeitinvariante Parameter (Seite 76) und erfordert die Jacobi-Matrix der Messgleichung. Diese kann analytisch, bedingt durch das nichtlineare Gleichungssystem, nicht exakt bestimmt werden. Um dennoch den PCRB berechnen zu können, wurde der Differentialquotient durch den Differenzenquotient approximiert. Die Rahmenbedingungen der Simulationsstudie sind in Tabelle 5.14 zusammengefasst, die Studienergebnisse sind in Abbildung 5.23 und in Tabelle 5.15 gezeigt. Die unteren Plots zeigen zusätzlich einige Schätztrajektorien des IUKF für verschiedene Fehlerausmaße (0, 2500, 5000) sowie ein 99%-iges Konfidenzintervall (Fehlerfall: 2500) für die Schätzung bei verschiedenen Einstellungen der Kovarianzmatrix $\mathbf{P}^{\tilde{\kappa}}$ des künstlichen Systemrauschens $\tilde{\kappa}$.

Rußmasse	$\tilde{m}_{RuB} [\frac{\mu g}{l}] = 0$
MC-Simulationen	$MC = 500$
Observationen	$K = 350$
Initialisierung	$\theta_0 [-] \sim \mathcal{U}(\theta_0; 0, 6800)$
Szenario 1	$\hat{\theta}_0 = \bar{\theta}_0 = \frac{6800}{2}$
Szenario 2	$\hat{\theta}_0 = 0$
Systemrauschen	-
Messrauschen	$\gamma_k [\text{hPa}] \sim \mathcal{N}(\gamma; 0, 6)$
POD	$p_D = 1$
Divergenzgrenze	$DIV_{Thres} [-] = \infty$
Filter	EKF, UKF, CDKF, IUKF

Tabelle 5.14: Studienparameter Diagnose Verschlussausbrüche.

Algorithmus	Szenario 1 - $\hat{\theta}_0 = \bar{\theta}_0 = \frac{6800}{2}$						Szenario 2 - $\hat{\theta}_0 = 0$					
	$\bar{\eta}$	RTAMSE	RI	$J_{Rob.}$	\bar{t}_{sim}	TI	$\bar{\eta}$	RTAMSE	RI	$J_{Rob.}$	\bar{t}_{sim}	TI
	[%]	[-]	[%]	[%]	[s]	[%]	[%]	[-]	[%]	[%]	[s]	[%]
EKF	16.5	224.4	0	99.6	17.6	0	1.01	4.03e3	0	100	16.8	0
CDKF	68.4	130.9	41.7	100	19.4	-10.2	52.5	288.6	92.8	100	19.5	-16.1
UKF	68.6	129.8	42.2	100	19.5	-10.8	59.0	264.5	93.4	100	19.7	-17.3
IUKF	68.6	129.8	42.2	100	226.7	-1.18e3	64.4	230.7	94.3	100	225.5	-1.24e3
PCRb	-	116.9	47.9	-	-	-	-	117.4	97.1	-	-	-

Tabelle 5.15: Quantitative Bewertungskriterien Diagnose Verschlussausbrüche.

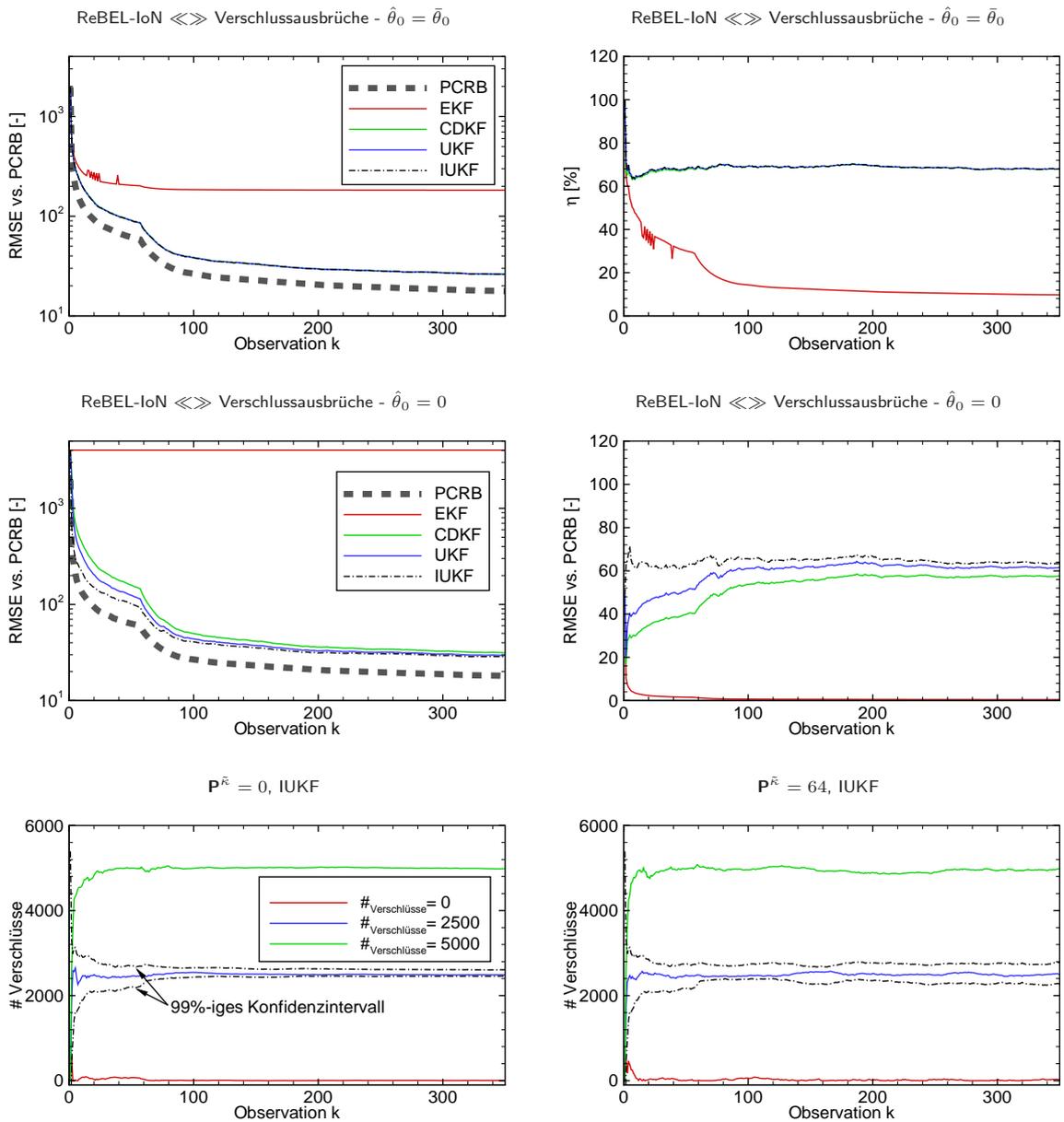


Abbildung 5.23: Einzeldiagnose Verschlussausbrüche.

Verifikation der Simulationsergebnisse anhand realer UDC-Zyklendaten

Um zu testen, ob der Algorithmus mit höchstem Wirkungsgrad (IUKF) auch anhand realer Fahrzeugdaten beste Diagnoseergebnisse liefert, wurden die Filter parallel zu den o.a. Fahrzyklusdaten mit den bereits dokumentierten Fehlerfällen betrieben. Dabei wurde die Kovarianzmatrix des Messrauschens mit $\mathbf{P}^\gamma = 6 \text{ hPa}^2$ und die Initialkovarianzmatrix mit $\mathbf{P}_0^\theta = \frac{6800^2}{12}$ (Varianz der Gleichverteilung) festgesetzt. Auch hier wurden die zwei Szenarien mit unterschiedlicher Initialisierung der Filter betrachtet. Die Abbildung 5.24 zeigt die geschätzte Anzahl an Einlasskanalausbrüchen für die zwei Szenarien sowie im unteren Plot

die Schätztrajektorien des IUKF mit verschiedenen Einstellungen der Kovarianzmatrix \mathbf{P}^k des künstlichen Systemrauschens mit Initialisierung $\hat{\theta}_0 = 0$.

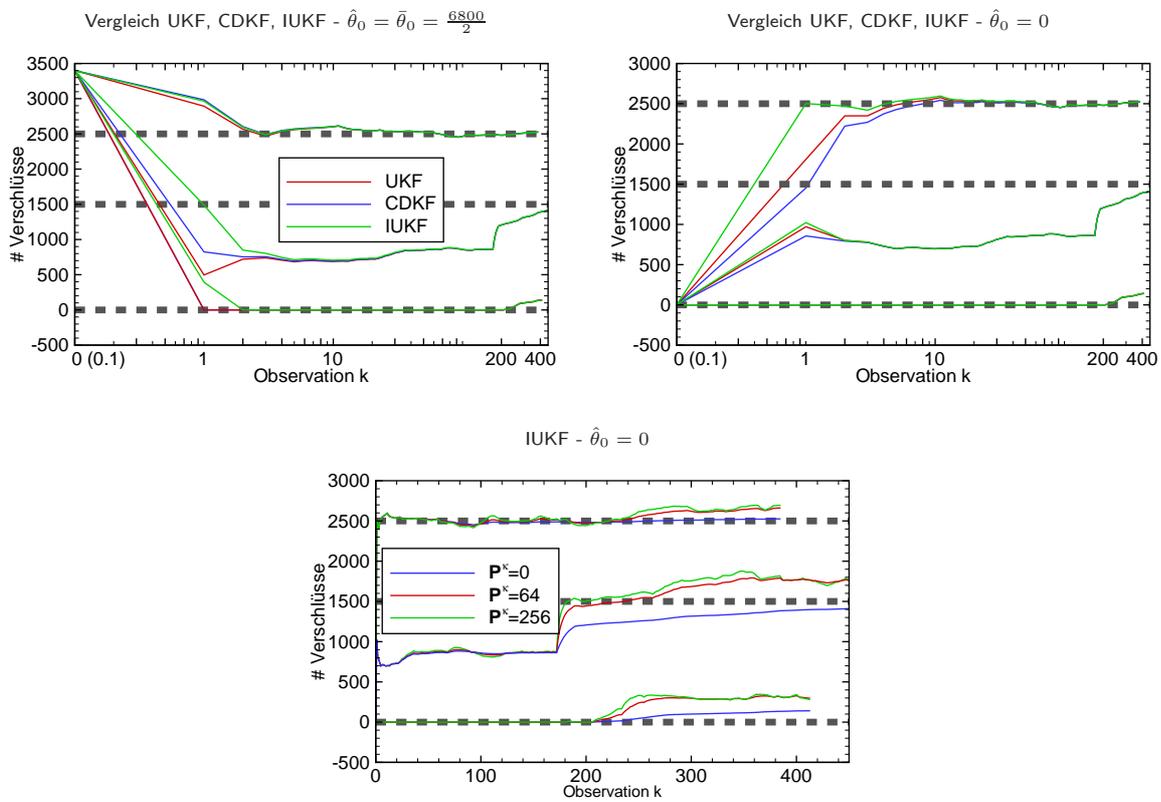


Abbildung 5.24: Einzeldiagnose Ausbrüche Einlasskanalverschlüsse anhand realer UDC-Fahrzyklusdaten.

Gemeinsame Schätzung Rußmasse und Ausbrüche Einlasskanalverschlüsse

Im realen Fahrzeugbetrieb ist es wünschenswert, sowohl die Rußmasse als auch die Ausbrüche der Einlasskanalverschlüsse gleichzeitig zu schätzen, um sowohl eine Fehlerdiagnose als auch eine Regenerationseinleitung zu realisieren. Um zu überprüfen, welches WI-Verfahren dazu am besten geeignet ist und ob die Rußmasse und die Verschlußausbrüche gemeinsam aus der Druckdifferenz schätzbar sind, wurde mit der MATLAB[®] Toolbox ReBEL-IoN eine Simulationsstudie mit den in Tabelle 5.16 aufgeführten Randbedingungen durchgeführt. Es sei dazu angemerkt, dass in dieser Studie nicht die Referenztrajektorien stochastisch initialisiert wurden, sondern die Startwerte der Filter. Bei Gauß-Filtern ist dies aufgrund der Äquivalenz [30]

$$p(\theta_0|\hat{\theta}_{0|0}) \equiv p(\hat{\theta}_{0|0}|\theta_0) \quad (5.67)$$

$$\stackrel{GRV}{=} \mathcal{N}(\theta_0; \hat{\theta}_{0|0}, \mathbf{P}_{0|0}^\theta) \quad (5.68)$$

$$= \mathcal{N}(\hat{\theta}_{0|0}; \theta_0, \mathbf{P}_{0|0}^\theta) \quad (5.69)$$

eine mathematisch äquivalente Studie (GRV: Gaussian Random Variable). Dies war hier notwendig, da das Modell bei stochastischer Initialisierung u.U. komplexe Resultate liefern kann, je nach Trajektorie. Damit würde eine Berechnung des PCRB unmöglich. Auch hier wurde zur Berechnung des PCRB der Differentialquotient durch den Differenzenquotienten ersetzt. Die Abbildung 5.25 zeigt die Ergebnisse der Simulationsstudie. Auf die tabellarische Darstellung der Ergebnisse wird auch hier aufgrund der geringen Abweichungen in der Schätzung verzichtet.

Rußmassenmodell	$\theta_{1,k} = \theta_{1,k-1} + \tilde{\kappa}_{k-1}$
MC-Simulationen	$MC = 500$
Observationen	$K = 350$
Initialisierung	$\hat{\theta}_{1,0} [\frac{g}{l}] \sim \mathcal{U}(\theta_{1,0}; 0.5, 1.5) \rightarrow \text{Ruß}$ $\hat{\theta}_{2,0} [-] \sim \mathcal{U}(\theta_{2,0}; 0, 2500) \rightarrow \text{Verschlussausbrüche}$
Systemrauschen	$\tilde{\kappa}_{k-1} [\frac{g}{l}] \sim \mathcal{N}(\tilde{\kappa}; 0.04, 0.01)$
Messrauschen	$\gamma_k [\text{hPa}] \sim \mathcal{N}(\gamma; 0, 6)$
POD	$p_D = 1$
Divergenzgrenze	$DIV_{Thres} [\frac{g}{l} -] = [\infty \infty]^T$
Filter	UKF, CDKF, IEKF ₁

Tabelle 5.16: Studienparameter Schätzung Rußmasse+Verschlussausbrüche.

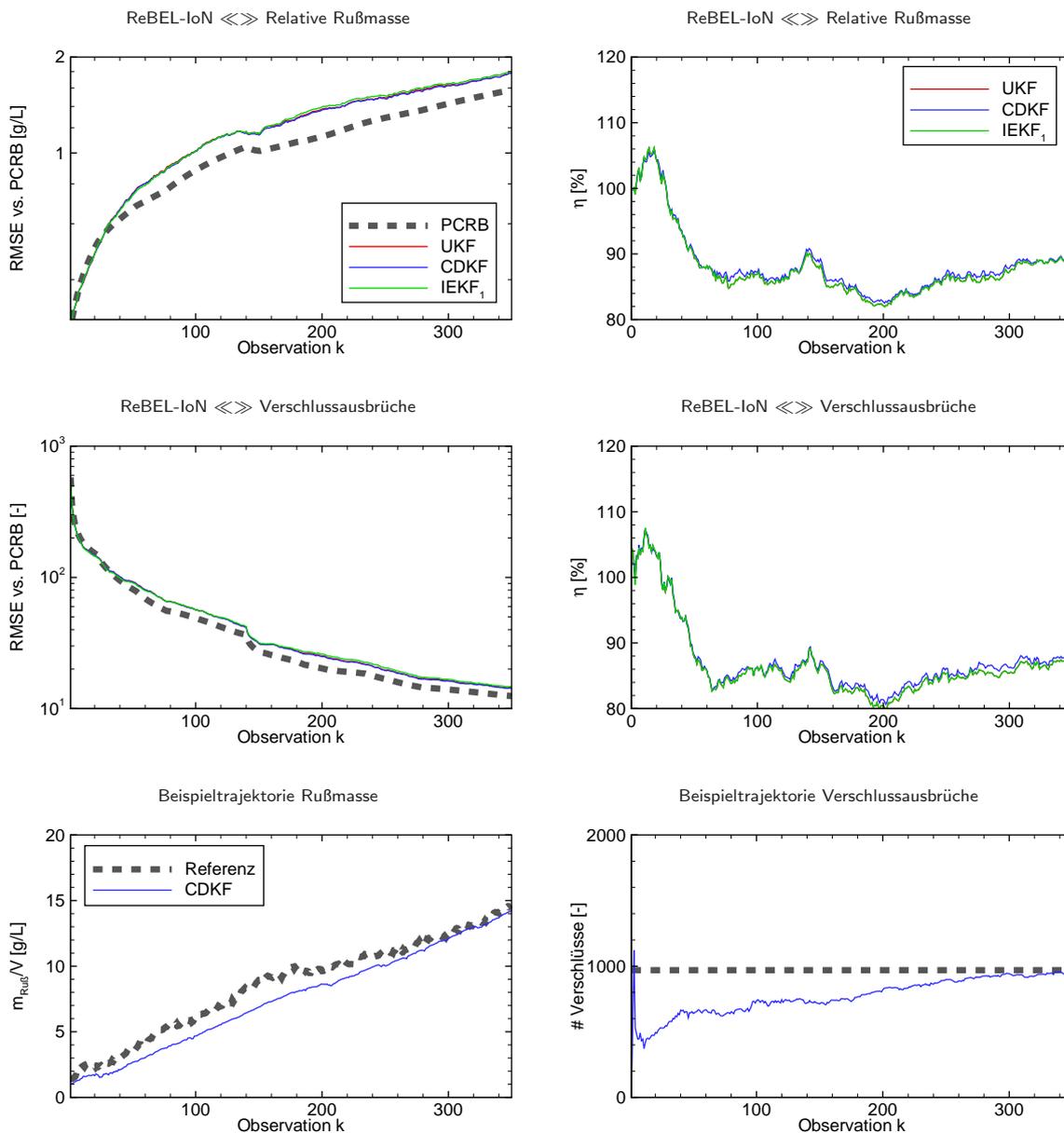


Abbildung 5.25: Gemeinsame Schätzung Ruß+Verschlussausbrüche.

5.5.5 Umgang mit Part-to-Part-Streuung mittels WI-Verfahren

Wie bereits in der Einleitung und zu Beginn dieses Kapitels erwähnt, ist ein Grund für die geringe Akzeptanz modellbasierter Diagnosestrategien im Serienfahrzeugbau die mangelnde Allgemeingültigkeit der Modelle. Zwar sind grundlegende physikalische Eigenschaften von Serienbauteilen identisch, aber in ihren Modellparametern können Schwankungen auftreten. Um eine modellbasierte Diagnosestrategie auf verschiedene Serienfahrzeuge (gilt nicht nur im Automotive-Bereich) anwenden zu können, muss das hinterlegte Modell jeweils auf das einzelne Fahrzeug kalibriert werden. Nachfolgend wird gezeigt, wie unter Verwendung von

WI-Verfahren dieses Problem gelöst werden kann.

Die Modellparameter ζ_{Exp} und ζ_{Kon} werden im Serienbetrieb kaum schwanken, da sie durch die Geometrie des DPF festgelegt sind. Der Parameter F hingegen vereinigt Reynoldszahl und einen Reibungsfaktor, welcher, bedingt durch die Kanalbeschichtung in der Serienfertigung, variieren kann. Weiterhin hat der Parameter einen starken Einfluss auf die Quadratfehlersumme des Modells, siehe Abbildung 5.21, und ist somit geeignet als zu adaptierender Modellparameter.

Simulative Evaluierung mit ReBEL-IoN

Auch hier wurde zunächst eine Simulationsstudie mit ReBEL-IoN durchgeführt, wobei die in Tabelle 5.17 aufgeführten Rahmenbedingungen festgelegt wurden. Die Berechnung des PCRB erfolgte wiederum über die in dieser Arbeit hergeleitete rekursive Bestimmungsgleichung der Filter-Informationsmatrix für zeitinvariante Parameter (siehe Seite 76). Erste Simulationsstudien haben gezeigt, dass kein gravierender

Ruß	$\bar{m}_{Ru\beta} [\frac{\mu}{T}] = 0$
Verschlussausbrüche	$N_{Defekt} [-] = 0$
MC-Simulationen	$MC = 500$
Observationen	$K = 350$
Initialisierung	$\theta_0 [-] \sim \mathcal{U}(\theta_0; \theta_{min}, \theta_{max})$
Systemrauschen	-
Messrauschen	$\gamma_k [\text{hPa}] \sim \mathcal{N}(\gamma; 0, 6)$
POD	$p_D = 1$
Divergenzgrenze	$DIV_{Thres} [-] = \infty$
Filter	CDKF ($\tilde{\kappa} = \mathcal{N}(\tilde{\kappa}; 0, \mathbf{P}^{\tilde{\kappa}} = 0)$, $\tilde{\kappa} = \mathcal{N}(\tilde{\kappa}; 0, \mathbf{P}^{\tilde{\kappa}} = 64)$, Robbins-Monro, λ -Decay)

Tabelle 5.17: Studienparameter Part-2-Part-Streuung.

Unterschied in der Filtergüte zwischen den Gauß-Filtern untereinander sowie zwischen den Gauß-Filtern und den Partikel-Filtern besteht. Daher wurden unter Einsatz eines CDKF verschiedene Einstellungen und Adaptionsmethoden für das künstliche Systemrauschen untersucht. Die Ergebnisse sind in der Abbildung 5.26 zu sehen und in Tabelle 5.18 zusammengefasst, wobei RM Robbins-Monro und LD λ -Decay abkürzt.

Algorithmus	$\bar{\eta}$	$RTAMSE$	RI	$J_{Robustness}$	\bar{t}_{sim}	TI
	[%]	[-]	[%]	[%]	[s]	[%]
$\mathbf{P}^{\tilde{\kappa}} = 0$	100	0.76	0	100	0.84	0
$\mathbf{P}^{\tilde{\kappa}} = 64$	12.6	2.3	-200.3	100	0.84	0
λ -Decay	32.4	1.0	-36.7	100	0.84	0
Robbins-Monro	53.7	0.88	-16.0	100	0.84	0
PCRB	-	0.76	0	-	-	-

Tabelle 5.18: Quantitative Bewertungskriterien On-Line Parameteridentifikation durch Schätzung von F .

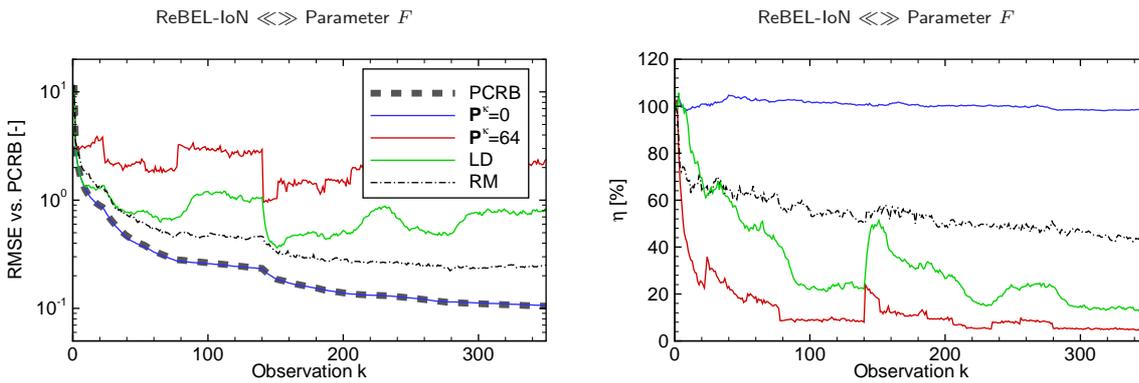


Abbildung 5.26: On-Line Parameteridentifikation durch Schätzung von F .

Verifikation der Simulationsergebnisse anhand realer UDC-Zyklendaten

Um auch hier zu testen, ob der simulativ ausgewählte Algorithmus auch anhand realer Fahrzeugdaten gute Schätzergebnisse liefert, wurde der CDKF parallel zu den o.a. Fahrzyklusdaten ohne Fehlerfall betrieben. Dabei wurde die Kovarianzmatrix des Messrauschens mit $\mathbf{P}^\gamma = 6 \text{ hPa}^2$ und die Initialkovarianzmatrix mit $\mathbf{P}_0^\theta = \frac{(\theta_{max} - \theta_{min})^2}{12}$ (Varianz der Gleichverteilung) festgesetzt. Die Abbildung 5.27 zeigt die On-Line Parameteridentifikation (Schätzung von F) für verschiedene Einstellungen des künstlichen Systemrauschens $\tilde{\kappa}$. Im rechten Bild ist zusätzlich ein 99 %-iges Konfidenzband für die Schätztrajektorie mit $\tilde{\kappa} = \mathcal{N}(\tilde{\kappa}; 0, \mathbf{P}^{\tilde{\kappa}} = 0)$ (Random-Constant-Ansatz) abgebildet.

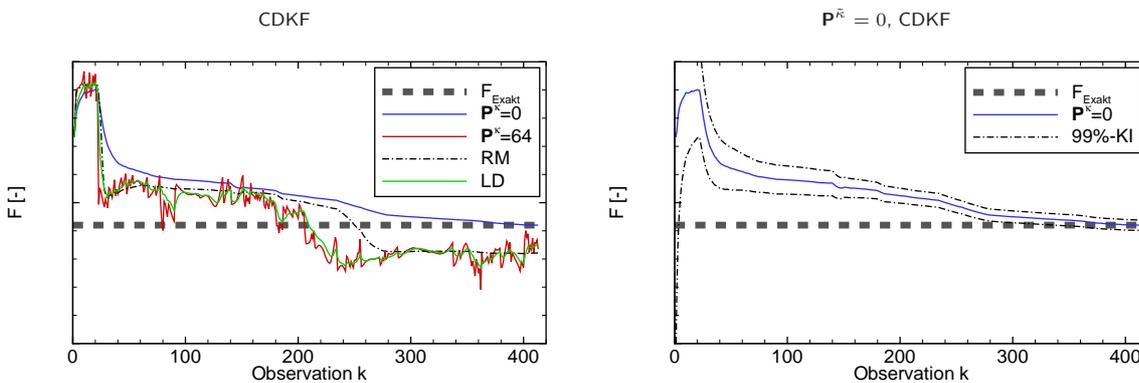


Abbildung 5.27: On-Line Parameteridentifikation durch Schätzung von F mit dem CDKF und verschiedenen Einstellungen für das künstliche Systemrauschen $\tilde{\kappa}$ anhand realer UDC-Fahrzyklusdaten.

Zur weiteren Demonstration der Leistungsfähigkeit von WI-Verfahren zur rekursiven Parameterschätzung wurden die Schätzungen des CDKF mit denen klassischer nichtlinearer (nichtrekursiver) Optimierungsroutinen zweiter Ordnung (Levenberg-Marquardt ([101], [102] und [103]) und Gauß-Newton [103]) verglichen. Die mittleren absoluten prozentualen Abweichungen des Schätzwerts für F , gemittelt über 100 Monte-

Carlo-Simulationen mit stochastischer Initialisierung, sowie die mittleren 99 %-igen Konfidenzintervalle für F sind in Tabelle 5.19 aufgeführt.

Algorithmus	Gütemaß	
	$ \Delta F $	Gemittelttes 99 %-iges KI
CDKF	0.019 %	± 0.96 %
LM	0.119 %	± 0.8 %
GN	0.109 %	± 0.8 %

Tabelle 5.19: Vergleich CDKF - Levenberg-Marquardt (LM) - Gauß-Newton (GN)

5.5.6 Bewertung der Methodik zur modellbasierten Diagnose mittels WI-Verfahren

Die in den vorherigen Unterkapiteln dargestellten Ergebnisse zur modellbasierten Diagnose durch Verwendung von (nicht)linearen Filteralgorithmen sollen im Folgenden diskutiert werden. Dazu wird die Strukturierung der Unterkapitel aufgegriffen, und es erfolgt eine getrennte Beurteilung der Blitzdiagnose, der nichtlinearen Modellierung mit Parameteridentifikation, der Diagnose sowie der Nominalparameterschätzung zum Umgang mit P2P-Streuung.

”Blitzdiagnose”

Die ”Blitzdiagnose” liefert eine zuverlässige Schätzung über den Verlauf der Kennlinie. Das neuronale Netz, welches Merkmale der Kennlinie einem Fehlerausmaß zuordnet, ermöglicht gute Diagnoseergebnisse. Anhand der realen UDC-Fahrzyklendaten mit Fehlerfällen (0, 1500, 2500) konnte die Praxistauglichkeit gezeigt werden. Die Aussage über das KNN lieferte signifikante Diagnoseaussagen, welche in der Größenordnung den realen Bedingungen/Fehlern entsprechen. Eine Diagnose ist somit auf diese Weise genau und vor allem robust möglich. Allerdings gibt es eine Einschränkung: Da die Kennlinie sowohl vom im Filter befindlichen Ruß als auch von den ausgebrochenen Einlasskanalverschlüssen beeinflusst wird, kann keine gemeinsame Schätzung erfolgen. Eine Fehlerdiagnose ist somit nur nach einer vollständigen Regeneration, eine Rußmassenschätzung nur unter der Annahme eines intakten Filters möglich. Für die Rußmassenschätzung aus der Kennlinie ist ebenso ein KNN oder ein Expertensystem notwendig.

Nichtlineare Modellierung

Ziel der Modellierung ist es, den Druckverlust entlang des DPF möglichst genau als Funktion externer, gegebener Eingangsdaten sowie der zu schätzenden Größen ”Rußmasse” und ”Anzahl Verschlussausbrüche” zu simulieren. Die Modellierungsergebnisse zeigen, dass das Systemverhalten in einem weiten Arbeitsbereich gut abgebildet werden kann. Die resultierenden Abweichungen sind als sehr gering einzustufen. Die größeren Abweichungen im intakten Zustand (ohne Verschlussausbrüche) sind auf die Unkenntnis der aktuellen Rußmasse zurückzuführen. Dies wird darin bestätigt, dass mit ansteigender Anzahl ausgebrochener

Verschlüsse die Rußmasse im DPF sowie ihr Einfluss auf den Druckverlust sinkt und folglich die Modellgüte steigt.

Diagnose

Rußmasse

Die Schätzung der Rußmasse aus Differenzdruckmessungen ist unter Verwendung von WI-Verfahren möglich. Die Simulationsstudie hat gezeigt, dass unter den verschiedenen Verfahren kein signifikanter Unterschied in der Filtergüte besteht. Alle Gauß-Filter sind effizient, folgen dem PCRB und liefern Wirkungsgrade um 100 %. Lediglich der GPF weist Wirkungsgrade kleiner als 100 % auf, was nur in der geringen Partikelanzahl begründet ist. Bereits leichte Erhöhung auf ca. 100 Partikel lässt auch den GPF effizient werden. Alle Filter weisen eine Robustheit von 100 % auf. Eine Auswahl basierend allein auf quantitativen Bewertungskriterien wird somit schwierig. Aufgrund der qualitativen Bewertungskriterien sollte ein CDKF zur Schätzung gewählt werden. Wichtigstes Fazit ist, dass die Rußmasse aus der verrauschten Messung genau herausgefiltert werden kann, was die Grundlage für das Einleiten einer Regeneration darstellt.

Ausbrüche Einlasskanalverschlüsse

Eine Diagnose der Anzahl der ausgebrochenen Einlasskanalverschlüsse ist ebenso durch WI-Verfahren realisierbar. Das zu lösende Gleichungssystem bedingt einen starken Rechenaufwand, wodurch die SMC-Methoden bei aktueller Rechenkapazität keine Alternative darstellen, obwohl sie durchaus bessere Filtergüte als die Gauß-Filter erzielen könnten. Die Simulationsstudien (Szenario 1 und 2) haben gezeigt, dass im Mittel ein IUKF mit Random-Constant-Ansatz ($\tilde{\kappa} = 0$) für das künstliche Systemrauschen der Schätzer mit geringstem RMSE ist. Besonders bei Szenario 2, d.h. Initialisierung $\hat{\theta}_0 = 0$, liefert der IUKF signifikant bessere Schätzergebnisse, allerdings auf Kosten erhöhter Rechenzeit.

Die Verifikation der Simulationsergebnisse anhand realer UDC-Fahrzyklusdaten zeigt, dass auch unter realen Bedingungen eine genaue Diagnose innerhalb eines Fahrzyklusses erfolgen kann. Dabei wird besonders deutlich, dass die Simulationsergebnisse aus ReBEL-IoN eine Vorhersage über die Schätzgüte im realen Betrieb zulassen. Besonders bei Szenario 2 mit Fehlerfall 2500 wird deutlich, dass die Rangfolge der Schätzgüte (aufsteigend: CDKF, UKF, IUKF) der Rangfolge aus ReBEL-IoN entspricht. Der IUKF erlaubt eine hochgenaue Schätzung des Fehlerparameters bereits nach der ersten Observation, die anderen Algorithmen benötigen deutlich länger zur Konvergenz, wobei der UKF dem CDKF in Konvergenzgeschwindigkeit überlegen ist. Es sei noch angemerkt, dass der EKF, Filter mit höchstem RMSE in den Simulationsstudien, auch bei den realen Daten schlechteste Schätzungen des Fehlers liefert. Durch die Ungleichheitsbeschränkung $0 \leq \theta \leq 6800$ schätzt der EKF bei allen drei Messreihen den Fehler zu $\hat{\theta} = 0$. Daher sind die Schätztrajektorien des EKF nicht in der entsprechenden Abbildung 5.24 dargestellt. Auch hier werden die Simulationsergebnisse aus ReBEL-IoN in der realen Anwendung wiedergespiegelt.

Analog zur Blitzdiagnose liefern die Schätzer (EKF ausgeschlossen) signifikante Unterschiede für die drei Fehlerfälle. Durch die Fehlerdetektion innerhalb eines Fahrzyklusses sind somit auch die vom Gesetzgeber geforderten Randbedingungen erfüllt. Welcher Algorithmus Verwendung finden sollte, hängt stark von der zur Verfügung stehenden Rechenleistung und benötigter Diagnosegeschwindigkeit ab. Reicht eine Detektion innerhalb eines Fahrzyklusses aus, so sollte, bedingt durch die qualitativen Kriterien, ein CDKF gewählt werden.

Rußmasse & Ausbrüche Einlasskanalverschlüsse

Die Simulationsstudie hat gezeigt, dass auch eine gemeinsame Schätzung der Rußmasse und der Ausbrüche möglich ist. Da auch hier das nichtlineare Gleichungssystem gelöst werden muss, wurden ein UKF, ein CDKF und ein IEKF₁ mit Random-Constant-Ansatz ($\tilde{\kappa} = 0$) für die Schätzung der Verschlussausbrüche getestet. Dabei zeigt der UKF die beste Filtergüte, leicht dem CDKF überlegen. Besonders deutlich wird, dass trotz schlechter Startwerte die Filter gegen die wahren Werte konvergieren. Die Detektion des Fehlers erfolgt auch hier innerhalb eines Fahrzyklusses.

P2P-Streuung

Die Simulationsstudie hat gezeigt, dass zwischen den Gauß-Filtern untereinander sowie zwischen Gauß- und Partikel-Filtern kein signifikanter Unterschied in der Filtergüte besteht. Aufgrund der qualitativen Bewertungskriterien sowie des erhöhten Rechenaufwands der PF wurde ein CDKF ausgewählt. In Simulationsstudien mit ReBEL-IoN wurde gezeigt, dass ein Random-Constant-Ansatz zu einem effizienten Filter mit Wirkungsgrad 100 % führt. Alle Einstellungen für das künstliche Systemrauschen sind robust, allerdings sinkt der Wirkungsgrad sowohl bei den Adaptionsmethoden als auch beim Random-Walk-Ansatz ($\tilde{\kappa} \neq 0$). Die Adaptionsmethoden haben lediglich den Vorteil, dass keine Einstellung von $\tilde{\kappa}$ erfolgen muss. Die Verifikation der Studienergebnisse erfolgte anhand realer UDC-Fahrzyklusdaten. Die bestmögliche Einstellung in Simulationsstudien liefert auch im realen Betrieb die besten Schätzungen mit geringstem RMSE. Der wahre Parameter F des DPF-Modells kann durch einen CDKF On-Line adaptiert werden. Der Vergleich mit den klassischen nichtlinearen, nichtrekursiven Optimierungsalgorithmen zweiter Ordnung zeigt, dass die Identifikation mit WI-Verfahren mindestens gleiche oder bessere Güte liefert. Im Vergleich zu bisherigen modellbasierten Methoden weist die vorgestellte Methodik somit eine erhöhte Praxistauglichkeit auf.

5.6 Kapitelzusammenfassung

Dieses Kapitel ist in fünf Hauptteile aufgeteilt:

Im ersten Teil wurde die in dieser Dissertation vorgestellte Methodik zur Evaluierung von wahrscheinlichkeitstheoretischen Inferenzverfahren an zwei illustrativen, aber eher akademischen Anwendungsbeispielen

demonstriert. Es wurde gezeigt, dass die Kombination aus simulativer, quantitativer Bewertung und qualitativen Bewertungskriterien eine fundierte Basis zur Auswahl von Filteralgorithmen darstellt.

Am Beispiel einer reversiblen Gasgleichgewichtsreaktion, ein Beispiel/System mit beschränkten Zustandsräumen, wurde demonstriert, wie sich die Integration von Informationen über diese Zustandsschranken auf die Schätzgüte und Robustheit von SMC-Methoden auswirken. Die neu entwickelten Algorithmen übertreffen ihre unbeschränkten Gegenparts in oben genannten quantitativen Kriterien.

Der dritte Teil befasste sich mit dem Problem terrainbasierter Navigation moderner Flugzeuge. Es konnte gezeigt werden, dass die entwickelte Methodik bei Terrain-Aided-Navigation angewendet werden kann, um den Anwender bei der Auswahl eines geeigneten Verfahrens zu unterstützen.

Die nichtlineare Regelung des KLATT-ENGELL-Reaktors durch exakte E/A-Linearisierung mit Zustandsschätzung durch WI-Verfahren wurde in Abschnitt vier vorgestellt. Zwar konnte kein signifikanter Unterschied in der Güte der Zustandsschätzungen festgestellt werden, jedoch ermöglicht eine Regelung mit Einsatz von WI-Verfahren einen stabilen Betrieb des Reaktors, wohingegen eine Regelung mit vollständiger Zustandsvektorrückführung (z.B. alle Zustände werden gemessen) instabil wird.

Im fünften Abschnitt wurde die modellbasierte Diagnose eines Dieselpartikelfilters unter Verwendung von WI-Verfahren dargestellt. Es wurde gezeigt, dass nichtlineare Filter zur Rußmassenschätzung, zur Fehlerdiagnose und zur kombinierten Rußmassen- und Fehlerparameterschätzung eingesetzt werden können. Des Weiteren können die Filterverfahren zur On-Line Parameteridentifikation/Modelladaption angewendet werden.

6 Zusammenfassung und Ausblick

Aufgabenstellung

Die vorliegende Arbeit behandelt die Evaluierung von wahrscheinlichkeitstheoretischen Inferenzverfahren für dynamische Systeme und die Applikation als Filter zur rekursiven Zustands- und Parameterschätzung. Die hohe Anzahl nichtlinearer Filteralgorithmen und der Mangel an einer praktischen Methode zur Bewertung und Auswahl inspirierten dieses Projekt. Neben der Aufbereitung stochastischer Filter für komplexe Zustands- und Parameterschätzaufgaben sollte eine Methodik zur praxisnahen Bewertung entwickelt werden, deren anwendungsfreundliche Umsetzung in einer MATLAB[®] basierten Toolbox erfolgen sollte. Als industrielles Anwendungsbeispiel sollte die Verwendung von WI-Verfahren als Filter in der modellbasierten Diagnose eines Dieselpartikelfilters gezeigt werden. Dabei war besonders darauf zu achten, dass vom Gesetzgeber vorgegebene Randbedingungen erfüllt werden und die Methodik für die Serienfertigung applizierbar ist.

Methodik

Evaluierung von WI-Verfahren

Die Bewertung von nichtlinearen Filteralgorithmen erfolgt durch simulatives Testen und quantitative Bewertung ergänzt durch qualitative Bewertungskriterien. Dazu war es notwendig das breite Spektrum an anspruchsvollen WI-Verfahren zu studieren, die Verfahren zu vereinheitlichen und für komplexe Zustands- und Parameterschätzaufgaben aufzubereiten.

Als quantitative Bewertungskriterien wurden *Filtergüte* (Wirkungsgrad + Relative Verbesserung), *Filterrobustheit* und *Rechenzeit* eingeführt. Die Berechnung der Filtergüte erfolgt über den Posterior Cramér-Rao Bound, der den Root Mean Squared Error erwartungstreuer Filter beschränkt und daher die Definition eines Filterwirkungsgrades erlaubt. Für die Schätzung zeitinvarianter Parameter wurde zudem eine neuartige rekursive Bestimmungsgleichung der Filter-Informationsmatrix (Inverse des PCRB) hergeleitet. Zur simulativen Evaluierung wurde eine MATLAB[®] basierte Testumgebung entwickelt.

Die untersuchten qualitativen Bewertungskriterien beurteilen die *Qualifikationsanforderungen an den Anwender*, den *Aufwand* und die *Unterstützung beim Filterentwurf* sowie den *Installations- und Implementierungsaufwand*. Um den abschließenden Entscheidungsprozess bei der Filterauswahl zu unterstützen oder vorab Filter auszuschließen, wurden für die qualitativen Kriterien Bewertungsmaßstäbe eingeführt und eine

hierarchische Anforderungsstruktur für die Gewichtung sämtlicher Kriterien vorgegeben.

Zusätzlich zur Evaluierungsmethodik wurden für die SMC-Methoden Algorithmen entwickelt, um Informationen über harte Zustandsbeschränkungen in den Schätzprozess zu integrieren. Grundidee dabei ist, die unbeschränkte priori und posteriori Verteilungen durch GMM-Modelle zu approximieren und über einen Projektionsansatz neue, beschränkte Punktmassenapproximationen zu generieren.

Modellbasierte Diagnose eines DPF mit WI-Verfahren

Es wurde eine modellbasierte Diagnose mit einem Grey-Box-Modell entwickelt. Das Modell ist in weiten Teilen physikalisch fundiert, beinhaltet jedoch vereinfachende Annahmen, wobei unsichere Modellparameter an Messdaten angepasst wurden. Die Motivation für die Diagnose mit grauen Modellen ist, ein im Fahrzeug On-Board rechenbares Modell mit physikalisch interpretierbaren System- und Fehlerparametern zu gewinnen, welche bei On-Line Schätzung eine direkte, detaillierte Aussage über Fehlerort und -ausmaß zulassen. Durch die zu Grunde liegenden physikalischen Beziehungen sind solche Modelle in einem weiten Arbeitsbereich gültig und durch die Adaption weniger Parameter an baugleiche Systeme anpassbar.

Zur Schätzung der im DPF befindlichen Rußmasse sowie der Ausbrüche von Einlasskanalverschlüssen wurden WI-Verfahren verwendet, wobei eine Simulationsstudie in der MATLAB[®] Toolbox ReBEL-IoN Basis für die Auswahl war. Zur Adaption der Modelle wurden ebenfalls WI-Verfahren eingesetzt, welche nominale Modellparameter On-Line adaptieren können. Auch hier wurde in Simulationsstudien unter Verwendung der o.a. Methodik zur Evaluierung zunächst ein geeigneter Algorithmus ausgewählt.

Ergebnisse & Fazit

Evaluierung von WI-Verfahren

Es konnte gezeigt werden, dass eine Kombination aus quantitativen und qualitativen Bewertungskriterien eine effiziente Beurteilung von Filterverfahren ermöglicht und dabei speziell ungeübte Anwender in der Auswahl unterstützt. Die in dieser Arbeit entwickelte MATLAB[®] basierte Toolbox ReBEL-IoN erlaubt es dem Anwender, sich voll auf die Modellierung des Systems zu konzentrieren und Zeit für die Implementierung von Filtern zu sparen. Durch den implementierten Testalgorithmus, der die reproduzierbaren Monte-Carlo-Simulationen organisiert, wird ein fairer numerischer Vergleich der Filterverfahren erreicht. Anhand der illustrativen Beispiele wurde gezeigt, dass die entwickelte Methodik in vielen technischen Bereichen Anwendung finden kann und einen wertvollen Beitrag zur verbesserten Akzeptanz und Anwendungsbreite moderner Filteralgorithmen liefert. Besonders hervorzuheben ist, dass die simulativ gewonnenen Ergebnisse auf reale Systeme direkt übertragbar sind, sofern ein verlässliches Referenzmodell und ein ausgewogener, praxisrelevanter Test vorliegt.

Die durchgeführten Simulationsstudien zu den neu entwickelten beschränkten SMC-Methoden zeigen, dass die gewonnenen Schätzverfahren bei beschränkten Schätzproblemen über bessere Schätzgüte und höhere Robustheit verfügen. Hier wurde ein wichtiger Beitrag zur Verbesserung moderner Schätzverfahren geleistet.

Modellbasierte Diagnose eines DPF mit WI-Verfahren

Die Applikation von nichtlinearen Filterverfahren als rekursive Parameterschätzverfahren zur modellbasierten Diagnose eines DPF konnte erfolgreich realisiert werden. Dabei können Filter zur (Fehler)diagnose sowohl der im Filter befindlichen Rußmasse als auch der Anzahl ausgebrochener Einlasskanalverschlüsse eingesetzt werden. Des Weiteren erlauben die WI-Verfahren eine Adaption von nominalen Systemparametern, welche in der Serienfertigung durch Fertigungstoleranzen schwanken, wobei die Güte der Parameteridentifikation mit denen nichtlinearer, nichtrekursiver Optimierungsroutinen zweiter Ordnung vergleichbar (oder besser) ist. Die Lösung dieses Problems ist ein wichtiger Aspekt in der Praxistauglichkeit von modellbasierter Diagnosemethoden, um Serienbauteile zu kalibrieren. Hier wurde ein wichtiger Beitrag zur verbesserten Akzeptanz und Einsatzfähigkeit dieser modernen Diagnosemethoden geleistet.

Ausblick auf weitere Arbeiten

Nichtlineare Filter sind ein Gebiet stetig voranschreitender Forschung mit vielen Anwendungsbereichen. Dadurch werden laufend neue Algorithmen entwickelt. Ebenso tragen steigende Rechenkapazität sowie kleiner werdende Recheneinheiten dazu bei, dass neue Anwendungsbereiche erschlossen und neue Methodiken zur Filterung entwickelt werden. Dies erfordert eine fortlaufende Weiterentwicklung der eingeführten Methodik und der Software ReBEL-IoN. Neue vielversprechende Konzepte wie Moving Horizon Estimation sind in diesem Zusammenhang zu nennen. Des Weiteren sollten für die Partikel-Filter und die Hybrid-Filter Methoden zur Integration von Gleichheits- und Ungleichheitsbeschränkungen (linear) entwickelt werden, um das Anwendungsspektrum weiter zu verbreitern. Ebenso ist die zeitabhängige Gewichtung der Filterwerte bzw. Messungen für die untersuchten Verfahren zu entwickeln, wodurch verhindert werden kann, dass zeitlich alte Filterwerte Einfluss auf aktuelle Schätzungen nehmen können (siehe z.B. [104]). Dies ist besonders dann sinnvoll, wenn die Prädiktion mit einem falschen Modell erfolgt. Nach endlicher Zeit stimmen selbst exakte Schätzwerte nach der Prädiktion nicht mehr mit der Realität überein und sind somit für den aktuellen Filterzyklus wertlos.

Die Anwendung der MBD für den DPF im Serienbetrieb erfordert die einmalige Bereitstellung von zuverlässigen Messdaten über die Rußentwicklung, um das Modell daraufhin zu optimieren. Dies könnte durch Lastkollektive und/oder Prüfstandsbeurteilungen realisiert werden. Eine anschließende Rußmassenschätzung kann dann durch ein WI-Verfahren erfolgen. Die vorgestellte Methodik zur MBD könnte ebenso auch auf andere Anwendungen übertragen werden. So ist der Einsatz zur Diagnose von Kühlsystemen denkbar, wobei Fehler wie Kühlerverschmutzung oder Thermostatdefekte zu detektieren sind. Aus ihrer Kenntnis und

Inversion des Kühlsystemmodells könnte so eine fehleradaptive Motorleistungsbegrenzung erfolgen, was die Verfügbarkeit des Fahrzeugs steigern kann. Ebenso interessant als weitere Anwendung ist die Betriebspunkt-optimierung moderner Hybridantriebe durch Schätzung des Batteriezustandes (State of Health/State of Charge) oder die Fehlerdiagnose solcher Antriebskonzepte.

Literaturverzeichnis

- [1] P. Eykhoff. *System Identifikation*. Wiley, 1974.
- [2] R. Isermann. Modellgestützte Überwachung und Fehlerdiagnose. *atp*, 38(5), 1996.
- [3] V. Krebs. *Nichtlineare Filterung*. R. Oldenbourg Verlag GmbH, München, 1980.
- [4] B. Øksendal. *Stochastic Differential Equations*. Springer Verlag, 2004.
- [5] R.E. Kalman. A new approach to linear filtering and prediction problems. *Trans. ASME, J. Basic Eng.*, 82D:35–45, March 1960.
- [6] A. Gelb. *Applied Optimal Estimation*. The MIT Press, Cambridge, Massachusetts, 1974.
- [7] R. van der Merwe. *Sigma-Point Kalman Filters for Probabilistic Inference in Dynamic State-Space Models*. Phd thesis, OGI School of Science and Engineering at Oregon Health and Science University, 2004.
- [8] S. Haykin. *Kalman Filtering and neural networks*. Wiley, New York, 2001.
- [9] A.H. Jazwinsky. *Stochastic Processes and Filtering Theory*. Academic Press, New York and London, 1970.
- [10] Oki Scientific Research. Nonlinearity-Compensation Extended Kalman Filter and Its Application to Target Motion Analysis. *Oki Electric Industry Co*, 63(159), 1997.
- [11] R.K. Mehra. On the Identification of Variances and Adaptive Kalman Filtering. *IEEE Transactions on Automatic Control*, 15(2):175–184, 1970.
- [12] M. Boutayeb, H. Rafaralahy, and M. Darouach. Convergence Analysis of the Extended Kalman Filter used as an Observer for Nonlinear Deterministic Discrete-Time Systems. *IEEE Transactions on Automatic Control*, 42(4), 1997.
- [13] S.J. Julier and J.K. Uhlmann. A general method for approximating nonlinear transformations of probability distributions. <http://phoebe.robots.ox.ac.uk/default.htm>, 1994.
- [14] M. Nørgaard, N.K. Poulsen, and O. Ravn. New developments in state estimation for nonlinear systems. *Automatica*, 36(11):1627–1638, November 2000.

- [15] K. Ito and K. Xiong. Gaussian filters for Nonlinear Filtering Problems. *IEEE Transactions on Automatic Control*, 45(5):910–927, May 2000.
- [16] J. Stoer. *Einführung in die Numerische Mathematik*. Springer Verlag, 1976.
- [17] G.H. Golub. Some modified matrix eigenvalue problems. *SIAM Review*, 15:318–334, 1973.
- [18] G. Sibley, G. Sukhatme, and L. Matthies. The Iterated Sigma Point Kalman Filter with Applications to Long Range Stereo. *Robotics Science and Systems*, 2006.
- [19] S.A. Banani and M.A. Masnadi-Shirazi. A New Version of the Unscented Kalman Filter. *Proceedings of World Academy of Science, Engineering and Technology*, 20, 2007.
- [20] R.S. Bucy. Digital synthesis of Nonlinear Filters. *Automatica*, 7:287–298, 1971.
- [21] A. Pole, M. West, and P.J. Harrison. *Non-normal and Nonlinear Dynamic Bayesian Modelling*. In *Bayesian Analysis of Time Series and Dynamic Models*. Ed. Marcel Dekker, New York, 1988.
- [22] B. Anderson and J. Moore. *Optimal Filtering*. Prentice Hall, 1979.
- [23] H.W. Sørensen and D.L. Alspach. Recursive Bayesian Estimation Using Gaussian Sums. *Automatica*, 7:465–479, 1971.
- [24] D.L. Alspach and H.W. Sørensen. Nonlinear Bayesian estimation using Gaussian sum approximations. *IEEE Transactions on Automatic Control*, 17(4):439–448, 1972.
- [25] H.W. Sørensen. On the development of practical nonlinear filters. *Information Sciences*, 7:253–270, 1974.
- [26] W. Koch and R. Klemm. Ground target tracking with STAP radar. *IEEE Proceedings - Radar, Sonar, Navigation*, 148(3):173–185, 2001.
- [27] A. Doucet, N. de Freitas, and N. Gordon. *Sequential Monte Carlo Methods in Practice*. Springer-Verlag, New York, 2001.
- [28] J. Liu and R. Chen. Sequential Monte Carlo Methods for Dynamic Systems. *Journal of the American Statistical Association*, 93:1032–1044, 1998.
- [29] J. Geweke. Bayesian inference in econometric models using monte carlo integration. *Econometrica*, 24:1317–1399, 1989.
- [30] Y. Bar-Shalom, X. Rong Li, and T. Kirubarajan. *Estimation with Applications to Tracking and Navigation*. Wiley, New York, 2001.
- [31] B. Efron. The Bootstrap, Jackknife and other Resampling Plans. *SIAM*, 1982.

- [32] E. Stadlober and R. Kremer. Sampling from discrete and continuous distributions with c-rand. <http://www.stat.tugraz.at/stadl/papers/sklec92.ps>.
- [33] G. Kitagawa. Monte Carlo filter and smoother for non-Gaussian nonlinear state space models. *Journal of Computational and Graphical Statistics*, 5:1–25, 1996.
- [34] M. Pitt and N. Shephard. Filtering via simulation: Auxiliary particle filters. *Journal of the American Statistical Society*, 94(446):590–599, 1999.
- [35] P. Fearnhead. *Sequential Monte Carlo Methods in Filter Theory*. Phd thesis, University of Oxford, 1998.
- [36] B. Ristic, S. Arulampalam, and N. Gordon. *Beyond the Kalman Filter*. Artech House, 2004.
- [37] W.R. Gilks and C. Berzuini. Following a moving target - Monte Carlo inference for dynamic Bayesian models. *Journal of the Royal Statistical Society*, 63:127–146, 2001.
- [38] C.P. Robert and G. Casella. *Monte Carlo Statistical Methods*. Springer Verlag, New York, 1999.
- [39] C. Chang. Kernel Particle Filter for Visual Tracking. *IEEE Transactions on Signal Processing*, 12(3):242–245, 2005.
- [40] Y. Cheng. Mean shift, mode seeking, and clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(8):790–799, 1995.
- [41] C. Beleznai, B. Frühstück, and H. Bischof. Human Tracking by Fast Mean Shift Mode Seeking. *Journal of Multimedia*, 1(1):1–8, 2006.
- [42] D. Comaniciu and P. Meer. Mean Shift: A Robust Approach Toward Feature Space Analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(5):603–619, 2002.
- [43] B. Zhang, W. Tian, and Z. Jin. Head tracking based on the integration of two different particle filters. *Measurement Science and Technology*, 17:2877–2883, 2006.
- [44] D. Fox. D. KLD-Sampling: Adaptive Particle Filters. *Advances in Neural Information Processing Systems*, 14:713–720, 2001.
- [45] D. Fox. D. kld-sampling: Adaptive particle filters and mobile robot localization. Tech. Rep. UW-CSE-01-08-02, University of Washington, 2002.
- [46] T.M. Cover and J.A. Thomas. *Elements of Information Theory*. Wiley, New York, 1991.
- [47] S. Kullback and R.A. Leibler. On Information and Sufficiency. *Annals of Mathematical Statistics*, 2(2):79–86, 1951.

- [48] A. Soto. Self Adaptive Particle Filter. *International Joint Conference on Artificial Intelligence*, pages 1398–1406, 2005.
- [49] J.H. Kotecha and P.M. Djurić. Gaussian Sum Particle Filtering. *IEEE Transactions on Signal Processing*, 51(10):2602–2612, 2003.
- [50] G. McLachlan and T. Krishnan. *The EM-Algorithm and Extensions*. Wiley, 1997.
- [51] R. van der Merwe and E.A. Wan. Gaussian Mixture Sigma-Point Particle Filters for Sequential Probabilistic Inference in Dynamic State-Space Models. *Proceedings of the International Conference on Acoustics, Speech and Signal Processing*, 2003.
- [52] S. Haykin. *Adaptive Filter Theory*, 3. ed. Prentice Hall, 1996.
- [53] A.T. Nelson. *Nonlinear Estimation and Modeling of Noisy Time-Series by Dual Kalman Filtering Methods*. Phd thesis, Oregon Graduate Institute, 2000.
- [54] H. Robbins and S. Monro. A Stochastic Approximation Method. *The Annals of Mathematical Statistics*, 22:400–407, 1951.
- [55] L. Ljung and T. Söderström. *Theory and Practice of Recursive Identification*. The MIT Press, Cambridge, Massachusetts, 1983.
- [56] W. Jakoby. *Rekursive Parameteridentifikationsalgorithmen für stochastische nichtlineare Systeme*. VDI-Verlag Düsseldorf, 1987.
- [57] M. Schreiber. *Getrennte Parameter- und Zustandsschätzung zur Echtzeit-Fehlerdiagnose an elektrohydraulischen Aktuatoren primärer Steuerflächen von Luftfahrzeugen*. Phd thesis, TU Darmstadt, 2003.
- [58] W. Wen and H. Durrant-Whyte. Model-based multi-sensor data fusion. *IEEE International Conference on Robotics and Automation*, 1992.
- [59] J. Porill. Optimal combination and constraints for geometrical sensor data. *International Journal of Robotics Research*, 7(6):66–77, 1988.
- [60] S. Hayward. Constrained Kalman filter for least-squares estimation of time-varying beamforming weights. *Mathematics in Signal Processing IV*, 7(6):66–77, 1998.
- [61] D. Simon. *Optimal State Estimation*. Wiley, 2006.
- [62] A. Bhattacharyya. On a measure of divergence between two statistical populations defined by probability distributions. *Bulletin of the Calcutta Mathematical Society*, 35:99–109, 1943.
- [63] L. Devroye. *Non-Uniform Random Variate Generation*. Springer-Verlag, 1986.

- [64] C.V. Rao, J.B. Rawlings, and D.Q. Mayne. Constrained State Estimation for Nonlinear Discrete-Time Systems: Stability and Moving Horizon Approximations. *IEEE Transactions on Automatic Control*, 48(2), 2003.
- [65] J.B. Rawlings. Model Predictive Control and Moving Horizon Estimation: Assessment of the Current Status. Danish Automation Society, Skaerbaek, 2004.
- [66] E.L. Haseltine and J.B. Rawlings. A Critical Evaluation of Extended Kalman Filtering and Moving Horizon Estimation. TWMCC Technical Report 2002-2003, 2003.
- [67] P.K. Findeisen. Moving Horizon State Estimation of Discrete Time Systems. University of Wisconsin-Madison, 1997.
- [68] F. Weller. *Numerische Mathematik für Ingenieure und Naturwissenschaftler*. Vieweg Verlag, Wiesbaden, 1996.
- [69] R. Verfürth. *Numerische Mathematik für Maschinenbauer und Umwelttechniker*. Vorlesungsskriptum Wintersemester 2006/2007, Ruhr-Universität Bochum, 2006.
- [70] W. Preuß and G. Wenisch. *Numerische Mathematik*. Fachbuchverlag Leipzig im Carl Hanser Verlag, 2001.
- [71] I. Karafyllis and C. Kravaris. On the observer problem for discrete-time control systems. Proceedings of the 44th Conference on Decision and Control, and the ECC, 2005.
- [72] H. Nijmeijer. Observability of autonomous discrete time nonlinear systems: A geometric approach. *International Journal of Control*, 36:867–873, 1982.
- [73] K. Reif, S. Günther, E. Yaz, and R. Unbehauen. Stochastic Stability of the Discrete-Time Extended Kalman Filter. *IEEE Transactions on Automatic Control*, 44:714–728, 1999.
- [74] T.E. Marlin. *Process Control, Designing Processes and Control Systems for Dynamic Performance*. McGraw-Hill, 1995.
- [75] B. Mahn. *Entwicklung einer Methode und Testumgebung für die Evaluierung von gehobenen Regelungsstrategien verfahrenstechnischer Anlagen*. Phd thesis, Ruhr-Universität Bochum, 2002.
- [76] T. Bréhard and Jean-Pierre Le Cadre. Initialization of particle filter and posterior cramer-rao bound for bearings-only tracking in modified polar coordinate system. Tech. Rep. 5074, INRIA, Rennes, 2004.
- [77] P. Tichavsky, C.H. Muravchik, and A. Nehorai. Posterior Cramér-Rao Bounds for Discrete-Time Nonlinear Filtering. *IEEE Transactions on Signal Processing*, 46(5):1386–1396, 1998.
- [78] A. Farina, B. Ristic, and L. Timmoneri. Cramér-Rao Bound for Nonlinear Filtering with $P_D < 1$ and Its Application to Target Tracking. *IEEE Transactions on Signal Processing*, 50(8):1916–1924, 2002.

- [79] S. Reece and D. Nicholson. Tighter alternatives to the cramer-rao lower bound for discrete-time filtering. <http://www.robots.ox.ac.uk/~sjrob/Pubs/fusion05b.pdf>.
- [80] L. Jung. Asymptotiv behaviour of the extended Kalman filter as a parameter estimator for linear systems. *IEEE Transactions on Automatic Control*, 24:36–50, 1979.
- [81] K. Reif and R. Unbehauen. The extended Kalman filter as an exponential observer for nonlinear systems. *submitted for publication*, 2006.
- [82] Y. Song and J.W. Grizzle. The extended Kalman filter as a local asymptotic observer for nonlinear discrete-time systems. *Proceedings of the American Control Conference*, pages 3365–3369, 1992.
- [83] K. Reif, S. Günther, E. Yaz, and R. Unbehauen. Stabilität des zeitkontinuierlichen erweiterten Kalman-Filters. *at*, 12:592–601, 1998.
- [84] R. van der Merwe and E. Wan. Matlab®Toolkit ReBEL. <http://choosh.bme.ogi.edu/rebel>.
- [85] O.J. Woodman. An introduction to inertial navigation. Technical Report No. 696, University of Cambridge, 2007.
- [86] N. Bergmann. *Recursive Bayesian Estimation: Navigation and Tracking Applications*. Phd thesis no. 579, Department of Electrical Engineering. Linköping Studies in Science and Technology, 1999.
- [87] N. Bergmann, L. Ljung, and F. Gustafsson. Terrain navigation using Bayesian statistics. *IEEE Control Systems Magazine*, 19(3):33–40, 1999.
- [88] M. Kayton and W. Fried. *Avionics Navigation Systems*. Wiley, 1997.
- [89] G.M. Siouris. *Aerospace Avionics Systems*. Academic Press, 1993.
- [90] S. Engell. *Entwurf nichtlinearer Regelungen*. Oldenbourg, 1995.
- [91] A. Isidori. *Nonlinear Control Systems*. Springer Verlag, 1995.
- [92] U. Schwenken. *Eine Methode zur Fehlerbewertung und zur adaptiven Motorleistungsbegrenzung auf der Basis einer modellbasierten Diagnose am Beispiel eines PKW-Kühlsystems*. Phd thesis, Ruhr-Universität Bochum, 2006.
- [93] H. Nguyen, N. Prasad, C. Walker, and E. Walker. *A First Course in Fuzzy and Neural Control*. Chapman&Hall, Norwell, USA, 2003.
- [94] G. Rigoll. *Neuronale Netze*. Expert Verlag, Weil der Stadt, 1994.
- [95] S. Leonhardt. *Modellgestützte Fehlererkennung mit neuronalen Netzen - Überwachung von Rad-aufhängungen mit Diesel-Einspritzanlagen*. VDI-Verlag, Düsseldorf, 1996.

- [96] M. Lattasch, B. Köppen-Seliger, P. Frank, and M. Estler. *Einsatz von neuronalen Netzen zur technischen Diagnose*. Oldenbourg Verlag, München, 1997.
- [97] A. G. Konstandopoulos, M. Kostoglou, E. Skaperdas, E. Papaioannou, D. Zarvalis, and E. Kladopoulo. Fundamental Studies of Diesel Particulate Filters: Transient Loading, Regeneration and Aging. *SAE Technical Paper 2000-01-1016*, 2000.
- [98] A. G. Konstandopoulos, E. Skaperdas, and M. Masoudi. Inertial Contributions to the Pressure Drop of Diesel Particulate Filters. *SAE Technical Paper 2001-01-0909*, 2001.
- [99] A. G. Konstandopoulos, E. Skaperdas, and M. Masoudi. Microstructural Properties of Soot Deposits in Diesel Particulate Traps. *SAE Technical Paper 2002-01-1015*, 2002.
- [100] A. G. Konstandopoulos. Flow Resistance Descriptors for Diesel Particulate Filters: Definitions, Measurements and Testing. *SAE Technical Paper 2003-01-0846*, 2003.
- [101] K. Levenberg. A Method for the Solution of Certain Problems in Least-Squares. *Quarterly Applied Mathematics* 2, pages 164–168, 1944.
- [102] D. Marquardt. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. *SIAM Journal of Applied Mathematics*, 11:431–441, 1963.
- [103] J.J. Moré. *Numerical Analysis*, chapter Lecture Notes in Mathematics 630, pages 105–116. Springer-Verlag.
- [104] S.L. Fagin. Recursive Linear Regression Theory: Optimal Filter Theory and Error Analysis of Optimal Systems. *IEEE Intern. Conv. Record*, 12, 1964.

Studentische Arbeiten

[StA 1] M.-A. Beyer. *Optimierung des Wasserpumpenmodells als Teil des Kühlerkreislauf-Simulationsmodells des M271 Verbrennungsmotors*. Interner Bericht, Lehrstuhl für Regelungssysteme und Steuerungstechnik, Ruhr-Universität Bochum, 2005.

[StA 2] M. Krämer. *Modellierung des Druckverlustes in einem Dieselpartikelfilter und Implementierung in Matlab*. Studienarbeit, Lehrstuhl für Regelungssysteme und Steuerungstechnik, Ruhr-Universität Bochum, 2007.

A Pseudo-Codes of selected inference algorithms

Introduction, Notations and Abbreviations

This supplement to the dissertation "Methodische Evaluierung von wahrscheinlichkeitstheoretischen Inferenzverfahren für dynamische Systeme" comprises the pseudo-codes of selected investigated probabilistic inference algorithms. Its purpose is to serve as reference book for PI-algorithms with a consistent notation. For more detailed information the main thesis is recommended.

In the following, basic notations and abbreviations are defined:

- The time discrete stochastic system is described by the general nonlinear, stochastic dynamic state space model:

$$\begin{aligned} x_k &= f(x_{k-1}, u_{k-1}) + K(x_{k-1}, u_{k-1})\kappa_{k-1} = \mathcal{F}(x_{k-1}, u_{k-1}, \kappa_{k-1}) \\ y_k &= h(x_k, u_k) + \Gamma(x_k, u_k)\gamma_k = \mathcal{H}(x_k, u_k, \gamma_k). \end{aligned}$$

- The differential matrix (Jacobian) \mathcal{J} of a function $g(x)$ with respect to x at $x = \mathcal{X}$ is given by the application of the NABLA-Operator to $g(x)$

$$\nabla_x g = \frac{\partial g_i}{\partial x_j} = \begin{bmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \dots & \frac{\partial g_1}{\partial x_n} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} & \dots & \frac{\partial g_2}{\partial x_n} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial g_n}{\partial x_1} & \frac{\partial g_n}{\partial x_2} & \dots & \frac{\partial g_n}{\partial x_n} \end{bmatrix},$$

i.e. $\mathcal{J}_g^x(\mathcal{X}) = \nabla_x g(x)|_{x=\mathcal{X}}$.

- The Cholesky factorization of a matrix \mathbf{P} is given by $\mathbf{S}\mathbf{S}^T = \mathbf{P}$, where \mathbf{S} is a lower triangular matrix. In the sequel the Cholesky factorization operator is denoted as $\text{CHOL}\{\cdot\}$, i.e. $\mathbf{S} = \text{CHOL}\{\mathbf{P}\}$.
- The QR-decomposition of a matrix $\mathbf{A} \in \mathbb{R}^{L \times N}$ is given by $\mathbf{A}^T = \mathbf{Q}\mathbf{R}$, where $\mathbf{Q} \in \mathbb{R}^{N \times L}$ is orthogonal, $\mathbf{R} \in \mathbb{R}^{L \times L}$ upper triangular and $N \geq L$. The upper triangular part of \mathbf{R} , $\tilde{\mathbf{R}}$ is the transpose of the Cholesky factor of $\mathbf{P} = \mathbf{A}\mathbf{A}^T$, i.e. $\tilde{\mathbf{R}}^T \tilde{\mathbf{R}} = \mathbf{A}\mathbf{A}^T = \mathbf{P}$. The operator $\text{QR}\{\cdot\}$ denotes a QR-decomposition, which gives $\tilde{\mathbf{R}}$.
- If \mathbf{S} is the original lower triangular Cholesky factor of $\mathbf{P} = \mathbf{A}\mathbf{A}^T$, then the Cholesky factor of the rank-1 update (or downdate) $\check{\mathbf{P}} = \mathbf{P} \pm \sqrt{\nu} \vartheta \vartheta^T$ is denoted in the sequel as $\check{\mathbf{S}} = \text{CHOLUPDATE}\{\mathbf{S}, \vartheta, \pm \nu\}$.

If ϑ is a matrix and not a vector, then the result is \mathfrak{M} consecutive updates of the Cholesky factor using the \mathfrak{M} columns of ϑ .

- A diagonal matrix \mathbf{A} with elements B, C, \dots

$$\mathbf{A} = \begin{bmatrix} B & & & \\ & C & & \\ & & \ddots & \\ & & & \end{bmatrix}$$

is denoted in the sequel as $\mathbf{A} = \text{DIAG}\{[B \ C \ \dots]\}$.

- The weighted covariance matrix \mathcal{CV} of a sample set $\{s^{(i)}, w^{(i)}\}_{i=1}^N$ given by

$$\mathcal{CV} = \sum_{i=1}^N w^{(i)} (s^{(i)} - \bar{s})(s^{(i)} - \bar{s})^T$$

is abbreviated as $\mathcal{CV} = \text{COV}\{s^{(i)}, w^{(i)}\}$.

- The vector outer product of column vector a is a matrix $\mathbf{A} = aa^T$ and is abbreviated in the sequel as $(\cdot)^2$, thus $\mathbf{A} = (a)^2$.

Gaussian Filters

Local Methods

Kalman Filter - KF

The Kalman Filter is only applicable to linear, gaussian systems, i.e. the stochastic, discrete state space model is given by

$$\begin{aligned}x_k &= \mathbf{A}(u_{k-1})x_{k-1} + \mathbf{B}(u_{k-1})u_{k-1} + K(u_{k-1})\kappa_{k-1} \\y_k &= \mathbf{C}(u_k)x_k + \mathbf{D}(u_k)u_k + \Gamma(u_k)\gamma_k,\end{aligned}$$

i.e. the matrices $\mathbf{A}, \mathbf{B}, K, \mathbf{C}, \mathbf{D}, \Gamma$ are all linear in x and only functions of the exogenous input u . For the sake of notational simplicity the abbreviated notations of the matrices $\mathbf{M}(u) = \mathbf{M}$ are used consecutively.

Algorithm 1 - Linear Kalman Filter - KF

$$[\{x_k, \mathbf{P}_k^x\}] = \text{KF}[\{x_{k-1}, \mathbf{P}_{k-1}^x\}, y_k]$$

Initialization:

$$\begin{aligned}\hat{x}_{0|0} &= \mathbb{E}[x_0] \\ \mathbf{P}_{0|0}^x &= \mathbb{E}[(x_0 - \mathbb{E}[x_0])(x_0 - \mathbb{E}[x_0])^T]\end{aligned}$$

FOR $k = 1, 2, \dots, K$:

$$\begin{aligned}\mathbf{P}_{k-1}^\kappa &= \mathbb{E}[(\kappa_{k-1} - \mathbb{E}[\kappa_{k-1}])(\kappa_{k-1} - \mathbb{E}[\kappa_{k-1}])^T] \\ \mathbf{P}_k^\gamma &= \mathbb{E}[(\gamma_k - \mathbb{E}[\gamma_k])(\gamma_k - \mathbb{E}[\gamma_k])^T]\end{aligned}$$

- Time Update

$$\begin{aligned}\hat{x}_{k|k-1} &= \mathbf{A}\hat{x}_{k-1|k-1} + \mathbf{B}u_{k-1} + K\mathbb{E}[\kappa_{k-1}] \\ \mathbf{P}_{k|k-1}^x &= \mathbf{A}\mathbf{P}_{k-1|k-1}^x\mathbf{A}^T + K\mathbf{P}_{k-1}^\kappa K^T\end{aligned}$$

- Measurement Update

$$\begin{aligned}\mathcal{K}_k &= \mathbf{P}_{k|k-1}^x\mathbf{A}^T(\mathbf{C}\mathbf{P}_{k|k-1}^x\mathbf{C}^T + \Gamma\mathbf{P}_k^\gamma\Gamma^T)^{-1} \\ \hat{x}_{k|k} &= \hat{x}_{k|k-1} + \mathcal{K}_k(y_k - \mathbf{C}\hat{x}_{k|k-1} - \mathbf{D}u_k - \Gamma\mathbb{E}[\gamma_k]) \\ \mathbf{P}_{k|k}^x &= (\mathbf{I} - \mathcal{K}_k\mathbf{C})\mathbf{P}_{k|k-1}^x\end{aligned}$$

END FOR

Parameters:

- -
-
-

Extended Kalman Filter - EKF

Algorithm 2 - Extended Kalman Filter - EKF

$$[\{x_k, \mathbf{P}_k^x\}] = \text{EKF}[\{x_{k-1}, \mathbf{P}_{k-1}^x\}, y_k]$$

Initialization:

$$\begin{aligned}\hat{x}_{0|0} &= \mathbb{E}[x_0] \\ \mathbf{P}_{0|0}^x &= \mathbb{E}[(x_0 - \mathbb{E}[x_0])(x_0 - \mathbb{E}[x_0])^T]\end{aligned}$$

FOR $k = 1, 2, \dots, K$:

$$\begin{aligned}\mathbf{P}_{k-1}^\kappa &= \mathbb{E}[(\kappa_{k-1} - \mathbb{E}[\kappa_{k-1}])(\kappa_{k-1} - \mathbb{E}[\kappa_{k-1}])^T] \\ \mathbf{P}_k^\gamma &= \mathbb{E}[(\gamma_k - \mathbb{E}[\gamma_k])(\gamma_k - \mathbb{E}[\gamma_k])^T]\end{aligned}$$

- Time Update

$$\begin{aligned}\mathcal{J}_{\mathcal{F}}^x &= \nabla_{x_{k-1}} \mathcal{F}(x_{k-1}, u_{k-1}, \mathbb{E}[\kappa_{k-1}])|_{x_{k-1}=\hat{x}_{k-1|k-1}} \\ \mathcal{J}_{\mathcal{F}}^\kappa &= \nabla_{\kappa_{k-1}} \mathcal{F}(\hat{x}_{k-1|k-1}, u_{k-1}, \kappa_{k-1})|_{\kappa_{k-1}=\mathbb{E}[\kappa_{k-1}]} \\ \hat{x}_{k|k-1} &= \mathcal{F}(\hat{x}_{k-1|k-1}, u_{k-1}, \mathbb{E}[\kappa_{k-1}]) \\ \mathbf{P}_{k|k-1}^x &= \mathcal{J}_{\mathcal{F}}^x \mathbf{P}_{k-1|k-1}^x (\mathcal{J}_{\mathcal{F}}^x)^T + \mathcal{J}_{\mathcal{F}}^\kappa \mathbf{P}_{k-1}^\kappa (\mathcal{J}_{\mathcal{F}}^\kappa)^T\end{aligned}$$

- Measurement Update

$$\begin{aligned}\mathcal{J}_{\mathcal{H}}^x &= \nabla_{x_k} \mathcal{H}(x_k, u_k, \mathbb{E}[\gamma_k])|_{x_k=\hat{x}_{k|k-1}} \\ \mathcal{J}_{\mathcal{H}}^\gamma &= \nabla_{\gamma_k} \mathcal{H}(\hat{x}_{k|k-1}, u_k, \gamma_k)|_{\gamma_k=\mathbb{E}[\gamma_k]} \\ \mathcal{K}_k &= \mathbf{P}_{k|k-1}^x (\mathcal{J}_{\mathcal{H}}^x)^T (\mathcal{J}_{\mathcal{H}}^x \mathbf{P}_{k|k-1}^x (\mathcal{J}_{\mathcal{H}}^x)^T + \mathcal{J}_{\mathcal{H}}^\gamma \mathbf{P}_k^\gamma (\mathcal{J}_{\mathcal{H}}^\gamma)^T)^{-1} \\ \hat{x}_{k|k} &= \hat{x}_{k|k-1} + \mathcal{K}_k (y_k - \mathcal{H}(\hat{x}_{k|k-1}, u_k, \mathbb{E}[\gamma_k])) \\ \mathbf{P}_{k|k}^x &= (\mathbf{I} - \mathcal{K}_k \mathcal{J}_{\mathcal{H}}^x) \mathbf{P}_{k|k-1}^x\end{aligned}$$

END FOR

Parameters:

- -
-
-

Square Root Extended Kalman Filter - SREKF

Algorithm 3 - Square Root Extended Kalman Filter - SREKF

$$[\{x_k, \mathbf{S}_k^x\}] = \text{SREKF}[\{x_{k-1}, \mathbf{S}_{k-1}^x\}, y_k]$$

Initialization:

$$\begin{aligned}\hat{x}_{0|0} &= \mathbb{E}[x_0] \\ \mathbf{S}_{0|0}^x &= \text{CHOL}\left\{\mathbb{E}\left[(x_0 - \mathbb{E}[x_0])(x_0 - \mathbb{E}[x_0])^T\right]\right\}\end{aligned}$$

FOR $k = 1, 2, \dots, K$:

$$\begin{aligned}\mathbf{S}_{k-1}^\kappa &= \text{CHOL}\left\{\mathbb{E}\left[(\kappa_{k-1} - \mathbb{E}[\kappa_{k-1}])(\kappa_{k-1} - \mathbb{E}[\kappa_{k-1}])^T\right]\right\} \\ \mathbf{P}_k^\gamma &= \mathbb{E}\left[(\gamma_k - \mathbb{E}[\gamma_k])(\gamma_k - \mathbb{E}[\gamma_k])^T\right]\end{aligned}$$

- Time Update

$$\begin{aligned}\mathcal{J}_{\mathcal{F}}^x &= \nabla_{x_{k-1}} \mathcal{F}(x_{k-1}, u_{k-1}, \mathbb{E}[\kappa_{k-1}])|_{x_{k-1}=\hat{x}_{k-1|k-1}} \\ \mathcal{J}_{\mathcal{F}}^\kappa &= \nabla_{\kappa_{k-1}} \mathcal{F}(\hat{x}_{k-1|k-1}, u_{k-1}, \kappa_{k-1})|_{\kappa_{k-1}=\mathbb{E}[\kappa_{k-1}]} \\ \hat{x}_{k|k-1} &= \mathcal{F}(\hat{x}_{k-1|k-1}, u_{k-1}, \mathbb{E}[\kappa_{k-1}]) \\ \mathbf{S}_{k|k-1}^x &= \text{QR}\left\{\left[\mathcal{J}_{\mathcal{F}}^x \mathbf{S}_{k-1|k-1}^x \quad \mathcal{J}_{\mathcal{F}}^\kappa \mathbf{S}_{k-1}^\kappa\right]\right\}^T\end{aligned}$$

- Measurement Update

$$\begin{aligned}\mathcal{J}_{\mathcal{H}}^x &= \nabla_{x_k} \mathcal{H}(x_k, u_k, \mathbb{E}[\gamma_k])|_{x_k=\hat{x}_{k|k-1}} \\ \mathcal{J}_{\mathcal{H}}^\gamma &= \nabla_{\gamma_k} \mathcal{H}(\hat{x}_{k|k-1}, u_k, \gamma_k)|_{\gamma_k=\mathbb{E}[\gamma_k]} \\ \mathcal{K}_k &= \mathbf{S}_{k|k-1}^x (\mathbf{S}_{k|k-1}^x)^T (\mathcal{J}_{\mathcal{H}}^x)^T (\mathcal{J}_{\mathcal{H}}^x \mathbf{S}_{k|k-1}^x (\mathbf{S}_{k|k-1}^x)^T (\mathcal{J}_{\mathcal{H}}^x)^T + \mathcal{J}_{\mathcal{H}}^\gamma \mathbf{P}_k^\gamma (\mathcal{J}_{\mathcal{H}}^\gamma)^T)^{-1} \\ \hat{x}_{k|k} &= \hat{x}_{k|k-1} + \mathcal{K}_k (y_k - \mathcal{H}(\hat{x}_{k|k-1}, u_k, \mathbb{E}[\gamma_k])) \\ \mathfrak{U} &= \text{CHOL}\{\mathcal{J}_{\mathcal{H}}^x \mathbf{S}_{k|k-1}^x (\mathbf{S}_{k|k-1}^x)^T (\mathcal{J}_{\mathcal{H}}^x)^T + \mathcal{J}_{\mathcal{H}}^\gamma \mathbf{P}_k^\gamma (\mathcal{J}_{\mathcal{H}}^\gamma)^T\} \\ \mathfrak{M} &= \text{CHOL}\{\mathcal{J}_{\mathcal{H}}^\gamma \mathbf{P}_k^\gamma (\mathcal{J}_{\mathcal{H}}^\gamma)^T\} \\ \mathbf{S}_{k|k}^x &= \mathbf{S}_{k|k-1}^x (\mathbf{I} - (\mathbf{S}_{k|k-1}^x)^T (\mathcal{J}_{\mathcal{H}}^x)^T (\mathfrak{U}^T)^{-1} (\mathfrak{U} + \mathfrak{M})^{-1} \mathcal{J}_{\mathcal{H}}^x \mathbf{S}_{k|k-1}^x)\end{aligned}$$

END FOR

Parameters:

- -

Decoupled Extended Kalman Filter - DEKF

Algorithm 4 - Decoupled Extended Kalman Filter - DEKF

$$\{x_k, \mathbf{P}_k^x, \theta_k, \mathbf{P}_k^\theta\} = \text{DEKF}[\{x_{k-1}, \mathbf{P}_{k-1}^x, \theta_{k-1}, \mathbf{P}_{k-1}^\theta\}, y_k]$$

Initialization:

$$\begin{aligned} \hat{x}_{0|0} &= \mathbb{E}[x_0] & \hat{\theta}_{0|0} &= \mathbb{E}[\theta_0] & \Lambda_0 &= 0 \\ \mathbf{P}_{0|0}^x &= \mathbb{E}[(x_0 - \mathbb{E}[x_0])(x_0 - \mathbb{E}[x_0])^T] & \mathbf{P}_{0|0}^\theta &= \mathbb{E}[(\theta_0 - \mathbb{E}[\theta_0])(\theta_0 - \mathbb{E}[\theta_0])^T] & \Pi_0 &= 0 \end{aligned}$$

FOR $k = 1, 2, \dots, K$:

$$\begin{aligned} \mathbf{P}_{k-1}^{\kappa} &= \mathbb{E}[(\kappa_{k-1} - \mathbb{E}[\kappa_{k-1}])(\kappa_{k-1} - \mathbb{E}[\kappa_{k-1}])^T] \\ \mathbf{P}_{k-1}^{\tilde{\kappa}} &= \mathbb{E}[(\tilde{\kappa}_{k-1} - \mathbb{E}[\tilde{\kappa}_{k-1}])(\tilde{\kappa}_{k-1} - \mathbb{E}[\tilde{\kappa}_{k-1}])^T] \\ \mathbf{P}_k^\gamma &= \mathbb{E}[(\gamma_k - \mathbb{E}[\gamma_k])(\gamma_k - \mathbb{E}[\gamma_k])^T] \end{aligned}$$

- Time Update Parameter Filter

$$\begin{aligned} \hat{\theta}_{k|k-1} &= \hat{\theta}_{k-1|k-1} + \mathbb{E}[\tilde{\kappa}_{k-1}] \\ \mathbf{P}_{k|k-1}^\theta &= \mathbf{P}_{k-1|k-1}^\theta + \mathbf{P}_{k-1}^{\tilde{\kappa}} \end{aligned}$$

- Time Update State Filter

$$\begin{aligned} \mathcal{J}_{\mathcal{F}}^x &= \nabla_{x_{k-1}} \mathcal{F}([x_{k-1}, \hat{\theta}_{k|k-1}], u_{k-1}, \mathbb{E}[\kappa_{k-1}])|_{x_{k-1}=\hat{x}_{k-1|k-1}} \\ \mathcal{J}_{\mathcal{F}}^{\kappa} &= \nabla_{\kappa_{k-1}} \mathcal{F}([\hat{x}_{k-1|k-1}, \hat{\theta}_{k|k-1}], u_{k-1}, \kappa_{k-1})|_{\kappa_{k-1}=\mathbb{E}[\kappa_{k-1}]} \\ \hat{x}_{k|k-1} &= \mathcal{F}([\hat{x}_{k-1|k-1}, \hat{\theta}_{k|k-1}], u_{k-1}, \mathbb{E}[\kappa_{k-1}]) \\ \mathbf{P}_{k|k-1}^x &= \mathcal{J}_{\mathcal{F}}^x \mathbf{P}_{k-1|k-1}^x (\mathcal{J}_{\mathcal{F}}^x)^T + \mathcal{J}_{\mathcal{F}}^{\kappa} \mathbf{P}_{k-1}^{\kappa} (\mathcal{J}_{\mathcal{F}}^{\kappa})^T \end{aligned}$$

- Measurement Update State Filter

$$\begin{aligned} \mathcal{J}_{\mathcal{H}}^x &= \nabla_{x_k} \mathcal{H}([x_k, \hat{\theta}_{k|k-1}], u_k, \mathbb{E}[\gamma_k])|_{x_k=\hat{x}_{k|k-1}} \\ \mathcal{J}_{\mathcal{H}}^\gamma &= \nabla_{\gamma_k} \mathcal{H}([\hat{x}_{k|k-1}, \hat{\theta}_{k|k-1}], u_k, \gamma_k)|_{\gamma_k=\mathbb{E}[\gamma_k]} \\ \mathcal{K}_k^x &= \mathbf{P}_{k|k-1}^x (\mathcal{J}_{\mathcal{H}}^x)^T (\mathcal{J}_{\mathcal{H}}^x \mathbf{P}_{k|k-1}^x (\mathcal{J}_{\mathcal{H}}^x)^T + \mathcal{J}_{\mathcal{H}}^\gamma \mathbf{P}_k^\gamma (\mathcal{J}_{\mathcal{H}}^\gamma)^T)^{-1} \\ \hat{x}_{k|k} &= \hat{x}_{k|k-1} + \mathcal{K}_k^x (y_k - \mathcal{H}([\hat{x}_{k|k-1}, \hat{\theta}_{k|k-1}], u_k, \mathbb{E}[\gamma_k])) \\ \mathbf{P}_{k|k}^x &= (\mathbf{I} - \mathcal{K}_k^x \mathcal{J}_{\mathcal{H}}^x) \mathbf{P}_{k|k-1}^x \end{aligned}$$

- Measurement Update Parameter Filter

$$\begin{aligned} \mathcal{J}_{\mathcal{F}}^\theta &= \nabla_{\theta_{k-1}} \mathcal{F}([\hat{x}_{k-1|k-1}, \theta_{k-1}], u_{k-1}, \mathbb{E}[\kappa_{k-1}])|_{\theta_{k-1}=\hat{\theta}_{k|k-1}} \\ \mathcal{J}_{\mathcal{H}}^\theta &= \nabla_{\theta_k} \mathcal{H}([\hat{x}_{k|k-1}, \theta_k], u_k, \mathbb{E}[\gamma_k])|_{\theta_k=\hat{\theta}_{k|k-1}} \\ \psi &= \Pi_{k-1} - \mathcal{K}_{k-1}^x \Lambda_{k-1} \\ \Pi_k &= \mathcal{J}_{\mathcal{F}}^\theta + \mathcal{J}_{\mathcal{F}}^x \psi \\ \Lambda_k &= \mathcal{J}_{\mathcal{H}}^\theta + \mathcal{J}_{\mathcal{H}}^x \Pi_k \\ \mathcal{K}_k^\theta &= \mathbf{P}_{k|k-1}^\theta \Lambda_k^T (\Lambda_k \mathbf{P}_{k|k-1}^\theta \Lambda_k^T + \mathcal{J}_{\mathcal{H}}^\gamma \mathbf{P}_k^\gamma (\mathcal{J}_{\mathcal{H}}^\gamma)^T)^{-1} \\ \hat{\theta}_{k|k} &= \hat{\theta}_{k|k-1} + \mathcal{K}_k^\theta (y_k - \mathcal{H}([\hat{x}_{k|k-1}, \hat{\theta}_{k|k-1}], u_k, \mathbb{E}[\gamma_k])) \\ \mathbf{P}_{k|k}^\theta &= (\mathbf{I} - \mathcal{K}_k^\theta \Lambda_k) \mathbf{P}_{k|k-1}^\theta \end{aligned}$$

END FOR

Parameters:

- -



Adaptive Extended Kalman Filter - AEKF

Algorithm 5 - Adaptive Extended Kalman Filter - AEKF

$$[\{x_k, \mathbf{P}_k^x\}] = \text{AEKF}[\{x_{k-1}, \mathbf{P}_{k-1}^x\}, y_k]$$

Initialization:

$$\begin{aligned}\hat{x}_{0|0} &= \mathbb{E}[x_0] \\ \mathbf{P}_{0|0}^x &= \mathbb{E}[(x_0 - \mathbb{E}[x_0])(x_0 - \mathbb{E}[x_0])^T]\end{aligned}$$

FOR $k = 1, 2, \dots, K$:

$$\begin{aligned}\mathbf{P}_{k-1}^\kappa &= \mathbb{E}[(\kappa_{k-1} - \mathbb{E}[\kappa_{k-1}])(\kappa_{k-1} - \mathbb{E}[\kappa_{k-1}])^T] \\ \mathbf{P}_k^\gamma &= \mathbb{E}[(\gamma_k - \mathbb{E}[\gamma_k])(\gamma_k - \mathbb{E}[\gamma_k])^T]\end{aligned}$$

- Time Update

$$\begin{aligned}\mathcal{J}_{\mathcal{F}}^x &= \nabla_{x_{k-1}} \mathcal{F}(x_{k-1}, u_{k-1}, \mathbb{E}[\kappa_{k-1}])|_{x_{k-1}=\hat{x}_{k-1|k-1}} \\ \mathcal{J}_{\mathcal{F}}^\kappa &= \nabla_{\kappa_{k-1}} \mathcal{F}(\hat{x}_{k-1|k-1}, u_{k-1}, \kappa_{k-1})|_{\kappa_{k-1}=\mathbb{E}[\kappa_{k-1}]} \\ \hat{x}_{k|k-1} &= \mathcal{F}(\hat{x}_{k-1|k-1}, u_{k-1}, \mathbb{E}[\kappa_{k-1}]) \\ \mathbf{P}_{k|k-1}^x &= \mathcal{J}_{\mathcal{F}}^x \mathbf{P}_{k-1|k-1}^x (\mathcal{J}_{\mathcal{F}}^x)^T + \mathcal{J}_{\mathcal{F}}^\kappa \mathbf{P}_{k-1}^\kappa (\mathcal{J}_{\mathcal{F}}^\kappa)^T\end{aligned}$$

- Measurement Update

$$\begin{aligned}\mathcal{J}_{\mathcal{H}}^x &= \nabla_{x_k} \mathcal{H}(x_k, u_k, \mathbb{E}[\gamma_k])|_{x_k=\hat{x}_{k|k-1}} \\ \mathcal{J}_{\mathcal{H}}^\gamma &= \nabla_{\gamma_k} \mathcal{H}(\hat{x}_{k|k-1}, u_k, \gamma_k)|_{\gamma_k=\mathbb{E}[\gamma_k]} \\ \mathcal{K}_k &= \mathbf{P}_{k|k-1}^x (\mathcal{J}_{\mathcal{H}}^x)^T (\mathcal{J}_{\mathcal{H}}^x \mathbf{P}_{k|k-1}^x (\mathcal{J}_{\mathcal{H}}^x)^T + \mathcal{J}_{\mathcal{H}}^\gamma (\varphi \mathcal{J}_{\mathcal{H}}^x \mathbf{P}_{k|k-1}^x (\mathcal{J}_{\mathcal{H}}^x)^T + \mathcal{J}_{\mathcal{H}}^\gamma \mathbf{P}_k^\gamma (\mathcal{J}_{\mathcal{H}}^\gamma)^T))^{-1} \\ \hat{x}_{k|k} &= \hat{x}_{k|k-1} + \mathcal{K}_k (y_k - \mathcal{H}(\hat{x}_{k|k-1}, u_k, \mathbb{E}[\gamma_k])) \\ \mathbf{P}_{k|k}^x &= (\mathbf{I} - \mathcal{K}_k \mathcal{J}_{\mathcal{H}}^x) \mathbf{P}_{k|k-1}^x\end{aligned}$$

END FOR

Parameters:

- φ

Parameter setup recommendation:

- $\varphi \geq 0, \varphi \approx 0.5$
-
-

Square Root Adaptive Extended Kalman Filter - SRAEKF

Algorithm 6 - Square Root Adaptive Extended Kalman Filter - SRAEKF

$$[\{x_k, \mathbf{S}_k^x\}] = \text{SRAEKF}[\{x_{k-1}, \mathbf{S}_{k-1}^x\}, y_k]$$

Initialization:

$$\begin{aligned}\hat{x}_{0|0} &= \mathbb{E}[x_0] \\ \mathbf{S}_{0|0}^x &= \text{CHOL}\left\{\mathbb{E}\left[(x_0 - \mathbb{E}[x_0])(x_0 - \mathbb{E}[x_0])^T\right]\right\}\end{aligned}$$

FOR $k = 1, 2, \dots, K$:

$$\begin{aligned}\mathbf{S}_{k-1}^\kappa &= \text{CHOL}\left\{\mathbb{E}\left[(\kappa_{k-1} - \mathbb{E}[\kappa_{k-1}])(\kappa_{k-1} - \mathbb{E}[\kappa_{k-1}])^T\right]\right\} \\ \mathbf{P}_k^\gamma &= \mathbb{E}\left[(\gamma_k - \mathbb{E}[\gamma_k])(\gamma_k - \mathbb{E}[\gamma_k])^T\right]\end{aligned}$$

- Time Update

$$\begin{aligned}\mathcal{J}_{\mathcal{F}}^x &= \nabla_{x_{k-1}} \mathcal{F}(x_{k-1}, u_{k-1}, \mathbb{E}[\kappa_{k-1}])|_{x_{k-1}=\hat{x}_{k-1|k-1}} \\ \mathcal{J}_{\mathcal{F}}^\kappa &= \nabla_{\kappa_{k-1}} \mathcal{F}(\hat{x}_{k-1|k-1}, u_{k-1}, \kappa_{k-1})|_{\kappa_{k-1}=\mathbb{E}[\kappa_{k-1}]} \\ \hat{x}_{k|k-1} &= \mathcal{F}(\hat{x}_{k-1|k-1}, u_{k-1}, \mathbb{E}[\kappa_{k-1}]) \\ \mathbf{S}_{k|k-1}^x &= \text{QR}\left\{\left[\mathcal{J}_{\mathcal{F}}^x \mathbf{S}_{k-1|k-1}^x \quad \mathcal{J}_{\mathcal{F}}^\kappa \mathbf{S}_{k-1}^\kappa\right]\right\}^T\end{aligned}$$

- Measurement Update

$$\begin{aligned}\mathcal{J}_{\mathcal{H}}^x &= \nabla_{x_k} \mathcal{H}(x_k, u_k, \mathbb{E}[\gamma_k])|_{x_k=\hat{x}_{k|k-1}} \\ \mathcal{J}_{\mathcal{H}}^\gamma &= \nabla_{\gamma_k} \mathcal{H}(\hat{x}_{k|k-1}, u_k, \gamma_k)|_{\gamma_k=\mathbb{E}[\gamma_k]} \\ \mathcal{K}_k &= \mathbf{S}_{k|k-1}^x (\mathbf{S}_{k|k-1}^x)^T (\mathcal{J}_{\mathcal{H}}^x)^T (\mathcal{J}_{\mathcal{H}}^x \mathbf{S}_{k|k-1}^x (\mathbf{S}_{k|k-1}^x)^T (\mathcal{J}_{\mathcal{H}}^x)^T + \mathcal{J}_{\mathcal{H}}^\gamma (\varphi \mathcal{J}_{\mathcal{H}}^x \mathbf{S}_{k|k-1}^x (\mathbf{S}_{k|k-1}^x)^T (\mathcal{J}_{\mathcal{H}}^x)^T + \mathcal{J}_{\mathcal{H}}^\gamma \mathbf{P}_k^\gamma (\mathcal{J}_{\mathcal{H}}^\gamma)^T)) (\mathcal{J}_{\mathcal{H}}^\gamma)^T)^{-1} \\ \hat{x}_{k|k} &= \hat{x}_{k|k-1} + \mathcal{K}_k (y_k - \mathcal{H}(\hat{x}_{k|k-1}, u_k, \mathbb{E}[\gamma_k])) \\ \mathfrak{L} &= \text{CHOL}\{\mathcal{J}_{\mathcal{H}}^x \mathbf{S}_{k|k-1}^x (\mathbf{S}_{k|k-1}^x)^T (\mathcal{J}_{\mathcal{H}}^x)^T + \mathcal{J}_{\mathcal{H}}^\gamma (\varphi \mathcal{J}_{\mathcal{H}}^x \mathbf{S}_{k|k-1}^x (\mathbf{S}_{k|k-1}^x)^T (\mathcal{J}_{\mathcal{H}}^x)^T + \mathcal{J}_{\mathcal{H}}^\gamma \mathbf{P}_k^\gamma (\mathcal{J}_{\mathcal{H}}^\gamma)^T)) (\mathcal{J}_{\mathcal{H}}^\gamma)^T\} \\ \mathfrak{M} &= \text{CHOL}\{\mathcal{J}_{\mathcal{H}}^\gamma (\varphi \mathcal{J}_{\mathcal{H}}^x \mathbf{S}_{k|k-1}^x (\mathbf{S}_{k|k-1}^x)^T (\mathcal{J}_{\mathcal{H}}^x)^T + \mathcal{J}_{\mathcal{H}}^\gamma \mathbf{P}_k^\gamma (\mathcal{J}_{\mathcal{H}}^\gamma)^T)) (\mathcal{J}_{\mathcal{H}}^\gamma)^T\} \\ \mathbf{S}_{k|k}^x &= \mathbf{S}_{k|k-1}^x (\mathbf{I} - (\mathbf{S}_{k|k-1}^x)^T (\mathcal{J}_{\mathcal{H}}^x)^T (\mathfrak{L}^T)^{-1} (\mathfrak{L} + \mathfrak{M})^{-1} \mathcal{J}_{\mathcal{H}}^x \mathbf{S}_{k|k-1}^x)\end{aligned}$$

END FOR

Parameters:

- φ

Parameter setup recommendation:

- $\varphi \geq 0, \varphi \approx 0.5$
-
-

Iterated Extended Kalman Filter - IEKF₁

Algorithm 7 - Iterated Extended Kalman Filter - IEKF₁

$$[\{x_k, \mathbf{P}_k^x\}] = \text{IEKF}_1[\{x_{k-1}, \mathbf{P}_{k-1}^x\}, y_k]$$

Initialization:

$$\begin{aligned}\hat{x}_{0|0} &= \mathbb{E}[x_0] \\ \mathbf{P}_{0|0}^x &= \mathbb{E}[(x_0 - \mathbb{E}[x_0])(x_0 - \mathbb{E}[x_0])^T]\end{aligned}$$

FOR $k = 1, 2, \dots, K$:

$$\begin{aligned}\mathbf{P}_{k-1}^\kappa &= \mathbb{E}[(\kappa_{k-1} - \mathbb{E}[\kappa_{k-1}])(\kappa_{k-1} - \mathbb{E}[\kappa_{k-1}])^T] \\ \mathbf{P}_k^\gamma &= \mathbb{E}[(\gamma_k - \mathbb{E}[\gamma_k])(\gamma_k - \mathbb{E}[\gamma_k])^T]\end{aligned}$$

- Time Update

$$\begin{aligned}\mathcal{J}_{\mathcal{F}}^x &= \nabla_{x_{k-1}} \mathcal{F}(x_{k-1}, u_{k-1}, \mathbb{E}[\kappa_{k-1}])|_{x_{k-1}=\hat{x}_{k-1|k-1}} \\ \mathcal{J}_{\mathcal{F}}^\kappa &= \nabla_{\kappa_{k-1}} \mathcal{F}(\hat{x}_{k-1|k-1}, u_{k-1}, \kappa_{k-1})|_{\kappa_{k-1}=\mathbb{E}[\kappa_{k-1}]} \\ \hat{x}_{k|k-1} &= \mathcal{F}(\hat{x}_{k-1|k-1}, u_{k-1}, \mathbb{E}[\kappa_{k-1}]) \\ \mathbf{P}_{k|k-1}^x &= \mathcal{J}_{\mathcal{F}}^x \mathbf{P}_{k-1|k-1}^x (\mathcal{J}_{\mathcal{F}}^x)^T + \mathcal{J}_{\mathcal{F}}^\kappa \mathbf{P}_{k-1}^\kappa (\mathcal{J}_{\mathcal{F}}^\kappa)^T\end{aligned}$$

- Measurement Update

$$\eta^{(1)} = \hat{x}_{k|k-1}$$

– FOR $i = 2, \dots, m$

$$\begin{aligned}\mathcal{J}_{\mathcal{H}}^x &= \nabla_{x_k} \mathcal{H}(x_k, u_k, \mathbb{E}[\gamma_k])|_{x_k=\eta^{(i-1)}} \\ \mathcal{J}_{\mathcal{H}}^\gamma &= \nabla_{\gamma_k} \mathcal{H}(\eta^{(i-1)}, u_k, \gamma_k)|_{\gamma_k=\mathbb{E}[\gamma_k]} \\ \mathcal{K}_k &= \mathbf{P}_{k|k-1}^x (\mathcal{J}_{\mathcal{H}}^x)^T (\mathcal{J}_{\mathcal{H}}^x \mathbf{P}_{k|k-1}^x (\mathcal{J}_{\mathcal{H}}^x)^T + \mathcal{J}_{\mathcal{H}}^\gamma \mathbf{P}_k^\gamma (\mathcal{J}_{\mathcal{H}}^\gamma)^T)^{-1} \\ \eta^{(i)} &= \hat{x}_{k|k-1} + \mathcal{K}_k (y_k - \mathcal{H}(\eta^{(i-1)}, u_k, \mathbb{E}[\gamma_k]) - \mathcal{J}_{\mathcal{H}}^x (\hat{x}_{k|k-1} - \eta^{(i-1)}))\end{aligned}$$

– END FOR

- Once after last iteration

$$\begin{aligned}\mathcal{J}_{\mathcal{H}}^x &= \nabla_{x_k} \mathcal{H}(x_k, u_k, \mathbb{E}[\gamma_k])|_{x_k=\eta^{(m)}} \\ \mathcal{J}_{\mathcal{H}}^\gamma &= \nabla_{\gamma_k} \mathcal{H}(\eta^{(m)}, u_k, \gamma_k)|_{\gamma_k=\mathbb{E}[\gamma_k]} \\ \mathcal{K}_k &= \mathbf{P}_{k|k-1}^x (\mathcal{J}_{\mathcal{H}}^x)^T (\mathcal{J}_{\mathcal{H}}^x \mathbf{P}_{k|k-1}^x (\mathcal{J}_{\mathcal{H}}^x)^T + \mathcal{J}_{\mathcal{H}}^\gamma \mathbf{P}_k^\gamma (\mathcal{J}_{\mathcal{H}}^\gamma)^T)^{-1} \\ \hat{x}_{k|k} &= \eta^{(m)} \\ \mathbf{P}_{k|k}^x &= (\mathbf{I} - \mathcal{K}_k \mathcal{J}_{\mathcal{H}}^x) \mathbf{P}_{k|k-1}^x\end{aligned}$$

END FOR

Parameters:

- m : Number of iterations

Parameter setup recommendation:

- $3 \leq m \leq 5$
-
-

Square Root Iterated Extended Kalman Filter - SRIEKF₁

Algorithm 8 - Square Root Iterated Extended Kalman Filter - SRIEKF₁

$$[\{x_k, \mathbf{S}_k^x\}] = \text{SRIEKF}_1[\{x_{k-1}, \mathbf{S}_{k-1}^x\}, y_k]$$

Initialization:

$$\begin{aligned}\hat{x}_{0|0} &= \mathbb{E}[x_0] \\ \mathbf{S}_{0|0}^x &= \text{CHOL}\left\{\mathbb{E}\left[(x_0 - \mathbb{E}[x_0])(x_0 - \mathbb{E}[x_0])^T\right]\right\}\end{aligned}$$

FOR $k = 1, 2, \dots, K$:

$$\begin{aligned}\mathbf{S}_{k-1}^{\kappa} &= \text{CHOL}\left\{\mathbb{E}\left[(\kappa_{k-1} - \mathbb{E}[\kappa_{k-1}])(\kappa_{k-1} - \mathbb{E}[\kappa_{k-1}])^T\right]\right\} \\ \mathbf{P}_k^\gamma &= \mathbb{E}\left[(\gamma_k - \mathbb{E}[\gamma_k])(\gamma_k - \mathbb{E}[\gamma_k])^T\right]\end{aligned}$$

- Time Update

$$\begin{aligned}\mathcal{J}_{\mathcal{F}}^x &= \nabla_{x_{k-1}} \mathcal{F}(x_{k-1}, u_{k-1}, \mathbb{E}[\kappa_{k-1}])|_{x_{k-1}=\hat{x}_{k-1|k-1}} \\ \mathcal{J}_{\mathcal{F}}^\kappa &= \nabla_{\kappa_{k-1}} \mathcal{F}(\hat{x}_{k-1|k-1}, u_{k-1}, \kappa_{k-1})|_{\kappa_{k-1}=\mathbb{E}[\kappa_{k-1}]} \\ \hat{x}_{k|k-1} &= \mathcal{F}(\hat{x}_{k-1|k-1}, u_{k-1}, \mathbb{E}[\kappa_{k-1}]) \\ \mathbf{S}_{k|k-1}^x &= \text{QR}\left\{\left[\mathcal{J}_{\mathcal{F}}^x \mathbf{S}_{k-1|k-1}^x \quad \mathcal{J}_{\mathcal{F}}^\kappa \mathbf{S}_{k-1}^\kappa\right]\right\}^T\end{aligned}$$

- Measurement Update

$$\eta^{(1)} = \hat{x}_{k|k-1}$$

– FOR $i = 2, \dots, m$

$$\begin{aligned}\mathcal{J}_{\mathcal{H}}^x &= \nabla_{x_k} \mathcal{H}(x_k, u_k, \mathbb{E}[\gamma_k])|_{x_k=\eta^{(i-1)}} \\ \mathcal{J}_{\mathcal{H}}^\gamma &= \nabla_{\gamma_k} \mathcal{H}(\eta^{(i-1)}, u_k, \gamma_k)|_{\gamma_k=\mathbb{E}[\gamma_k]} \\ \mathcal{K}_k &= \mathbf{S}_{k|k-1}^x (\mathbf{S}_{k|k-1}^x)^T (\mathcal{J}_{\mathcal{H}}^x)^T (\mathcal{J}_{\mathcal{H}}^x \mathbf{S}_{k|k-1}^x (\mathbf{S}_{k|k-1}^x)^T (\mathcal{J}_{\mathcal{H}}^x)^T + \mathcal{J}_{\mathcal{H}}^\gamma \mathbf{P}_k^\gamma (\mathcal{J}_{\mathcal{H}}^\gamma)^T)^{-1} \\ \eta^{(i)} &= \hat{x}_{k|k-1} + \mathcal{K}_k (y_k - \mathcal{H}(\eta^{(i-1)}, u_k, \mathbb{E}[\gamma_k]) - \mathcal{J}_{\mathcal{H}}^x (\hat{x}_{k|k-1} - \eta^{(i-1)}))\end{aligned}$$

– END FOR

- Once after last iteration

$$\begin{aligned}\mathcal{J}_{\mathcal{H}}^x &= \nabla_{x_k} \mathcal{H}(x_k, u_k, \mathbb{E}[\gamma_k])|_{x_k=\eta^{(m)}} \\ \mathcal{J}_{\mathcal{H}}^\gamma &= \nabla_{\gamma_k} \mathcal{H}(\eta^{(m)}, u_k, \gamma_k)|_{\gamma_k=\mathbb{E}[\gamma_k]} \\ \hat{x}_{k|k} &= \eta^{(m)} \\ \mathfrak{U} &= \text{CHOL}\{\mathcal{J}_{\mathcal{H}}^x \mathbf{S}_{k|k-1}^x (\mathbf{S}_{k|k-1}^x)^T (\mathcal{J}_{\mathcal{H}}^x)^T + \mathcal{J}_{\mathcal{H}}^\gamma \mathbf{P}_k^\gamma (\mathcal{J}_{\mathcal{H}}^\gamma)^T\} \\ \mathfrak{M} &= \text{CHOL}\{\mathcal{J}_{\mathcal{H}}^\gamma \mathbf{P}_k^\gamma (\mathcal{J}_{\mathcal{H}}^\gamma)^T\} \\ \mathbf{S}_{k|k}^x &= \mathbf{S}_{k|k-1}^x (\mathbf{I} - (\mathbf{S}_{k|k-1}^x)^T (\mathcal{J}_{\mathcal{H}}^x)^T (\mathfrak{U}^T)^{-1} (\mathfrak{U} + \mathfrak{M})^{-1} \mathcal{J}_{\mathcal{H}}^x \mathbf{S}_{k|k-1}^x)\end{aligned}$$

END FOR

Parameters:

- m : Number of iterations

Parameter setup recommendation:

- $3 \leq m \leq 5$
-
-

Iterated Extended Kalman Filter - IEKF₂

Algorithm 9 - Iterated Extended Kalman Filter - IEKF₂

$$[\{x_k, \mathbf{P}_k^x\}] = \text{IEKF}_2[\{x_{k-1}, \mathbf{P}_{k-1}^x\}, y_k]$$

Initialization:

$$\begin{aligned}\hat{x}_{0|0} &= \mathbb{E}[x_0] \\ \mathbf{P}_{0|0}^x &= \mathbb{E}[(x_0 - \mathbb{E}[x_0])(x_0 - \mathbb{E}[x_0])^T]\end{aligned}$$

FOR $k = 1, 2, \dots, K$:

$$\begin{aligned}\mathbf{P}_{k-1}^{\kappa} &= \mathbb{E}[(\kappa_{k-1} - \mathbb{E}[\kappa_{k-1}])(\kappa_{k-1} - \mathbb{E}[\kappa_{k-1}])^T] \\ \mathbf{P}_k^{\gamma} &= \mathbb{E}[(\gamma_k - \mathbb{E}[\gamma_k])(\gamma_k - \mathbb{E}[\gamma_k])^T]\end{aligned}$$

- One-time prediction

$$\begin{aligned}\xi^{(1)} &= \hat{x}_{k-1|k-1} \\ \eta^{(1)} &= \mathcal{F}(\xi^{(1)}, u_{k-1}, \mathbb{E}[\kappa_{k-1}])\end{aligned}$$

- FOR $i = 1, \dots, m$

$$\begin{aligned}\mathcal{J}_{\mathcal{F}}^x &= \nabla_{x_{k-1}} \mathcal{F}(x_{k-1}, u_{k-1}, \mathbb{E}[\kappa_{k-1}])|_{x_{k-1}=\xi^{(i)}} \\ \mathcal{J}_{\mathcal{F}}^{\kappa} &= \nabla_{\kappa_{k-1}} \mathcal{F}(\xi^{(i)}, u_{k-1}, \kappa_{k-1})|_{\kappa_{k-1}=\mathbb{E}[\kappa_{k-1}]} \\ \check{x} &= \mathcal{F}(\xi^{(i)}, u_{k-1}, \mathbb{E}[\kappa_{k-1}]) \\ \hat{x}_{k|k-1} &= \check{x} + \mathcal{J}_{\mathcal{F}}^x(\hat{x}_{k-1|k-1} - \xi^{(i)}) \\ \mathbf{P}_{k|k-1}^x &= \mathcal{J}_{\mathcal{F}}^x \mathbf{P}_{k-1|k-1}^x (\mathcal{J}_{\mathcal{F}}^x)^T + \mathcal{J}_{\mathcal{F}}^{\kappa} \mathbf{P}_{k-1}^{\kappa} (\mathcal{J}_{\mathcal{F}}^{\kappa})^T \\ \mathcal{S}_k &= \mathbf{P}_{k-1|k-1}^x (\mathcal{J}_{\mathcal{F}}^x)^T (\mathbf{P}_{k|k-1}^x)^{-1} \\ \mathcal{J}_{\mathcal{H}}^x &= \nabla_{x_k} \mathcal{H}(x_k, u_k, \mathbb{E}[\gamma_k])|_{x_k=\eta^{(i)}} \\ \mathcal{J}_{\mathcal{H}}^{\gamma} &= \nabla_{\gamma_k} \mathcal{H}(\eta^{(i)}, u_k, \gamma_k)|_{\gamma_k=\mathbb{E}[\gamma_k]} \\ \mathcal{K}_k &= \mathbf{P}_{k|k-1}^x (\mathcal{J}_{\mathcal{H}}^x)^T (\mathcal{J}_{\mathcal{H}}^x \mathbf{P}_{k|k-1}^x (\mathcal{J}_{\mathcal{H}}^x)^T + \mathcal{J}_{\mathcal{H}}^{\gamma} \mathbf{P}_k^{\gamma} (\mathcal{J}_{\mathcal{H}}^{\gamma})^T)^{-1} \\ \eta^{(i+1)} &= \hat{x}_{k|k-1} + \mathcal{K}_k (y_k - \mathcal{H}(\eta^{(i)}, u_k, \mathbb{E}[\gamma_k]) - \mathcal{J}_{\mathcal{H}}^x (\hat{x}_{k|k-1} - \eta^{(i)})) \\ \xi^{(i+1)} &= \hat{x}_{k-1|k-1} + \mathcal{S}_k (\eta^{(i+1)} - \hat{x}_{k|k-1})\end{aligned}$$

- END FOR
- Once after last iteration

$$\begin{aligned}\hat{x}_{k|k} &= \eta^{(m)} \\ \mathbf{P}_{k|k}^x &= (\mathbf{I} - \mathcal{K}_k \mathcal{J}_{\mathcal{H}}^x) \mathbf{P}_{k|k-1}^x\end{aligned}$$

END FOR

Parameters:

- m : Number of iterations

Parameter setup recommendation:

- $3 \leq m \leq 5$
-
-

Quadrature Methods

Unscented Kalman Filter - UKF

Algorithm 10 - Unscented Kalman Filter - UKF

$$[\{x_k, \mathbf{P}_k^x\}] = \text{UKF}[\{x_{k-1}, \mathbf{P}_{k-1}^x\}, y_k]$$

Weights:

$$\begin{aligned} w^{(1),m} &= \frac{\lambda}{L+\lambda} \\ w^{(1),c} &= \frac{\lambda}{L+\lambda} + (1 - \alpha^2 + \beta) \\ w^{(i),m} &= \frac{1}{2(L+\lambda)} \quad i = 2, \dots, 2L+1 \\ w^{(i),c} &= \frac{1}{2(L+\lambda)} \quad i = 2, \dots, 2L+1 \end{aligned}$$

Initialization:

$$\begin{aligned} \hat{x}_{0|0} &= \mathbb{E}[x_0] \\ \mathbf{P}_{0|0}^x &= \mathbb{E}[(x_0 - \mathbb{E}[x_0])(x_0 - \mathbb{E}[x_0])^T] \end{aligned}$$

FOR $k = 1, 2, \dots, K$:

$$\begin{aligned} \mathbf{P}_{k-1}^\kappa &= \mathbb{E}[(\kappa_{k-1} - \mathbb{E}[\kappa_{k-1}])(\kappa_{k-1} - \mathbb{E}[\kappa_{k-1}])^T] \\ \mathbf{P}_{k-1}^\gamma &= \mathbb{E}[(\gamma_k - \mathbb{E}[\gamma_k])(\gamma_k - \mathbb{E}[\gamma_k])^T] \\ \hat{x}_{k-1|k-1}^a &= [\hat{x}_{k-1|k-1} \quad \mathbb{E}[\kappa_{k-1}] \quad \mathbb{E}[\gamma_k]]^T \\ \mathbf{P}_{k-1|k-1}^a &= \text{DIAG}\{\mathbf{P}_{k-1|k-1}^x \quad \mathbf{P}_{k-1}^\kappa \quad \mathbf{P}_{k-1}^\gamma\} \end{aligned}$$

• Time-Update:

$$\begin{aligned} \mathcal{X}_{k-1}^a &= [\hat{x}_{k-1|k-1}^a \quad \hat{x}_{k-1|k-1}^a + \varphi \text{CHOL}\{\mathbf{P}_{k-1|k-1}^a\} \quad \hat{x}_{k-1|k-1}^a - \varphi \text{CHOL}\{\mathbf{P}_{k-1|k-1}^a\}] \\ \mathcal{X}_{k|k-1}^x &= \mathcal{F}(\mathcal{X}_{k-1}^x, u_{k-1}, \mathcal{X}_{k-1}^\kappa) \\ \hat{x}_{k|k-1} &= \sum_{i=1}^{2L+1} w^{(i),m} \mathcal{X}_{k|k-1}^{x,(i)} \\ \mathbf{P}_{k|k-1}^x &= \sum_{i=1}^{2L+1} w^{(i),c} (\mathcal{X}_{k|k-1}^{x,(i)} - \hat{x}_{k|k-1})(\mathcal{X}_{k|k-1}^{x,(i)} - \hat{x}_{k|k-1})^T \end{aligned}$$

• Measurement-Update:

$$\begin{aligned} \mathcal{Y}_{k|k-1} &= \mathcal{H}(\mathcal{X}_{k|k-1}^x, u_k, \mathcal{X}_{k-1}^\gamma) \\ \hat{y}_{k|k-1} &= \sum_{i=1}^{2L+1} w^{(i),m} \mathcal{Y}_{k|k-1}^{(i)} \\ \mathbf{P}_k^{\hat{y}} &= \sum_{i=1}^{2L+1} w^{(i),c} (\mathcal{Y}_{k|k-1}^{(i)} - \hat{y}_{k|k-1})(\mathcal{Y}_{k|k-1}^{(i)} - \hat{y}_{k|k-1})^T \\ \mathbf{P}_k^{xy} &= \sum_{i=1}^{2L+1} w^{(i),c} (\mathcal{X}_{k|k-1}^{x,(i)} - \hat{x}_{k|k-1})(\mathcal{Y}_{k|k-1}^{(i)} - \hat{y}_{k|k-1})^T \\ \mathcal{K}_k &= \mathbf{P}_k^{xy} (\mathbf{P}_k^{\hat{y}})^{-1} \\ \hat{x}_{k|k} &= \hat{x}_{k|k-1} + \mathcal{K}_k (y_k - \hat{y}_{k|k-1}) \\ \mathbf{P}_{k|k}^x &= \mathbf{P}_{k|k-1}^x - \mathcal{K}_k \mathbf{P}_k^{\hat{y}} \mathcal{K}_k^T \end{aligned}$$

END FOR

Compound Matrices:

- $x^a = [x^T \quad \kappa^T \quad \gamma^T]^T$
- $\mathcal{X}^a = [(\mathcal{X}^x)^T \quad (\mathcal{X}^\kappa)^T \quad (\mathcal{X}^\gamma)^T]^T$

Parameters:

- $L = n + v + z$: augmented state dimension
- κ, α and β : scaled unscented transformation scaling parameters
- $\lambda = \alpha^2(L + \kappa) - L$
- $\varphi = \sqrt{L + \lambda}$

Parameter setup recommendation:

- $\kappa \geq 0$
 - $0 \leq \alpha \leq 1$
 - $\beta \geq 0, \beta = 2$ for gaussian priors
-
-

Unscented Kalman Filter - Version "Additive Noise Case"

Algorithm 11 - Unscented Kalman Filter - Version "Additive Noise Case"

$$[\{x_k, \mathbf{P}_k^x\}] = \text{UKF}[\{x_{k-1}, \mathbf{P}_{k-1}^x\}, y_k]$$

Weights:

$$\begin{aligned} w^{(1),m} &= \frac{\lambda}{L+\lambda} \\ w^{(1),c} &= \frac{\lambda}{L+\lambda} + (1 - \alpha^2 + \beta) \\ w^{(i),m} &= \frac{1}{2(L+\lambda)} \quad i = 2, \dots, 2L+1 \\ w^{(i),c} &= \frac{1}{2(L+\lambda)} \quad i = 2, \dots, 2L+1 \end{aligned}$$

Initialization:

$$\begin{aligned} \hat{x}_{0|0} &= \mathbb{E}[x_0] \\ \mathbf{P}_{0|0}^x &= \mathbb{E}[(x_0 - \mathbb{E}[x_0])(x_0 - \mathbb{E}[x_0])^T] \end{aligned}$$

FOR $k = 1, 2, \dots, K$:

$$\begin{aligned} \mathbf{P}_{k-1}^\kappa &= \mathbb{E}[(\kappa_{k-1} - \mathbb{E}[\kappa_{k-1}])(\kappa_{k-1} - \mathbb{E}[\kappa_{k-1}])^T] \\ \mathbf{P}_k^\gamma &= \mathbb{E}[(\gamma_k - \mathbb{E}[\gamma_k])(\gamma_k - \mathbb{E}[\gamma_k])^T] \end{aligned}$$

- Time Update

$$\begin{aligned} \mathcal{X}_{k-1} &= \left[\hat{x}_{k-1|k-1} \quad \hat{x}_{k-1|k-1} + \varphi \text{CHOL}\{\mathbf{P}_{k-1|k-1}^x\} \quad \hat{x}_{k-1|k-1} - \varphi \text{CHOL}\{\mathbf{P}_{k-1|k-1}^x\} \right] \\ \mathcal{X}_{k|k-1}^* &= \mathcal{F}(\mathcal{X}_{k-1}, u_{k-1}, \mathbb{E}[\kappa_{k-1}]) \\ \hat{x}_{k|k-1} &= \sum_{i=1}^{2L+1} w^{(i),m} \mathcal{X}_{k|k-1}^{*(i)} \\ \mathbf{P}_{k|k-1}^x &= \mathbf{P}_{k-1}^\kappa + \sum_{i=1}^{2L+1} w^{(i),c} (\mathcal{X}_{k|k-1}^{*(i)} - \hat{x}_{k|k-1})(\mathcal{X}_{k|k-1}^{*(i)} - \hat{x}_{k|k-1})^T \end{aligned}$$

- Measurement Update

$$\begin{aligned} \mathcal{X}_{k|k-1} &= \left[\hat{x}_{k|k-1} \quad \hat{x}_{k|k-1} + \varphi \text{CHOL}\{\mathbf{P}_{k|k-1}^x\} \quad \hat{x}_{k|k-1} - \varphi \text{CHOL}\{\mathbf{P}_{k|k-1}^x\} \right] \\ \mathcal{Y}_{k|k-1} &= \mathcal{H}(\mathcal{X}_{k|k-1}, u_k, \mathbb{E}[\gamma_k]) \\ \hat{y}_{k|k-1} &= \sum_{i=1}^{2L+1} w^{(i),m} \mathcal{Y}_{k|k-1}^{(i)} \\ \mathbf{P}_k^{\hat{y}} &= \mathbf{P}_k^\gamma + \sum_{i=1}^{2L+1} w^{(i),c} (\mathcal{Y}_{k|k-1}^{(i)} - \hat{y}_{k|k-1})(\mathcal{Y}_{k|k-1}^{(i)} - \hat{y}_{k|k-1})^T \\ \mathbf{P}_k^{xy} &= \sum_{i=1}^{2L+1} w^{(i),c} (\mathcal{X}_{k|k-1}^{(i)} - \hat{x}_{k|k-1})(\mathcal{Y}_{k|k-1}^{(i)} - \hat{y}_{k|k-1})^T \\ \mathcal{K}_k &= \mathbf{P}_k^{xy} (\mathbf{P}_k^{\hat{y}})^{-1} \\ \hat{x}_{k|k} &= \hat{x}_{k|k-1} + \mathcal{K}_k (y_k - \hat{y}_{k|k-1}) \\ \mathbf{P}_{k|k}^x &= \mathbf{P}_{k|k-1}^x - \mathcal{K}_k \mathbf{P}_k^{\hat{y}} \mathcal{K}_k^T \end{aligned}$$

END FOR

Parameters:

- $L = n$: augmented state dimension
- κ , α and β : scaled unscented transformation scaling parameters
- $\lambda = \alpha^2(L + \kappa) - L$
- $\varphi = \sqrt{L + \lambda}$

Parameter setup recommendation:

- $\kappa \geq 0$
 - $0 \leq \alpha \leq 1$
 - $\beta \geq 0$, $\beta = 2$ for gaussian priors
-

Unscented Kalman Filter - Version "Parameter Estimation"

Algorithm 12 - Unscented Kalman Filter - Version "Parameter Estimation"

$$[\{\theta_k, \mathbf{P}_k^\theta\}] = \text{UKF}[\{\theta_{k-1}, \mathbf{P}_{k-1}^\theta\}, y_k]$$

Weights:

$$\begin{aligned} w^{(1),m} &= \frac{\lambda}{L+\lambda} \\ w^{(1),c} &= \frac{\lambda}{L+\lambda} + (1 - \alpha^2 + \beta) \\ w^{(i),m} &= \frac{1}{2(L+\lambda)} \quad i = 2, \dots, 2L+1 \\ w^{(i),c} &= \frac{1}{2(L+\lambda)} \quad i = 2, \dots, 2L+1 \end{aligned}$$

Initialization:

$$\begin{aligned} \hat{\theta}_{0|0} &= \mathbb{E}[\theta_0] \\ \mathbf{P}_{0|0}^\theta &= \mathbb{E}[(\theta_0 - \mathbb{E}[\theta_0])(\theta_0 - \mathbb{E}[\theta_0])^T] \end{aligned}$$

FOR $k = 1, 2, \dots, K$:

$$\begin{aligned} \mathbf{P}_{k-1}^{\tilde{\kappa}} &= \mathbb{E}[(\tilde{\kappa}_{k-1} - \mathbb{E}[\tilde{\kappa}_{k-1}])(\tilde{\kappa}_{k-1} - \mathbb{E}[\tilde{\kappa}_{k-1}])^T] \\ \mathbf{P}_k^\gamma &= \mathbb{E}[(\gamma_k - \mathbb{E}[\gamma_k])(\gamma_k - \mathbb{E}[\gamma_k])^T] \end{aligned}$$

- Time Update

$$\begin{aligned} \hat{\theta}_{k|k-1} &= \hat{\theta}_{k-1|k-1} + \mathbb{E}[\tilde{\kappa}_{k-1}] \\ \mathbf{P}_{k|k-1}^\theta &= \mathbf{P}_{k-1|k-1}^\theta + \mathbf{P}_{k-1}^{\tilde{\kappa}} \end{aligned}$$

- Measurement Update

$$\begin{aligned} \mathcal{X}_{k|k-1} &= \left[\hat{\theta}_{k|k-1} \quad \hat{\theta}_{k|k-1} + \varphi \text{CHOL}\{\mathbf{P}_{k|k-1}^\theta\} \quad \hat{\theta}_{k|k-1} - \varphi \text{CHOL}\{\mathbf{P}_{k|k-1}^\theta\} \right] \\ \mathcal{Y}_{k|k-1} &= \mathcal{H}(\mathcal{X}_{k|k-1}, u_k, \mathbb{E}[\gamma_k]) = \mathcal{KF}(\mathcal{X}_{k|k-1}, u_k) + \mathbb{E}[\gamma_k] \\ \text{Option 1: } \hat{y}_{k|k-1} &= \sum_{i=1}^{2L+1} w^{(i),m} \mathcal{Y}_{k|k-1}^{(i)} \\ \text{Option 2: } \hat{y}_{k|k-1} &= \mathcal{H}(\hat{\theta}_{k|k-1}, u_k, \mathbb{E}[\gamma_k]) = \mathcal{KF}(\hat{\theta}_{k|k-1}, u_k) + \mathbb{E}[\gamma_k] \\ \mathbf{P}_k^{\tilde{y}} &= \mathbf{P}_k^\gamma + \sum_{i=1}^{2L+1} w^{(i),c} (\mathcal{Y}_{k|k-1}^{(i)} - \hat{y}_{k|k-1})(\mathcal{Y}_{k|k-1}^{(i)} - \hat{y}_{k|k-1})^T \\ \mathbf{P}_k^{\theta y} &= \sum_{i=1}^{2L+1} w^{(i),c} (\mathcal{X}_{k|k-1}^{(i)} - \hat{\theta}_{k|k-1})(\mathcal{Y}_{k|k-1}^{(i)} - \hat{y}_{k|k-1})^T \\ \mathcal{K}_k &= \mathbf{P}_k^{\theta y} (\mathbf{P}_k^{\tilde{y}})^{-1} \\ \hat{\theta}_{k|k} &= \hat{\theta}_{k|k-1} + \mathcal{K}_k (y_k - \hat{y}_{k|k-1}) \\ \mathbf{P}_{k|k}^\theta &= \mathbf{P}_{k|k-1}^\theta - \mathcal{K}_k \mathbf{P}_k^{\tilde{y}} \mathcal{K}_k^T \end{aligned}$$

END FOR

Parameters:

- L : "state" dimension (number of estimated parameters)
- κ , α and β : scaled unscented transformation scaling parameters
- $\lambda = \alpha^2(L + \kappa) - L$
- $\varphi = \sqrt{L + \lambda}$

Parameter setup recommendation:

- $\kappa \geq 0$
 - $0 \leq \alpha \leq 1$
 - $\beta \geq 0$, $\beta = 2$ for gaussian priors
-

Square Root Unscented Kalman Filter - SRUKF

Algorithm 13 - Square Root Unscented Kalman Filter - SRUKF

$$[\{x_k, \mathbf{S}_k^x\}] = \text{SRUKF}[\{x_{k-1}, \mathbf{S}_{k-1}^x\}, y_k]$$

Weights:

$$\begin{aligned} w^{(1),m} &= \frac{\lambda}{L+\lambda} \\ w^{(1),c} &= \frac{\lambda}{L+\lambda} + (1 - \alpha^2 + \beta) \\ w^{(i),m} &= \frac{1}{2(L+\lambda)} \quad i = 2, \dots, 2L+1 \\ w^{(i),c} &= \frac{1}{2(L+\lambda)} \quad i = 2, \dots, 2L+1 \end{aligned}$$

Initialization:

$$\begin{aligned} \hat{x}_{0|0} &= \mathbb{E}[x_0] \\ \mathbf{S}_{0|0}^x &= \text{CHOL} \left\{ \mathbb{E} \left[(x_0 - \mathbb{E}[x_0])(x_0 - \mathbb{E}[x_0])^T \right] \right\} \end{aligned}$$

FOR $k = 1, 2, \dots, K$:

$$\begin{aligned} \mathbf{S}_{k-1}^\kappa &= \text{CHOL} \left\{ \mathbb{E} \left[(\kappa_{k-1} - \mathbb{E}[\kappa_{k-1}])(\kappa_{k-1} - \mathbb{E}[\kappa_{k-1}])^T \right] \right\} \\ \mathbf{S}_k^\gamma &= \text{CHOL} \left\{ \mathbb{E} \left[(\gamma_k - \mathbb{E}[\gamma_k])(\gamma_k - \mathbb{E}[\gamma_k])^T \right] \right\} \\ \hat{x}_{k-1|k-1}^a &= [\hat{x}_{k-1|k-1} \quad \mathbb{E}[\kappa_{k-1}] \quad \mathbb{E}[\gamma_k]]^T \\ \mathbf{S}_{k-1|k-1}^a &= \text{DIAG} \{ [\mathbf{S}_{k-1|k-1}^x \quad \mathbf{S}_{k-1}^\kappa \quad \mathbf{S}_k^\gamma] \} \end{aligned}$$

- Time Update

$$\begin{aligned} \mathcal{X}_{k-1}^a &= \left[\hat{x}_{k-1|k-1}^a \quad \hat{x}_{k-1|k-1}^a + \varphi \mathbf{S}_{k-1|k-1}^a \quad \hat{x}_{k-1|k-1}^a - \varphi \mathbf{S}_{k-1|k-1}^a \right] \\ \mathcal{X}_{k|k-1}^x &= \mathcal{F}(\mathcal{X}_{k-1}^x, u_{k-1}, \mathcal{X}_{k-1}^\kappa) \\ \hat{x}_{k|k-1} &= \sum_{i=1}^{2L+1} w^{(i),m} \mathcal{X}_{k|k-1}^{x,(i)} \\ \mathbf{S}_{k|k-1}^x &= \text{QR} \{ [\sqrt{w^{(2),c}} (\mathcal{X}_{k|k-1}^{x,(2:2L+1)} - \hat{x}_{k|k-1})] \} \\ \mathbf{S}_{k|k-1}^x &= \text{CHOLUPDATE} \{ \mathbf{S}_{k|k-1}^x, \mathcal{X}_{k|k-1}^{x,(1)} - \hat{x}_{k|k-1}, w^{(1),c} \} \end{aligned}$$

- Measurement Update

$$\begin{aligned} \mathcal{Y}_{k|k-1} &= \mathcal{H}(\mathcal{X}_{k|k-1}^x, u_k, \mathcal{X}_{k-1}^\gamma) \\ \hat{y}_{k|k-1} &= \sum_{i=1}^{2L+1} w^{(i),m} \mathcal{Y}_{k|k-1}^{(i)} \\ \mathbf{S}_k^{\tilde{y}} &= \text{QR} \{ [\sqrt{w^{(2),c}} (\mathcal{Y}_{k|k-1}^{(2:2L+1)} - \hat{y}_{k|k-1})] \} \\ \mathbf{S}_k^{\tilde{y}} &= \text{CHOLUPDATE} \{ \mathbf{S}_k^{\tilde{y}}, \mathcal{Y}_{k|k-1}^{(1)} - \hat{y}_{k|k-1}, w^{(1),c} \}^T \\ \mathbf{P}_k^{xy} &= \sum_{i=1}^{2L+1} w^{(i),c} (\mathcal{X}_{k|k-1}^{x,(i)} - \hat{x}_{k|k-1})(\mathcal{Y}_{k|k-1}^{(i)} - \hat{y}_{k|k-1})^T \\ \mathcal{K}_k &= (\mathbf{P}_k^{xy} / (\mathbf{S}_k^{\tilde{y}})^T) / \mathbf{S}_k^{\tilde{y}} \\ \hat{x}_{k|k} &= \hat{x}_{k|k-1} + \mathcal{K}_k (y_k - \hat{y}_{k|k-1}) \\ \mathfrak{U} &= \mathcal{K}_k \mathbf{S}_k^{\tilde{y}} \\ \mathbf{S}_{k|k}^x &= \text{CHOLUPDATE} \{ \mathbf{S}_{k|k-1}^x, \mathfrak{U}, -1 \}^T \end{aligned}$$

END FOR

Compound Matrices:

- $x^a = [x^T \quad \kappa^T \quad \gamma^T]^T$
- $\mathcal{X}^a = [(\mathcal{X}^x)^T \quad (\mathcal{X}^\kappa)^T \quad (\mathcal{X}^\gamma)^T]^T$

Parameters:

- $L = n + v + z$: augmented state dimension
- κ, α and β : scaled unscented transformation scaling parameters
- $\lambda = \alpha^2(L + \kappa) - L$
- $\varphi = \sqrt{L + \lambda}$

Parameter setup recommendation:

- $\kappa \geq 0$
 - $0 \leq \alpha \leq 1$
 - $\beta \geq 0, \beta = 2$ for gaussian priors
-
-

Square Root Unscented Kalman Filter - Version "Additive Noise Case"

Algorithm 14 - Square Root Unscented Kalman Filter - Version "Additive Noise Case"

$$[\{x_k, \mathbf{S}_k^x\}] = \text{SRUKF}[\{x_{k-1}, \mathbf{S}_{k-1}^x\}, y_k]$$

Weights:

$$\begin{aligned} w^{(1),m} &= \frac{\lambda}{L+\lambda} \\ w^{(1),c} &= \frac{\lambda}{L+\lambda} + (1 - \alpha^2 + \beta) \\ w^{(i),m} &= \frac{1}{2(L+\lambda)} & i = 2, \dots, 2L+1 \\ w^{(i),c} &= \frac{1}{2(L+\lambda)} & i = 2, \dots, 2L+1 \end{aligned}$$

Initialization:

$$\begin{aligned} \hat{x}_{0|0} &= \mathbb{E}[x_0] \\ \mathbf{S}_{0|0}^x &= \text{CHOL} \left\{ \mathbb{E} \left[(x_0 - \mathbb{E}[x_0])(x_0 - \mathbb{E}[x_0])^T \right] \right\} \end{aligned}$$

FOR $k = 1, 2, \dots, K$:

$$\begin{aligned} \mathbf{S}_{k-1}^\kappa &= \text{CHOL} \left\{ \mathbb{E} \left[(\kappa_{k-1} - \mathbb{E}[\kappa_{k-1}])(\kappa_{k-1} - \mathbb{E}[\kappa_{k-1}])^T \right] \right\} \\ \mathbf{S}_k^\gamma &= \text{CHOL} \left\{ \mathbb{E} \left[(\gamma_k - \mathbb{E}[\gamma_k])(\gamma_k - \mathbb{E}[\gamma_k])^T \right] \right\} \end{aligned}$$

- Time Update

$$\begin{aligned} \mathcal{X}_{k-1} &= \left[\hat{x}_{k-1|k-1} \quad \hat{x}_{k-1|k-1} + \varphi \mathbf{S}_{k-1|k-1}^x \quad \hat{x}_{k-1|k-1} - \varphi \mathbf{S}_{k-1|k-1}^x \right] \\ \mathcal{X}_{k|k-1}^* &= \mathcal{F}(\mathcal{X}_{k-1}, u_{k-1}, \mathbb{E}[\kappa_{k-1}]) \\ \hat{x}_{k|k-1} &= \sum_{i=1}^{2L+1} w^{(i),m} \mathcal{X}_{k|k-1}^{*(i)} \\ \mathbf{S}_{k|k-1}^x &= \text{QR} \left\{ \left[\sqrt{w^{(2),c}} (\mathcal{X}_{k|k-1}^{*(2:2L+1)} - \hat{x}_{k|k-1}) \quad \mathbf{S}_{k-1}^\kappa \right] \right\} \\ \mathbf{S}_{k|k-1}^x &= \text{CHOLUPDATE} \left\{ \mathbf{S}_{k|k-1}^x, \mathcal{X}_{k|k-1}^{*(1)} - \hat{x}_{k|k-1}, w^{(1),c} \right\} \end{aligned}$$

- Measurement Update

$$\begin{aligned} \mathcal{X}_{k|k-1} &= \left[\hat{x}_{k|k-1} \quad \hat{x}_{k|k-1} + \varphi (\mathbf{S}_{k|k-1}^x)^T \quad \hat{x}_{k|k-1} - \varphi (\mathbf{S}_{k|k-1}^x)^T \right] \\ \mathcal{Y}_{k|k-1} &= \mathcal{H}(\mathcal{X}_{k|k-1}, u_k, \mathbb{E}[\gamma_k]) \\ \hat{y}_{k|k-1} &= \sum_{i=1}^{2L+1} w^{(i),m} \mathcal{Y}_{k|k-1}^{(i)} \\ \mathbf{S}_{k|k-1}^{\tilde{y}} &= \text{QR} \left\{ \left[\sqrt{w^{(2),c}} (\mathcal{Y}_{k|k-1}^{(2:2L+1)} - \hat{y}_{k|k-1}) \quad \mathbf{S}_k^\gamma \right] \right\} \\ \mathbf{S}_k^{\tilde{y}} &= \text{CHOLUPDATE} \left\{ \mathbf{S}_{k|k-1}^{\tilde{y}}, \mathcal{Y}_{k|k-1}^{(1)} - \hat{y}_{k|k-1}, w^{(1),c} \right\}^T \\ \mathbf{P}_k^{xy} &= \sum_{i=1}^{2L+1} w^{(i),c} (\mathcal{X}_{k|k-1}^{(i)} - \hat{x}_{k|k-1})(\mathcal{Y}_{k|k-1}^{(i)} - \hat{y}_{k|k-1})^T \\ \mathcal{K}_k &= (\mathbf{P}_k^{xy} / (\mathbf{S}_k^{\tilde{y}})^T) / \mathbf{S}_k^{\tilde{y}} \\ \hat{x}_{k|k} &= \hat{x}_{k|k-1} + \mathcal{K}_k (y_k - \hat{y}_{k|k-1}) \\ \mathfrak{U} &= \mathcal{K}_k \mathbf{S}_k^{\tilde{y}} \\ \mathbf{S}_{k|k}^x &= \text{CHOLUPDATE} \left\{ \mathbf{S}_{k|k-1}^x, \mathfrak{U}, -1 \right\}^T \end{aligned}$$

END FOR

Parameters:

- $L = n$: augmented state dimension
- κ , α and β : scaled unscented transformation scaling parameters
- $\lambda = \alpha^2(L + \kappa) - L$
- $\varphi = \sqrt{L + \lambda}$

Parameter setup recommendation:

- $\kappa \geq 0$
 - $0 \leq \alpha \leq 1$
 - $\beta \geq 0$, $\beta = 2$ for gaussian priors
-

Square Root Unscented Kalman Filter - Version "Parameter Estimation"

Algorithm 15 - Square Root Unscented Kalman Filter - Version "Parameter Estimation"

$$[\{\theta_k, \mathbf{S}_k^\theta\}] = \text{SRUKF}[\{\theta_{k-1}, \mathbf{S}_{k-1}^\theta\}, y_k]$$

Weights:

$$\begin{aligned} w^{(1),m} &= \frac{\lambda}{L+\lambda} \\ w^{(1),c} &= \frac{\lambda}{L+\lambda} + (1 - \alpha^2 + \beta) \\ w^{(i),m} &= \frac{1}{2(L+\lambda)} \quad i = 2, \dots, 2L+1 \\ w^{(i),c} &= \frac{1}{2(L+\lambda)} \quad i = 2, \dots, 2L+1 \end{aligned}$$

Initialization:

$$\begin{aligned} \hat{\theta}_{0|0} &= \mathbb{E}[\theta_0] \\ \mathbf{S}_{0|0}^\theta &= \text{CHOL} \left\{ \mathbb{E} \left[(\theta_0 - \mathbb{E}[\theta_0]) (\theta_0 - \mathbb{E}[\theta_0])^T \right] \right\} \end{aligned}$$

FOR $k = 1, 2, \dots, K$:

$$\begin{aligned} \tilde{\mathbf{S}}_{k-1} &= \text{CHOL} \left\{ \mathbb{E} \left[(\tilde{\kappa}_{k-1} - \mathbb{E}[\tilde{\kappa}_{k-1}]) (\tilde{\kappa}_{k-1} - \mathbb{E}[\tilde{\kappa}_{k-1}])^T \right] \right\} \\ \mathbf{S}_k^\gamma &= \text{CHOL} \left\{ \mathbb{E} \left[(\gamma_k - \mathbb{E}[\gamma_k]) (\gamma_k - \mathbb{E}[\gamma_k])^T \right] \right\} \end{aligned}$$

- Time Update

$$\begin{aligned} \hat{\theta}_{k|k-1} &= \hat{\theta}_{k-1|k-1} + \mathbb{E}[\tilde{\kappa}_{k-1}] \\ \mathbf{S}_{k|k-1}^\theta &= \text{QR} \{ [\mathbf{S}_{k-1|k-1}^\theta \quad \tilde{\mathbf{S}}_{k-1}^\gamma] \} \end{aligned}$$

- Measurement Update

$$\begin{aligned} \mathcal{X}_{k|k-1} &= \left[\hat{\theta}_{k|k-1} \quad \hat{\theta}_{k|k-1} + \varphi(\mathbf{S}_{k|k-1}^\theta)^T \quad \hat{\theta}_{k|k-1} - \varphi(\mathbf{S}_{k|k-1}^\theta)^T \right] \\ \mathcal{Y}_{k|k-1} &= \mathcal{H}(\mathcal{X}_{k|k-1}, u_k, \mathbb{E}[\gamma_k]) = \mathcal{KF}(\mathcal{X}_{k|k-1}, u_k) + \mathbb{E}[\gamma_k] \\ \text{Option 1 : } \hat{y}_{k|k-1} &= \sum_{i=1}^{2L+1} w^{(i),m} \mathcal{Y}_{k|k-1}^{(i)} \\ \text{Option 2 : } \hat{y}_{k|k-1} &= \mathcal{H}(\hat{\theta}_{k|k-1}, u_k, \mathbb{E}[\gamma_k]) = \mathcal{KF}(\hat{\theta}_{k|k-1}, u_k) + \mathbb{E}[\gamma_k] \\ \mathbf{S}_k^{\tilde{y}} &= \text{QR} \{ [\sqrt{w^{(2),c}} \mathcal{Y}_{k|k-1}^{(2;2L+1)} - \hat{y}_{k|k-1} \quad \mathbf{S}_k^\gamma] \} \\ \mathbf{S}_k^{\tilde{y}} &= \text{CHOLUPDATE} \{ \mathbf{S}_k^{\tilde{y}}, \mathcal{Y}_{k|k-1}^{(1)} - \hat{y}_{k|k-1}, w^{(1),c} \}^T \\ \mathbf{P}_k^{\theta y} &= \sum_{i=1}^{2L+1} w^{(i),c} (\mathcal{X}_{k|k-1}^{(i)} - \hat{\theta}_{k|k-1}) (\mathcal{Y}_{k|k-1}^{(i)} - \hat{y}_{k|k-1})^T \\ \mathcal{K}_k &= (\mathbf{P}_k^{\theta y} / (\mathbf{S}_k^{\tilde{y}})^T) / \mathbf{S}_k^{\tilde{y}} \\ \hat{\theta}_{k|k} &= \hat{\theta}_{k|k-1} + \mathcal{K}_k (y_k - \hat{y}_{k|k-1}) \\ \mathfrak{U} &= \mathcal{K}_k \mathbf{S}_k^{\tilde{y}} \\ \mathbf{S}_{k|k}^\theta &= \text{CHOLUPDATE} \{ \mathbf{S}_{k|k-1}^\theta, \mathfrak{U}, -1 \}^T \end{aligned}$$

END FOR

Parameters:

- L : augmented state dimension (number of parameters)
- κ , α and β : scaled unscented transformation scaling parameters
- $\lambda = \alpha^2(L + \kappa) - L$
- $\varphi = \sqrt{L + \lambda}$

Parameter setup recommendation:

- $\kappa \geq 0$
 - $0 \leq \alpha \leq 1$
 - $\beta \geq 0$, $\beta = 2$ for gaussian priors
-

Central Difference Kalman Filter - CDKF

Algorithm 16 - Central Difference Kalman Filter - CDKF

$$[x_k, \mathbf{P}_k^x] = \text{CDKF}[\{x_{k-1}, \mathbf{P}_{k-1}^x\}, y_k]$$

Weights:

$$\begin{aligned} w^{(1),m} &= \frac{h^2 - L}{h^2} \\ w^{(i),m} &= \frac{1}{2h^2} & i = 2, \dots, 2L + 1 \\ w^{(i),c1} &= \frac{1}{4h^2} & i = 2, \dots, 2L + 1 \\ w^{(i),c2} &= \frac{h^2 - 1}{4h^4} & i = 2, \dots, 2L + 1 \end{aligned}$$

Initialization:

$$\begin{aligned} \hat{x}_{0|0} &= \mathbb{E}[x_0] \\ \mathbf{P}_{0|0}^x &= \mathbb{E}[(x_0 - \mathbb{E}[x_0])(x_0 - \mathbb{E}[x_0])^T] \end{aligned}$$

FOR $k = 1, 2, \dots, K$:

$$\begin{aligned} \mathbf{P}_{k-1}^\kappa &= \mathbb{E}[(\kappa_{k-1} - \mathbb{E}[\kappa_{k-1}])(\kappa_{k-1} - \mathbb{E}[\kappa_{k-1}])^T] \\ \mathbf{P}_k^\gamma &= \mathbb{E}[(\gamma_k - \mathbb{E}[\gamma_k])(\gamma_k - \mathbb{E}[\gamma_k])^T] \end{aligned}$$

- Time Update

$$\begin{aligned} \hat{x}_{k-1|k-1}^{\alpha\kappa} &= [\hat{x}_{k-1|k-1} \quad \mathbb{E}[\kappa_{k-1}]] \\ \mathbf{P}_{k-1}^{\alpha\kappa} &= \text{DIAG}\{\{\mathbf{P}_{k-1}^x \quad \mathbf{P}_{k-1}^\kappa\}\} \\ \mathcal{X}_{k-1}^{\alpha\kappa} &= [\hat{x}_{k-1|k-1}^{\alpha\kappa} \quad \hat{x}_{k-1|k-1}^{\alpha\kappa} + h\text{CHOL}\{\mathbf{P}_{k-1}^{\alpha\kappa}\} \quad \hat{x}_{k-1|k-1}^{\alpha\kappa} - h\text{CHOL}\{\mathbf{P}_{k-1}^{\alpha\kappa}\}] \\ \mathcal{X}_{k|k-1}^x &= \mathcal{F}(\mathcal{X}_{k-1}^x, u_{k-1}, \mathcal{X}_{k-1}^{\alpha\kappa}) \\ \hat{x}_{k|k-1} &= \sum_{i=1}^{2L+1} w^{(i),m} \mathcal{X}_{k|k-1}^{x,(i)} \\ \mathbf{P}_{k|k-1}^x &= \sum_{i=2}^{L+1} [w^{(i),c1} (\mathcal{X}_{k|k-1}^{x,(i)} - \mathcal{X}_{k|k-1}^{x,(L+i)})^2 + w^{(i),c2} (\mathcal{X}_{k|k-1}^{x,(i)} + \mathcal{X}_{k|k-1}^{x,(L+i)} - 2\mathcal{X}_{k|k-1}^{x,(1)})^2] \end{aligned}$$

- Measurement Update

$$\begin{aligned} \hat{x}_{k|k-1}^{\alpha\gamma} &= [\hat{x}_{k-1|k-1} \quad \mathbb{E}[\gamma_k]] \\ \mathbf{P}_{k|k-1}^{\alpha\gamma} &= \text{DIAG}\{\{\mathbf{P}_{k|k-1}^x \quad \mathbf{P}_k^\gamma\}\} \\ \mathcal{X}_{k|k-1}^{\alpha\gamma} &= [\hat{x}_{k|k-1}^{\alpha\gamma} \quad \hat{x}_{k|k-1}^{\alpha\gamma} + h\text{CHOL}\{\mathbf{P}_{k|k-1}^{\alpha\gamma}\} \quad \hat{x}_{k|k-1}^{\alpha\gamma} - h\text{CHOL}\{\mathbf{P}_{k|k-1}^{\alpha\gamma}\}] \\ \mathcal{Y}_{k|k-1} &= \mathcal{H}(\mathcal{X}_{k|k-1}^x, u_k, \mathcal{X}_{k|k-1}^{\alpha\gamma}) \\ \hat{y}_{k|k-1} &= \sum_{i=1}^{2L+1} w^{(i),m} \mathcal{Y}_{k|k-1}^{(i)} \\ \mathbf{P}_k^{\tilde{y}} &= \sum_{i=2}^{L+1} [w^{(i),c1} (\mathcal{Y}_{k|k-1}^{(i)} - \mathcal{Y}_{k|k-1}^{(L+i)})^2 + w^{(i),c2} (\mathcal{Y}_{k|k-1}^{(i)} + \mathcal{Y}_{k|k-1}^{(L+i)} - 2\mathcal{Y}_{k|k-1}^{(1)})^2] \\ \mathbf{P}_k^{xy} &= \text{CHOL}\{w^{(2),c1} \mathbf{P}_{k|k-1}^x\} [\mathcal{Y}_{k|k-1}^{(2:L+1)} - \mathcal{Y}_{k|k-1}^{(L+2:2L+1)}]^T \\ \mathcal{K}_k &= \mathbf{P}_k^{xy} (\mathbf{P}_k^{\tilde{y}})^{-1} \\ \hat{x}_{k|k} &= \hat{x}_{k|k-1} + \mathcal{K}_k (y_k - \hat{y}_{k|k-1}) \\ \mathbf{P}_{k|k}^x &= \mathbf{P}_{k|k-1}^x - \mathcal{K}_k \mathbf{P}_k^{\tilde{y}} \mathcal{K}_k^T \end{aligned}$$

END FOR

Compound Matrices:

- $x^{a\kappa} = [x^T \quad \kappa^T]^T$
- $x^{a\gamma} = [x^T \quad \gamma^T]^T$
- $\mathcal{X}^{a\kappa} = [(\mathcal{X}^x)^T \quad (\mathcal{X}^\kappa)^T]^T$
- $\mathcal{X}^{a\gamma} = [(\mathcal{X}^x)^T \quad (\mathcal{X}^\gamma)^T]^T$

Parameters:

- $L = n + v$: augmented state dimension time update
- $L = n + z$: augmented state dimension measurement update
- h : central difference interval parameter

Parameter setup recommendation:

- $h \geq 1$, $h = \sqrt{3}$ for gaussian priors
-
-

Central Difference Kalman Filter - Version "Additive Noise Case"

Algorithm 17 - Central Difference Kalman Filter - Version "Additive Noise Case"

$$[\{x_k, \mathbf{P}_k^x\}] = \text{CDKF}[\{x_{k-1}, \mathbf{P}_{k-1}^x\}, y_k]$$

Weights:

$$\begin{aligned} w^{(1),m} &= \frac{h^2 - L}{h^2} \\ w^{(i),m} &= \frac{1}{2h^2} & i = 2, \dots, 2L + 1 \\ w^{(i),c1} &= \frac{1}{4h^2} & i = 2, \dots, 2L + 1 \\ w^{(i),c2} &= \frac{h^2 - 1}{4h^4} & i = 2, \dots, 2L + 1 \end{aligned}$$

Initialization:

$$\begin{aligned} \hat{x}_{0|0} &= \mathbb{E}[x_0] \\ \mathbf{P}_{0|0}^x &= \mathbb{E}[(x_0 - \mathbb{E}[x_0])(x_0 - \mathbb{E}[x_0])^T] \end{aligned}$$

FOR $k = 1, 2, \dots, K$:

$$\begin{aligned} \mathbf{P}_{k-1}^{\kappa} &= \mathbb{E}[(\kappa_{k-1} - \mathbb{E}[\kappa_{k-1}])(\kappa_{k-1} - \mathbb{E}[\kappa_{k-1}])^T] \\ \mathbf{P}_k^{\gamma} &= \mathbb{E}[(\gamma_k - \mathbb{E}[\gamma_k])(\gamma_k - \mathbb{E}[\gamma_k])^T] \end{aligned}$$

- Time Update

$$\begin{aligned} \mathcal{X}_{k-1} &= [\hat{x}_{k-1|k-1} \quad \hat{x}_{k-1|k-1} + h \text{CHOL}\{\mathbf{P}_{k-1|k-1}^x\} \quad \hat{x}_{k-1|k-1} - h \text{CHOL}\{\mathbf{P}_{k-1|k-1}^x\}] \\ \mathcal{X}_{k|k-1} &= \mathcal{F}(\mathcal{X}_{k-1}, u_{k-1}, \mathbb{E}[\kappa_{k-1}]) \\ \hat{x}_{k|k-1} &= \sum_{i=1}^{2L+1} w^{(i),m} \mathcal{X}_{k|k-1}^{(i)} \\ \mathbf{P}_{k|k-1}^x &= \mathbf{P}_{k-1}^{\kappa} + \sum_{i=2}^{L+1} [w^{(i),c1} (\mathcal{X}_{k|k-1}^{(i)} - \mathcal{X}_{k|k-1}^{(L+i)})^2 + w^{(i),c2} (\mathcal{X}_{k|k-1}^{(i)} + \mathcal{X}_{k|k-1}^{(L+i)} - 2\mathcal{X}_{k|k-1}^{(1)})^2] \end{aligned}$$

- Measurement Update

$$\begin{aligned} \mathcal{X}_{k|k-1} &= [\hat{x}_{k|k-1} \quad \hat{x}_{k|k-1} + h \text{CHOL}\{\mathbf{P}_{k|k-1}^x\} \quad \hat{x}_{k|k-1} - h \text{CHOL}\{\mathbf{P}_{k|k-1}^x\}] \\ \mathcal{Y}_{k|k-1} &= \mathcal{H}(\mathcal{X}_{k|k-1}, u_k, \mathbb{E}[\gamma_k]) \\ \hat{y}_{k|k-1} &= \sum_{i=1}^{2L+1} w^{(i),m} \mathcal{Y}_{k|k-1}^{(i)} \\ \mathbf{P}_k^{\tilde{y}} &= \mathbf{P}_k^{\gamma} + \sum_{i=2}^{L+1} [w^{(i),c1} (\mathcal{Y}_{k|k-1}^{(i)} - \mathcal{Y}_{k|k-1}^{(L+i)})^2 + w^{(i),c2} (\mathcal{Y}_{k|k-1}^{(i)} + \mathcal{Y}_{k|k-1}^{(L+i)} - 2\mathcal{Y}_{k|k-1}^{(1)})^2] \\ \mathbf{P}_k^{xy} &= \text{CHOL}\{w^{(2),c1} \mathbf{P}_{k|k-1}^x\} [\mathcal{Y}_{k|k-1}^{(2:L+1)} - \mathcal{Y}_{k|k-1}^{(L+2:2L+1)}]^T \\ \mathcal{K}_k &= \mathbf{P}_k^{xy} (\mathbf{P}_k^{\tilde{y}})^{-1} \\ \hat{x}_{k|k} &= \hat{x}_{k|k-1} + \mathcal{K}_k (y_k - \hat{y}_{k|k-1}) \\ \mathbf{P}_{k|k}^x &= \mathbf{P}_{k|k-1}^x - \mathcal{K}_k \mathbf{P}_k^{\tilde{y}} \mathcal{K}_k^T \end{aligned}$$

END FOR

Parameters:

- $L = n$: augmented state dimension
- h : central difference interval parameter

Parameter setup recommendation:

- $h \geq 1$, $h = \sqrt{3}$ for gaussian priors
-
-

Central Difference Kalman Filter - Version "Parameter Estimation"

Algorithm 18 - Central Difference Kalman Filter - Version "Parameter Estimation"

$$[\{\theta_k, \mathbf{P}_k^\theta\}] = \text{CDKF}[\{\theta_{k-1}, \mathbf{P}_{k-1}^\theta\}, y_k]$$

Weights:

$$\begin{aligned} w^{(1),m} &= \frac{h^2 - L}{h^2} \\ w^{(i),m} &= \frac{1}{2h^2} & i = 2, \dots, 2L + 1 \\ w^{(i),c1} &= \frac{1}{4h^2} & i = 2, \dots, 2L + 1 \\ w^{(i),c2} &= \frac{h^2 - 1}{4h^4} & i = 2, \dots, 2L + 1 \end{aligned}$$

Initialization:

$$\begin{aligned} \hat{\theta}_{0|0} &= \mathbb{E}[\theta_0] \\ \mathbf{P}_{0|0}^\theta &= \mathbb{E}[(\theta_0 - \mathbb{E}[\theta_0])(\theta_0 - \mathbb{E}[\theta_0])^T] \end{aligned}$$

FOR $k = 1, 2, \dots, K$:

$$\begin{aligned} \mathbf{P}_{k-1}^{\tilde{\kappa}} &= \mathbb{E}[(\tilde{\kappa}_{k-1} - \mathbb{E}[\tilde{\kappa}_{k-1}])(\tilde{\kappa}_{k-1} - \mathbb{E}[\tilde{\kappa}_{k-1}])^T] \\ \mathbf{P}_k^\gamma &= \mathbb{E}[(\gamma_k - \mathbb{E}[\gamma_k])(\gamma_k - \mathbb{E}[\gamma_k])^T] \end{aligned}$$

- Time Update

$$\begin{aligned} \hat{\theta}_{k|k-1} &= \hat{\theta}_{k-1|k-1} + \mathbb{E}[\tilde{\kappa}_{k-1}] \\ \mathbf{P}_{k|k-1}^\theta &= \mathbf{P}_{k-1|k-1}^\theta + \mathbf{P}_{k-1}^{\tilde{\kappa}} \end{aligned}$$

- Measurement Update

$$\begin{aligned} \mathcal{X}_{k|k-1} &= \left[\hat{\theta}_{k|k-1} \quad \hat{\theta}_{k|k-1} + h \text{CHOL}\{\mathbf{P}_{k|k-1}^\theta\} \quad \hat{\theta}_{k|k-1} - h \text{CHOL}\{\mathbf{P}_{k|k-1}^\theta\} \right] \\ \mathcal{Y}_{k|k-1} &= \mathcal{H}(\mathcal{X}_{k|k-1}, u_k, \mathbb{E}[\gamma_k]) = \mathcal{KF}(\mathcal{X}_{k|k-1}, u_k) + \mathbb{E}[\gamma_k] \\ \text{Option 1: } \hat{y}_{k|k-1} &= \sum_{i=1}^{2L+1} w^{(i),m} \mathcal{Y}_{k|k-1}^{(i)} \\ \text{Option 2: } \hat{y}_{k|k-1} &= \mathcal{H}(\hat{\theta}_{k|k-1}, u_k, \mathbb{E}[\gamma_k]) = \mathcal{KF}(\hat{\theta}_{k|k-1}, u_k) + \mathbb{E}[\gamma_k] \\ \mathbf{P}_k^{\tilde{y}} &= \mathbf{P}_k^\gamma + \sum_{i=2}^{L+1} \left[w^{(i),c1} (\mathcal{Y}_{k|k-1}^{(i)} - \mathcal{Y}_{k|k-1}^{(L+i)})^2 \right. \\ &\quad \left. + w^{(i),c2} (\mathcal{Y}_{k|k-1}^{(i)} + \mathcal{Y}_{k|k-1}^{(L+i)} - 2\mathcal{Y}_{k|k-1}^{(1)})^2 \right] \\ \mathbf{P}_k^{\theta y} &= \text{CHOL}\{w^{(2),c1} \mathbf{P}_{k|k-1}^\theta\} \left[\mathcal{Y}_{k|k-1}^{(2:L+1)} - \mathcal{Y}_{k|k-1}^{(L+2:2L+1)} \right]^T \\ \mathcal{K}_k &= \mathbf{P}_k^{\theta y} (\mathbf{P}_k^{\tilde{y}})^{-1} \\ \hat{\theta}_{k|k} &= \hat{\theta}_{k|k-1} + \mathcal{K}_k (y_k - \hat{y}_{k|k-1}) \\ \mathbf{P}_{k|k}^\theta &= \mathbf{P}_{k|k-1}^\theta - \mathcal{K}_k \mathbf{P}_k^{\tilde{y}} \mathcal{K}_k^T \end{aligned}$$

END FOR

Parameters:

- L : "state" dimension (number of estimated parameters)
- h : central difference interval parameter

Parameter setup recommendation:

- $h \geq 1$, $h = \sqrt{3}$ for gaussian priors
-
-

Square Root Central Difference Kalman Filter - SRCDKF

Algorithm 19 - Square Root Central Difference Kalman Filter - SRCDKF

$$\{x_k, \mathbf{S}_k^x\} = \text{SRCDKF}\{x_{k-1}, \mathbf{S}_{k-1}^x, y_k\}$$

Weights:

$$\begin{aligned} w^{(1),m} &= \frac{h^2 - L}{h^2} \\ w^{(i),m} &= \frac{1}{2h^2} & i = 2, \dots, 2L + 1 \\ w^{(i),c1} &= \frac{1}{4h^2} & i = 2, \dots, 2L + 1 \\ w^{(i),c2} &= \frac{h^2 - 1}{4h^4} & i = 2, \dots, 2L + 1 \end{aligned}$$

Initialization:

$$\begin{aligned} \hat{x}_{0|0} &= \mathbb{E}[x_0] \\ \mathbf{S}_{0|0}^x &= \text{CHOL}\left\{\mathbb{E}\left[(x_0 - \mathbb{E}[x_0])(x_0 - \mathbb{E}[x_0])^T\right]\right\} \end{aligned}$$

FOR $k = 1, 2, \dots, K$:

$$\begin{aligned} \mathbf{S}_{k-1}^\kappa &= \text{CHOL}\left\{\mathbb{E}\left[(\kappa_{k-1} - \mathbb{E}[\kappa_{k-1}])(\kappa_{k-1} - \mathbb{E}[\kappa_{k-1}])^T\right]\right\} \\ \mathbf{S}_k^\gamma &= \text{CHOL}\left\{\mathbb{E}\left[(\gamma_k - \mathbb{E}[\gamma_k])(\gamma_k - \mathbb{E}[\gamma_k])^T\right]\right\} \end{aligned}$$

- Time Update

$$\begin{aligned} \hat{x}_{k-1|k-1}^{\alpha\kappa} &= [\hat{x}_{k-1|k-1} \quad \mathbb{E}[\kappa_{k-1}]] \\ \mathbf{S}_{k-1|k-1}^{\alpha\kappa} &= \text{DIAG}\{[\mathbf{S}_{k-1|k-1}^x \quad \mathbf{S}_{k-1}^\kappa]\} \\ \mathcal{X}_{k-1}^{\alpha\kappa} &= [\hat{x}_{k-1|k-1}^{\alpha\kappa} \quad \hat{x}_{k-1|k-1}^{\alpha\kappa} + h\mathbf{S}_{k-1|k-1}^{\alpha\kappa} \quad \hat{x}_{k-1|k-1}^{\alpha\kappa} - h\mathbf{S}_{k-1|k-1}^{\alpha\kappa}] \\ \mathcal{X}_{k|k-1}^x &= \mathcal{F}(\mathcal{X}_{k-1}^x, u_{k-1}, \mathcal{X}_{k-1}^\kappa) \\ \hat{x}_{k|k-1} &= \sum_{i=1}^{2L+1} w^{(i),m} \mathcal{X}_{k|k-1}^{x,(i)} \\ \mathbf{S}_{k|k-1}^x &= \text{QR}\left\{\left[\sqrt{w^{(2),c1}}(\mathcal{X}_{k|k-1}^{x,(2:L+1)} - \mathcal{X}_{k|k-1}^{x,(L+2:2L+1)}) \quad \sqrt{w^{(2),c2}}(\mathcal{X}_{k|k-1}^{x,(2:L+1)} + \mathcal{X}_{k|k-1}^{x,(L+2:2L+1)} - 2\mathcal{X}_{k|k-1}^{x,(1)})\right]\right\} \end{aligned}$$

- Measurement Update

$$\begin{aligned} \hat{x}_{k|k-1}^{\alpha\gamma} &= [\hat{x}_{k-1|k-1} \quad \mathbb{E}[\gamma_k]] \\ \mathbf{S}_{k|k-1}^{\alpha\gamma} &= \text{DIAG}\{[(\mathbf{S}_{k|k-1}^x)^T \quad \mathbf{S}_k^\gamma]\} \\ \mathcal{X}_{k|k-1}^{\alpha\gamma} &= [\hat{x}_{k|k-1}^{\alpha\gamma} \quad \hat{x}_{k|k-1}^{\alpha\gamma} + h\mathbf{S}_{k|k-1}^{\alpha\gamma} \quad \hat{x}_{k|k-1}^{\alpha\gamma} - h\mathbf{S}_{k|k-1}^{\alpha\gamma}] \\ \mathcal{Y}_{k|k-1} &= \mathcal{H}(\mathcal{X}_{k|k-1}^x, u_k, \mathcal{X}_{k|k-1}^\gamma) \\ \hat{y}_{k|k-1} &= \sum_{i=1}^{2L+1} w^{(i),m} \mathcal{Y}_{k|k-1}^{(i)} \\ \mathbf{S}_k^{\tilde{y}} &= \text{QR}\left\{\left[\sqrt{w^{(2),c1}}(\mathcal{Y}_{k|k-1}^{(2:L+1)} - \mathcal{Y}_{k|k-1}^{(L+2:2L+1)}) \quad \sqrt{w^{(2),c2}}(\mathcal{Y}_{k|k-1}^{(2:L+1)} - \mathcal{Y}_{k|k-1}^{(L+2:2L+1)} - 2\mathcal{Y}_{k|k-1}^{(1)})\right]\right\}^T \\ \mathbf{P}_k^{xy} &= \sqrt{w^{(2),c1}}(\mathbf{S}_{k|k-1}^x)^T [\mathcal{Y}_{k|k-1}^{(2:L+1)} - \mathcal{Y}_{k|k-1}^{(L+2:2L+1)}]^T \\ \mathcal{K}_k &= (\mathbf{P}_k^{xy} / (\mathbf{S}_k^{\tilde{y}})^T) / \mathbf{S}_k^{\tilde{y}} \\ \hat{x}_{k|k} &= \hat{x}_{k|k-1} + \mathcal{K}_k(y_k - \hat{y}_{k|k-1}) \\ \mathfrak{U} &= \mathcal{K}_k \mathbf{S}_k^{\tilde{y}} \\ \mathbf{S}_{k|k}^x &= \text{CHOLUPDATE}\{\mathbf{S}_{k|k-1}^x, \mathfrak{U}, -1\}^T \end{aligned}$$

END FOR

Compound Matrices:

- $x^{a\kappa} = [x^T \quad \kappa^T]^T$
- $x^{a\gamma} = [x^T \quad \gamma^T]^T$
- $\mathcal{X}^{a\kappa} = [(\mathcal{X}^x)^T \quad (\mathcal{X}^\kappa)^T]^T$
- $\mathcal{X}^{a\gamma} = [(\mathcal{X}^x)^T \quad (\mathcal{X}^\gamma)^T]^T$

Parameters:

- $L = n + v$: augmented state dimension time update
- $L = n + z$: augmented state dimension measurement update
- h : central difference interval parameter

Parameter setup recommendation:

- $h \geq 1$, $h = \sqrt{3}$ for gaussian priors
-
-

Square Root Central Difference Kalman Filter - Version "Additive Noise Case"

Algorithm 20 - Square Root Central Difference Kalman Filter - Version "Additive Noise Case"

$$\{x_k, \mathbf{S}_k^x\} = \text{SRCDKF}[\{x_{k-1}, \mathbf{S}_{k-1}^x\}, y_k]$$

Weights:

$$\begin{aligned} w^{(1),m} &= \frac{h^2 - L}{h^2} \\ w^{(i),m} &= \frac{1}{2h^2} & i = 2, \dots, 2L + 1 \\ w^{(i),c1} &= \frac{1}{4h^2} & i = 2, \dots, 2L + 1 \\ w^{(i),c2} &= \frac{h^2 - 1}{4h^4} & i = 2, \dots, 2L + 1 \end{aligned}$$

Initialization:

$$\begin{aligned} \hat{x}_{0|0} &= \mathbb{E}[x_0] \\ \mathbf{S}_{0|0}^x &= \text{CHOL} \left\{ \mathbb{E} \left[(x_0 - \mathbb{E}[x_0])(x_0 - \mathbb{E}[x_0])^T \right] \right\} \end{aligned}$$

FOR $k = 1, 2, \dots, K$:

$$\begin{aligned} \mathbf{S}_{k-1}^{\kappa} &= \text{CHOL} \left\{ \mathbb{E} \left[(\kappa_{k-1} - \mathbb{E}[\kappa_{k-1}])(\kappa_{k-1} - \mathbb{E}[\kappa_{k-1}])^T \right] \right\} \\ \mathbf{S}_k^\gamma &= \text{CHOL} \left\{ \mathbb{E} \left[(\gamma_k - \mathbb{E}[\gamma_k])(\gamma_k - \mathbb{E}[\gamma_k])^T \right] \right\} \end{aligned}$$

- Time Update

$$\begin{aligned} \mathcal{X}_{k-1} &= \begin{bmatrix} \hat{x}_{k-1|k-1} & \hat{x}_{k-1|k-1} + h\mathbf{S}_{k-1|k-1}^x & \hat{x}_{k-1|k-1} - h\mathbf{S}_{k-1|k-1}^x \end{bmatrix} \\ \mathcal{X}_{k|k-1} &= \mathcal{F}(\mathcal{X}_{k-1}, u_{k-1}, \mathbb{E}[\kappa_{k-1}]) \\ \hat{x}_{k|k-1} &= \sum_{i=1}^{2L+1} w^{(i),m} \mathcal{X}_{k|k-1}^{(i)} \\ \mathbf{S}_{k|k-1}^x &= \text{QR} \left\{ \left[\sqrt{w^{(2),c1}} (\mathcal{X}_{k|k-1}^{x,(2:L+1)} - \mathcal{X}_{k|k-1}^{x,(L+2:2L+1)}) \quad \sqrt{w^{(2),c2}} (\mathcal{X}_{k|k-1}^{x,(2:L+1)} + \mathcal{X}_{k|k-1}^{x,(L+2:2L+1)} - 2\mathcal{X}_{k|k-1}^{x,(1)}) \quad \mathbf{S}_{k-1}^{\kappa} \right] \right\} \end{aligned}$$

- Measurement Update

$$\begin{aligned} \mathcal{X}_{k|k-1}^* &= \begin{bmatrix} \hat{x}_{k|k-1} & \hat{x}_{k|k-1} + h(\mathbf{S}_{k|k-1}^x)^T & \hat{x}_{k|k-1} - h(\mathbf{S}_{k|k-1}^x)^T \end{bmatrix} \\ \mathcal{Y}_{k|k-1} &= \mathcal{H}(\mathcal{X}_{k|k-1}^*, u_k, \mathbb{E}[\gamma_k]) \\ \hat{y}_{k|k-1} &= \sum_{i=1}^{2L+1} w^{(i),m} \mathcal{Y}_{k|k-1}^{(i)} \\ \mathbf{S}_{k|k-1}^{\hat{y}} &= \text{QR} \left\{ \left[\sqrt{w^{(2),c1}} (\mathcal{Y}_{k|k-1}^{(2:L+1)} - \mathcal{Y}_{k|k-1}^{(L+2:2L+1)}) \quad \sqrt{w^{(2),c2}} (\mathcal{Y}_{k|k-1}^{(2:L+1)} - \mathcal{Y}_{k|k-1}^{(L+2:2L+1)} - 2\mathcal{Y}_{k|k-1}^{(1)}) \quad \mathbf{S}_{k-1}^\gamma \right] \right\}^T \\ \mathbf{P}_k^{xy} &= \sqrt{w^{(2),c1}} (\mathbf{S}_{k|k-1}^x)^T \left[\mathcal{Y}_{k|k-1}^{(2:L+1)} - \mathcal{Y}_{k|k-1}^{(L+2:2L+1)} \right]^T \\ \mathcal{K}_k &= (\mathbf{P}_k^{xy} / (\mathbf{S}_{k|k-1}^{\hat{y}})^T) / \mathbf{S}_{k|k-1}^{\hat{y}} \\ \hat{x}_{k|k} &= \hat{x}_{k|k-1} + \mathcal{K}_k (y_k - \hat{y}_{k|k-1}) \\ \mathfrak{U} &= \mathcal{K}_k \mathbf{S}_{k|k-1}^{\hat{y}} \\ \mathbf{S}_{k|k}^x &= \text{CHOLUPDATE} \{ \mathbf{S}_{k|k-1}^x, \mathfrak{U}, -1 \}^T \end{aligned}$$

END FOR

Parameters:

- $L = n$: augmented state dimension
- h : central difference interval parameter

Parameter setup recommendation:

- $h \geq 1$, $h = \sqrt{3}$ for gaussian priors
-
-

Square Root Central Difference Kalman Filter - Version "Parameter Estimation"

Algorithm 21 - Square Root Central Difference Kalman Filter - Version "Parameter Estimation"

$$[\{\theta_k, \mathbf{S}_k^\theta\}] = \text{SRCDFK}[\{\theta_{k-1}, \mathbf{S}_{k-1}^\theta\}, y_k]$$

Weights:

$$\begin{aligned} w^{(1),m} &= \frac{h^2 - L}{h^2} \\ w^{(i),m} &= \frac{1}{2h^2} & i = 2, \dots, 2L + 1 \\ w^{(i),c1} &= \frac{1}{4h^2} & i = 2, \dots, 2L + 1 \\ w^{(i),c2} &= \frac{h^2 - 1}{4h^4} & i = 2, \dots, 2L + 1 \end{aligned}$$

Initialization:

$$\begin{aligned} \hat{\theta}_{0|0} &= \mathbb{E}[\theta_0] \\ \mathbf{S}_{0|0}^\theta &= \text{CHOL} \left\{ \mathbb{E} \left[(\theta_0 - \mathbb{E}[\theta_0])(\theta_0 - \mathbb{E}[\theta_0])^T \right] \right\} \end{aligned}$$

FOR $k = 1, 2, \dots, K$:

$$\begin{aligned} \tilde{\mathbf{S}}_{k-1} &= \text{CHOL} \left\{ \mathbb{E} \left[(\tilde{\kappa}_{k-1} - \mathbb{E}[\tilde{\kappa}_{k-1}])(\tilde{\kappa}_{k-1} - \mathbb{E}[\tilde{\kappa}_{k-1}])^T \right] \right\} \\ \mathbf{S}_k^\gamma &= \text{CHOL} \left\{ \mathbb{E} \left[(\gamma_k - \mathbb{E}[\gamma_k])(\gamma_k - \mathbb{E}[\gamma_k])^T \right] \right\} \end{aligned}$$

- Time Update

$$\begin{aligned} \hat{\theta}_{k|k-1} &= \hat{\theta}_{k-1|k-1} + \mathbb{E}[\tilde{\kappa}_{k-1}] \\ \mathbf{S}_{k|k-1}^\theta &= \text{QR} \{ [\mathbf{S}_{k-1|k-1}^\theta \quad \tilde{\mathbf{S}}_{k-1}^\gamma] \} \end{aligned}$$

- Measurement Update

$$\begin{aligned} \mathcal{X}_{k|k-1} &= \left[\hat{\theta}_{k|k-1} \quad \hat{\theta}_{k|k-1} + h(\mathbf{S}_{k|k-1}^\theta)^T \quad \hat{\theta}_{k|k-1} - h(\mathbf{S}_{k|k-1}^\theta)^T \right] \\ \mathcal{Y}_{k|k-1} &= \mathcal{H}(\mathcal{X}_{k|k-1}, u_k, \mathbb{E}[\gamma_k]) = \mathcal{KF}(\mathcal{X}_{k|k-1}, u_k) + \mathbb{E}[\gamma_k] \\ \text{Option 1: } \hat{y}_{k|k-1} &= \sum_{i=1}^{2L+1} w^{(i),m} \mathcal{Y}_{k|k-1}^{(i)} \\ \text{Option 2: } \hat{y}_{k|k-1} &= \mathcal{H}(\hat{\theta}_{k|k-1}, u_k, \mathbb{E}[\gamma_k]) = \mathcal{KF}(\hat{\theta}_{k|k-1}, u_k) + \mathbb{E}[\gamma_k] \\ \mathbf{S}_k^{\hat{y}} &= \text{QR} \left\{ \left[\sqrt{w^{(2),c1}} (\mathcal{Y}_{k|k-1}^{(2:L+1)} - \mathcal{Y}_{k|k-1}^{(L+2:2L+1)}) \right. \right. \\ &\quad \left. \left. \sqrt{w^{(2),c2}} (\mathcal{Y}_{k|k-1}^{(2:L+1)} - \mathcal{Y}_{k|k-1}^{(L+2:2L+1)} - 2\mathcal{Y}_{k|k-1}^{(1)}) \quad \mathbf{S}_k^\gamma \right] \right\}^T \\ \mathbf{P}_k^{\theta y} &= \sqrt{w^{(2),c1}} (\mathbf{S}_{k|k-1}^\theta)^T \left[\mathcal{Y}_{k|k-1}^{(2:L+1)} - \mathcal{Y}_{k|k-1}^{(L+2:2L+1)} \right]^T \\ \mathcal{K}_k &= (\mathbf{P}_k^{\theta y} / (\mathbf{S}_k^{\hat{y}})^T) / \mathbf{S}_k^{\hat{y}} \\ \hat{\theta}_{k|k} &= \hat{\theta}_{k|k-1} + \mathcal{K}_k (y_k - \hat{y}_{k|k-1}) \\ \mathfrak{U} &= \mathcal{K}_k \mathbf{S}_k^{\hat{y}} \\ \mathbf{S}_{k|k}^\theta &= \text{CHOLUPDATE} \{ \mathbf{S}_{k|k-1}^\theta, \mathfrak{U}, -1 \}^T \end{aligned}$$

END FOR

Parameters:

- L : "state" dimension (number of estimated parameters)
- h : central difference interval parameter

Parameter setup recommendation:

- $h \geq 1$, $h = \sqrt{3}$ for gaussian priors
-
-

Gauss Hermite Filter - GHF (Only Additive Noise Case)

Algorithm 22 - Gauss Hermite Filter - GHF

$$[\{x_k, \mathbf{P}_k^x\}] = \text{GHF}[\{x_{k-1}, \mathbf{P}_{k-1}^x\}, y_k]$$

Quadrature Points and Weights:

Let \mathbf{J} be a symmetric tridiagonal matrix with zero diagonals and $\mathbf{J}^{(i),(i+1)} = \sqrt{i/2}$, $1 \leq i \leq m-1$:

$$\mathbf{J} = \begin{bmatrix} 0 & \sqrt{1/2} & & & \\ \sqrt{1/2} & \ddots & & & \\ & \ddots & \ddots & & \\ & & \ddots & \sqrt{(m-1)/2} & \\ & & & \sqrt{(m-1)/2} & 0 \end{bmatrix}$$

Then the quadrature points are $q^{(i)} = \lambda^{(i)} \sqrt{2}$ ($\mathcal{QP} = [q^{(1)} \ q^{(2)} \ \dots \ q^{(m)}]$) where $\lambda^{(i)}$ are the eigenvalues of \mathbf{J} and the weights are $w^{(i)} = \left| \left(v^{(i)} \right)_1 \right|^2$ where $\left(v^{(i)} \right)_1$ denotes the first element of the i_{th} normalized eigenvector of \mathbf{J} .

Initialization:

$$\begin{aligned} \hat{x}_{0|0} &= \mathbb{E}[x_0] \\ \mathbf{P}_{0|0}^x &= \mathbb{E}[(x_0 - \mathbb{E}[x_0])(x_0 - \mathbb{E}[x_0])^T] \end{aligned}$$

FOR $k = 1, 2, \dots, K$:

$$\begin{aligned} \mathbf{P}_{k-1}^{\kappa} &= \mathbb{E}[(\kappa_{k-1} - \mathbb{E}[\kappa_{k-1}])(\kappa_{k-1} - \mathbb{E}[\kappa_{k-1}])^T] \\ \mathbf{P}_k^{\gamma} &= \mathbb{E}[(\gamma_k - \mathbb{E}[\gamma_k])(\gamma_k - \mathbb{E}[\gamma_k])^T] \end{aligned}$$

- Time Update

$$\begin{aligned} \mathcal{X} &= \hat{x}_{k-1|k-1} + \text{CHOL}\{\mathbf{P}_{k-1|k-1}^x\} \mathcal{QP} \\ \hat{x}_{k|k-1} &= \sum_{i_1=1}^m \cdots \sum_{i_L=1}^m \mathcal{F}([\mathcal{X}_{i_1}, \mathcal{X}_{i_2}, \dots, \mathcal{X}_{i_L}]^T, u_{k-1}, \mathbb{E}[\kappa_{k-1}]) w_{i_1} w_{i_2} \cdots w_{i_L} \\ \mathbf{P}_{k|k-1}^x &= \mathbf{P}_{k-1}^x + \sum_{i_1=1}^m \cdots \sum_{i_L=1}^m (\mathcal{F}([\mathcal{X}_{i_1}, \mathcal{X}_{i_2}, \dots, \mathcal{X}_{i_L}]^T, u_{k-1}, \mathbb{E}[\kappa_{k-1}]) - \hat{x}_{k|k-1})^2 w_{i_1} w_{i_2} \cdots w_{i_L} \end{aligned}$$

- Measurement Update

$$\begin{aligned} \mathcal{V} &= \hat{x}_{k|k-1} + \text{CHOL}\{\mathbf{P}_{k|k-1}^x\} \mathcal{QP} \\ \hat{y}_{k|k-1} &= \sum_{i_1=1}^m \cdots \sum_{i_L=1}^m \mathcal{H}([\mathcal{V}_{i_1}, \mathcal{V}_{i_2}, \dots, \mathcal{V}_{i_L}]^T, u_k, \mathbb{E}[\gamma_k]) w_{i_1} w_{i_2} \cdots w_{i_L} \\ \mathbf{P}_k^{\hat{y}} &= \mathbf{P}_k^x + \sum_{i_1=1}^m \cdots \sum_{i_L=1}^m (\mathcal{H}([\mathcal{V}_{i_1}, \mathcal{V}_{i_2}, \dots, \mathcal{V}_{i_L}]^T, u_k, \mathbb{E}[\gamma_k]) - \hat{y}_{k|k-1})^2 w_{i_1} w_{i_2} \cdots w_{i_L} \\ \mathbf{P}_k^{xy} &= \sum_{i_1=1}^m \cdots \sum_{i_L=1}^m ([\mathcal{V}_{i_1}, \mathcal{V}_{i_2}, \dots, \mathcal{V}_{i_L}]^T - \hat{x}_{k|k-1}) \\ &\quad \times (\mathcal{H}([\mathcal{V}_{i_1}, \mathcal{V}_{i_2}, \dots, \mathcal{V}_{i_L}]^T, u_k, \mathbb{E}[\gamma_k]) - \hat{y}_{k|k-1})^T w_{i_1} w_{i_2} \cdots w_{i_L} \\ \mathcal{K}_k &= \mathbf{P}_k^{xy} (\mathbf{P}_k^{\hat{y}})^{-1} \\ \hat{x}_{k|k} &= \hat{x}_{k|k-1} + \mathcal{K}_k (y_k - \hat{y}_{k|k-1}) \\ \mathbf{P}_{k|k}^x &= \mathbf{P}_{k|k-1}^x - \mathcal{K}_k \mathbf{P}_k^{\hat{y}} \mathcal{K}_k^T \end{aligned}$$

END FOR

Parameters:

- $L = n$: state dimension
- m : order of gauss hermite quadrature rule

Parameter setup recommendation:

- $m \geq 1, 3 \leq m \leq 7$

Note, that the GHF is only suitable for additive noise case systems, since the augmentation of the state vector with the noises (non additive noise case) would lead to a severe increase of the computational effort and the algorithm would become computationally intractable very soon.

Gauss Hermite Filter - Version "Parameter Estimation"

Algorithm 23 - Gauss Hermite Filter - Version "Parameter Estimation"

$$[\{\theta_k, \mathbf{P}_k^\theta\}] = \text{GHF}[\{\theta_{k-1}, \mathbf{P}_{k-1}^\theta\}, y_k]$$

Quadrature Points and Weights:

Let \mathbf{J} be a symmetric tridiagonal matrix with zero diagonals and $\mathbf{J}^{(i), (i+1)} = \sqrt{i/2}$, $1 \leq i \leq m-1$:

$$\mathbf{J} = \begin{bmatrix} 0 & \sqrt{1/2} & & & \\ \sqrt{1/2} & \ddots & & & \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ & & & \sqrt{(m-1)/2} & 0 \end{bmatrix}$$

Then the quadrature points are $q^{(i)} = \lambda^{(i)}\sqrt{2}$ ($\mathcal{QP} = [q^{(1)} \ q^{(2)} \ \dots \ q^{(m)}]$) where $\lambda^{(i)}$ are the eigenvalues of \mathbf{J} and the weights are $w^{(i)} = \left| \left(v^{(i)} \right)_1 \right|^2$ where $\left(v^{(i)} \right)_1$ denotes the first element of the i_{th} normalized eigenvector of \mathbf{J} .

Initialization:

$$\begin{aligned} \hat{\theta}_{0|0} &= \mathbb{E}[\theta_0] \\ \mathbf{P}_{0|0}^\theta &= \mathbb{E}[(\theta_0 - \mathbb{E}[\theta_0])(\theta_0 - \mathbb{E}[\theta_0])^T] \end{aligned}$$

FOR $k = 1, 2, \dots, K$:

$$\begin{aligned} \mathbf{P}_{k-1}^{\tilde{\kappa}} &= \mathbb{E}[(\tilde{\kappa}_{k-1} - \mathbb{E}[\tilde{\kappa}_{k-1}])(\tilde{\kappa}_{k-1} - \mathbb{E}[\tilde{\kappa}_{k-1}])^T] \\ \mathbf{P}_k^\gamma &= \mathbb{E}[(\gamma_k - \mathbb{E}[\gamma_k])(\gamma_k - \mathbb{E}[\gamma_k])^T] \end{aligned}$$

- Time Update

$$\begin{aligned} \hat{\theta}_{k|k-1} &= \hat{\theta}_{k-1|k-1} + \mathbb{E}[\tilde{\kappa}_{k-1}] \\ \mathbf{P}_{k|k-1}^\theta &= \mathbf{P}_{k-1|k-1}^\theta + \mathbf{P}_{k-1}^{\tilde{\kappa}} \end{aligned}$$

- Measurement Update

$$\begin{aligned} \mathcal{V} &= \hat{\theta}_{k|k-1} + \text{CHOL}\{\mathbf{P}_{k|k-1}^\theta\} \mathcal{QP} \\ \hat{y}_{k|k-1} &= \sum_{i_1=1}^m \dots \sum_{i_L=1}^m \mathcal{H}([x_k^T, [\mathcal{V}_{i_1}, \mathcal{V}_{i_2}, \dots, \mathcal{V}_{i_L}]]^T, u_k, \mathbb{E}[\gamma_k]) w_{i_1} w_{i_2} \dots w_{i_L} \\ \mathbf{P}_k^{\hat{y}} &= \mathbf{P}_k^\gamma + \sum_{i_1=1}^m \dots \sum_{i_L=1}^m (\mathcal{H}([x_k^T, [\mathcal{V}_{i_1}, \mathcal{V}_{i_2}, \dots, \mathcal{V}_{i_L}]]^T, u_k, \mathbb{E}[\gamma_k]) - \hat{y}_{k|k-1})^2 w_{i_1} w_{i_2} \dots w_{i_L} \\ \mathbf{P}_k^{\theta y} &= \sum_{i_1=1}^m \dots \sum_{i_L=1}^m ([\mathcal{V}_{i_1}, \mathcal{V}_{i_2}, \dots, \mathcal{V}_{i_L}]^T - \hat{\theta}_{k|k-1}) \\ &\quad \times (\mathcal{H}([x_k^T, [\mathcal{V}_{i_1}, \mathcal{V}_{i_2}, \dots, \mathcal{V}_{i_L}]]^T, u_k, \mathbb{E}[\gamma_k]) - \hat{y}_{k|k-1})^T w_{i_1} w_{i_2} \dots w_{i_L} \\ \mathcal{K}_k &= \mathbf{P}_k^{\theta y} (\mathbf{P}_k^{\hat{y}})^{-1} \\ \hat{\theta}_{k|k} &= \hat{\theta}_{k|k-1} + \mathcal{K}_k (y_k - \hat{y}_{k|k-1}) \\ \mathbf{P}_{k|k}^\theta &= \mathbf{P}_{k|k-1}^\theta - \mathcal{K}_k \mathbf{P}_k^{\hat{y}} \mathcal{K}_k^T \end{aligned}$$

END FOR

Parameters:

- L : "state" dimension (number of estimated parameters)
- m : order of gauss hermite quadrature rule

Parameter setup recommendation:

- $m \geq 1, 3 \leq m \leq 7$
-
-

Gaussian Sum Kalman Filter

Gaussian Sum Kalman Filter - GSKF

Algorithm 24 - Gaussian Sum Kalman Filter - GSKF

$$[\{x_k, \bar{x}_k^{(g)}, \mathbf{P}_k^{x,(g)}\}_{g=1}^G] = \text{GSKF}[\{\bar{x}_{k-1}^{(g)}, \mathbf{P}_{k-1}^{x,(g)}\}_{g=1}^G, y_k]$$

Initialization:

$$\begin{aligned} \hat{x}_{0|0} &= \mathbb{E}[x_0] \\ p(x_0) &\approx \sum_{g=1}^G \alpha_0^{(g)} \mathcal{N}(x_0; \bar{x}_0^{(g)}, \mathbf{P}_0^{x,(g)}) \end{aligned}$$

FOR $k = 1, \dots, K$:

- GMM-Models:

- Process noise density: $p_{\mathcal{GMM}}(\kappa_{k-1}) = \sum_{m=1}^M \beta_{k-1}^{(m)} \mathcal{N}(\kappa_{k-1}; \bar{\kappa}_{k-1}^{(m)}, \mathbf{P}_{k-1}^{\kappa,(m)})$
- Observation noise density: $p_{\mathcal{GMM}}(\gamma_k) = \sum_{j=1}^J \varphi_{k-1}^{(j)} \mathcal{N}(\gamma_k; \bar{\gamma}_k^{(j)}, \mathbf{P}_k^{\gamma,(j)})$

- Option 1:

- Time Update

- * FOR $g = 1, \dots, G$
 - FOR $m = 1, \dots, M$

$$\begin{aligned} gm &= g \cdot m \\ [\{\bar{x}_{k|k-1}^{(gm)}, \mathbf{P}_{k|k-1}^{x,(gm)}\}] &= \text{TimeUpd}_{EKF,UKF,\dots}[\{\bar{x}_{k-1}^{(g)}, \mathbf{P}_{k-1}^{x,(g)}, \bar{\kappa}_{k-1}^{(m)}, \mathbf{P}_{k-1}^{\kappa,(m)}\}] \\ \alpha^{(gm)} &= \frac{\alpha_{k-1}^{(g)} \cdot \beta_{k-1}^{(m)}}{\sum_{g=1}^G \sum_{m=1}^M \alpha_{k-1}^{(g)} \cdot \beta_{k-1}^{(m)}} \end{aligned}$$

· END FOR

- * END FOR

$$* \alpha^{(gm)} = \frac{\alpha^{(gm)}}{\sum_{gm=1}^{g \cdot m} \alpha^{(gm)}}$$

- Measurement Update

- * FOR $gm = 1, \dots, g \cdot m$
 - FOR $j = 1, \dots, J$

$$\begin{aligned} gmj &= g \cdot m \cdot j \\ [\{\bar{x}_{k|k}^{(gmj)}, \mathbf{P}_{k|k}^{x,(gmj)}\}] &= \text{MeasUpd}_{EKF,UKF,\dots}[\{\bar{x}_{k|k-1}^{(gm)}, \mathbf{P}_{k|k-1}^{x,(gm)}, \bar{\gamma}_k^{(j)}, \mathbf{P}_k^{\gamma,(j)}\}] \\ \alpha^{(gmj)} &= \frac{\alpha^{(gm)} \cdot \varphi_k^{(j)} \cdot p(y_k | \bar{x}_{k|k}^{(gm)})^{(j)}}{\sum_{gm=1}^{g \cdot m} \sum_{j=1}^J \alpha^{(gm)} \cdot \varphi_k^{(j)} \cdot p(y_k | \bar{x}_{k|k-1}^{(gm)})^{(j)}} \end{aligned}$$

· END FOR

- * END FOR

$$- \alpha^{(gmj)} = \frac{\alpha^{(gmj)}}{\sum_{gmj=1}^{g \cdot m \cdot j} \alpha^{(gmj)}}$$

- Option 2:

– FOR $g = 1, \dots, G$

* FOR $m = 1, \dots, M$

· FOR $j = 1, \dots, J$

$$gmj = g \cdot m \cdot j$$

$$\{ \{ \bar{x}_{k|k}^{(gmj)}, \mathbf{P}_{k|k}^{x,(gmj)} \} \} = \text{EKF/UKF} / \dots / \{ \{ \bar{x}_{k-1}^{(g)}, \mathbf{P}_{k-1}^{x,(g)}, \bar{\kappa}_{k-1}^{(m)}, \mathbf{P}_{k-1}^{\kappa,(m)}, \bar{\gamma}_k^{(j)}, \mathbf{P}_k^{\gamma,(j)} \} \}$$

$$\alpha^{(gmj)} = \alpha_{k-1}^{(g)} \cdot \beta_{k-1}^{(m)} \cdot \varphi_k^{(j)} \cdot p(y_k | \bar{x}_{k|k}^{(gmj)})$$

· END FOR

* END FOR

– END FOR

$$\alpha^{(gmj)} = \frac{\alpha^{(gmj)}}{\sum_{gmj=1}^{g \cdot m \cdot j} \alpha^{(gmj)}}$$

- Inference (MMSE)

$$\hat{x}_{k|k} = \sum_{gmj=1}^{g \cdot m \cdot j} w^{(gmj)} \bar{x}_{k|k}^{(gmj)}$$

- GMM-Recovery:

– For $g = 1, 2, \dots, G$ draw a number $d \in \{1, \dots, g \cdot m \cdot j\}$ with probabilities proportional to $\{\alpha^{(gmj)}\}$, assign $\{ \bar{x}_k^{(g)}, \mathbf{P}_k^{x,(g)} \} = \{ \bar{x}_{k|k}^{(d)}, \mathbf{P}_{k|k}^{x,(d)} \}$ and set $\alpha_k^{(g)} = \frac{1}{G}$.

$$\text{GMM-Model: } p_{\mathcal{GMM}}(x_k | Y_k) = \sum_{g=1}^G \alpha_k^{(g)} \mathcal{N}(x_k; \bar{x}_k^{(g)}, \mathbf{P}_k^{x,(g)})$$

END FOR

Parameters:

- G, M, J : Number of mixing components for state and system noise density
- Depending on the choice of gaussian filter (EKF, UKF, ...) the respective parameters accrue additionally

Sequential Monte Carlo Methods

Sampling Importance Resampling (Systematic Resampling)

Sampling Importance Resampling (Systematic Resampling)

$$[\{x_k^{(j)*}, w_k^{(j)}, ParI^{(j)}\}_{j=1}^N] = \text{SIR-RESAMPLE}[\{x_k^{(i)}, w_k^{(i)}\}_{i=1}^N]$$

- FOR $i = 2, \dots, N$
 - Construct (Cumulative Sum of Weights)-Vector: $CSW^{(i)} = CSW^{(i-1)} + w_k^{(i)}$, $CSW^{(1)} = w_k^{(1)}$
- END FOR
- $\vartheta^{(1)} \sim \mathcal{U}(\vartheta; 0, N^{-1})$
- FOR $j = 1, 2, \dots, N$
 - $\vartheta^{(j)} = \vartheta^{(1)} + N^{-1}(j - 1)$
 - WHILE $\vartheta^{(j)} > CSW^{(i)}$
 - * $i = i + 1$
 - END WHILE
 - Assign Particle: $x_k^{(j)*} = x_k^{(i)}$
 - Assign Weight: $w_k^{(j)} = N^{-1}$
 - Assign Parent Index: $ParI^{(j)} = i$
- END FOR

Residual Resampling

Residual Resampling

- $$[\{x_k^{(j)*}, w_k^{(j)}, ParI^{(j)}\}_{j=1}^N] = \text{RR-RESAMPLE}[\{x_k^{(i)}, w_k^{(i)}\}_{i=1}^N]$$
- FOR $i = 1, 2, \dots, N$
 - Assign "Children A": $N_a^{(i)} = \lfloor N w_k^{(i)} \rfloor$
 - END FOR
 - $\bar{N}_k = N - \sum_{i=1}^N N_a^{(i)}$
 - $w_{res} = (w_k^{N-N_a}) / \bar{N}_k$
 - IF $\bar{N}_k \neq 0$
 - $[\{-, -, ParI^{(i)}\}_{i=1}^N] = \text{SIR-RESAMPLE}[\{-, w_{res}^{(i)}\}_{i=1}^N]$
 - Assign "Children B": $N_b = ParI^{(i)}$
 - $N_{Child} = N_a + N_b$
 - ELSE
 - $N_{Child} = N_a$
 - END IF
 - $j = 1$
 - FOR $i = 1, 2, \dots, N$
 - IF $N_{Child}^{(i)} \neq 0$
 - * FOR $m = j : j + N_{Child}^{(i)} - 1$
 - Assign Particle: $x_k^{(m)*} = x_k^{(i)}$
 - Assign Weight: $w_k^{(m)} = N^{-1}$
 - Assign Parent Index: $ParI^{(m)} = i$
 - * END FOR
 - END IF
 - $j = j + N_{Child}^{(i)}$
 - END FOR
-
-

Markov Chain Monte Carlo Step

Markov Chain Monte Carlo Step

$$[\{x_k^{(i)}\}_{i=1}^N] = \text{MCMC}[\{\tilde{x}_k^{(i)}, \tilde{x}_{k-1}^{(i)}\}_{i=1}^N, y_k]$$

Definitions:

- \tilde{x}_k : Resampled particles at time k
- \tilde{x}_{k-1} : Resampled particles at time $k - 1$

Algorithm:

- FOR $i = 1, 2, \dots, N$
 - $x_k^{(i)*} \sim \pi(x_k | \tilde{x}_{k-1}^{(i)}, y_k)$ (sampling from instrumental proposal distribution)
 - $\vartheta \sim \mathcal{U}(\vartheta; 0, 1)$
 - $\alpha = \min\left\{1, \frac{p(y_k | x_k^{(i)*})p(x_k^{(i)*} | \tilde{x}_{k-1}^{(i)})\pi(\tilde{x}_k^{(i)} | \tilde{x}_{k-1}^{(i)}, y_k)}{p(y_k | \tilde{x}_k^{(i)})p(\tilde{x}_k^{(i)} | \tilde{x}_{k-1}^{(i)})\pi(x_k^{(i)*} | \tilde{x}_{k-1}^{(i)}, y_k)}\right\}$
 - IF $\vartheta \leq \alpha$
 - * $x_k^{(i)} = x_k^{(i)*}$
 - ELSE
 - * $x_k^{(i)} = \tilde{x}_k^{(i)}$
 - END IF
- END FOR

Generic Particle Filter - GPF

Algorithm 25 - Generic Particle Filter - GPF

$$[\{x_k^{(i)}, w_k^{(i)}\}_{i=1}^N] = \text{GPF}[\{x_{k-1}^{(i)}, w_{k-1}^{(i)}\}_{i=1}^N, y_k]$$

Initialization:

$$\begin{aligned} \hat{x}_{0|0} &= \mathbb{E}[x_0] \\ x_0^{(i)} &\sim p(x_0) & i &= 1, 2, \dots, N \\ w_0^{(i)} &= N^{-1} & i &= 1, 2, \dots, N \end{aligned}$$

FOR $k = 1, 2, \dots, K$:

- FOR $i = 1, 2, \dots, N$
 - $\kappa_{k-1}^{(i)} \sim p(\kappa_{k-1})$
 - $x_k^{(i)} = \mathcal{F}(x_{k-1}^{(i)}, u_{k-1}, \kappa_{k-1}^{(i)})$
 - $w_k^{(i)} = w_{k-1}^{(i)} p(y_k | x_k^{(i)})$
- END FOR
- $\tilde{w}_k = w_k / \sum_{i=1}^N w_k^{(i)}$
- $N_{eff} = 1 / \sum_{i=1}^N (\tilde{w}_k^{(i)})^2$
- IF $N_{eff} < N_{thr}$
 - $[\{x_k^{(i)}, w_k^{(i)} = \frac{1}{N}, ParI^{(i)}\}_{i=1}^N] = \text{RESAMPLE}[\{x_k^{(i)}, \tilde{w}_k^{(i)}\}_{i=1}^N]$
 - $w_k^{(i)} = \tilde{w}_k^{(i)} = \frac{1}{N}$
 - Optional MCMC-Step:
 - * Assign: $\{\tilde{x}_{k-1}^{(i)} = x_{k-1}^{ParI^{(i)}}\}_{i=1}^N$
 - * $[\{x_k^{(i)}\}_{i=1}^N] = \text{MCMC}[\{x_k^{(i)}, \tilde{x}_{k-1}^{(i)}\}_{i=1}^N, y_k]$
- END IF
- Inference (MMSE)
 - $\hat{x}_{k|k} = \sum_{i=1}^N \tilde{w}_k^{(i)} x_k^{(i)}$

END FOR

Parameters:

- N : Number of particles
 - N_{thr} : Resample threshold
-
-

Auxiliary Particle Filter - ASIR

Algorithm 26 - Auxiliary Particle Filter - ASIR

$$[\{x_k^{(i)}, w_k^{(i)}\}_{i=1}^N] = \text{ASIR}[\{x_{k-1}^{(i)}, w_{k-1}^{(i)}\}_{i=1}^N, y_k]$$

Initialization:

$$\begin{aligned} \hat{x}_{0|0} &= \mathbb{E}[x_0] \\ x_0^{(i)} &\sim p(x_0) & i &= 1, 2, \dots, N \\ w_0^{(i)} &= N^{-1} & i &= 1, 2, \dots, N \end{aligned}$$

FOR $k = 1, 2, \dots, K$:

- FOR $i = 1, 2, \dots, N$

$$\begin{aligned} \text{Option 1: } \kappa_{k-1}^{(i)} &\sim p(\kappa_{k-1}) \\ \mu_k^{(i)} &= \mathcal{F}(x_{k-1}^{(i)}, u_{k-1}, \kappa_{k-1}^{(i)}) \\ \text{Option 2: } \mu_k^{(i)} &= \mathcal{F}(x_{k-1}^{(i)}, u_{k-1}, \mathbb{E}[\kappa_{k-1}]) \end{aligned}$$

$$w_k^{(i)} = w_{k-1}^{(i)} p(y_k | \mu_k^{(i)})$$

- END FOR

- $\tilde{w}_k = w_k / \sum_{i=1}^N w_k^{(i)}$

- Resampling

$$- [\{-, -, \text{ParI}^{(i)}\}_{i=1}^N] = \text{RESAMPLE}[\{\mu_k^{(i)}, \tilde{w}_k^{(i)}\}_{i=1}^N]$$

- FOR $i = 1, 2, \dots, N$

$$\begin{aligned} - \kappa_{k-1}^{(i)} &\sim p(\kappa_{k-1}) \\ - x_k^{(i)} &= \mathcal{F}(x_{k-1}^{\text{ParI}^{(i)}}, u_{k-1}, \kappa_{k-1}^{(i)}) \\ - \tilde{w}_k^{(i)} &= \frac{p(y_k | x_k^{(i)})}{p(y_k | \mu_k^{\text{ParI}^{(i)}})} \end{aligned}$$

- END FOR

- $w_k = \tilde{w}_k / \sum_{i=1}^N \tilde{w}_k^{(i)}$

- Inference (MMSE)

$$- \hat{x}_{k|k} = \sum_{i=1}^N w_k^{(i)} x_k^{(i)}$$

END FOR

Parameters:

- N : Number of particles
-
-

Regularized Particle Filter - RPF

Algorithm 27 - Regularized Particle Filter - RPF

$$[\{x_k^{(i)}, w_k^{(i)}\}_{i=1}^N] = \text{RPF}[\{x_{k-1}^{(i)}, w_{k-1}^{(i)}\}_{i=1}^N, y_k]$$

Initialization:

$$\begin{aligned} \hat{x}_{0|0} &= \mathbb{E}[x_0] \\ x_0^{(i)} &\sim p(x_0) & i &= 1, 2, \dots, N \\ w_0^{(i)} &= N^{-1} & i &= 1, 2, \dots, N \end{aligned}$$

FOR $k = 1, 2, \dots, K$:

- FOR $i = 1, 2, \dots, N$
 - $\kappa_{k-1}^{(i)} \sim p(\kappa_{k-1})$
 - $x_k^{(i)} = \mathcal{F}(x_{k-1}^{(i)}, u_{k-1}, \kappa_{k-1}^{(i)})$
 - $w_k^{(i)} = w_{k-1}^{(i)} p(y_k | x_k^{(i)})$
- END FOR
- $\tilde{w}_k = w_k / \sum_{i=1}^N w_k^{(i)}$
- $N_{eff} = 1 / \sum_{i=1}^N (\tilde{w}_k^{(i)})^2$
- IF $N_{eff} < N_{thr}$
 - $\mathcal{CV} = \text{COV}\{x_k^{(i)}, \tilde{w}_k^{(i)}\}$
 - $\mathcal{SV} = \text{CHOL}\{\mathcal{CV}\}$
 - $[\{x_k^{(i)*}, w_k^{(i)} = \frac{1}{N}, \text{ParI}^{(i)}\}_{i=1}^N] = \text{RESAMPLE}[\{x_k^{(i)}, w_k^{(i)}\}_{i=1}^N]$
 - $\psi^{(i)} \sim \text{KEpanechnikov/Gauss}$
 - $x_k^{(i)} = x_k^{(i)*} + \lambda_{Opt} \mathcal{SV} \psi^{(i)}$
 - $w_k^{(i)} = \tilde{w}_k^{(i)} = \frac{1}{N}$
 - Optional MCMC-Step:
 - * Assign: $\{\tilde{x}_{k-1}^{(i)} = x_{k-1}^{(i)*}\}_{i=1}^N$
 - * $[\{x_k^{(i)}\}_{i=1}^N] = \text{MCMC}[\{x_k^{(i)}, \tilde{x}_{k-1}^{(i)}, y_k\}_{i=1}^N]$
- END IF
- Inference (MMSE)
 - $\hat{x}_{k|k} = \sum_{i=1}^N \tilde{w}_k^{(i)} x_k^{(i)}$

END FOR

Parameters:

- N : Number of particles
- N_{thr} : Resample threshold
- λ_{Opt} : Kernel bandwidth

Parameter setup recommendation:

- $\lambda_{Opt} = 0.5 \left(\frac{4}{N(n+2)} \right)^{\frac{1}{n+4}}$ for gaussian kernel
-
-

Gaussian Filter Bank Particle Filter - EKPF, SPPF

Algorithm 28 - Gaussian Filter Bank Particle Filter - EKPF, SPPF $\{(SR)UKF, (SR)CDKF\}$

$$\{\{x_k^{(i)}, w_k^{(i)}\}_{i=1}^N\} = \text{GFBBPF}[\{\{x_{k-1}^{(i)}, w_{k-1}^{(i)}\}_{i=1}^N, y_k\}]$$

Initialization:

$$\begin{aligned} \hat{x}_{0|0} &= \mathbb{E}[x_0] \\ x_0^{(i)} &\sim p(x_0) & i &= 1, 2, \dots, N \\ w_0^{(i)} &= N^{-1} & i &= 1, 2, \dots, N \end{aligned}$$

FOR $k = 1, 2, \dots, K$:

- FOR $i = 1, 2, \dots, N$
 - $\{\{\hat{x}_k^{(i)}, \hat{\mathbf{P}}_k^{x, (i)}\}\} = \text{EKF/UKF/CDKF}/\dots[\{x_{k-1}^{(i)}, \hat{\mathbf{P}}_{k-1}^{x, (i)}, y_k\}]$
 - $x_k^{(i)} \sim \mathcal{N}(x_k^{(i)}; \hat{x}_k^{(i)}, \hat{\mathbf{P}}_k^{x, (i)})$
 - $w_k^{(i)} = w_{k-1}^{(i)} \frac{p(y_k | x_k^{(i)}) p(x_k^{(i)} | x_{k-1}^{(i)})}{\mathcal{N}(x_k^{(i)}; \hat{x}_k^{(i)}, \hat{\mathbf{P}}_k^{x, (i)})} = w_{k-1}^{(i)} \frac{\text{Likelihood Prior}}{\text{Proposal}}$
- END FOR
- $\tilde{w}_k = w_k / \sum_{i=1}^N w_k^{(i)}$
- $N_{eff} = 1 / \sum_{i=1}^N (\tilde{w}_k^{(i)})^2$
- IF $N_{eff} < N_{thr}$
 - $\{\{x_k^{(i)}, w_k^{(i)} = \frac{1}{N}, \text{Par}I^{(i)}\}_{i=1}^N\} = \text{RESAMPLE}[\{\{x_k^{(i)}, \tilde{w}_k^{(i)}\}_{i=1}^N\}]$
 - $w_k^{(i)} = \tilde{w}_k^{(i)} = \frac{1}{N}$
 - Optional MCMC-Step:
 - * Assign: $\{\{\tilde{x}_{k-1}^{(i)} = x_{k-1}^{\text{Par}I^{(i)}}\}_{i=1}^N\}$
 - * $\{\{x_k^{(i)}\}_{i=1}^N\} = \text{MCMC}[\{\{x_k^{(i)}, \tilde{x}_{k-1}^{(i)}, y_k\}_{i=1}^N\}]$
- END IF
- Inference (MMSE)
 - $\hat{x}_{k|k} = \sum_{i=1}^N \tilde{w}_k^{(i)} x_k^{(i)}$

END FOR

Parameters:

- N : Number of particles
- N_{thr} : Resample threshold
- Depending on the choice of proposal (EKF, UKF, ...) the respective parameters accrue additionally

Gaussian Sum Particle Filter - GSPF

Algorithm 29 - Gaussian Sum Particle Filter - GSPF

$$[\{x_k^{(i)}, w_k^{(i)}\}_{i=1}^N] = \text{GSPF}[\{x_{k-1}^{(i)}, w_{k-1}^{(i)}\}_{i=1}^N, y_k]$$

Initialization:

$$\begin{aligned} \hat{x}_{0|0} &= \mathbb{E}[x_0] \\ x_0^{(i)} &\sim p(x_0) \quad i = 1, 2, \dots, N \\ p(x_0) &\approx \sum_{g=1}^G \alpha_0^{(g)} \mathcal{N}(x_0; \bar{x}_0^{(g)}, \mathbf{P}_0^{x, (g)}) \end{aligned}$$

FOR $k = 1, 2, \dots, K$:

- GMM-Model process noise:

$$- \text{Process noise density: } p_{\mathcal{GM}}(\kappa_{k-1}) = \sum_{m=1}^M \beta_{k-1}^{(m)} \mathcal{N}(\kappa_{k-1}; \bar{\kappa}_{k-1}^{(m)}, \mathbf{P}_{k-1}^{\kappa, (m)})$$

- Time Update

$$\begin{aligned} - \text{For } m = 1, 2, \dots, M, \text{ set } p(\kappa_{k-1})^{(m)} &= \mathcal{N}(\kappa_{k-1}; \bar{\kappa}_{k-1}^{(m)}, \mathbf{P}_{k-1}^{\kappa, (m)}) \\ \text{For } g = 1, 2, \dots, G, \text{ set } p(x_{k-1}|Y_{k-1})^{(g)} &= \mathcal{N}(x_{k-1}; \bar{x}_{k-1}^{(g)}, \mathbf{P}_{k-1}^{x, (g)}) \\ - \text{FOR } g = 1, 2, \dots, G \end{aligned}$$

$$\begin{aligned} * \text{ FOR } i = 1, 2, \dots, N \\ \cdot x_{k-1}^{(g), (i)} &\sim \mathcal{N}(x_{k-1}; \bar{x}_{k-1}^{(g)}, \mathbf{P}_{k-1}^{x, (g)}) \\ * \text{ END FOR} \end{aligned}$$

- END FOR

- FOR $m = 1, 2, \dots, M$

$$\begin{aligned} * \text{ FOR } g = 1, 2, \dots, G \\ \cdot gm = g + (k-1)G \\ \cdot \text{ FOR } i = 1, 2, \dots, N \end{aligned}$$

$$\begin{aligned} \kappa_{k-1}^{(m), (i)} &\sim \mathcal{N}(\kappa_{k-1}; \bar{\kappa}_{k-1}^{(m)}, \mathbf{P}_{k-1}^{\kappa, (m)}) \\ x_{k|k-1}^{(gm), (i)} &= \mathcal{F}(x_{k-1}^{(g), (i)}, u_{k-1}, \kappa_{k-1}^{(m), (i)}) \end{aligned}$$

$$\begin{aligned} \cdot \text{ END FOR} \\ \cdot w^{(gm)} &= \alpha_{k-1}^{(g)} \cdot \beta_{k-1}^{(m)} \end{aligned}$$

* END FOR

$$* \bar{w}^{(gm)} = \frac{w^{(gm)}}{\sum_{gm=1}^{GM} w^{(gm)}}$$

* FOR $gm = 1, 2, \dots, GM$

$$\begin{aligned} \cdot \bar{x}_{k|k-1}^{(gm)} &= \frac{1}{N} \sum_{i=1}^N x_{k|k-1}^{(gm), (i)} \\ \cdot \mathbf{P}_{k|k-1}^{x, (gm)} &= \text{COV}\{x_{k|k-1}^{(gm), (i)}, \frac{1}{N}\} \end{aligned}$$

* END FOR

- END FOR

- Measurement Update

$$\begin{aligned} - \text{For } gm = 1, 2, \dots, GM, \text{ set } p_{\mathcal{GM}}(x_k|Y_{k-1}) &= \sum_{gm=1}^{GM} \bar{w}^{(gm)} \mathcal{N}(x_k; \bar{x}_{k|k-1}^{(gm)}, \mathbf{P}_{k|k-1}^{x, (gm)}) \\ - \text{FOR } gm = 1, 2, \dots, GM \end{aligned}$$

$$\begin{aligned} * \text{ FOR } i = 1, 2, \dots, N \\ \cdot x_{k|k}^{(gm), (i)} &\sim \mathcal{N}(x_k; \bar{x}_{k|k-1}^{(gm)}, \mathbf{P}_{k|k-1}^{x, (gm)}) \\ \cdot \tilde{w}^{(gm), (i)} &= p(y_k | x_{k|k}^{(gm), (i)}) \end{aligned}$$

* END FOR

$$* \bar{x}_{k|k}^{(gm)} = \frac{\sum_{i=1}^N \tilde{w}^{(gm), (i)} x_{k|k}^{(gm), (i)}}{\sum_{i=1}^N \tilde{w}^{(gm), (i)}}$$

$$\begin{aligned}
& * \mathbf{P}_{k|k}^{x,(gm)} = \text{COV}\{x_{k|k}^{(gm),(i)}, \tilde{w}^{(gm),(i)} / \sum_{i=1}^N \tilde{w}^{(gm),(i)}\} \\
& * \tilde{w}^{(gm)} = \tilde{w}^{(gm)} \sum_{i=1}^N \tilde{w}^{(gm),(i)} / \sum_{gm=1}^{GM} \sum_{i=1}^N \tilde{w}^{(gm),(i)} \\
& * w^{(gm)} = \tilde{w}^{(gm)} / \sum_{gm=1}^{GM} \tilde{w}^{(gm)}
\end{aligned}$$

– END FOR

- Inference (MMSE)

$$- \hat{x}_{k|k} = \sum_{gm=1}^{GM} w^{(gm)} \bar{x}_{k|k}^{x,(gm)}$$

- GMM-Recovery:

– For $g = 1, 2, \dots, G$ draw a number $j \in \{1, \dots, GM\}$ with probabilities proportional to $\{w^{(gm)}\}$, assign $\{\bar{x}_k^{(g)}, \mathbf{P}_k^{x,(g)}\} = \{\bar{x}_{k|k}^{(j)}, \mathbf{P}_{k|k}^{x,(j)}\}$ and set $\alpha_k^{(g)} = \frac{1}{G}$.

$$\text{GMM-Model: } p_{\mathcal{G}, \mathcal{M}, \mathcal{M}}(x_k | Y_k) = \sum_{g=1}^G \alpha_k^{(g)} \mathcal{N}(x_k; \bar{x}_k^{(g)}, \mathbf{P}_k^{x,(g)})$$

END FOR

Parameters:

- N : Number of particles
- G, M : Number of mixing components for state and process noise density

Gaussian Mixture Sigma Point Particle Filter - GMSPPF

Algorithm 30 - Gaussian Mixture Sigma Point Particle Filter - GMSPPF

$$[\{x_k^{(i)}, w_k^{(i)}\}_{i=1}^N] = \text{GMSPPF}[\{x_{k-1}^{(i)}, w_{k-1}^{(i)}\}_{i=1}^N, y_k]$$

Initialization:

$$\begin{aligned} \hat{x}_{0|0} &= \mathbb{E}[x_0] \\ x_0^{(i)} &\sim p(x_0) \quad i = 1, 2, \dots, N \\ p(x_0) &\approx \sum_{g=1}^G \alpha_0^g \mathcal{N}(x_0; \bar{x}_0^{(g)}, \mathbf{P}_0^{x, (g)}) \end{aligned}$$

FOR $k = 1, 2, \dots, K$:

- GMM-Models:

- Process noise density: $p_{\mathcal{GMM}}(\kappa_{k-1}) = \sum_{m=1}^M \beta_{k-1}^{(m)} \mathcal{N}(\kappa_{k-1}; \bar{\kappa}_{k-1}^{(m)}, \mathbf{P}_{k-1}^{\kappa, (m)})$
- Observation noise density: $p_{\mathcal{GMM}}(\gamma_k) = \sum_{j=1}^J \varphi_k^{(j)} \mathcal{N}(\gamma_k; \bar{\gamma}_k^{(j)}, \mathbf{P}_k^{\gamma, (j)})$

- Time Update

- For $j = 1, 2, \dots, J$, set $p(\gamma_k)^{(j)} = \mathcal{N}(\gamma_k; \bar{\gamma}_k^{(j)}, \mathbf{P}_k^{\gamma, (j)})$
For $m = 1, 2, \dots, M$, set $p(\kappa_{k-1})^{(m)} = \mathcal{N}(\kappa_{k-1}; \bar{\kappa}_{k-1}^{(m)}, \mathbf{P}_{k-1}^{\kappa, (m)})$
For $g = 1, 2, \dots, G$, set $p(x_{k-1}|Y_{k-1})^{(g)} = \mathcal{N}(x_{k-1}; \bar{x}_{k-1}^{(g)}, \mathbf{P}_{k-1}^{x, (g)})$
- For $g' = 1, 2, \dots, G'$ ($g' = g + (m-1)G$) use the time update step of a gaussian filter (employing the DSSM process equation $x_k = \mathcal{F}(x_{k-1}, u_{k-1}, \kappa_{k-1})$ and densities $p(x_{k-1}|Y_{k-1})^{(g)}$ and $p(\kappa_{k-1})^{(m)}$ from above) to calculate a gaussian approximate prior density $p(x_k|Y_{k-1})^{(g')} = \mathcal{N}(x_k; \bar{x}_{k|k-1}^{(g')}, \mathbf{P}_{k|k-1}^{x, (g')})$ and update the mixing weights $\alpha_{k|k-1}^{(g')} = \alpha_{k-1}^{(g)} \beta_{k-1}^{(m)} / \sum_{g=1}^G \sum_{m=1}^M \alpha_{k-1}^{(g)} \beta_{k-1}^{(m)}$.
- For $g'' = 1, 2, \dots, G''$ ($g'' = g' + (j-1)G'$) use the measurement update step of a gaussian filter (employing the DSSM measurement equation $y_k = \mathcal{H}(x_k, u_k, \gamma_k)$, the current observation y_k , and densities $p(x_k|Y_{k-1})^{(g')}$ and $p(\gamma_k)^{(j)}$ from above) to calculate a gaussian approximate posterior density $p(x_k|Y_k)^{(g'')} = \mathcal{N}(x_k; \bar{x}_{k|k}^{(g'')}, \mathbf{P}_{k|k}^{x, (g'')})$ and update the mixing weights $\alpha_{k|k}^{(g'')} = \alpha_{k|k-1}^{(g')} \varphi_k^{(j)} p(y_k|\bar{x}_{k|k-1}^{(g')}) / \sum_{g'=1}^{G'} \sum_{j=1}^J \alpha_{k|k-1}^{(g')} \varphi_k^{(j)} p(y_k|\bar{x}_{k|k-1}^{(g')})$.

- GMM-Models:

- Prior state density: $p_{\mathcal{GMM}}(x_k|Y_{k-1}) = \sum_{g'=1}^{G'} \alpha_{k|k-1}^{(g')} \mathcal{N}(x_k; \bar{x}_{k|k-1}^{(g')}, \mathbf{P}_{k|k-1}^{x, (g')})$
- Posterior state density: $p_{\mathcal{GMM}}(x_k|Y_k) = \sum_{g''=1}^{G''} \alpha_{k|k}^{(g'')} \mathcal{N}(x_k; \bar{x}_{k|k}^{(g'')}, \mathbf{P}_{k|k}^{x, (g'')})$

- FOR $i = 1, 2, \dots, N$

- $x_k^{(i)} \sim p_{\mathcal{GMM}}(x_k|Y_k)$
- $w_k^{(i)} = \frac{p(y_k|x_k^{(i)}) p_{\mathcal{GMM}}(x_k^{(i)}|Y_{k-1})}{p_{\mathcal{GMM}}(x_k^{(i)}|Y_k)}$

- END FOR

- $\tilde{w}_k = w_k / \sum_{i=1}^N w_k^{(i)}$

- GMM-Recovery:

- Option 1: Expectation-Maximization Recovery
 - * $[\{x_k^{(i)}, \frac{1}{N}, \text{Par}I^{(i)}\}_{i=1}^N] = \text{RESAMPLE}[\{x_k^{(i)}, \tilde{w}_k^{(i)}\}_{i=1}^N]$
 - * $[\{\alpha_k^{(g)}, \bar{x}_k^{(g)}, \mathbf{P}_k^{x, (g)}\}_{g=1}^G] = \text{EM}[\{x_k^{(i)}\}_{i=1}^N]$
- Option 2: Weighted Expectation-Maximization Recovery
 - * $[\{\alpha_k^{(g)}, \bar{x}_k^{(g)}, \mathbf{P}_k^{x, (g)}\}_{g=1}^G] = \text{WEM}[\{x_k^{(i)}, \tilde{w}_k^{(i)}\}_{i=1}^N]$

GMM-Model: $p_{\mathcal{GMM}}(x_k|Y_k) = \sum_{g=1}^G \alpha_k^{(g)} \mathcal{N}(x_k; \bar{x}_k^{(g)}, \mathbf{P}_k^{x, (g)})$

- Inference (MMSE)

- Option 1: Infer before the EM/WEM Step - $\hat{x}_{k|k} = \sum_{i=1}^N \tilde{w}_k^{(i)} x_k^{(i)}$
- Option 2: Infer after the EM/WEM Step - $\hat{x}_{k|k} = \sum_{g=1}^G \alpha_k^{(g)} \tilde{x}_k^{(g)}$

END FOR

Parameters:

- N : Number of particles
 - G, M, J : Number of mixing components for state, process noise and observation noise density
 - Depending on the choice of proposal (EKF, UKF, ...) the respective parameters accrue additionally
-
-

Kurzlebenslauf

Dipl.-Ing. Marc-André Beyer

Geburtsdatum und -ort 19. November 1979, Essen
Nationalität deutsch



Schulausbildung

08/1986 - 07/1990 Grundschole, Essen-Stoppenberg
08/1990 - 05/1999 Bischöfliches Gymnasium am Stoppenberg
Abitur (Durchschnittsnote: 2,0)

Wehrdienst

07/1999 - 06/2000 Bundeswehrdienst bei der 1. Gebirgsdivision

Berufsausbildung

10/2000 - 09/2005 Maschinenbaustudium an der Ruhr-Universität Bochum
Konstruktionstechnik und Produktautomatisierung
Grundstudium/Hauptstudium (Durchschnittsnoten: 1,6/1,0)
01/2004 - 05/2004 Auslandsstudium an der Texas A&M University

Stipendien und Preise

04/2003 - 09/2005 Stipendiat der Prof. Dr. Koepchen-Studienstiftung
01/2004 - 05/2004 Auslandsstipendium des DAAD
05/2006 Adam Opel Preis 2006

Praktika und Nebentätigkeiten

02/2002 - 10/2003 Fachpraktikum in vier Teilen bei der RWE Power AG
Bereich Regenerative Energien, Solar- und Windkraftwerke
01/2002 - 12/2003 Studentische Hilfskraft am Lehrstuhl für Maschinenelemente,
Getriebe und Kraftfahrzeuge an der Ruhr-Universität Bochum
07/2004 - 09/2005 Studentische Hilfskraft am Lehrstuhl für Regelungssysteme
und Steuerungstechnik an der Ruhr-Universität Bochum

Berufstätigkeit

Seit 10/2005 Wissenschaftlicher Mitarbeiter am Lehrstuhl für Regelungssysteme
und Steuerungstechnik an der Ruhr-Universität Bochum

Schriftenreihe des Lehrstuhls

In der Schriftenreihe des Lehrstuhls für Regelungssysteme und Steuerungstechnik der Ruhr-Universität Bochum (vormals Lehrstuhl für Meß- und Regelungstechnik) sind bisher folgende Titel erschienen:

- Heft 1 (1973) *Roman Krempl*: Anwendung von dreiwertigen pseudozufälligen Signalen zur Identifikation nichtlinearer Regelungssysteme.
- Heft 2 (1972) *Peter Vingron*: Ein Beitrag zur Theorie sequentieller Schaltungen (Das Transduktionsverfahren zur Synthese asynchroner Automaten).
- Heft 3 (1974) *Hanns P. Jörgl*: Untersuchung des Eigenschwingverhaltens von Ein- und Zweigrößen - Relais -Regelsystemen mit asymmetrischen Linearteilen.
- Heft 4 (1974) *Munir Sohrwardy*: Entwicklung einer Theorie der Einschnitt Prädiktivregelung.
- Heft 5 (1975) *Werner Hübl*: Ein Beitrag zur Synthese asynchroner Grund- und Standardschaltwerke.
- Heft 6 (1975) *Luigi Tuis*: Anwendung von mehrwertigen pseudozufälligen Signalen zur Identifikation von nichtlinearen Regelungssystemen.
- Heft 7 (1976) *Manfred Gronau*: Digitale Simulation eines Windverteilungssystems mit Pumpgrenz- und Mengenregelung für eine Hochofenanlage mit Hilfe der Simulationssprache ASIM 2.
- Heft 8 (1977) *Georg Grübel*: Reglersynthese durch Verknüpfung von Zustandsraum und Frequenzbereich.
- Heft 9 (1977) *Georg Grübel*: Beobachter zur Reglersynthese.
- Heft 10 (1977) *Werner Roddeck*: Mehrfacher Signalwechsel in Schaltnetzen und asynchronen Schaltwerken.
- Heft 11 (1978) *Willi Gründer*: Die Wirbelblende - ein neuartiges Element zur Erzeugung volumenstromproportionaler Oszillationen in Wirbelstraßen - Durchflussmessgeräten.
- Heft 12 (1978) *Jürgen Fricke*: Dekomposition mehrwertiger logischer Funktionen - Ein Beitrag zur Realisierung mehrwertiger digitaler Schaltungen.
- Heft 13 (1979) *Dieter Streppel*: Entwurf von Schaltnetzen und asynchronen Schaltwerken mit Hilfe der Partitionentheorie.
- Heft 14 (1980) *Anton B. Höfler*: Gradientenkettensoperatoren und ihre Anwendung bei der Reglerparameteroptimierung.
- Heft 15 (1981) *Bernd Gebhardt; Rainer Mückenhoff*: Systemidentifikation mittels Multifrequenz-Testsignalen.

- Heft 16 (1981) *Manfred Hoppe*: Die Regelung von Systemen mit Allpasseigenschaften - Dargestellt durch theoretische und experimentelle Untersuchung einer Wasserkraftanlage.
- Heft 17 (1982) *Johannes Blanke-Bohne*: Ein Beitrag zur rechnergestützten Analyse und Synthese von nichtlinearen Regelungssystemen.
- Heft 18 (1982) *Hans-Werner Müller*: Überlegungen zur digitalen Drehzahlregelung von Wasserturbinen.
- Heft 19 (1983) *Manfred Gronau*: Ein Beitrag zur theoretischen Modellbildung von Verdichteranlagen.
- Heft 20 (1983) *Manfred Hoppe; Seid Tesnjak*: Modellbildung und Simulation des dynamischen Verhaltens von Wasserkraftanlagen.
- Heft 21 (1984) *Ulrich Grebhofer*: Entwicklung eines vollautomatischen Konturfolgers zum Zwecke der Informationsumsetzung von Zeichnungsvorlagen in Ansteuerprogrammen für NC-Maschinen.
- Heft 22 (1984) *Raymond Figura*: Zur energieoptimalen Einsatzplanung parallel betriebener Turboverdichter nach mathematischen Verfahren.
- Heft 23 (1984) *Hans-Werner Müller*: Algorithmen für die digitale Drehzahlregelung von Wasserturbinen.
- Heft 24 (1985) *Martin Reike*: Die Regelung qualitätsbeeinflussender Wasserverteilnetzwerke.
- Heft 25 (1985) *Helmut Hehmann*: Ein Systemkonzept zur Prozesssteuerung mit Mikrorechnerkomponenten.
- Heft 26 (1985) *Elmar Heinemann*: Ein Beitrag zur berührungslosen Erfassung und Überwachung der kinematischen Größen eines bewegten Objektes.
- Heft 27 (1985) *Bernhard Gebhardt*: Simulation und rechnergestützter Reglerentwurf für eine Verdichteranlage.
- Heft 28 (1986) *Reinhardt Vonnoe*: Modellbildung und Simulation eines Fahrzeugantriebs mit stufenlos wirkendem hydrostatischen Lastschaltgetriebe.
- Heft 29 (1986) *Ralf Dreibold*: Reglersynthese für einen Fahrzeugantrieb mit stufenlos wirkendem hydrostatischen Lastschaltgetriebe.
- Heft 30 (1988) *Michael Pohl*: Digitale Regelung von Maschinensätzen einer Wasserkraftanlage mit robusten und gesteuert-parameteradaptiven Algorithmen.
- Heft 31 (1988) *Wilfried Blotenberg*: Ein Beitrag zur digitalen Pumpschutzregelung von Turbokompressoren.
- Heft 32 (1989) *Heinz-Günther Gehre*: Regelungstechnische Modellapproximation in Zeit- und Frequenzbereich.
- Heft 33 (1989) *Reinhard Grybowski*: Synthese von Regelungssystemen unter Nutzung von Entwurfsfreiheiten bei Eigenwertgebietvorgabe.

- Heft 34 (1990) *Gregor Krampe*: Automatische Erkennung von Prozessveränderungen durch rekursive Parameterschätzverfahren.
- Heft 35 (1991) *Ralf Sondermann*: Hard- und Softwareentwicklung für ein frei konfigurierbares Mehrprozessor-Turbinenleitgerät.
- Heft 36 (1991) *Peter Stammen*: Experimentelle Prozeßanalyse großer Systeme mit dem Gauß-Seidel Verfahren.
- Heft 37 (1991) *Joachim Barbe*: Ein System zur modellgestützten Führung und Überwachung kontinuierlicher industrieller Prozesse.
- Heft 38 (1992) *Liping Wang*: Berechnung von kontinuierlichen Modellen aus diskreten Identifikationsergebnissen.
- Heft 39 (1993) *Stefan Sommer*: Prädiktive und adaptive Regelung nichtlinearer sowie bilinearer Systeme anhand parametrischer Prozeßmodelle.
- Heft 40 (1993) *Roland Knof*: Ein Konzept für Entwurf und Optimierung von Fuzzy-Regelungssystemen.
- Heft 41 (1993) *Klaus Diekmann*: Prozeßüberwachung mit regelungstechnischen Verfahren anhand von dynamischen Modellen.
- Heft 42 (1994) *Karl Heinz Fasol*: Regelungstechnik - Von den Anfängen bis heute. Abschiedsvorlesung vom 29.1.1992.
- Heft 43 (1994) *Michael Montag*: Strategien zur optimierten digitalen Pumpschutzregelung für Verdichteranlagen.
- Heft 44 (1994) *Thomas Greim*: Simulation und Analyse von Steuerungssystemen auf der Basis von Petri-Netzen.
- Heft 45 (1995) *Peter Dellwig*: Entwicklung eines Systems zur Parallelsimulation komplexer Anlagenmodelle in heterogener Rechnerumgebung.
- Heft 46 (1996) *Uwe von Döllen*: Konzepte für die übergeordnete Prozeßführung in der Energiewirtschaft.
- Heft 47 (1999) *Michael Rempe*: Entwicklung und Verifikation von Steuerungssystemen auf der Basis hierarchischer Zustandsmodelle.
- Heft 48 (1999) *Michael Schlothane*: Objektorientierte Modellierung regelungstechnischer Daten.
- Heft 49 (1999) *Karl Heinz Fasol*: Zur Entwicklungsgeschichte der Regelung der Wasserturbinen.
- Heft 50 (2002) *Murat Ünlü*: Toolgestützte Validierungsmethode für verteilte eingebettete Steuerungssysteme im Passagierflugzeugbau.
- Heft 51 (2003) *Bernd Mahn*: Entwicklung einer Methode und Testumgebung für die Evaluierung von gehobenen Regelungsstrategien für verfahrenstechnische Anlagen.
- Heft 52 (2005) *Mirko Bebar*: Regelgütebewertung in kontinuierlichen verfahrenstechnischen Anlagen anhand vorliegender Messreihen.

- Heft 53 (2006) *Ulrich Schwenken*: Eine Methode zur Fehlerbewertung und zur adaptiven Motorleistungsbegrenzung auf der Basis einer modellbasierten Diagnose am Beispiel eines PKW-Kühlsystems.
- Heft 54 (2008) *Marc-André Beyer*: Methodische Evaluierung von wahrscheinlichkeitstheoretischen Inferenzverfahren für dynamische Systeme.