

TECHNISCHE UNIVERSITÄT DRESDEN

**Entwurf modellprädiktiver Regler in der Gebäudetechnik
auf der Basis datengetriebener Modelle**

Fabian Paschke

von der Fakultät Elektrotechnik und Informationstechnik
der Technischen Universität Dresden

zur Erlangung des akademischen Grades eines

Doktoringenieurs

(Dr.-Ing.)

genehmigte Dissertation

Vorsitzender: Univ.-Prof. Dr. techn. Klaus Janschek (TU Dresden)

Gutachter: Prof. Dr.-Ing. habil. Dipl.-Math. Klaus Röbenack (TU Dresden)
Prof. Dr.-Ing. Christian Bohn (TU Clausthal)

Tag der Einreichung: 17. Dezember 2021

Tag der Verteidigung: 18. Juli 2022

Copyright © Fabian Paschke 2023
Alle Rechte vorbehalten.

Paschke, Fabian: *Entwurf modellprädiktiver Regler in der Gebäudetechnik auf der Basis datengetriebener Modelle*

Gesetzt vom Author mit L^AT_EX.
Dokument zuletzt erstellt am: 5. Februar 2023, 14:25

Vorwort

Die Erstellung dieser Arbeit war von freudigen Phasen der Erkenntnis, aber auch von Momenten der Ernüchterung geprägt. Während dieser Zeit erhielt ich die Hilfe vieler Kollegen und Personen, die mich in unterschiedlicher Weise unterstützt haben:

An erster Stelle danke ich ganz besonders Herrn Prof. Dr.-Ing. habil. Dipl.-Math. Klaus Röbenack, dem Leiter des Instituts für Regelungs- und Steuerungstheorie der TU Dresden, für die Betreuung und Unterstützung dieser Arbeit. Nebst seiner wertvollen Anregungen bin ich ihm vor allem auch für seine Vorlesungen während des Studiums dankbar, welche mein Interesse an der Regelungstheorie geweckt und somit diese Arbeit erst möglich gemacht haben. Recht herzlich möchte ich mich auch bei Prof. Dr.-Ing. Christian Bohn der TU Clausthal für die Übernahme des Zweitgutachtens bedanken.

Die vorliegende Arbeit entstand während meiner Tätigkeit als wissenschaftlicher Mitarbeiter am Institutsteil Entwicklung Adaptiver Systeme des Fraunhofer Instituts für Integrierte Schaltungen in Dresden. Die offene Arbeitsatmosphäre am Institut und die hervorragenden Arbeitsbedingungen haben die Grundlage für die Entstehung dieser Arbeit geschaffen. Persönlich möchte ich mich hierbei insbesondere bei meinem Gruppen- und Abteilungsleiter Dr.-Ing. Jan Bräunig bzw. Dr.-Ing. Dirk Mayer für die gewährten Freiräume und das entgegengebrachte Vertrauen bedanken. Weiterer Dank gebührt zudem Prof. Dr.-Ing. Matthias Franke, welcher mich zum Anfertigen dieser Arbeit ermutigt hat, sowie Prof. Dr.-Ing. Tobias Zaiczek, welcher trotz seines enormen Einsatzes in der täglichen Projektarbeit stets ein offenes Ohr für fachliche Probleme hatte. Weiterhin bedanke ich mich bei Dr. Jürgen Haufe, welcher durch sein Engagement bei der Projektakquise die finanzielle Grundlage dieser Arbeit geschaffen hat. Den weiteren Mitgliedern meiner Arbeitsgruppe und allen anderen Kollegen danke ich für den fachlichen Austausch und die angenehme Zusammenarbeit im Institut.

Ein besonderer Dank gilt weiterhin meinen lieben Freunden, die durch vielfältige Aktivitäten für den notwendigen Ausgleich auf dem oft anstrengenden Weg zur Promotion gesorgt haben. Persönlich bedanken möchte ich mich vor allem auch bei meiner Schwester und engen Freunden, denen ich die kritische Durchsicht der Arbeit und die Korrektur von Rechtschreibfehlern zu verdanken habe.

Mein wichtigster Rückhalt war jedoch stets meine geliebte Familie. Den Dank für ihr entgegengebrachtes Verständnis, ihren Beistand sowie die großzügige Unterstützung während der Schul- und Studienzeit kann ich nicht annähernd in Worte fassen!

Dresden, im November 2021

Fabian Paschke

Kurzzusammenfassung

Die vorliegende Arbeit befasst sich mit datengetriebenen Modellierungsansätzen in der Gebäudetechnik zum Zweck des Entwurfs modellprädiktiver Regler. Dabei widmet sich der erste Teil (Kapitel 2 und 3) im Wesentlichen den theoretischen Grundlagen der Systemidentifikation wohingegen im zweiten Teil (Kapitel 4 bis 6) praktische Anwendungen der Identifikation sowie die Implementierung modellprädiktiver Regler im Gebäudebereich im Vordergrund stehen.

Die Grundlage der Arbeit stellt die Systemidentifikation mittels der Prädiktionsfehlerminimierung (Prediction Error Minimization – PEM) dar, deren Basis stochastische Modelle dynamischer Systeme bilden. Daher wird ausgehend von den unterschiedlichen Darstellungsformen stochastischer linearer zeitinvarianter Systeme (LTI-Systeme) die Problematik der optimalen Prädiktion thematisiert, wobei ein allgemeiner Ausdruck zur Berechnung der Mehrschrittprädiktion von LTI-Systemen hergeleitet wird. Anschließend werden Implementierungsaspekte von Mehr- bzw. Multischritt-PEM-Verfahren diskutiert und der Schätzer im Frequenzbereich analysiert, wobei sich verglichen mit der 1-Schritt-PEM eine erwartungsgemäß andere Wichtung des geschätzten Modells im Frequenzbereich ergibt. Die aus diesen Untersuchungen gewonnenen Identifizierbarkeitsbedingungen entsprechen im offenen Regelkreis den bekannten Bedingungen der 1-Schritt-PEM, wohingegen sich im geschlossenen Regelkreis eine restriktivere Bedingung ergibt. Die weiterhin durchgeführte Genauigkeitsanalyse führt letztlich zu einem Ausdruck für die Abschätzung der Schätzfehlerkovarianzmatrix der Modellparameter, welche wiederum für die Varianz- bzw. Konfidenzschätzung anderer Systemmerkmale genutzt werden kann.

In der Gebäudetechnik spielen nichtlineare Modellansätze eine wichtige Rolle. Für diese ist die explizite Berechnung der optimalen Prädiktion, und damit eine Identifikation mittels PEM, in der Regel problematisch bzw. nicht möglich. In der Arbeit werden daher zwei nichtlineare Modellstrukturen vorgeschlagen, für welche entsprechende Ausdrücke angegeben werden können, sodass eine direkte Anwendung der PEM ermöglicht wird. Dabei hat sich im Anwendungsteil der Arbeit herausgestellt, dass einer dieser Ansätze sinnvoll in der Gebäudetechnik genutzt werden kann. Aufgrund der Ähnlichkeit zum linearen Fall können zudem die Ideen, die typischerweise für die Parametrierung stochastischer LTI Systeme genutzt werden, auf den nichtlinearen Fall übertragen werden.

Im Anwendungsteil der Arbeit wird dann die Nutzung der PEM im Gebäudebereich anhand praktisch aufgenommenen Messdaten einiger Demonstratoren veranschaulicht. Es werden Modell- und Identifikationsansätze für die Schätzung von thermischen Raum-

modellen sowie einer Gastherme und einer Wärmepumpe vorgeschlagen, wobei plausible Identifikationsergebnisse erzielt werden konnten. Weiterhin wird im Rahmen der Identifikation von thermischen Raummodellen auch auf die wichtige Problematik nicht messbarer Störungen (z. B. Raumbelegung sowie Wärmeströme aus Nachbarräumen) eingegangen und mögliche Lösungsansätze aufgezeigt. Aufgrund der umfangreichen Sensorausstattung einiger Demonstratoren konnte zudem ein direkter Vergleich eines LTI-Modells mit dem obig genannten nichtlinearen Modellansatz durchgeführt werden, wobei eine vergleichbare Prädiktionsgüte beider Modellierungsansätze festgestellt wurde.

Im letzten Teil wird dann der Entwurf modellprädiktiver Regler anhand zweier typischer Beispiele aus der Gebäudetechnik veranschaulicht. Für die Demonstration des Vorgehens werden Regelstreckenmodelle mittels der MODELICA BUILDINGS Bibliothek entworfen und dann für die Generierung synthetischer Daten verwendet, welche zur Identifikation von Modellen für den MPC Entwurf verwendet werden. Die identifizierten Modelle sind in beiden Fällen in der Lage das dynamische Verhalten des MODELICA Prozessmodells hinreichend gut vorherzusagen, was die Sinnhaftigkeit des gewählten Modellierungs- bzw. Identifikationsansatzes untermauert. Aufbauend wird dann für beide Beispiele auf die Formulierung einer prädiktiven Regelungsstrategie eingegangen, wobei spezielle Modelleigenschaften ausgenutzt werden. Für beide Beispiele konnten simulativ Vorteile gegenüber einer konventionellen Regelungsstrategie nachgewiesen werden.

Abstract

This thesis deals with data-driven modeling approaches in building systems engineering for the purpose of model predictive control design. The first part is dedicated to theoretical aspects of system identification (chapter 2 and 3), whereas in the second part practical applications of identification and model predictive control for building systems are of interest (chapter 4 to 6).

The basis of this work is system identification by means of prediction error minimization (PEM). Since PEM is based on stochastic models of dynamic systems, the concept of optimal prediction is introduced first, resulting in a general expression for the derivation of multiple-step predictions for linear time-invariant (LTI) systems. Subsequently, implementation aspects of multiple- and multi-step PEM methods are discussed. Compared to well known 1-step PEM the frequency domain analysis of the estimator reveals a different frequency weighting of the estimated model. Based on this analysis identifiability conditions of multi-step PEM are obtained, which correspond to the well known conditions for 1-step PEM for open loop identification experiments, whereas for closed loop identification a more restrictive condition is obtained. Ultimately, an expression for the estimation of error covariance matrix of the model parameters is derived, which in turn can be used for variance or confidence estimation of other system features.

Nonlinear modeling approaches are of great importance for building energy systems. The explicit calculation of the optimal prediction, and thus an identification by means of PEM, is usually problematic or even impossible for nonlinear systems. Thus, two model structures are proposed for which corresponding expressions can be derived such that a direct application of the PEM method is possible. In the application part of the thesis it turned out that one of these approaches can be applied for modeling building energy systems. Moreover, due to the similarity to the linear case, the approaches typically used for the parameterization of stochastic LTI systems can be applied in the nonlinear case as well.

In the application part of the thesis system identification in the building sector by means of PEM is illustrated with recorded measurement data of practical demonstrators. Therefore, modeling and identification approaches for the estimation of thermal room models as well as a gas boiler and a heat pump are proposed. The performed identification experiments showed that reasonable identification results can be obtained. Furthermore, the important issue of non-measurable disturbances (e. g. occupancy and heat flows from neighbouring rooms) is discussed in the context of identification and possible solutions are shown. Due to the good sensor equipment of some demonstrators a direct comparison of an LTI model with the above mentioned nonlinear model structure was

possible. As an outcome a comparable prediction quality was found for both modeling approaches.

In the last part the design of model predictive controllers is illustrated for two typical examples from the domain of building automation. To demonstrate the approach, process models are designed using the `MODELICA BUILDINGS` library, which are then utilized to generate synthetic data for identification of reduced order models used for MPC design. The identified models are able to predict the dynamic behaviour of the `MODELICA` process model sufficiently well, which supports the reasonableness of the chosen modeling and identification approach. Taking advantage of the identified model properties the design of a predictive control law is demonstrated. Advantages of MPC compared to a conventional control strategy could be simulatively proven in both cases.

Inhaltsverzeichnis

1	Einführung	1
1.1	Motivation der Arbeit	1
1.2	Aufbau der Arbeit und eigene Beiträge	2
2	Modelle dynamischer Systeme	5
2.1	Klassifikation von Systemmodellen	6
2.2	Zeitdiskrete lineare zeitinvariante Systeme	7
2.2.1	Differenzgleichung und Übertragungsfunktion	7
2.2.2	Impulsantwort	10
2.2.3	Zustandsraumdarstellung	11
2.3	Optimale Prädiktion zeitdiskreter LTI-Systeme	14
2.4	Parametrierung zeitdiskreter LTI-Systeme	18
2.4.1	Spezielle Parametrierungen von LTI-Systemen	20
2.4.2	Multivariable Modellstrukturen	22
2.5	Nichtlineare Modellstrukturen	22
3	Identifikation dynamischer Systeme	29
3.1	Klassifikation von Identifikationsverfahren für LTI-Systeme	29
3.1.1	Nichtparametrische Identifikationsverfahren	30
3.1.2	Parametrische Identifikationsverfahren	32
3.2	Prädiktionsfehlerminimierung	33
3.2.1	Implementierungsaspekte von PEM-Verfahren	35
3.2.2	Analyse der Mehrschritt-PEM	39
3.3	Methode der kleinsten Fehlerquadrate	48
3.3.1	Einschritt-PEM für ARX-Modellstrukturen	48
3.3.2	Hilfsvariablenverfahren	50
3.3.3	Pseudolineare-Regression	51
3.4	PEM-Verfahren für nichtlineare Modellstrukturen	52
4	Regelungstechnik in der Gebäudeautomation	57
4.1	Funktionsprinzipien und Regelungsstrategien in der Gebäudetechnik	57
4.2	Modellprädiktive Regelung in der Gebäudetechnik	62
4.2.1	Einführung	62
4.2.2	Literaturüberblick	63

5	Anwendung der Systemidentifikation in der Gebäudetechnik	67
5.1	Thermische Raummodelle	70
5.1.1	Anwendungsbeispiel 1: Besprechungsraum	74
5.1.2	Anwendungsbeispiel 2: Büroraum mit Fußbodenheizung	84
5.1.3	Anwendungsbeispiel 3: Turnhalle mit Deckenheizsystem	86
5.1.4	Anwendungsbeispiel 4: Hörsaalgebäude	90
5.1.5	Zusammenfassung	91
5.2	Wärmeerzeuger	93
5.2.1	Heizkessel	93
5.2.2	Wärmepumpe	96
6	Beispielimplementierungen modellprädiktiver Regler	103
6.1	Modellprädiktiver Einzelraumregler	104
6.1.1	Prozessmodell	105
6.1.2	Identifikation	106
6.1.3	Reglerentwurf	108
6.1.4	Simulationsergebnisse	112
6.2	Modellprädiktiver Mehrzonenregler	115
6.2.1	Prozessmodell	116
6.2.2	Identifikation	118
6.2.3	Reglerentwurf	119
6.2.4	Simulationsergebnisse	122
7	Zusammenfassung und Ausblick	129
7.1	Zusammenfassung	129
7.2	Ausblick und offene Fragen	130
A	Begriffe und Definitionen	133
A.1	Diskrete Fourier-Transformation	133
A.2	Begriffe der Stochastik	133
A.2.1	Unabhängigkeit und Unkorreliertheit stochastischer Variablen	134
A.2.2	Konvergenz stochastischer Variablen	134
A.2.3	Stochastische Prozesse	134
A.2.4	Begriffe der Schätztheorie	138
A.3	Lemmas	140
B	Ergänzungen zur PEM	143
B.1	Gradient des Einzschritt Prädiktionsfehlers	143
B.2	Schätzung von Anfangswerten	144
B.2.1	Methode der kleinsten Fehlerquadrate	144
B.2.2	Backforecasting	146
B.3	Herleitung des Autoleistungsdichtespektrums der Mehrschritt-PEM	147
B.4	Berechnung der Schätzfehlerverteilung	148

C	Thermodynamische Zusammenhänge in der Gebäudetechnik	151
C.1	Statische Wärmeabgabe von Heizflächen	151
C.2	Dichte feuchter Luft	152
D	Software	155
D.1	Kurzbeschreibung	155
D.2	Beispiele	158
D.2.1	LTI-System zweiter Ordnung	158
D.2.2	Bilineares System mit periodischer Störung	159
D.2.3	Nichtlineares MIMO System	160
	Abkürzungsverzeichnis	163
	Symbolverzeichnis	165
	Literatur	167
	Abbildungsverzeichnis	185
	Tabellenverzeichnis	187

Kapitel 1

Einführung

1.1 Motivation der Arbeit

Im Jahr 2030 soll in Deutschland entsprechend des 2021 novellierten Bundes-Klimaschutzgesetzes der Ausstoß von Treibhausgasen um 65% gegenüber 1990 reduziert und bis 2045 Treibhausgasneutralität erreicht werden. Da etwa 15% der Emissionen [BMU21] und $\approx 33\%$ des Endenergiebedarfs¹ [BMWi20] auf den Gebäudesektor entfallen, muss bis 2030 im Gebäudebereich die Treibhausgasemission auf etwa ein Drittel gegenüber 1990 und um weitere 45% gegenüber 2020 reduziert werden.

Bei der Realisierung der Einsparungen spielen nicht nur konstruktive Maßnahmen und moderne Anlagentechnik eine wichtige Rolle, sondern auch leistungsfähige Steuer- und Regelungskonzepte. Dabei werden in der Praxis heutzutage überwiegend mittels Erfahrungswissen parametrisierte PI(D)-Regler in Kombination mit regelbasierten Ansätzen verwendet (siehe Abschnitt 4.1). Insbesondere die auch zukünftig stark zunehmende Nutzung regenerativer Quellen, die Verwendung von Speichertechnologien und die Vernetzung der Energiesysteme auf lokaler Ebene erfordern jedoch intelligente und vorausschauende Steuer- und Regelungsansätze um deren volles Potenzial auszuschöpfen.

Seit etwa 10 Jahren wird in der Wissenschaft daher verstärkt der Einsatz höherer Regelungsstrategien, u. a. der modellprädiktive Regelung (Model predictive control – MPC), im Gebäudebereich untersucht. Obwohl deren Nutzen bereits an zahlreichen Beispielen simulativ nachgewiesen werden konnte, ist es bemerkenswert, dass praktische Anwendungen relativ selten sind, und sich fast ausschließlich auf Demonstratoren im universitären Umfeld beschränken (siehe Abschnitt 4.2). Der Technologie- und Wissenstransfer in den Markt befindet sich daher noch in einem sehr frühen Stadium, sodass das Forschungsfeld nach wie vor als Gegenstand der insbesondere anwendungsorientierten Forschung betrachtet werden kann.

Eine wesentliche Problematik bei der Adaption fortschrittlicher Regelungskonzepte im Gebäudebereich ist die anspruchsvolle Skalierbarkeit, welche durch die enorme Individualität der Gebäude- und Anlagenkonfiguration entsteht. Aus rein ökonomischer Sicht lässt sich der Einsatz von MPC in der industriellen Praxis heutzutage daher nur selten rechtfertigen [SGM16]. Dies lässt sich zum einen durch den deutlich höheren Entwurfs-, Test- und Investitionsaufwand, welchem ein hoher Kostendruck sowie i. d. R. wenig

¹Nach Energiewandlungs- und Übertragungsverlusten übrig gebliebener Teil der Primärenergie.

Erfahrung der Ingenieure im Bereich moderner modellbasierter Regelungsverfahren gegenübersteht, begründen. Des Weiteren sei erwähnt, dass bei der praktischen Implementierung, insbesondere im Bestand, weitere Hürden wie bspw. eine nicht ausreichende oder unzuverlässige Sensorik, existieren.

Demgegenüber muss allerdings auch berücksichtigt werden, dass vermutlich die wirtschaftliche Attraktivität MPC-basierter Regelungsansätze in der Gebäudetechnik zukünftig zunehmen wird. Dies lässt sich zum einen durch stetig steigende Energiekosten sowie deren zunehmende Volatilität, die durch MPC ausgenutzt werden kann, begründen. Auf der anderen Seite ist zu erwarten, dass die anfänglichen Investitionskosten aufgrund fallender Preise für die zusätzlich notwendige Sensorik und Rechentechnik weiter abnehmen werden. Zuletzt sei noch erwähnt, dass auch der vermehrte Einsatz von Energiespeichern eine steigende Notwendigkeit nach intelligenten Regelungsansätzen schaffen wird.

Die größte Unsicherheit bei der Kosten-Nutzen-Analyse liegt beim Entwurfs- und Testaufwand für den modellprädiktiven Regler. Die Verfügbarkeit spezialisierter und flexibler Softwarewerkzeuge, welche die Ingenieure bei der Modellierung und Validierung unterstützen, wird daher als entscheidender Faktor für die Wirtschaftlichkeit und die industrielle Akzeptanz modellprädiktiver Regelungskonzepte im Gebäudebereich betrachtet [GG13; SGM16; SFC18].

Als besonders zeitaufwendiger und kritischer Aspekt bei der Implementierung wird die Modellierung der relevanten Systemkomponenten betrachtet [Pri13; PCZ13; SGM16; SFC18]. Häufig wird dabei der Ansatz einer physikalisch motivierten Modellbildung mittels Wärme- bzw. Massenbilanzen verfolgt, welche mittels der Konstruktions- bzw. Anlagendaten des Gebäudes parametrisiert werden [OPJ12; SGS14; SGM16]. Aufgrund der Verfügbarkeit von Laufzeitdaten aus dem Gebäudebetrieb, sind in der Praxis jedoch ebenfalls datenbasierte Modellierungsansätze naheliegend. Eine wesentliche Motivation bei der Anfertigung dieser Arbeit bestand daher darin, die Anwendung der Methoden der Systemidentifikation im Gebäudebereich zu untersuchen und damit den Einsatz modellprädiktiver Regelungskonzepte in der Praxis voranzutreiben.

1.2 Aufbau der Arbeit und eigene Beiträge

Die vorliegende Arbeit ist in einen theoretischen und in einen praktischen Anwendungsteil gegliedert. Im theoretischen Teil (Kapitel 2 und 3) werden wichtige mathematische Grundlagen der Systemidentifikation vermittelt. In den übrigen Kapiteln wird dann die Verwendung dieser Methoden im Gebäudebereich demonstriert.

Das Kapitel 2 befasst sich mit der Modellierung stochastischer zeitdiskreter dynamischer Systeme. Dabei wird u. a. auf die verschiedenen Darstellungsformen und Parametrisierungen sowie deren theoretische Zusammenhänge eingegangen und weiterhin das Konzept der optimalen Prädiktion eingeführt. Zuletzt werden einige in der Literatur verbreitete nichtlineare Modellansätze genannt und eine nach Kenntnis des Autors neue Modellstruktur vorgeschlagen, welche insbesondere im Praxisteil der Arbeit Anwendung

findet.

Das nachfolgende Kapitel widmet sich der Problematik der Identifikation dynamischer Systeme. Dabei wird zunächst auf die Klassifikation verbreiteter Schätzmethoden eingegangen. Die übrigen Abschnitte beschäftigen sich dann mit theoretischen Merkmalen und Implementierungsaspekten der *Prädiktionsfehlerminimierung*. Hierbei werden die wesentlichen Ergebnisse aus der Literatur zusammengetragen und auf den Fall der Mehrschritt bzw. Multischritt-PEM erweitert. Weiterhin wird u. a. ein Ausdruck zur Berechnung der Schätzgenauigkeit der Mehrschritt-PEM hergeleitet. Die wesentlichen Ergebnisse des Abschnittes werden zudem anhand einiger theoretischer Beispiele veranschaulicht und diskutiert.

Das Kapitel 4 dient als Einleitung in den Anwendungsteil der Arbeit. Dabei soll dem Leser ein kurzer Einblick in die wesentlichen Funktionsprinzipien der Gebäudetechnik vermittelt werden. Des Weiteren werden mögliche Anwendungsfälle der modellprädiktiven Regelung in der Gebäudetechnik genannt und klassifiziert, sowie die entsprechenden Referenzen aus der wissenschaftlichen Literatur zugeordnet.

Im Kapitel 5 wird die Anwendung der Systemidentifikation in der Gebäudetechnik anhand praktisch aufgenommener Messdaten demonstriert. Dabei werden konkrete Modellstrukturen und Identifikationsansätze für ausgewählte Beispiele vorgeschlagen und es wird auf typische in der Praxis auftretende Probleme, insbesondere auf nicht messbare Störgrößen (Wärmeflüsse aus Nachbarräumen, interne Gewinne durch Gebäude- bzw. Raumnutzung), eingegangen.

Das letzte Kapitel widmet sich der Implementierung modellprädiktiver Regelungsansätze für zwei konkrete Anwendungsbeispiele. Ersteres veranschaulicht den Entwurf eines Einzelraumreglers für einen Raum mit Fußbodenheizung, wohingegen zweiteres die Realisierung eines Mehrzonenreglers für ein Hörsaalgebäude mit einer zentralen Lüftungsanlage demonstriert. Die dabei verwendeten Regelstreckenmodelle wurden mit Hilfe der Modellierungssprache MODELICA unter Verwendung der BUILDINGS Bibliothek entworfen. Diese werden zum einen zur Generierung „synthetischer“ Messdaten verwendet, welche für die Identifikation der für den MPC-Entwurf notwendigen Modelle eingesetzt werden. Zum anderen werden die MODELICA Modelle als Regelstreckenmodell über die FMI-Schnittstelle in die Simulationsumgebung MATLAB/SIMULINK integriert und für die Validierung und den Vergleich des MPC-Reglers mit einer konventionellen Regelungsstrategie verwendet.

Das abschließende Kapitel 7 fasst die Ergebnisse und den Inhalt der Arbeit zusammen und geht auf mögliche weiterführende Fragestellungen ein.

Software

Im Rahmen der Arbeit ist eine objektorientierte Softwarebibliothek zur Systemidentifikation entstanden, welche in der Programmiersprache MATLAB implementiert wurde. Deren Funktionsumfang überschneidet sich mit Teilen der System Identification Toolbox, insbesondere mit der Funktion `pem()`. Zusätzlich sind jedoch weitere nützliche Eigenschaften implementiert worden:

- PEM-Identifikation von Hammerstein- und nichtlinearen Modellen des Typs (2.51)
- Schätzung (linear) faktorisierter Modelle (siehe Bemerkung 2.9)
- PEM-Identifikation mittels Mehrschritt- bzw. Multischrittkriterium (LTI- und Hammerstein-Modelle)
- Berechnung des Gradienten (6.23) zum Zwecke der Implementierung nichtlinearer modellprädiktiver Regler, die durch Modelle des Typs (2.51) beschrieben werden können

Eine kurze anwendungsorientierte Einführung in die Nutzung der Softwarebibliothek kann im Anhang D nachgeschlagen werden.

Die Beispiele in der Arbeit basieren auf synthetisch generierten oder auf real gemessenen Daten, weswegen ein klassisches Nachrechnen der Ergebnisse nicht möglich ist. Um dennoch ein möglichst hohes Maß an Nachvollziehbarkeit zu erreichen, wurde der Quellcode und die Daten zum Erzeugen der Ergebnisse digital unter

`https://github.com/fpaschke/phd_identification_mpc.git`

veröffentlicht.

Kapitel 2

Modelle dynamischer Systeme

Die Basis für den Entwurf und die Analyse von Regelungs- und Steuerungssystemen sind Modelle, welche das Verhalten eines betrachteten Systems bzw. Prozesses hinreichend gut beschreiben können. Hierbei wird der Begriff „System“ bzw. „Prozess“ als Ausdruck für die in der Wirklichkeit ablaufenden physikalischen, chemischen oder ökonomischen Vorgänge verwendet¹. Aus der Sicht der Identifikation stellt das System experimentelle Daten $\{\mathbf{u}[t], \mathbf{y}[t]\}$ zur Verfügung, die als Ein- bzw. Ausgangssignale bezeichnet werden, und mathematisch als Realisierungen stochastischer Prozesse aufgefasst werden können (siehe Anhang A.2.3). Als ein Modell versteht man dann ein mehr oder weniger stark vereinfachtes Abbild des Systems in graphischer oder tabellarischer Form oder in der Gestalt mathematischer Gleichungen.

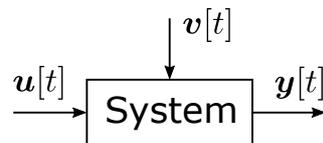


Abbildung 2.1: Veranschaulichung eines Systems mit Ein- und Ausgangssignalen $\mathbf{u}[t]$ bzw. $\mathbf{y}[t]$ und einer auf das System einwirkenden Störung $\mathbf{v}[t]$

Modelle sind die Basis für die Lösung vieler technisch relevanter Problemstellungen. Sie sind oftmals Voraussetzung zur Entwicklung eines tieferen Systemverständnisses und sind zudem die Grundlage für den Entwurf von Regelungs- und Steuerungseinrichtungen [Lud95; Rei06; Lun16a] sowie modellbasierter Fehlerdiagnose- und Überwachungssysteme [CP98; Ise06]. Eine der Kernaufgaben des Ingenieurs besteht daher darin, ein geeignetes Systemmodell aufzustellen. Dabei wird in der Literatur oftmals zwischen drei unterschiedlichen Vorgehensweisen unterschieden [Nel01; Buc10; Lju12]:

- **White-Box-Modellierung:** Die Modellgleichungen werden vollständig durch die ablaufenden physikalischen, chemischen oder ökonomischen Gesetzmäßigkeiten festgelegt. Die innerhalb des Modells verwendeten Parameter, welche mithilfe theoretischer Überlegungen oder durch Identifikation aus Messdaten gewonnen werden, besitzen eine direkte Interpretation in der jeweiligen physikalischen Domäne.

¹Für eine formale Definition des Begriffes „(technischer) Prozess“ sei auf die DIN66201 und die DIN IEC 60050-351 verwiesen.

- **Black-Box-Modellierung:** Sowohl die meist nicht interpretierbaren Parameter, als auch die Modellstruktur werden ausschließlich mittels der Messwerte der Eingangs- und Ausgangsgrößen des Systems bestimmt. Dabei wird kein oder nur sehr wenig Vorwissen über die ablaufenden Vorgänge verwendet.
- **Gray-Box-Modellierung:** Gray-Box-Modelle stellen einen Kompromiss der oben genannten Modellierungsansätze dar. Für die Modellbildung werden sowohl Messdaten als auch verfügbare Informationen über die theoretischen Eigenschaften des Systems herangezogen. Hierzu gehören u. a. Kenntnisse in Form systemtheoretischer Eigenschaften wie bspw. die Linearität oder Stabilität des Übertragungsverhaltens.

2.1 Klassifikation von Systemmodellen

Modelle dynamischer Systeme können auf der Basis ihrer mathematischen Eigenschaften unterschieden werden. Dabei können unter anderem folgende Merkmale zur Klassifikation herangezogen werden:

- **Linearität:** Ein Model ist linear, wenn die Superpositionseigenschaft gilt (siehe z. B. [Kai80, Abschnitt 1.1]). Man beachte, dass sich der Begriff auf das Eingangs-Ausgangs-Verhalten des Modells bezieht und nicht mit der in der Systemidentifikation wichtigen Eigenschaft der Parameterlinearität verwechselt werden darf.
- **Ein- bzw. Mehrgrößensysteme (SISO-/MIMO-System):** Bezieht sich auf die betrachtete Anzahl der Ein- bzw. Ausgangsgrößen des Systems.
- **Parametrische bzw. nichtparametrische Modelle:** Parametrische Modelle werden durch eine endlichen Anzahl von Parametern beschrieben [Nel01, Abschn. 16.1] und sind die in der modernen Systemidentifikation verbreitete Beschreibungsform. Da die meisten Anwendungen auf parametrischen Modellformen aufbauen, und nichtparametrische Modelle (bspw. Impulsantworten oder Frequenzgänge) aus parametrischen Modellen berechnet werden können, werden im Rahmen dieser Arbeit ausschließlich parametrische Modelle verwendet. Für eine Einführung in die nichtparametrische Identifikation sei auf [Ise92a; SS89; Lju12] verwiesen.
- **Deterministische bzw. stochastische Modelle:** Bei der deterministischen Modellierung wird angenommen, dass bei bekannten Eingangsgrößen und Anfangswerten der Modellausgang exakt vorhergesagt werden kann. Im Gegensatz dazu geht die stochastische Betrachtungsweise davon aus, dass das System zusätzlich von nicht vorhersagbaren Störungen beeinflusst wird, welche mathematisch durch Zufallsgrößen mit definierten Eigenschaften erfasst werden. Aus mathematischer Sicht kann das System daher als stochastischer Prozess, und die gemessenen Eingangs- und Ausgangssignale als dessen Realisierungen betrachtet werden. Es sollte jedoch beachtet werden, dass der Begriff „stochastisches Modell“ ebenfalls für gemischt deterministisch-stochastische Modellformen verwendet wird (siehe Abschn. 2.2.1).

- **Zeitkontinuierliche bzw. zeitdiskrete Modelle:** Ein zeitdiskretes Modell² beschreibt die Beziehung zwischen den Ein- und den Ausgangsgrößen des Systems an diskreten Zeitpunkten $t \in \mathbb{Z}$, wohingegen zeitkontinuierliche Modelle von einer reellen Zeitachse $t \in \mathbb{R}$ oder einem Zeitintervall $[0, \infty)$ ausgehen. Da der Ausgangspunkt für die Identifikation meist an definierten Zeitpunkten aufgezeichnete Messdaten sind und in vielen Anwendungsfällen ohnehin von zeitdiskreten Modellen ausgegangen wird, werden nachfolgend ausschließlich zeitdiskrete Modelle mit einer äquidistanten Abtastschrittweite thematisiert. Es sei weiterhin erwähnt, dass die Analyse zeitkontinuierlicher stochastischer Systeme weiterführende Konzepte der stochastischen Analysis erfordert. Für eine kurze Einführung sei auf [Jaz70; Ast70] verwiesen.

2.2 Zeitdiskrete lineare zeitinvariante Systeme

Lineare zeitinvariante Systeme (LTI-Systeme) besitzen sowohl in der Theorie als auch in der praktischen Anwendung eine herausragende Rolle. Dies liegt zum einen daran, dass die Annäherung des typischerweise nichtlinearen Systemverhaltens durch ein LTI-System aus praktischer Sicht in vielen Fällen gerechtfertigt werden kann. Zum anderen existiert eine umfangreiche Literatur, wobei insbesondere die im Bereich der Identifikation dynamischer Systeme üblichen Standardwerke [Ast70; SS89; Ise92a; OM96; BJR08; Lju12] hervorgehoben seien.

In den folgenden Abschnitten soll auf die mathematische Beschreibung deterministisch und stochastisch erregter LTI-Systeme eingegangen werden.

2.2.1 Differenzgleichung und Übertragungsfunktion

Ein zeitdiskretes LTI-System lässt sich allgemein durch ein System linearer Differenzgleichungen beschreiben. Diese ergeben sich in natürlicher Weise aus gewöhnlichen linearen Differentialgleichungen, indem die darin vorkommenden Differentialquotienten durch entsprechende Näherungen³ ersetzt werden (zum Beispiel durch Vorwärts- oder Rückwärtsdifferenzenquotienten). Im Allgemeinen ergibt sich dadurch ein lineares Differenzgleichungssystem der Form

$$\sum_{i=0}^{n_A} \mathbf{a}^{[i]} \mathbf{y}[t - i] = \sum_{i=0}^{n_B} \mathbf{b}^{[i]} \mathbf{u}[t - i], \quad (2.1)$$

²In einigen Referenzen wird auch von diskontinuierlichen, Abtast- oder Impulssystemen gesprochen [Gue86, Kap. 0].

³Die Transformation von zeitkontinuierlichen in zeitdiskrete Systeme muss nicht unbedingt als eine Näherung verstanden werden, sondern kann unter der Annahme stückweise konstanter Eingangssignale auch exakt sein [Lun16b, Abschn. 11.1.6].

wobei $\mathbf{y} \in \mathbb{R}^{n_y}$ und $\mathbf{u} \in \mathbb{R}^{n_u}$ die n_y - bzw. n_u -dimensionalen Aus- bzw. Eingangsfolgen des Systems darstellen. Weiterhin kennzeichnet $t \in \mathbb{N}_{>0}$ die diskretisierte Zeit⁴, und $\mathbf{a}^{[i]} \in \mathbb{R}^{n_y \times n_y}$ und $\mathbf{b}^{[i]} \in \mathbb{R}^{n_y \times n_u}$ die Koeffizientenmatrizen des Modells, wobei $\mathbf{a}^{[0]} = \mathbf{I}$ festgelegt wird, um die Eindeutigkeit der Darstellung zu erzwingen. Gilt weiterhin $\mathbf{b}^{[0]} = \mathbf{0}$, dann wird von einem System ohne Durchgriff gesprochen.

Bei der Analyse von LTI-Systemen ist zudem die Darstellung im Bildbereich üblich. Dabei wird für die Umwandlung im diskreten Zeitbereich die \mathcal{Z} -Transformation verwendet [HW14, Kap. 4]. Unter der Annahme verschwindender Anfangswerte lässt sich das System (2.1) durch Anwendung des Verschiebungssatzes, d. h. $\mathcal{Z}\{\mathbf{y}[k-i]\} = z^{-i}\mathbf{Y}(z)$, in die Darstellung

$$\mathbf{A}(z)\mathbf{Y}(z) = \mathbf{B}(z)\mathbf{U}(z) \quad (2.2)$$

transformieren. Dabei stellen $z \in \mathbb{C}$ die komplexe Variable der \mathcal{Z} -Transformation, $\mathbf{U}(z)$ und $\mathbf{Y}(z)$ die \mathcal{Z} -transformierten der Ein- bzw. Ausgangsfolgen und $\mathbf{A}(z)$ und $\mathbf{B}(z)$ die Polynommatrizen

$$\begin{aligned} \mathbf{A}(z) &= \mathbf{I} + \mathbf{a}^{[1]}z^{-1} + \dots + \mathbf{a}^{[n_A]}z^{-n_A} \quad \text{bzw.} \\ \mathbf{B}(z) &= \mathbf{b}^{[0]} + \mathbf{b}^{[1]}z^{-1} + \dots + \mathbf{b}^{[n_B]}z^{-n_B} \end{aligned} \quad (2.3)$$

dar. Das Übertragungsverhalten des Systems lässt sich dann durch

$$\mathbf{Y}(z) = \mathbf{A}^{-1}(z)\mathbf{B}(z)\mathbf{U}(z) = \mathbf{G}(z)\mathbf{U}(z) \quad (2.4)$$

darstellen, wobei $\mathbf{G}(z)$ multivariable Übertragungsfunktion [Kai80, Abschn. 6.1], Übertragungsfunktionsmatrix oder kurz Übertragungsmatrix [Lju12, Abschn. 2.5] genannt wird. Entsprechend werden die Elemente $G_{ij}(z)$ dann als Übertragungsfunktionen des Systems bezeichnet.

Bei der Umwandlung von (2.1) in die Darstellung (2.2) wird z als die komplexe Variable der \mathcal{Z} -Transformation interpretiert. In den üblichen Werken der Systemidentifikation werden zudem häufig die Symbole q bzw. q^{-1} verwendet [SS89; Lju12], welche allerdings als Zeitverschiebungsoperatoren ($q \mathbf{y}[t] = \mathbf{y}[t+1]$ bzw. $q^{-1} \mathbf{y}[t] = \mathbf{y}[t-1]$) aufgefasst werden müssen. Für die Eingangs-Ausgangs-Beziehung (2.4) wird dann oftmals

$$\mathbf{y}[t] = \mathbf{G}(q^{-1})\mathbf{u}[t] \quad (2.5)$$

geschrieben. Streng genommen, darf in diesem Kontext dann nicht von der Übertragungsmatrix gesprochen werden, weswegen in Anlehnung an die Bezeichnung in [SS89] der Begriff des Filters bzw. der Filtermatrix verwendet wird.

Bemerkung 2.1 (Wahl des Argumentes von \mathbf{G}): Obwohl das Filter \mathbf{G} meist als Funktion des Verzögerungsoperators eingeführt wird, und q^{-1} daher als die natürliche Wahl

⁴Um explizit die kontinuierliche von der diskretisierten Zeit zu unterscheiden, wird insbesondere in der deutschsprachigen Literatur oftmals das Symbol k für die diskrete Zeit verwendet [Gue86; Ise92a; Ise92b; BU16]. Da in der Arbeit ausschließlich zeitdiskrete Modelle verwendet werden, wird hier im Wesentlichen den englischsprachigen Standardwerken [SS89; BJR08; Lju12] gefolgt.

des Argumentes von \mathbf{G} erscheint, wird in einigen Referenzen dennoch $\mathbf{G}(q)$ favorisiert, um die formale Übereinstimmung mit den Bezeichnungen der \mathcal{Z} -Transformation zu erreichen [Lju12, Abschn. 2.1]. \triangleleft

In den meisten praktischen Anwendungen muss davon ausgegangen werden, dass das betrachtete System durch Störungen beeinflusst wird. Hierzu gehören unter anderem Messrauschen, unbekannte bzw. nicht messbare Störsignale oder sonstige Modellierungsungenauigkeiten. Obwohl diese oftmals eines unterschiedlichen Ursprungs sind und daher eine differenzierte Sichtweise nahelegen, ist es zweckmäßig, diese in einem additiven Störsignal $\mathbf{v}[t]$ zusammenzufassen. Innerhalb einer deterministischen Betrachtungsweise sind dabei oftmals konstante, polynomiale oder sinusförmige Störgrößenansätze von Interesse (siehe [GP00] oder [Lju12, Abschn. 14.1]). Im Bereich der stochastischen Regelungstheorie und in der Systemidentifikation hingegen werden Störungen als Realisierungen stochastischer Prozesse verstanden (siehe Anhang A.2.3). Dabei hat es sich als zweckmäßig erwiesen, Störungen als gefiltertes weißes Rauschen (siehe Anhang A.2.3.4), d. h. durch den Ansatz

$$\mathbf{v}[t] = \mathbf{H}(q^{-1})\mathbf{e}[t] \quad (2.6)$$

zu modellieren, wobei \mathbf{H} eine $n_y \times n_y$ - Filtermatrix und $\mathbf{e}[t]$ n_y -dimensionales mittelwertfreies weißes Rauschen mit Kovarianzmatrix $\mathbf{\Lambda}_e \in \mathbb{R}^{n_y \times n_y}$ darstellen. Da die Varianzen von \mathbf{e} beliebig angepasst werden können, lässt sich ohne Einschränkung der Allgemeinheit $\mathbf{H}(0) = \mathbf{I}$ fordern⁵. Das lineare stochastische Gesamtmodell ergibt sich dann zu

$$\begin{aligned} \mathbf{y}[t] &= \mathbf{G}(q^{-1})\mathbf{u}[t] + \mathbf{v}[t] \\ &= \mathbf{G}(q^{-1})\mathbf{u}[t] + \mathbf{H}(q^{-1})\mathbf{e}[t], \quad \text{mit} \quad \mathbb{E}\{\mathbf{e}[t]\mathbf{e}^\top[s]\} = \mathbf{\Lambda}_e\delta_{t,s}, \end{aligned} \quad (2.7)$$

wobei $\mathbf{G}(q^{-1})$ im Folgenden als deterministisches und $\mathbf{H}(q^{-1})$ als stochastisches Teilsystem bezeichnet werden. Die Modellstruktur nach Gleichung (2.7) ist in der Abbildung 2.2 dargestellt.

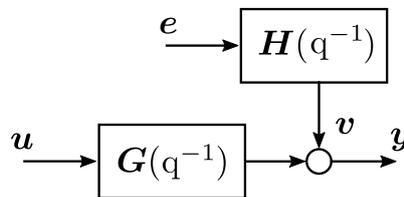


Abbildung 2.2: Lineares stochastisches Modell nach Gleichung (2.7)

Bemerkung 2.2 (Einschränkung der Modellstruktur (2.7)): Obwohl an dieser Stelle noch keine Beschränkungen der Modellstruktur (2.7) notwendig sind, sei bereits hier

⁵Diese Festlegung orientiert sich an den Referenzen [SS89; Lju12]. In [Ast70] wird \mathbf{e} als Zufallsgröße mit normierter Varianz \mathbf{I} betrachtet.

darauf hingewiesen, dass im Rahmen der Identifikation lediglich Modelle mit asymptotisch stabilen Filtern $\mathbf{H}^{-1}(q^{-1})$ und $\mathbf{H}^{-1}(q^{-1})\mathbf{G}(q^{-1})$ zugelassen werden (siehe Abschn. 2.3 und 3.2). \triangleleft

Bemerkung 2.3 (Verteilungsdichte des Rauschsignals \mathbf{e}): Für das weiße Rauschen $\mathbf{e}[t]$ sind zunächst keine weiteren Einschränkungen notwendig, sodass beliebige Verteilungsfunktionen zugelassen sind. Zudem sei angemerkt, dass sich durch die geforderte Mittelwertfreiheit des Rauschens keine Einschränkung der Allgemeingültigkeit ergibt. Der praktisch bedeutsame Fall einer nicht mittelwertfreien Störung (beispielsweise durch einen unbekanntem Sensoroffset), kann beispielsweise durch eine Konstante im deterministischen Teilmodell berücksichtigt werden. Für eine ausführliche Diskussion dieser Problematik sei auf [Lju12, Abschn. 14.1] verwiesen. \triangleleft

2.2.2 Impulsantwort

Neben der Darstellung als Filtermatrix (2.7) kann das Übertragungsverhalten zeitdiskreter LTI-Systeme ebenfalls mittels der Faltungssumme, d. h. durch

$$\mathbf{y}[t] = \sum_{\tau=0}^{\infty} \mathbf{g}[\tau] \mathbf{u}[t - \tau] + \sum_{\tau=0}^{\infty} \mathbf{h}[\tau] \mathbf{e}[t - \tau], \quad (2.8)$$

beschrieben werden, wobei die Elemente $\mathbf{g}[\tau] \in \mathbb{R}^{n_y \times n_u}$ und $\mathbf{h}[\tau] \in \mathbb{R}^{n_y \times n_y}$ als Impulsantwort oder als Impuls- bzw. Gewichtsfolge der Filter $\mathbf{G}(q^{-1})$ und $\mathbf{H}(q^{-1})$ bezeichnet werden⁶.

Eine weitere nützliche Beschreibungsform, die als Kombination der Impulsantwort und der Filterdarstellung betrachtet werden kann, erhält man, wenn die Reihendarstellung der Filter \mathbf{G} bzw. \mathbf{H} nach dem Glied $k - 1$ unterbrochen wird, was im Folgenden am Beispiel eines SISO-Filters veranschaulicht werden soll: Sei $G(q^{-1})$ ein Filter mit dem Zähler- und Nennerpolynom $Z_G(q^{-1})$ bzw. $N_G(q)$ sowie dem Ein- und Ausgangssignal $u[t]$ bzw. $y[t]$, d. h.

$$y[t] = G(q^{-1})u[t] = \frac{Z_G(q^{-1})}{N_G(q^{-1})}u[t]. \quad (2.9)$$

Seien weiterhin die Polynome

$$P_k(q^{-1}) = p^{[0]} + p^{[1]}q^{-1} + \dots + p^{[k-1]}q^{-k+1} \quad \text{und} \quad (2.10a)$$

$$R_k(q^{-1}) = r^{[0]} + r^{[1]}q^{-1} + \dots + r^{[n_R]}q^{-n_R} \quad (2.10b)$$

mit $k \geq 1$ und dem Grad $n_R = \max(n_N - 1, n_Z - k)$ Lösungen der *Diophantischen Gleichung*

$$Z_G(q^{-1}) = P_k(q^{-1})N_G(q^{-1}) + q^{-k}R_k(q^{-1}), \quad (2.11)$$

⁶In vielen Referenzen wird die Impulsantwort allgemein durch $\mathbf{g}[t]$ gekennzeichnet, wohingegen $\mathbf{h}[t]$ im Allgemeinen für die Sprungantwort des gleichen Systems steht (siehe z. B. [Rei06]). In dieser Arbeit hingegen wird den Konventionen der Referenzwerke [SS89; Lju12] gefolgt, wobei Kleinbuchstaben stets für Impulsantworten der entsprechend bezeichneten Filter reserviert sind.

dann kann das Eingangs-Ausgangs-Verhalten des Filters (2.9) ebenfalls durch

$$y[t] = P_k(q^{-1})u[t] + q^{-k} \frac{R_k(q^{-1})}{N_G(q^{-1})}u[t] \quad (2.12)$$

ausgedrückt werden. In der Darstellung (2.12) ist hervorzuheben, dass die Systemantwort y zum Zeitpunkt t in einen vom Eingangssignal u bis zum Zeitschritt $t - k$ und einen ab $t - k + 1$ abhängigen Anteil zerlegt wird. Entsprechend kann für das Filter G ebenfalls

$$G(q^{-1}) := G^{[0,k-1]}(q^{-1}) + G^{[k,\infty]}(q^{-1}), \quad (2.13)$$

mit

$$G^{[0,k-1]}(q^{-1}) := P_k(q^{-1}) = g[0] + g[1]q^{-1} + \dots + g[k-1]q^{-k+1} \quad \text{und} \quad (2.14a)$$

$$G^{[k,\infty]}(q^{-1}) := q^{-k} \bar{G}^{[k,\infty]}(q^{-1}) := q^{-k} \frac{R_k(q^{-1})}{N_G(q^{-1})} \quad (2.14b)$$

geschrieben werden⁷, wobei $P_k(q^{-1})$ die Impulsantwort von $G(q^{-1})$ bis zum Glied $k - 1$ darstellt.

Bemerkung 2.4 (Lösung der Diophantischen Gleichung (2.11)): Durch Koeffizientenvergleich in Gl. (2.11) ergibt sich ein Gleichungssystem für die gesuchten Koeffizienten der Polynome P_k und Q_k , woraus sich wiederum die Rekursionsvorschrift

$$p^{[j]} = g^{[j]} = z^{[j]} - \sum_{i=1}^{\min(j, n_N)} n^{[i]} g^{[j-i]} \quad (2.15)$$

gewinnen lässt, wobei $z^{[j]}$ und $n^{[i]}$ die Koeffizienten der Zähler- und Nennerpolynome Z_G bzw. N_G kennzeichnen (siehe [Mac00, Abschn. 4.2] und [Soe92]). \triangleleft

2.2.3 Zustandsraumdarstellung

Eine weitere übliche Darstellungsform linearer Systeme ist das Zustandsraummodell (ZRM), welches formal als ein System linearer Differenzgleichungen erster Ordnung definiert ist. Wird ein ausschließlich deterministisches System betrachtet, dann ist der Zustand $\mathbf{x} \in \mathbb{R}^{n_x}$ die minimale Information des vergangenen Systemverhaltens, welche notwendig ist, um mithilfe der Eingangssignale den Systemausgang zu präzisieren.

Bei einem stochastischen System kann nicht davon ausgegangen werden, dass das zukünftige Systemverhalten exakt präzisiert werden kann. In diesem Fall wird gefordert, dass die Wahrscheinlichkeitsverteilungen der zukünftigen Zustände lediglich vom aktuellen Systemzustand und gegebenenfalls den zukünftigen Eingangssignalen abhängen. Stochastische Prozesse, welche dieser Eigenschaft genügen, heißen *MARKOV-Prozesse* [Ast70, Abschn. 3.2].

⁷Die in Gl. (2.13) eingeführte Schreibweise gilt für Mehrgrößenfilter $\mathbf{G}(q^{-1})$ in analoger Weise, indem die Filtermatrix entsprechend (2.12) elementweise zerlegt wird.

Im Rahmen der Systemidentifikation und der Theorie der Optimalfilterung werden mindestens zwei strukturell unterschiedliche, in einem gewissen Sinne jedoch gleichwertige Zustandsraumdarstellungen stochastischer LTI-Systeme unterschieden⁸:

Eine in der Literatur übliche und sehr allgemeine Zustandsbeschreibung, welche oftmals auch als Ausgangspunkt für die Herleitung des weit verbreiteten *KÁLMÁN-Filters* verwendet wird [Kal60a; KB61; IM11]⁹, ist die sogenannte *Prozessform*

$$\begin{aligned}\mathbf{x}[t+1] &= \mathbf{A}\mathbf{x}[t] + \mathbf{B}\mathbf{u}[t] + \mathbf{v}[t] \\ \mathbf{y}[t] &= \mathbf{C}\mathbf{x}[t] + \mathbf{D}\mathbf{u}[t] + \mathbf{w}[t],\end{aligned}\tag{2.16}$$

wobei $\mathbf{A} \in \mathbb{R}^{n_x \times n_x}$, $\mathbf{B} \in \mathbb{R}^{n_x \times n_u}$, $\mathbf{C} \in \mathbb{R}^{n_y \times n_x}$ und $\mathbf{D} \in \mathbb{R}^{n_y \times n_u}$ Systemmatrizen genannt werden [AM79; SS89; Lju12]. Die als das Prozess- bzw. Messrauschen bezeichneten n_x - bzw. n_y -dimensionalen Signale \mathbf{v} und \mathbf{w} werden als mittelwertfreie unabhängig identisch verteilte (u. i. v.) Zufallsgrößen mit den Kovarianzmatrizen $\mathbf{\Lambda}_v \in \mathbb{R}^{n_x \times n_x}$ und $\mathbf{\Lambda}_w \in \mathbb{R}^{n_y \times n_y}$ modelliert.

Die Prozessform ist eng verknüpft mit der *Innovationsform* [SS89; OM96; Lju12]

$$\begin{aligned}\mathbf{x}[t+1] &= \mathbf{A}\mathbf{x}[t] + \mathbf{B}\mathbf{u}[t] + \mathbf{K}\mathbf{e}[t] \\ \mathbf{y}[t] &= \mathbf{C}\mathbf{x}[t] + \mathbf{D}\mathbf{u}[t] + \mathbf{e}[t],\end{aligned}\tag{2.17}$$

welche als Zustandsraumdarstellung der Modellstruktur (2.7) betrachtet werden kann. Entsprechend kennzeichnet \mathbf{e} wiederum ein n_y -dimensionales mittelwertfreies, weißes Rauschen mit entsprechender Kovarianzmatrix $\mathbf{\Lambda}_e$. Die Matrix $\mathbf{K} \in \mathbb{R}^{n_x \times n_y}$ wird *KÁLMÁN-Matrix* des Systems genannt.

Prozessform und Innovationsform sind in einem gewissen Sinne gleichwertige Darstellungen stochastischer LTI-Systeme und können unter gewissen Voraussetzungen ineinander überführt werden. Dabei besteht die wesentliche Schwierigkeit darin, dass der stochastische Modellanteil bei der Prozessform durch zwei Rauschquellen beschrieben wird, wohingegen bei der Innovationsform lediglich ein Rauschsignal $\mathbf{e}[t]$ der Dimension n_y vorkommt. Eine exakte Übereinstimmung im deterministischen Sinne ist daher nicht möglich.

Aufgrund des Rauschens müssen die Ausgangssignale der Systeme (2.16) bzw. (2.17) als stochastische Prozesse betrachtet werden. Daher besteht der wesentliche Gedanke bei der Transformation darin, die stochastischen Eigenschaften der jeweiligen Ausgangssignale, insbesondere deren Kovarianzfunktionen $\mathbf{r}_y(\tau)$ bzw. die spektralen Leistungsdichten $\phi_y(\omega)$, zur Übereinstimmung zu bringen. In der Literatur wird dieses Problem im

⁸Die Zustände \mathbf{x} sind nicht eindeutig definierte innere Größen des Systems. Genau genommen existieren für jedes System daher auch unendlich viele Zustandsraumbeschreibungen, die aus einem gegebenen Zustandsmodell mittels einer beliebigen regulären Transformationsmatrix erzeugt werden können [Lun16a; Lun16b].

⁹Das Schätzverfahren ist benannt nach dem US-amerikanischen Mathematiker ungarischer Herkunft Rudolf Emil Kálmán. Da allerdings auch Richard Snowden Bucy und Ruslan Leontjewitsch Stratonovich wichtige Beiträge lieferten, wird oftmals auch die Bezeichnung *KÁLMÁN-BUCY-* oder *STRATONOVICH-KÁLMÁN-BUCY-Filter* verwendet.

Zusammenhang mit der spektralen Faktorisierung stochastischer Prozesse diskutiert¹⁰ (siehe u. a. [Ast70, Abschn. 4.3], [SS89, Appendix A6.1] und [SM05, Kap. 4]). Dabei wird gezeigt, dass sich für

$$\begin{aligned} \mathbf{K} &= \mathbf{A}\mathbf{P}\mathbf{C}^\top (\mathbf{C}\mathbf{P}\mathbf{C}^\top + \mathbf{\Lambda}_w)^{-1} \quad \text{und} \\ \mathbf{\Lambda}_e &= \mathbf{C}\mathbf{P}\mathbf{C}^\top + \mathbf{\Lambda}_w \end{aligned} \quad (2.18)$$

Gleichheit im Sinne der Spektraldichte ergibt, wobei $\mathbf{P} \in \mathbb{R}^{n \times n}$ die eindeutige (positiv semidefinite) stabilisierende Lösung der *diskreten algebraischen RICCATI Gleichung* (DARE)

$$\mathbf{0} = \mathbf{A}\mathbf{P}\mathbf{A}^\top - \mathbf{P} + \mathbf{\Lambda}_v - \mathbf{A}\mathbf{P}\mathbf{C}^\top (\mathbf{C}\mathbf{P}\mathbf{C}^\top + \mathbf{\Lambda}_w)^{-1} \mathbf{C}\mathbf{P}\mathbf{A}^\top \quad (2.19)$$

darstellt.

Bemerkung 2.5 (Stabilität): Der Zusatz „stabilisierend“ bezieht sich auf die Eigenwerte der Matrix $\mathbf{A} - \mathbf{K}\mathbf{C}$. In Anlehnung an [CG84; SGG86] wird die Lösung \mathbf{P} der DARE als stabilisierend bezeichnet, wenn alle Eigenwerte von $\mathbf{A} - \mathbf{K}\mathbf{C}$ innerhalb des Einheitskreises liegen. Weiterhin sei angemerkt, dass die Eigenwerte von $\mathbf{A} - \mathbf{K}\mathbf{C}$ den Polstellen der Übertragungsfunktionen $\mathbf{H}^{-1}(z)$ und $\mathbf{H}^{-1}(z)\mathbf{G}(z)$ entsprechen [SS89, Abschn. 6.2]. Die Existenz einer stabilisierenden Lösung der DARE impliziert daher die asymptotische Stabilität der Übertragungsfunktionen $\mathbf{H}^{-1}(z)$ und $\mathbf{H}^{-1}(z)\mathbf{G}(z)$ (Vgl. Bemerkung 2.2). \triangleleft

Bemerkung 2.6 (Lösungen der DARE): Im Allgemeinen besitzt die DARE (2.19) maximal eine positiv definite Lösung [SS89, Appendix A6.1]. Es existieren jedoch auch Fälle, in welchen keine positiv definite, stattdessen allerdings eine positiv semidefinite stabilisierende Lösung existiert. Zudem können Situationen auftreten, in denen mehrere positiv semidefinite Lösungen existieren, was die zusätzliche Einschränkung „stabilisierend“ rechtfertigt. Für ein illustrierendes Beispiel sei auf [SS89, Bsp. A6.1.2] verwiesen. \triangleleft

Bemerkung 2.7 (Existenz einer stabilisierenden Lösung): Seien $\mathbf{\Lambda}_v \geq \mathbf{0}$ und $\mathbf{\Lambda}_w > \mathbf{0}$. Es existiert eine stabilisierende Lösung dann und nur dann, wenn die folgenden Bedingungen erfüllt sind [SGG86]:

1. Das Paar (\mathbf{A}, \mathbf{C}) ist detektierbar.
2. Das Paar $(\mathbf{A}, \mathbf{\Sigma}_v)$ hat keine nicht steuerbaren Eigenwerte auf dem Einheitskreis, wobei $\mathbf{\Lambda}_v = \mathbf{\Sigma}_v^\top \mathbf{\Sigma}_v$ gilt.

\triangleleft

¹⁰Im Rahmen der spektralen Faktorisierung wird gezeigt, dass eine beliebige spektrale Leistungsdichte $\phi_y(\omega)$ mit $\det(\phi_y(\omega)) \neq 0$, welche allgemein eine gebrochen rationale Funktion in $\cos(\omega)$ ist, als Produkt $\mathbf{H}(e^{-j\omega})\mathbf{\Lambda}_e\mathbf{H}^\top(e^{j\omega})$ dargestellt werden kann. Als Konsequenz dieses Resultates lässt sich schlussfolgern, dass ein Signal mit spektraler Leistungsdichte $\phi_y(\omega)$ durch $\mathbf{y}[t] = \mathbf{H}(q^{-1})\mathbf{e}[t]$ mit $\mathbb{E}\{\mathbf{e}[t]\mathbf{e}^\top[s]\} = \delta_{t,s}\mathbf{\Lambda}_e$, modelliert werden kann, was wiederum den Modellansatz (2.6) rechtfertigt.

2.3 Optimale Prädiktion zeitdiskreter LTI-Systeme

Mithilfe der im Abschnitt 2.2 eingeführten Modelle lassen sich Lösungsansätze für eine Vielzahl technischer Problemstellungen erarbeiten. Der in der Praxis wohl üblichste Anwendungsfall ist jedoch die Vorhersage des zukünftigen Systemverhaltens. Dabei werden insbesondere im Kontext der Identifikation und der optimalen Regelung dynamischer Systeme zwei eng verwandte Konzepte unterschieden:

- **Simulation:** Mithilfe des Modells und gegebener Eingangsgrößen wird der zeitliche Verlauf der Ausgangsgrößen bei fest vorgegebenen Anfangswerten berechnet. Simulationen sind oftmals Grundlage für die Auslegung, Analyse und Optimierung des Systems zur Entwurfszeit.
- **Prädiktion:** Mittels der gemessenen Ein- und Ausgangsgrößen bis zur Zeit t und gegebenenfalls zukünftiger Eingangsgrößen $\mathbf{u}[t]$, wird der zeitliche Verlauf der zukünftigen Ausgangsgrößen $\mathbf{y}[t]$ für eine gegebene Anzahl von Prädiktionsschritten $k \in \mathbb{N}_{>0}$ vorhergesagt. Da bei der Prädiktion die Information über den Verlauf vergangener Prozessgrößen verwendet wird, bieten Prädiktionsmodelle im Vergleich zu Simulationsmodellen üblicherweise eine bessere Vorhersagegenauigkeit.

Prädiktionsmodelle sind die Basis der für diese Arbeit grundlegenden Prädiktionsfehlerminimierung, zudem aber auch verbreiteter Reglerentwurfverfahren wie beispielsweise der *Minimum Varianz Regelung* [Ast70, Kap. 6], [LLM98, Abschn. 7.6], [LZ06, Abschn. 4.2]. Es soll daher auf das Problem der *optimalen Prädiktion* zeitdiskreter LTI-Systeme der Form (2.7) eingegangen werden, wobei folgende Annahmen getroffen werden:

- (A1) Eine *k-Schritt Prädiktion* zur Zeit t ist eine beliebige Funktion der beobachteten Ausgangsfolge $\mathbf{y}[t]$ bis zum Zeitpunkt t und gegebenenfalls der Eingangsfolge $\mathbf{u}[t]$ bis $t + k$.
- (A2) Die *optimale k-Schritt Prädiktion* $\hat{\mathbf{y}}[t + k|t]$ zum Zeitpunkt t minimiert die mittlere quadratische Abweichung (Mean squared error – MSE) der *k-Schritt Prädiktionsfehler*

$$\boldsymbol{\varepsilon}[t + k|t] = \mathbf{y}[t + k] - \hat{\mathbf{y}}[t + k|t]. \quad (2.20)$$

- (A3) Die Signale $\mathbf{u}[t]$ und $\mathbf{e}[s]$ seien unabhängig für alle t und s .
- (A4) Die Filter $\mathbf{H}^{-1}(q^{-1})$ und $\mathbf{H}^{-1}(q^{-1})\mathbf{G}(q^{-1})$ sind asymptotisch stabil (siehe Bemerkung 2.2).

Im Folgenden soll ein Ausdruck für die optimale *k-Schritt Prädiktion* $\hat{\mathbf{y}}[t + k|t]$ für das Modell (2.7) hergeleitet werden. Der dabei grundlegende Gedanke ist es, die Störung \mathbf{e} erst ab dem Prädiktionszeitpunkt t als unbekannt zu betrachten, da diese unter der Annahme asymptotischer Stabilität des Filters $\mathbf{H}^{-1}(q^{-1})$ und $\mathbf{H}^{-1}(q^{-1})\mathbf{G}(q^{-1})$ durch

$$\mathbf{e}[t] = \mathbf{H}^{-1}(q^{-1})(\mathbf{y}[t] - \mathbf{G}(q^{-1})\mathbf{u}[t]) \quad (2.21)$$

rekonstruiert werden kann. Der Systemausgang $\mathbf{y}[t+k]$ von (2.7) lässt sich daher mithilfe der in (2.13) eingeführten Zerlegung

$$\mathbf{H}(q^{-1}) = \mathbf{H}^{[0,k-1]}(q^{-1}) + q^{-k} \bar{\mathbf{H}}^{[k,\infty]}(q^{-1}) \quad (2.22)$$

durch

$$\begin{aligned} \mathbf{y}[t+k] &= \mathbf{G}\mathbf{u}[t+k] + \mathbf{H}\mathbf{e}[t+k] \\ &= \mathbf{G}\mathbf{u}[t+k] + q^{-k} \bar{\mathbf{H}}^{[k,\infty]} \mathbf{e}[t+k] + \mathbf{H}^{[0,k-1]} \mathbf{e}[t+k] \\ &= \mathbf{G}\mathbf{u}[t+k] + q^{-k} \bar{\mathbf{H}}^{[k,\infty]} \mathbf{H}^{-1} (\mathbf{y}[t+k] - \mathbf{G}\mathbf{u}[t+k]) \\ &\quad + \mathbf{H}^{[0,k-1]} \mathbf{e}[t+k] \\ &= (\mathbf{H} - q^{-k} \bar{\mathbf{H}}^{[k,\infty]}) \mathbf{H}^{-1} \mathbf{G}\mathbf{u}[t+k] + \bar{\mathbf{H}}^{[k,\infty]} \mathbf{H}^{-1} \mathbf{y}[t] \\ &\quad + \mathbf{H}^{[0,k-1]} \mathbf{e}[t+k] \\ &= \mathbf{H}^{[0,k-1]} \mathbf{H}^{-1} \mathbf{G}\mathbf{u}[t+k] + \bar{\mathbf{H}}^{[k,\infty]} \mathbf{H}^{-1} \mathbf{y}[t] + \mathbf{H}^{[0,k-1]} \mathbf{e}[t+k] \end{aligned} \quad (2.23)$$

ausdrücken, wobei das Argument q^{-1} aufgrund der Übersichtlichkeit unterdrückt wurde. Mithilfe dieser Modellumformung lässt sich der optimale k -Schritt Prädiktor aus der letzten Zeile von (2.23) ablesen:

$$\hat{\mathbf{y}}[t+k|t] = \mathbf{H}^{[0,k-1]} \mathbf{H}^{-1} \mathbf{G}\mathbf{u}[t+k] + \bar{\mathbf{H}}^{[k,\infty]} \mathbf{H}^{-1} \mathbf{y}[t] \quad (2.24)$$

Die angegebene Gleichung bringt zum Ausdruck, dass $\hat{\mathbf{y}}[t+k]$ tatsächlich nur von Messwerten des Ausgangs \mathbf{y} bis zur Zeit t abhängig ist. Wird weiterhin $q^{-k} \bar{\mathbf{H}}^{[k,\infty]}$ durch $\mathbf{H} - \mathbf{H}^{[0,k-1]}$ substituiert, so ergibt sich die übliche Darstellung der Optimalprädiktion (vgl. [HW99; GPS02]):

$$\hat{\mathbf{y}}[t+k|t] = \mathbf{H}^{[0,k-1]} \mathbf{H}^{-1} \mathbf{G}\mathbf{u}[t+k] + (\mathbf{I} - \mathbf{H}^{[0,k-1]} \mathbf{H}^{-1}) \mathbf{y}[t+k] \quad (2.25)$$

Für den k -Schritt Prädiktionsfehler $\boldsymbol{\varepsilon}[t+k|t]$ lässt sich dann

$$\begin{aligned} \boldsymbol{\varepsilon}[t+k|t] &= \mathbf{y}[t+k] + \hat{\mathbf{y}}[t+k|t] \\ &= \mathbf{y}[t+k] - \mathbf{H}^{[0,k-1]} \mathbf{H}^{-1} \mathbf{G}\mathbf{u}[t+k] - \bar{\mathbf{H}}^{[k,\infty]} \mathbf{H}^{-1} \mathbf{y}[t] \\ &= (\mathbf{H} - q^{-k} \bar{\mathbf{H}}^{[k,\infty]}) \mathbf{H}^{-1} \mathbf{y}[t+k] - \mathbf{H}^{[0,k-1]} \mathbf{H}^{-1} \mathbf{G}\mathbf{u}[t+k] \end{aligned} \quad (2.26)$$

notieren, sodass sich das *Residuenfilter* zu

$$\boldsymbol{\varepsilon}[t+k|t] = \mathbf{H}^{[0,k-1]} \mathbf{H}^{-1} (\mathbf{y}[t+k] - \mathbf{G}\mathbf{u}[t+k]) \quad (2.27)$$

ergibt.

Aufgrund der Linearität des Erwartungswertes und der Annahme $E\{\mathbf{e}[t]\} = \mathbf{0}$ ergibt sich $E\{\mathbf{H}^{[0,k-1]}\mathbf{e}[t]\} = \mathbf{0}$, sodass man für die Kovarianzmatrix des Prädiktionsfehlers

$$\begin{aligned} \text{Cov}\{\boldsymbol{\varepsilon}[t|t-k]\} &= E\left\{\mathbf{H}^{[0,k-1]}\mathbf{e}[t]\left(\mathbf{H}^{[0,k-1]}\mathbf{e}[t]\right)^\top\right\} \\ &= \sum_{i=0}^{k-1} \mathbf{h}[i]\boldsymbol{\Lambda}_e\mathbf{h}^\top[i] \end{aligned} \quad (2.28)$$

erhält, wobei die Unabhängigkeit von $\mathbf{e}[t], \mathbf{e}[t-1], \dots, \mathbf{e}[t-k+1]$ berücksichtigt wurde und $\mathbf{h}[i]$ die Koeffizienten der Impulsantwort des Rauschfilters $\mathbf{H}(q^{-1})$ darstellen.

Beweis 2.1: Der Beweis ist eine Verallgemeinerung der in [Ast70, Abschn. 6.3] und [SS89, Abschn. 7.3] verwendeten Beweisidee und veranschaulicht, dass das Filter (2.24) bzw. (2.25) eine im Sinne des MSE optimale Prädiktion darstellt (Annahme (A2)):

Sei $\mathbf{z}[t] := \mathbf{H}^{[0,k-1]}\mathbf{H}^{-1}\mathbf{G}\mathbf{u}[t] + \bar{\mathbf{H}}^{[k,\infty]}\mathbf{H}^{-1}\mathbf{y}[t-k]$ und sei $\mathbf{y}^*[t]$ ein beliebiger k -Schritt Prädiktor des Systemausgangs $\mathbf{y}[t]$. Beachte, dass die Terme $\mathbf{H}^{[0,k-1]}\mathbf{e}[t]$ und $\mathbf{z}[t]$ bzw. $\mathbf{y}^*[t]$ aufgrund der Annahme der Unabhängigkeit von \mathbf{u} und \mathbf{e} unkorreliert sind. Für die Schätzfehlerkovarianz des k -Schritt Prädiktionsfehlers $\boldsymbol{\varepsilon}$ gilt daher

$$\begin{aligned} &E\left\{\left(\mathbf{y}[t] - \mathbf{y}^*[t]\right)\left(\mathbf{y}[t] - \mathbf{y}^*[t]\right)^\top\right\} \\ &= E\left\{\left(\mathbf{z}[t] + \mathbf{H}^{[0,k-1]}\mathbf{e}[t] - \mathbf{y}^*[t]\right)\left(\mathbf{z}[t] + \mathbf{H}^{[0,k-1]}\mathbf{e}[t] - \mathbf{y}^*[t]\right)^\top\right\} \\ &= E\left\{\mathbf{H}^{[0,k-1]}\mathbf{e}[t]\left(\mathbf{H}^{[0,k-1]}\mathbf{e}[t]\right)^\top\right\} + E\left\{\left(\mathbf{z}[t] - \mathbf{y}^*[t]\right)\left(\mathbf{z}[t] - \mathbf{y}^*[t]\right)^\top\right\} \\ &\quad + E\left\{\mathbf{H}^{[0,k-1]}\mathbf{e}[t]\left(\mathbf{z}[t] - \mathbf{y}^*[t]\right)^\top + \left(\mathbf{z}[t] - \mathbf{y}^*[t]\right)\left(\mathbf{H}^{[0,k-1]}\mathbf{e}[t]\right)^\top\right\} \\ &= E\left\{\mathbf{H}^{[0,k-1]}\mathbf{e}[t]\left(\mathbf{H}^{[0,k-1]}\mathbf{e}[t]\right)^\top\right\} + E\left\{\left(\mathbf{z}[t] - \mathbf{y}^*[t]\right)\left(\mathbf{z}[t] - \mathbf{y}^*[t]\right)^\top\right\}, \end{aligned} \quad (2.29)$$

wobei sich das Minimum für $\mathbf{y}^*[t] \equiv \mathbf{z}[t]$ ergibt. \square

Korollar 2.1: Für den wichtigen Spezialfall der 1-Schritt Prädiktion gilt $\mathbf{H}^{[0,0]}(q^{-1}) = \mathbf{h}[0] = \mathbf{I}$, sodass sich

$$\hat{\mathbf{y}}[t|t-1] = \mathbf{H}^{-1}(q^{-1})\mathbf{G}(q^{-1})\mathbf{u}[t] + \left(\mathbf{I} - \mathbf{H}^{-1}(q^{-1})\right)\mathbf{y}[t] \quad (2.30)$$

und

$$\boldsymbol{\varepsilon}[t|t-1] = \mathbf{H}^{-1}(q^{-1})\left(\mathbf{y}[t] - \mathbf{G}(q^{-1})\mathbf{u}[t]\right) \quad (2.31)$$

ergibt (vgl. [SS89, Abschn. 7.3] und [Lju12, Abschn. 3.2]). Es ist zu beachten, dass sich aus (2.21) der Zusammenhang (2.31) ergibt, wenn $\mathbf{e}[t]$ durch $\boldsymbol{\varepsilon}[t|t-1]$ ersetzt wird. Da das Filter (2.31) asymptotisch stabil ist, gilt für $t \rightarrow \infty$ die Beziehung $\mathbf{e}[t] \equiv \boldsymbol{\varepsilon}[t|t-1]$, sodass $\boldsymbol{\varepsilon}[t|t-1]$ als Schätzer von $\mathbf{e}[t]$ betrachtet werden kann. Anschaulich gesprochen entspricht der 1-Schritt Prädiktionsfehler demnach dem Anteil des Systemausgangs $\mathbf{y}[t]$ der zum Zeitpunkt $t-1$ nicht vorhergesagt werden kann, weswegen $\mathbf{e}[t]$ oftmals als *Innovation* oder als *Residuum* bezeichnet wird.

Durch Vergleich von (2.27) und (2.31) ist zudem ersichtlich, dass der k -Schritt Prädiktionsfehler durch Filterung der 1-Schritt Prädiktionsfehlerfolge, d. h. mittels

$$\boldsymbol{\varepsilon}[t|t-k] = \mathbf{H}^{[0,k-1]}(q^{-1})\boldsymbol{\varepsilon}[t|t-1] \quad (2.32)$$

dargestellt werden kann (vgl. [HW99]). Man beachte auch, dass diese Gleichung zum Ausdruck bringt, dass die Folgeelemente $\boldsymbol{\varepsilon}[t|t-k]$, $\boldsymbol{\varepsilon}[t-1|t-k-1]$, usw. als zeitlich korreliert betrachtet werden müssen. \triangleleft

Bemerkung 2.8 (Optimale Prädiktion im geschlossenen Regelkreis): Die Unabhängigkeit des Eingangssignals $\mathbf{u}[t]$ und der Störung $\mathbf{e}[s]$ für $s > t$ ist aufgrund der Kausalität in der Praxis in der Regel gewährleistet, wohingegen dies für $t \geq s$ im Allgemeinen nicht garantiert ist. Im Fall des geschlossenen Regelkreises ist diese Forderung beispielsweise verletzt, da der Eingang der Regelstrecke durch die Ausgangsrückführung mit der Störung korreliert ist (siehe Abb. 2.3). Im Allgemeinen stellt die Prädiktion nach Gleichung (2.25) dann keine optimale Prädiktion dar. Gilt die Unabhängigkeit für $s > t$, und zudem $\mathbf{G}(0) = \mathbf{0}$ (kein Durchgriff) sowie $k = 1$ (1-Schritt Prädiktion), dann hängt die Folge $\mathbf{z}[t]$ nur von $\mathbf{u}[t-1]$, $\mathbf{u}[t-2]$, ... ab und ist daher unkorreliert mit $\mathbf{H}^{[0,0]}\mathbf{e}[t] = \mathbf{e}[t]$. Das Filter (2.25) stellt dann auch für den Fall des geschlossenen Regelkreises die optimale 1-Schritt Prädiktion dar. Mit einer analogen Überlegung lässt sich schlussfolgern, dass (2.25) die optimale Prädiktion des Systemausgangs \mathbf{y} ist, wenn die Anzahl der Verzögerungsglieder des geöffneten Regelkreises größer oder gleich k ist. \triangleleft

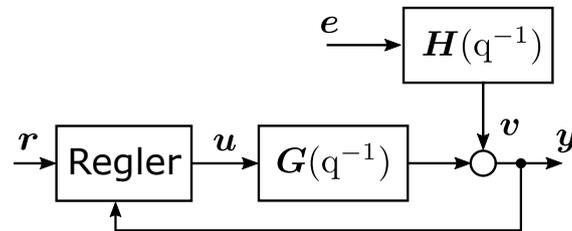


Abbildung 2.3: Lineares stochastisches Modell im geschlossenen Regelkreis

Optimale Prädiktion für ein Zustandsraummodell

Der optimale 1-Schritt Prädiktor des Zustandsraummodells (2.16) bzw. (2.17) ist durch das *KÁLMÁN-Filter*

$$\begin{aligned} \hat{\mathbf{x}}[t+1|t] &= \mathbf{A}_K \hat{\mathbf{x}}[t|t-1] + \mathbf{B}\mathbf{u}[t] + \mathbf{K}\mathbf{y}[t] \\ \hat{\mathbf{y}}[t|t-1] &= \mathbf{C}\hat{\mathbf{x}}[t|t-1] + \mathbf{D}\mathbf{u}[t], \end{aligned} \quad (2.33)$$

mit

$$\mathbf{A}_K = \mathbf{A} - \mathbf{K}\mathbf{C} \quad (2.34)$$

gegeben, wobei $\hat{\mathbf{x}}$ und $\hat{\mathbf{y}}$ die Schätzungen der Zustandsgrößen, bzw. des Systemausgangs darstellen (siehe [Lju12, Abschn. 4.3] und [SS89, Abschn. 7.3]). Für den Prädiktionsfehler gilt dann die Beziehung

$$\boldsymbol{\varepsilon}[t|t-1] = \mathbf{y}[t] - \hat{\mathbf{y}}[t|t-1] = \mathbf{y}[t] - \mathbf{C}\hat{\mathbf{x}}[t|t-1] - \mathbf{D}\mathbf{u}[t]. \quad (2.35)$$

Es sei darauf hingewiesen, dass (2.33) im Rahmen der unterraumbasierten Identifikation (Subspace-Identification) auch als *Prädiktorform* bezeichnet wird [Qin06; Buc10]. Da die Matrix \mathbf{A}_K Eigenwerte besitzt, deren Beträge stets kleiner als 1 sind (siehe Bemerkung 2.5), ist das Modell (2.33) stets asymptotisch stabil, sodass dieses insbesondere für die Identifikation instabiler Prozesse geeignet ist [Qin06].

2.4 Parametrierung zeitdiskreter LTI-Systeme

Eine bei der Identifikation wichtige Problematik ist die Frage nach einer sinnvollen Parametrierung der allgemeinen polynomialen Modellstruktur (2.7). Diese beeinflusst zum einen die Flexibilität des Modells, zum anderen aber auch die Komplexität der Identifikationsaufgabe. Im Bereich der Systemidentifikation wird für stochastische LTI-Eingrößensysteme typischerweise die allgemeine Darstellung

$$A(q^{-1})y[t] = \frac{B(q^{-1})}{F(q^{-1})}u[t] + \frac{C(q^{-1})}{D(q^{-1})}e[t] \quad \mathbb{E}\{e^2[t]\} = \lambda_e \delta_{t,s} \quad (2.36)$$

mit den Polynomen

$$A(q^{-1}) = 1 + a^{[1]}q^{-1} + \dots + a^{[n_A]}q^{-n_A}, \quad (2.37a)$$

$$B(q^{-1}) = b^{[0]} + b^{[1]}q^{-1} + \dots + b^{[n_B]}q^{-n_B}, \quad (2.37b)$$

$$C(q^{-1}) = 1 + c^{[1]}q^{-1} + \dots + c^{[n_C]}q^{-n_C}, \quad (2.37c)$$

$$D(q^{-1}) = 1 + d^{[1]}q^{-1} + \dots + d^{[n_D]}q^{-n_D} \quad \text{und} \quad (2.37d)$$

$$F(q^{-1}) = 1 + f^{[1]}q^{-1} + \dots + c^{[n_F]}q^{-n_F} \quad (2.37e)$$

zugrunde gelegt [Lju12, Abschn. 4.2], wobei deren Koeffizienten allgemein als reelle Modellparameter betrachtet werden. Bezogen auf die im Abschnitt 2.2.1 eingeführte Darstellung (2.7) ergeben sich daher die Zusammenhänge

$$G(q^{-1}) = \frac{B(q^{-1})}{A(q^{-1})F(q^{-1})} \quad \text{und} \quad H(q^{-1}) = \frac{C(q^{-1})}{A(q^{-1})D(q^{-1})}. \quad (2.38)$$

Es sei angemerkt, dass die allgemeine Struktur (2.36) in der Praxis nur selten verwendet wird. Stattdessen ist es meistens ausreichend eines oder mehrere der Polynome in (2.37) zu Eins zu setzen. Einige häufig verwendete SISO-Modellstrukturen, sind in der Tabelle 1 zusammengefasst. Für die Bedeutung der in der Tabelle verwendeten Akronyme sei u. a. auf [SS89; Nel01; Lju12; BU16] verwiesen.

Akronym	Polynome	1-Schritt Prädiktor	Bemerkung
FIR	B	$\hat{y}[t] = Bu[t]$	(1),(2)
ARX	A, B	$\hat{y}[t] = Bu[t] + (1 - A)y[t]$	(1),(2)
ARMAX	A, B, C	$\hat{y}[t] = \frac{B}{C}u[t] + \frac{C-A}{C}y[t]$	
ARARX	A, B, D	$\hat{y}[t] = BDu[t] + (1 - AD)y[t]$	
ARARMAX	A, B, C, D	$\hat{y}[t] = \frac{BD}{C}u[t] + \frac{C-AD}{C}y[t]$	
OE	B, F	$\hat{y}[t] = \frac{B}{F}u[t]$	(2),(3)
BJ	B, C, D, F	$\hat{y}[t] = \frac{BD}{CF}u[t] + \frac{C-D}{C}y[t]$	(2),(3),(4)

Tabelle 1: Typische SISO-Modellstrukturen als Spezialfälle von (2.36)

Bemerkungen in Tabelle 1:

- (1) Die 1-Schritt Prädiktion ist eine lineare Funktion der in (2.37) eingeführten Polynomkoeffizienten. Die Schätzung der Modellparameter durch Prädiktionsfehlerminimierung ist in diesem Fall besonders einfach (siehe Abschn. 3.3). Es sei jedoch angemerkt, dass diese Aussage lediglich für den Fall der 1-Schritt Prädiktion gilt. Für ein ARX-Modell, ist die k -Schritt Prädiktion beispielsweise eine nichtlineare Funktion der Koeffizienten von $A(q^{-1})$ und $B(q^{-1})$.
- (2) Das deterministische und das stochastische Teilsystem sind unabhängig parametrisiert. Diese Eigenschaft kann insbesondere bei der Identifikation von Systemen nützlich sein, bei denen die Störung nicht durch den Ansatz $H(q^{-1})e[t]$ beschrieben werden kann (siehe Abschn. 3.2.2). Weiterhin kann aufgrund des unabhängig parametrisierten Rauschmodells durch Vorfilterung der Ein- und Ausgangssignale bzw. durch gezielte Festlegung von $H(q^{-1})$ eine Anpassung von $G(q^{-1})$ in einem vorgegebenen Frequenzbereich erzwungen werden (siehe Gln. (3.10) und (3.25)).
- (3) Ist instabiles Verhalten des Prozesses $G(q^{-1})$ zu erwarten, dann ist der Modellansatz aufgrund der Instabilität des optimalen Prädiktors zur Identifikation nicht geeignet (vgl. Bem. 2.2).

Bemerkung 2.9 (Parametrierung der Polynome $A(q^{-1})$ bis $F(q^{-1})$): Wie in (2.37) angedeutet, wird bei der Systemidentifikation typischerweise ein additiver Ansatz für die Parametrierung der Polynome $A(q^{-1})$ bis $F(q^{-1})$ verwendet¹¹. Es ist jedoch auch eine faktorisierte Darstellung mittels Linear- und/oder quadratischer Faktoren denkbar [Gro14, Kap. 2]. Dies kann beispielsweise dann sinnvoll sein, wenn bei der Identifikation die Modellmenge auf die Klasse der stabilen und/oder nicht schwingungsfähigen Modelle eingeschränkt werden soll. Für eine ausführliche Diskussion der Problematik der Integration von Vorwissen sei auf [Gro14] verwiesen. \triangleleft

¹¹Das innerhalb der MATLAB SYSTEM IDENTIFICATION TOOLBOX für die Darstellung polynomialer Modelle des Typs (2.36) definierte `idpoly()`-Objekt basiert beispielsweise ausschließlich auf einer Summendarstellung.

2.4.1 Spezielle Parametrierungen von LTI-Systemen

In der Literatur werden weitere Modellansätze verwendet, die als Spezialfälle der in Tabelle 1 angegebenen Strukturen aufgefasst werden können:

Modelle ohne exogene Eingangsgrößen

Im Bereich der Zeitreihenanalyse werden Modelle ohne exogene Eingangsgrößen $u[t]$ verwendet. Die sich aus der Darstellung (2.36) ergebenden Modellstrukturen werden als

- AR-Modell (Autoregressive): $y[t] = \frac{1}{A(q^{-1})}e[t]$
- MA-Modell (Moving average): $y[t] = C(q^{-1})e[t]$
- ARMA-Modell (Autoregressive moving average): $y[t] = \frac{C(q^{-1})}{A(q^{-1})}e[t]$

bezeichnet [BJR08].

Stochastisches Modell mit integrierendem Verhalten

In der Praxis ist die Annahme der (quasi)-Stationarität des zu modellierenden Prozesses oftmals verletzt (siehe Anhang A.2.3). Aus regelungstechnischer Sicht ist dies beispielsweise dann der Fall, wenn nicht messbare Störungen auf das System einwirken. Kann diese gegenüber der zu modellierenden Systemdynamik als tieffrequent betrachtet werden, dann wird die Störung oftmals als *Drift* oder als *Trend* bezeichnet [KN06; BJR08]. Eine Möglichkeit, diese Problematik durch den Modellansatz zu berücksichtigen, ist es ein Rauschmodell $H(z)$ mit einem oder mehreren Polen bei $z = 1$ zu verwenden. Weil dies einer Integration des Rauschsignals entspricht, werden die entsprechenden Modelle in der Regel mit Akronymen der Form „ARI(X)“, „ARIMA(X)“ usw. abgekürzt, wobei „I“ abkürzend für die Bezeichnung „integrated noise“ verwendet wird.

Saisonale Modelle

Ist bei der Modellierung mit periodischen Störungen zu rechnen, welche gegenüber der Systemdynamik nicht als tieffrequent betrachtet werden können, dann ist möglicherweise die Verwendung saisonaler Modellansätze sinnvoll. Dabei können die folgenden beiden Fälle unterschieden werden:

Sinusförmige Störsignale: Sinusförmige Störgrößen treten beispielsweise bei mechanischen Systemen auf, die durch rotierende Teile verursacht werden. Dieser Fall kann durch Übertragungsfunktionen $H(z)$ mit einem oder mehreren konjugiert komplexen Polpaaren auf dem Einheitskreis, d. h. bei $z = \cos(\omega_0) \pm j \sin(\omega_0)$, berücksichtigt werden, wobei ω_0 der zu erwartenden Kreisfrequenz der Störung entspricht [GEL86].

Beliebige periodische Störsignale: Die prinzipielle Herangehensweise bei der Parametrierung saisonaler Modelle mit periodischen Störungen soll zunächst anhand des Beispiels einer Zeitreihe, d. h. eines Prozesses ohne deterministische Eingangsgrößen, motiviert werden: Die wesentliche Besonderheit saisonaler Zeitreihen mit Periodendauer

p ist die Eigenschaft, dass sich Beobachtungen mit einem Abstand von p Zeiteinheiten ähneln. Ein deterministisches periodisches Signal kann daher beispielsweise durch das Modell $y[t] = y[t - p]$ beschrieben werden. Wird hingegen erwartet, dass die Werte näherungsweise linear korreliert sind, dann ist ein AR-Modell $(1 + a^{[p]} q^{-p})y[t] = e[t]$ mit dem Parameter $a^{[p]} \in \mathbb{R}$ und der Störung $e[t]$ naheliegend. Um weiterhin eine „Mittelung“ über die vergangenen Beobachtungen zu ermöglichen, ist es denkbar diesen Ansatz durch

$$(1 + a^{[p_{\min}]} q^{-p_{\min}} + \dots + a^{[p_{\max}]} q^{-p_{\max}})y[t] = e[t] \quad (2.39)$$

zu erweitern ($p_{\max} \geq p_{\min}$). Wird zudem davon ausgegangen, dass auch die Werte $y[t], y[t - 1], \dots$ untereinander linear korreliert sind, dann lässt dies ein AR-Modell der Form

$$y[t] = \frac{1}{A(q^{-1})} e[t] = \frac{1}{\tilde{A}(q^{-1}) + \bar{A}(q^{-1})} e[t] \quad (2.40)$$

mit den Polynomen

$$\begin{aligned} \tilde{A}(q^{-1}) &= 1 + \tilde{a}^{[1]} q^{-1} + \dots + \tilde{a}^{[n_{\tilde{a}}]} q^{-n_{\tilde{a}}}, \quad \text{und} \\ \bar{A}(q^{-1}) &= \bar{a}^{[p_{\min}]} q^{-p_{\min}} + \dots + \bar{a}^{[p_{\max}]} q^{-p_{\max}} \end{aligned} \quad (2.41)$$

rechtfertigen ($n_{\tilde{a}} < p_{\min}$). Ein saisonales ARMA-Modell (S-ARMA) lässt sich entsprechend durch

$$y[t] = \frac{C(q^{-1})}{A(q^{-1})} e[t] = \frac{\tilde{C}(q^{-1}) + \bar{C}(q^{-1})}{\tilde{A}(q^{-1}) + \bar{A}(q^{-1})} e[t] \quad (2.42)$$

parametrieren.

Bemerkung 2.10 (Parametrierung saisonaler Modelle): Entgegen der hier genutzten additiven Parametrierung (2.42) werden im Rahmen der Zeitreihenanalyse meist multiplikative Modellansätze der Form

$$\tilde{A}(q^{-1})\bar{A}(q^{-p})y[t] = \tilde{C}(q^{-1})\bar{C}(q^{-p})e[t] \quad (2.43)$$

für die Beschreibung saisonaler ARMA-Prozesse verwendet [BJR08, Kap. 9]. \triangleleft

Die Methodik der Parametrierung saisonaler Zeitreihen kann auf Prozesse mit exogenen Eingangsgrößen übertragen werden. Die dabei wesentliche Idee besteht darin, das Rauschmodell $H(q^{-1})$ derart anzusetzen, dass der periodische Charakter der Störung durch das Rauschfilter erfasst werden kann. Ein entsprechender Box-Jenkins (BJ) Modellansatz kann daher durch

$$y[t] = \frac{B(q^{-1})}{F(q^{-1})} u[t] + \frac{\tilde{C}(q^{-1}) + \bar{C}(q^{-1})}{\tilde{D}(q^{-1}) + \bar{D}(q^{-1})} e[t], \quad (2.44)$$

parametriert werden, wobei die Polynome des Rauschmodells analog zu (2.41) definiert werden.

Bemerkung 2.11 (Parametrierung saisonaler ARMAX-Modelle): Bei der Parametrierung saisonaler ARMAX-Modelle kann die periodische Störung durch das Polynom

$A(q^{-1})$ oder durch $D(q^{-1})$ (ARARMAX-Modell) modelliert werden. Es sei jedoch darauf hingewiesen, dass das Polynom $A(q^{-1})$ sowohl im Nenner des deterministischen als auch des stochastischen Teilsystems auftaucht (siehe Gl. (2.38)), sodass im ersteren Fall oftmals zusätzliche Parameter im Zähler von $G(q^{-1})$ vorgesehen werden müssen. Im Hinblick auf eine möglichst sparsame Parametrierung des Modells ist es daher sinnvoll, einen ARARMAX-Ansatz zu verwenden, wobei der saisonale Anteil durch das Polynom $D(q^{-1})$ zu berücksichtigen ist. \triangleleft

2.4.2 Multivariable Modellstrukturen

Es ist naheliegend, die Parametrierung des SISO-Falls auf den multivariaten Fall zu übertragen, indem die Gleichung (2.36) als vektorielles Differenzgleichungssystem mit entsprechend dimensionierten Polynommatrizen angesehen wird. Betrachtet man die darin vorkommenden Koeffizientenmatrizen in analoger Weise als freie Parameter des Modells, dann kann dies jedoch zu Fällen führen, in denen die Identifikationsbedingung (A.19) verletzt ist, sodass die Einschränkung der Struktur einer oder mehrerer Polynommatrizen notwendig ist. Die Frage nach einer identifizierbaren Parametrierung von (2.7) ist daher mit der Problemstellung kanonischer MIMO-Modellstrukturen eng verknüpft. Für eine weiterführende Diskussion dieser Problematik sei auf [Lju12, App. 4A], [SS89, Abschn. 6.2], [Kai80, Abschn. 6] und die darin enthaltenen Referenzen verwiesen.

2.5 Nichtlineare Modellstrukturen

In den bisherigen Abschnitten wurden allgemeine Beschreibungsformen für zeitdiskrete stochastische LTI-Systeme diskutiert. In vielen Fällen lässt sich das dynamische Verhalten aber nur durch einen nichtlinearen Modellansatz hinreichend gut beschreiben. Im Bereich der Systemidentifikation wurden in der Vergangenheit daher eine Vielzahl nichtlinearer Modellstrukturen vorgeschlagen, welche oftmals als eine Verallgemeinerung linearer Modelle betrachtet werden können.

Ein typisches Beispiel hierfür ist die zeitdiskrete *VOLTERRA-Reihe*, welche über die sogenannten Volterra-Kerne g_i parametrierbar wird. Für ein deterministisches kausales SISO-System lässt sich diese beispielsweise zu

$$y[t] = g_0 + \sum_{\tau_1=0}^{\infty} g_1[\tau_1]u[t - \tau_1] + \sum_{\tau_1=0}^{\infty} \sum_{\tau_2=\tau_1}^{\infty} g_2[\tau_1, \tau_2]u[t - \tau_1]u[t - \tau_2] + \dots \quad (2.45)$$

notieren, wobei hier die Darstellung mit sogenannten Dreieckkernen angegeben wurde [BU16, Abschn. 7.3.1]. Für ein LTI-System entspricht die Volterra-Reihe offenbar der Darstellung (2.8), sodass die Volterra-Kerne als Verallgemeinerung der Impulsantwort betrachtet werden können. Aus (2.45) wird ersichtlich, dass es sich um eine eher unhandliche nichtparametrische Beschreibungsform nichtlinearer Systeme handelt, weswegen die Volterra-Reihe im weiteren Verlauf der Arbeit nicht weiter betrachtet wird. Der interessierte Leser sei unter anderem auf [Sch80; Rug81; Ste04; BU16] und auf die in diesen Werken enthaltenen Referenzen verwiesen.

Die Beschreibung (2.7) stellt ein System linearer stochastischer Differenzgleichungen dar. Eine naheliegende Möglichkeit diese auf den nichtlinearen Fall zu übertragen ist das sogenannte *NARMAX-Modell*¹²

$$\mathbf{y}[t] = \mathbf{F}(\mathbf{y}[t-1], \dots, \mathbf{y}[t-n_A], \mathbf{u}[t], \dots, \mathbf{u}[t-n_B], \mathbf{e}[t-1], \dots, \mathbf{e}[t-n_C]) + \mathbf{e}[t], \quad (2.46)$$

wobei \mathbf{F} eine allgemeine nichtlineare Funktion darstellt. Ist \mathbf{F} nicht von der Störung \mathbf{e} , bzw. zusätzlich nicht von \mathbf{y} abhängig, dann wird (2.46) in Analogie zum linearen Fall als *NARX-* bzw. *NFIR-Modell* bezeichnet (vgl. Abschnitt 2.4). Nichtlineare Zeitreihenmodelle, für welche oftmals die Akronyme *NAR*, *NMA* und *NARMA* verwendet werden, können in analoger Weise definiert werden. Für die Beschreibung der nichtlinearen Funktion \mathbf{F} werden unter anderem polynomiale [Oza81; BF85; Bil85; BFS88; ZL05; CER01], gebrochen rationale [ZB93; BZ94], exponentielle Ansatzfunktionen [HO81] sowie Wavelets [WB04; BW05] verwendet.

Eine weitere allgemeine Systembeschreibung, die als Verallgemeinerung des LTI-Modells in Prozessform (2.16) betrachtet werden kann, ist das *nichtlineare Zustandsraummodell*

$$\begin{aligned} \mathbf{x}[t+1] &= \mathbf{f}(\mathbf{x}[t], \mathbf{u}[t], \mathbf{v}[t]) \\ \mathbf{y}[t] &= \mathbf{h}(\mathbf{x}[t], \mathbf{u}[t], \mathbf{w}[t]), \end{aligned} \quad (2.47)$$

wobei die Bezeichnungen der Prozessdarstellung (2.16) gelten, und \mathbf{f} sowie \mathbf{h} nichtlineare Funktionen darstellen. Die nichtlineare Zustandsdarstellung wird sowohl im Bereich der Optimalfilterung [Jaz70; AM79] als auch in der Systemidentifikation [ADS04; SWN11; SLD15] angewendet.

Blockorientierte Modelle

Die genannten nichtlinearen Modelle sind sehr allgemeine Beschreibungen nichtlinearer Systeme. Im Bereich der Systemidentifikation sind aber oftmals, z. B. aufgrund vorhandenen physikalischen Vorwissens, speziellere Modellstrukturen von Interesse. Eine dabei häufig anzutreffende Modellklasse sind *blockorientierte Modelle*, wobei *vorwärts-* und *rückgekoppelte* Modellstrukturen unterschieden werden können. Diese lassen sich als Verknüpfungen von LTI-Systemen mit statischen Nichtlinearitäten darstellen. Da diese im Rahmen der Arbeit von besonderem Interesse sind, sollen einige häufig anzutreffende Modellansätze kurz vorgestellt werden.

Die bekannteste vorwärtsgerichtete blockorientierte Struktur ist das *Hammerstein-Wiener-Modell* (HW- oder NLN-Modell). Dieses besteht aus einem dynamischen LTI-System, welchem eine statische Eingangsnichtlinearität \mathbf{f}_u vor- und eine statische Ausgangsnichtlinearität \mathbf{f}_y nachgeschaltet ist (siehe Abb. 2.4 (a)). Die Spezialfälle

¹²Es sei darauf hingewiesen, dass die Abkürzung NARMAX (Nonlinear autoregressive moving average exogenous input) sinngemäß unzutreffend ist, da es sich im Allgemeinen nicht um eine gleitende Mittelwertbildung (MA) des Signals $\mathbf{e}[t]$ handelt. Aufgrund der Analogie zum linearen Fall hat sich diese Bezeichnung im Bereich der nichtlinearen Systemidentifikation aber dennoch durchgesetzt.

bei denen entweder die Eingangs- oder die Ausgangsnichtlinearität der identischen Abbildung entsprechen, werden als *Wiener-Modell* bzw. als *Hammerstein-Modell* bezeichnet. Parallelschaltungen mehrerer Wiener- bzw. Hammerstein-Modelle werden *Projection-Pursuit-* bzw. *Uryson-Modell* genannt [Buc10; BU16].

Das HW-Modell sollte nicht mit dem *Wiener-Hammerstein-Modell* (WH-Modell oder LNL-Modell) verwechselt werden, welches aus zwei LTI-Blöcken und einer zwischengeschalteten statischen Nichtlinearität \mathbf{f} besteht [HK86; SRM09; BU16; ZSS18]. Schematische Darstellungen des HW- bzw. WH-Modells sind in der Abbildung 2.4 dargestellt.

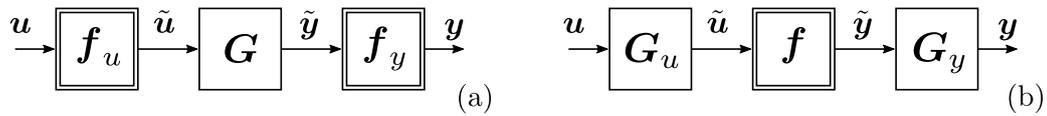


Abbildung 2.4: *Deterministisches Hammerstein-Wiener-Modell (a) bzw. Wiener-Hammerstein-Modell (b). Dabei stellen \mathbf{f}_u , \mathbf{f}_y bzw. \mathbf{f} statische Nichtlinearitäten und $\mathbf{G}(q^{-1})$, $\mathbf{G}_u(q^{-1})$ bzw. $\mathbf{G}_y(q^{-1})$ Übertragungsfunktionsmatrizen entsprechender Dimension dar.*

Im Gegensatz zu vorwärtsgerichteten Modellen enthalten rückgekoppelte blockorientierte Modelle einen oder mehrere Rückkopplungszweige. Hierzu gehört beispielsweise das *LURE-Modell*, welches aus einem LTI-Teilsystem besteht, dessen Ausgang über eine statische Nichtlinearität negativ auf den Eingang des LTI-Blocks rückgekoppelt ist (siehe Abb. 2.5 (a)). Dieses kann jedoch umgekehrt auch als statische Nichtlinearität im Vorwärtszweig, deren Ausgang über das LTI-Teilsystem rückgekoppelt wird, dargestellt werden [SGM08; ST17]. In der Literatur wird das Lure-Modell insbesondere im Rahmen der Stabilitätsanalyse [Kal63; WIO98; NMI19] aber auch der Systemidentifikation [SGM08; SPP19] ausführlich diskutiert.

Eine sehr allgemeine rückgekoppelte blockorientierte Struktur, die bis auf das HW-Modell alle genannten blockorientierten Ansätze als Spezialfälle enthält, stellt das „*Linear fractional representation*“-Modell (LFR-Modell) dar (siehe Abbildung 2.5 (b)). Das LFR-Modell wird unter anderem in den Referenzen [Sin80; Van96; VS99; MSV13] diskutiert.

Man beachte, dass sich die genannten blockorientierten Modellstrukturen auf den Fall deterministischer Systeme beziehen. In der praktischen Anwendung treten jedoch immer Störungen auf, sodass oftmals eine stochastische Betrachtungsweise sinnvoll ist. Im Rahmen der Diskussion verschiedener Identifikationsansätze lassen sich dabei verschiedene Typen stochastischer HW- bzw. WH-Modelle untersuchen, welche sich durch die Stelle des Eingriffs und die stochastischen Eigenschaften der Störung unterscheiden. Eine Auswahl unterschiedlicher in der Literatur betrachteter Fälle ist in der Tabelle 2 angegeben. Für eine ausführliche Darstellung von Identifikationsansätzen für blockorientierte Modelle sei unter anderem auf [GB10; ST17] verwiesen.

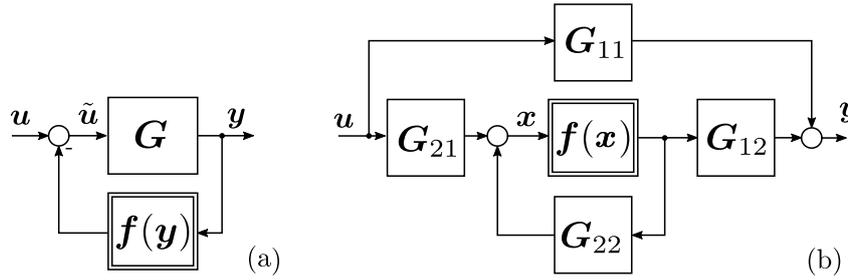


Abbildung 2.5: Deterministisches Lure- (a) bzw. LFR-Modell (b). Dabei stellen f eine statische Nichtlinearität und $G(q^{-1})$ bzw. $G_{ij}(q^{-1})$ ($i, j \in \{1, 2\}$) Übertragungsfunktionsmatrizen entsprechender Dimension dar.

Referenz	Modell	Störung	LTI	Methode
[RZV05]	W-MIMO	-	FIR	Subspace
[GPS05]	H-MIMO	add. weißes Prozess- und Messrauschen	ZRM	Subspace
[Bai02]	H-SISO	add. nicht spezifizierte Störung nach G	Ü-Fkt.	MkQ, Korrelationsanal.
[WV96]	W-MIMO	add. lin. stoch. Prozess nach f_y	ZRM	Subspace
[Wig93]	W-SISO	-	Ü-Fkt.	Rek. PEM
[PZR19]	H-SISO	add. lin. stat. stoch. Prozess nach G	Ü-Fkt.	PEM
[HLW08]	W-SISO	add. lin. stat. stoch. Prozesse vor und nach f_y	Ü-Fkt.	ML
[WSL13]	HW-MIMO	add. lin. stat. stoch. Prozesse vor und nach f_y	beliebig	ML
[HK86]	WH-SISO	add. stat. stoch. Prozesse nach G_y und am Prozesseingang u	-	-
[ZSS18]	WH-SISO	add. stat. stoch. Prozesse nach G_y sowie vor oder nach f	-	-
[SPP19]	Lure	add. weißes Messrauschen	ZRM	PEM

Tabelle 2: Übersicht der in der Literatur diskutierten blockorientierten Modellstrukturen und der dabei verwendete Störgrößenansatz. Die Spalte LTI und Methode geben den zur Beschreibung des LTI-Subsystems verwendeten Modellansatz und die vorgeschlagene Identifikationsmethode an (sofern in der Referenz angegeben).

Optimale Prädiktion für nichtlineare Modellstrukturen

Das Konzept der optimalen Prädiktion ist sowohl für die Identifikation als auch für den Entwurf von Optimalregelungen von besonderem Interesse. Im Gegensatz zu stochastischen LTI-Systemen existiert im Falle nichtlinearer Systeme der Form (2.46) bzw.

(2.47) keine allgemeingültige explizite¹³ Darstellung des optimalen Prädiktors der Form (2.25). Die im folgenden Kapitel diskutierte und für den weiteren Verlauf der Arbeit grundlegende Identifikation mittels PEM ist für nichtlineare Modellstrukturen daher meist nicht direkt implementierbar. In diesem Fall kann jedoch auf numerische Ansätze, wie beispielsweise *Sequentielle Monte-Carlo-Verfahren* (SMC-Verfahren) zurückgegriffen werden [SWN11; SLD15].

Obwohl für nichtlineare Systeme im Allgemeinen keine explizite Darstellung des optimalen Prädiktors existiert, gibt es Spezialfälle, für welche ein entsprechender geschlossener Ausdruck angegeben werden kann. Für ein beliebiges deterministisches System, dessen Ausgang additiv mit einem Störsignal $\mathbf{v}[t] = \mathbf{H}(q^{-1})\mathbf{e}[t]$ überlagert ist (siehe Abbildung 2.6 (a)), kann die optimale Prädiktion in Analogie zu (2.25) hergeleitet werden. Ist das Modell beispielsweise in der Form

$$\begin{aligned} \mathbf{x}[t+1] &= \mathbf{f}(\mathbf{x}[t], \mathbf{u}[t]) \\ \mathbf{y}[t] &= \underbrace{\mathbf{h}(\mathbf{x}[t], \mathbf{u}[t])}_{\tilde{\mathbf{y}}[t]} + \mathbf{H}(q^{-1})\mathbf{e}[t], \end{aligned} \quad (2.48)$$

gegeben, dann ergeben sich die Beziehungen

$$\hat{\mathbf{y}}[t|t-k] = \mathbf{H}^{[0,k-1]} \mathbf{H}^{-1} \tilde{\mathbf{y}}[t] + (\mathbf{I} - \mathbf{H}^{[0,k-1]} \mathbf{H}^{-1}) \mathbf{y}[t] \quad \text{bzw.} \quad (2.49)$$

$$\boldsymbol{\varepsilon}[t|t-k] = \mathbf{H}^{[0,k-1]} \mathbf{H}^{-1} (\mathbf{y}[t] - \tilde{\mathbf{y}}[t]), \quad (2.50)$$

wobei $\tilde{\mathbf{y}}[t]$ den Ausgang des deterministischen Teilsystems kennzeichnet. Man beachte, dass der Modellansatz (2.48) impliziert, dass auf das nichtlineare Teilsystem keine Störungen einwirken, was aus praktischer Sicht eine möglicherweise einschränkende Annahme darstellt.

Ein weiterer nützlicher Modellansatz besteht aus einem gewöhnlichen stochastischen LTI-Modell, dessen Ausgang über eine statische Nichtlinearität $\mathbf{f}(\mathbf{u}[t], \mathbf{y}[t])$ auf den Eingang des LTI-Teilsystems rückgekoppelt ist¹⁴:

$$\mathbf{y}[t] = \underbrace{q^{-d} \bar{\mathbf{G}}(q^{-1})}_{\mathbf{G}(q^{-1})} \mathbf{f}(\mathbf{u}[t], \mathbf{y}[t]) + \mathbf{H}(q^{-1})\mathbf{e}[t] \quad (d \geq 1) \quad (2.51)$$

Man beachte, dass dieser Modellansatz formal der Situation der Prädiktion im geschlossenen Regelkreis entspricht. Entsprechend der in Bemerkung 2.8 dargelegten

¹³Damit ist gemeint, dass der optimale Prädiktor $\hat{\mathbf{y}}[t|t-k]$ nur von $\mathbf{u}[t]$ und $\mathbf{y}[t-k]$ abhängig ist (siehe Bedingung (A1) in Abschnitt 2.3). Einige Identifikationsalgorithmen hingegen basieren auf impliziten Darstellungen, die zusätzlich von dem unbekanntem Störsignal \mathbf{e} abhängig sind (Vergleiche Abschn. 3.3.3). Wird beispielsweise das NARMAX-Modell (2.46) zugrunde gelegt, dann lässt sich die implizite Form des 1-Schritt Prädiktors

$$\hat{\mathbf{y}}[t|t-1] = \mathbf{F}(\mathbf{y}[t-1], \dots, \mathbf{y}[t-n_A], \mathbf{u}[t], \dots, \mathbf{u}[t-n_B], \mathbf{e}[t-1], \dots, \mathbf{e}[t-n_C])$$

direkt aus der Modellgleichung ablesen.

¹⁴Da der Systemausgang $\mathbf{y}[t]$ auf der rechten Seite der Gl. (2.51) vorkommt, kann $\mathbf{G}(q^{-1})\mathbf{f}(\mathbf{u}[t], \mathbf{y}[t])$ nicht als lineare Filterung betrachtet werden. Stattdessen ist das Modell (2.51) als herrührende stochastische Differenzgleichung zu interpretieren.

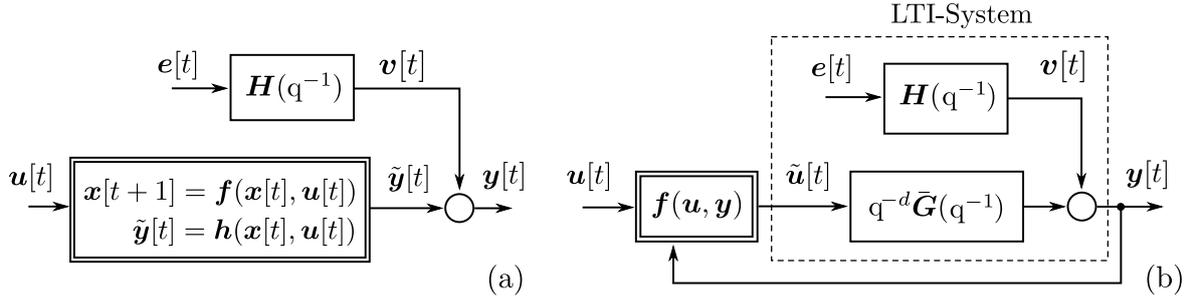


Abbildung 2.6: Schematische Darstellung der durch (2.48) (a), bzw. (2.51) (b) beschriebenen Systeme.

Argumentation, kann die optimale Prädiktion und die Prädiktionsfehlerfolge für $d \geq k$ deshalb zu

$$\hat{\mathbf{y}}[t|t-k] = \mathbf{q}^{-d} \mathbf{H}^{[0,k-1]} \mathbf{H}^{-1} \bar{\mathbf{G}} \mathbf{f}(\mathbf{u}[t], \mathbf{y}[t]) + (\mathbf{I} - \mathbf{H}^{[0,k-1]} \mathbf{H}^{-1}) \mathbf{y}[t] \quad (2.52)$$

$$\boldsymbol{\varepsilon}[t|t-k] = \mathbf{H}^{[0,k-1]} \mathbf{H}^{-1} (\mathbf{y}[t] - \mathbf{q}^{-d} \bar{\mathbf{G}} \mathbf{f}(\mathbf{u}[t], \mathbf{y}[t])) \quad (2.53)$$

angegeben werden.

Obwohl die Forderung $d \geq k$ restriktiv erscheint, kann die vorgeschlagene Modellstruktur praktisch sehr nützlich sein. Dies ist insbesondere dann der Fall, wenn es möglich ist ein virtuelles Eingangssignal $\tilde{\mathbf{u}}[t] = \mathbf{f}(\mathbf{u}[t], \mathbf{y}[t])$ zu konstruieren, für welches ein näherungsweise lineares Übertragungsverhalten auf den Systemausgang erwartet werden kann. Müssen zudem Störungen innerhalb der Rückkopplungsschleife berücksichtigt werden, welche durch lineare stochastische Prozesse beschrieben werden können, dann ist es naheliegend die Methoden bzw. Ideen für die Parametrierung einer geeigneten Modellstruktur aus der linearen Theorie zu übertragen. Saisonales oder instationäres Verhalten kann daher durch ein entsprechend parametrisiertes Filter $\mathbf{H}(q^{-1})$ berücksichtigt werden (siehe Abschnitt 2.4.1 und Beispiel 3.3). Ein weiterer Vorteil der in der Abbildung 2.6 (b) dargestellten Modellstruktur ist, dass die optimale Prädiktion bzw. die Prädiktionsfehlerfolge entsprechend der Gleichungen (2.52) bzw. (2.53) effizient durch lineare Filterung erzeugt werden können. Die Anpassung der für LTI-Systeme üblichen Identifikationsalgorithmen ist daher mit relativ geringem Aufwand möglich.

Bemerkung 2.12 (Mehrschrittprädiktion im Fall $k > d$): In den meisten praktischen Fällen werden Modellansätze mit $d = 1$ gewählt. Dies bedeutet wiederum, dass diese zumindest mit der 1-Schritt-PEM identifiziert werden können (siehe Kapitel 3.2). Soll das Modell jedoch zur Mehrschrittprädiktion verwendet werden, dann muss beachtet werden, dass im ersten Summanden der Gleichung (2.52) Messwerte des Ausgangssignals \mathbf{y} vorkommen, welche zum Prädiktionszeitpunkt $t - k$ noch unbekannt sind. Eine naheliegende Möglichkeit besteht daher darin die zukünftigen Messwerte $\mathbf{y}[\tau]$ durch die prädizierten Werte $\hat{\mathbf{y}}[\tau]$ zu ersetzen ($\tau = t - k + d, \dots, t$). Es sollte jedoch beachtet werden, dass die Prädiktion dann im Sinne des MSE im Allgemeinen nicht optimal ist. \triangleleft

Kapitel 3

Identifikation dynamischer Systeme

Das Problem der experimentellen bzw. datengetriebenen Ermittlung relevanter Systemeigenschaften, als welche die System- bzw. Prozessidentifikation im weitesten Sinne verstanden werden kann, ist eine historisch bereits lange bestehende Problemstellung. Die mathematischen Grundlagen dieser Verfahren, gehen mindestens bis zu den Arbeiten von *CARL FRIEDRICH GAUSS* am Ende des 18. und Anfang des 19. Jahrhunderts zurück. Hier ist insbesondere die von ihm begründete *Methode der kleinsten Fehlerquadrate*¹ (MkQ) zu nennen, welche die mathematische Grundlage der meisten modernen Identifikations- und Optimierungsverfahren darstellt.

Seit Beginn der 1960er Jahre hat sich insbesondere durch die verstärkte Entwicklung digitaler Rechentechnik, die Nutzung von Identifikationsverfahren in automatisierungstechnischen Anwendungen etabliert (siehe unter anderem [BB63; Eyk63; Ast64; AB65]). Das daraus entstandene Fachgebiet wird insbesondere in den ingenieurwissenschaftlichen Disziplinen der Regelungs- und Automatisierungstechnik als System- oder Prozessidentifikation bezeichnet.

Die theoretischen Grundlagen und die mathematischen Verfahren überschneiden sich jedoch stark mit denen der *Zeitreihenanalyse*, deren Anwendung parallel zur Entwicklung der Theorie stochastischer Prozesse bereits in den 1920er und 1930er Jahren begann (siehe z. B. [Yul27] und [Wol38]). Im Gegensatz zur Systemidentifikation, welche sich im Wesentlichen auf technische Applikationen bezieht, werden innerhalb der Zeitreihenanalyse auch Anwendungen in der Ökonometrie, Astronomie, Biologie und Meteorologie diskutiert. Zudem sei darauf hingewiesen, dass sich die Bezeichnungen und die verwendete Notation in den Referenzwerken der Systemidentifikation [SS89; Lju12] und der Zeitreihenanalyse [BJR08] teilweise sehr stark unterscheiden.

3.1 Klassifikation von Identifikationsverfahren für LTI-Systeme

Aufgrund der langen Historie und der vielfältigen Anwendungsgebiete ist die moderne Systemidentifikation ein vergleichsweise unübersichtliches wissenschaftliches Themenfeld geworden. In der Vergangenheit wurden eine Vielzahl unterschiedlicher, teilweise eng

¹Least squares method (LS-method)

miteinander verwandter Algorithmen mit nicht einheitlicher Bezeichnung und sich stark überschneidenden Akronymen entwickelt. In diesem Abschnitt soll daher auf die wesentlichen Identifikationsverfahren für zeitdiskrete LTI-Systeme und deren grundlegenden Vor- und Nachteile eingegangen werden.

Eine auf der verwendeten Modellstruktur basierende Klassifikation der Identifikationsalgorithmen ist in der Abbildung 3.1 dargestellt. Dabei wird zunächst grundlegend in parametrische sowie nichtparametrische Methoden unterschieden.

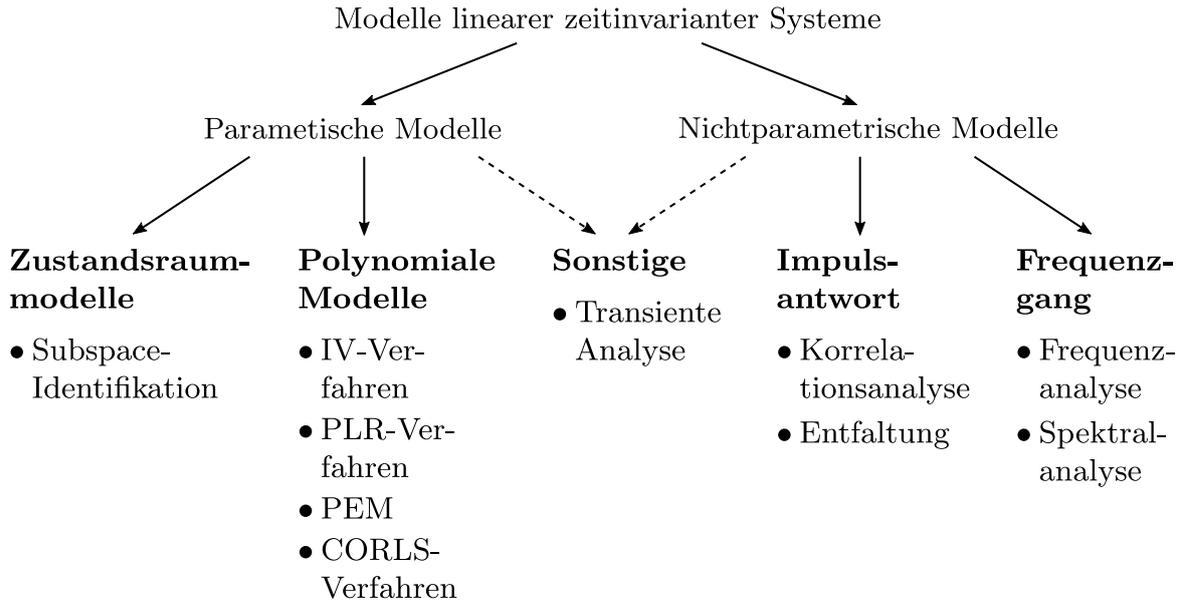


Abbildung 3.1: Auf der verwendeten Modellstruktur basierende Klassifikation typischer Identifikationsverfahren für zeitdiskrete LTI-Systeme

3.1.1 Nichtparametrische Identifikationsverfahren

Die nichtparametrische Identifikation umfasst im Wesentlichen Verfahren zur Bestimmung von Frequenzgängen, Impulsantworten oder sonstiger praktisch relevanter Systemcharakteristiken. Hierzu gehören beispielsweise Zeitkonstanten, Verstärkungen oder Totzeiten, welche mithilfe von Experimenten durch Messung der Impuls- oder der Sprungantwort bestimmt werden (siehe [Rak80] und [LW14, Abschn. 9.3]). Obwohl aus den entsprechenden Verfahren oftmals parametrische Modelle, wie zum Beispiel zeitkontinuierliche Übertragungsfunktionen gewonnen werden, sind die Methoden der *transienten Analyse* in der Literatur meist den nichtparametrischen Identifikationsverfahren zugeordnet [SS89, Abschn. 3.2].

In Analogie zur transienten Analyse, die auf der Messung der Impuls- bzw. der Sprungantwort basiert, werden bei der *Frequenzanalyse* harmonische Anregungssignale verwendet. Dabei wird die Tatsache ausgenutzt, dass ein LTI-System bei harmonischer Anregung mit einer bestimmten Kreisfrequenz ω_i und zugehöriger Amplitude $A_u(\omega_i)$ durch ein um

die Phase $\varphi_y(\omega_i)$ verschobenes und verstärktes Ausgangssignal mit Amplitude $A_y(\omega_i)$ antwortet, sodass aus dem Amplitudenverhältnis und der entsprechenden Phasenverschiebung eine Näherung des Frequenzgangs für verschiedene ω_i berechnet werden kann. Da dieses Verfahren oftmals empfindlich auf Störungen reagiert, wurden in der Vergangenheit ebenfalls korrelationsbasierte Methoden zur Störunterdrückung entwickelt (siehe [Lju12, Abschn. 6.2] und [Rak80]).

Die *Fourier- bzw. Spektralanalyse*, welche aus statistischen Methoden der Spektraldichteschätzung hervorgegangen ist, steht in enger Beziehung zur Frequenzanalyse. Dabei ist insbesondere der *empirische Frequenzgangschätzer*

$$\hat{G}(e^{-j\omega}) = \frac{\hat{\phi}_{yu}[\omega]}{\hat{\phi}_u[\omega]} = \frac{Y_N[\omega]}{U_N[\omega]} \quad (3.1)$$

zu nennen, welcher im englischen Sprachraum als *ETFE (Empirical transfer function estimate)* bezeichnet wird. Dabei kennzeichnen $U_N[\omega]$ und $Y_N[\omega]$ die diskrete Fouriertransformation (DFT) des Ein- bzw. Ausgangssignals sowie $\hat{\phi}_u$ und $\hat{\phi}_{yu}$ entsprechende Schätzungen des Auto- bzw. Kreuzleistungsdichtespektrums (siehe Anhang A.2.3). Hierbei ist zu erwähnen, dass (3.1) im Allgemeinen² ein asymptotisch erwartungstreuer, allerdings inkonsistenter Schätzer ist [Lju12, Abschn. 6.3]. In der Praxis führt die Inkonsistenzproblematik meist zu stark „verrauschten“ Schätzungen des Frequenzgangs, weswegen oftmals zusätzliche Fensterfunktionen zur Schätzung der Leistungsdichtespektren verwendet werden [SM05, Kap. 2]. Eine weitere häufig verwendete Methode zur Vermeidung der Inkonsistenzproblematik ist die Mittelung mehrerer mithilfe separater Datensätze geschätzter Frequenzgänge $\hat{G}_i(e^{j\omega})$, wobei insbesondere unerwünschte Leckeffekte („spectral leakage“) bei der Berechnung der Fourier-Transformierten beachtet werden sollten [PKJ11].

Weitere in der Literatur häufig anzutreffende Methoden der nichtparametrischen Identifikation befassen sich mit der Ermittlung der Impulsantwort (2.8), wobei die Impulsfolge $\mathbf{g}[t]$ durch ein FIR-Modell endlicher Ordnung approximiert wird:

$$\mathbf{y}[t] = \sum_{i=0}^n \mathbf{g}[i] \mathbf{u}[t - i] \quad (3.2)$$

Da diese parameterlineare Modellstrukturen darstellen, können die Parameter mittels der Eingangs- und Ausgangsfolgen durch Lösung eines gegebenenfalls überbestimmten linearen Gleichungssystems im Sinne der Methode der kleinsten Fehlerquadrate berechnet werden, was in der deutschsprachigen Literatur oftmals als *Entfaltung* bezeichnet wird (siehe u. a. [Ise92a, Abschn. 6.2] und [BU16, Abschn. 2.2.3]). Zudem ist insbesondere die *Korrelationsanalyse* zu nennen, die prinzipiell auf dem gleichen Ansatz beruht, bei welcher jedoch die Schätzungen der Auto- bzw. Kreuzkorrelationsfolgen $\hat{r}_{yu}[\tau]$ bzw. $\hat{r}_u[\tau]$ verwendet werden (siehe [IM11, Abschn. 7.2] und [SS89, Abschn. 3.4]).

²Hiermit ist der Fall gemeint, dass $u[t]$ eine nicht endliche Anzahl an Frequenzen enthält. Für den Fall der harmonischen Anregung mit einer Frequenz ω_i ist (3.1) ein sowohl erwartungstreuer als auch konsistenter Schätzer für $\hat{G}(e^{j\omega_i})$.

Der wesentliche Nachteil bei der Impulsfolgeschätzung besteht darin, dass für die Approximation in der Regel sehr viele Parameter $\mathbf{g}[i]$ benötigt werden. Eine Möglichkeit dies zu vermeiden, ist es in (3.2) an Stelle der $\mathbf{u}[t - i]$ alternative Aufbaufunktionen zu verwenden. Diese gewinnt man meist durch geeignete Filterung der Eingangssignale, wobei oftmals *LAGUERRE-* oder *KAUTZ-Filter* verwendet werden [Wah91; Wah94].

Obwohl nichtparametrische Methoden insbesondere im Bereich der Nachrichtentechnik und in der Signalverarbeitung weit verbreitet sind, besitzen sie für regelungstechnische Anwendungen oftmals erhebliche Nachteile. Zum einen führt die meist große Anzahl an Parametern, die generell dem Prinzip der Sparsamkeit (*parsimony principle*) widerspricht, zu Schätzungen mit hoher Varianz und somit zu ungenauen Modellen. Zum anderen ist zu beachten, dass viele Verfahren wie beispielsweise der empirische Frequenzgangschätzer zu verzerrten und somit auch inkonsistenten Schätzungen im geschlossenen Regelkreis führen (siehe [Lju12, Abschn. 13.4] und [SS89, Kap. 10]). Weiterhin sei angemerkt, dass bei den meisten nichtparametrischen Verfahren die Störung nicht explizit modelliert wird. Stattdessen wird oftmals ein allgemeines Störsignal $\mathbf{v}[t]$ mit nicht genau spezifizierten stochastischen Eigenschaften angenommen, sodass das Prädiktorfilter (2.25), welches in vielen Anwendungsfällen von Interesse ist, nicht berechnet werden kann.

3.1.2 Parametrische Identifikationsverfahren

Die genannten Nachteile nichtparametrischer Verfahren, die gute Verfügbarkeit leistungsfähiger Identifikationsalgorithmen sowie die Tatsache, dass die meisten Reglerentwurfverfahren parametrische Modelle voraussetzen, sind Gründe dafür, dass in der Regelungstechnik üblicherweise parametrische Methoden bevorzugt werden.

Die auf der Darstellung des Systems als Zustandsraummodell beruhenden Identifikationsmethoden gehen auf die Realisierungstheorie zurück [Sch00]. Dabei sind insbesondere die unterraumbasierten Verfahren der *Subspace-Identifikation* hervorzuheben. Diese sind vergleichsweise moderne Verfahren, die verstärkt seit den 90er Jahren entwickelt wurden [OM93; OM94; OM95; OM96], deren Anfänge jedoch bis in die 80er Jahre zurückreichen [Lar83]. Dabei handelt es sich um Schätzverfahren, die im Wesentlichen auf geometrischen Überlegungen und der Anwendung der Methoden der linearen Algebra, insbesondere der Singulärwertzerlegung, basieren. Besondere Vorteile bestehen unter anderem darin, dass es sich um nicht iterative Verfahren handelt, die zusätzlich zur Schätzung der Systemmatrizen auch eine Schätzung der Systemordnung liefern. Weiterhin sei erwähnt, dass sie aufgrund des verwendeten Modellansatzes besonders für die Identifikation von Mehrgrößensystemen geeignet sind. Da die Subspace-Identifikation innerhalb dieser Arbeit nur eine untergeordnete Bedeutung hat, sei für einen ausführlichen deutschsprachigen Überblick auf die Dissertationsschrift [Buc10] verwiesen.

Die in der Praxis am häufigsten eingesetzten Identifikationsverfahren basieren auf der polynomialen Darstellung (2.7), bzw. auf deren Spezialfällen (siehe Tabelle 1). Hier sind insbesondere die *Prädiktionsfehlerminimierung* (PEM) [AB65; Ast70; SS89;

[Lju12], die auf der *pseudolinearen Regression* beruhenden Verfahren (PLR³) und die *Hilfsvariablenverfahren* (IV⁴) zu nennen [Row70; Jak79; SS83a; SS89]. Weiterhin sei darauf hingewiesen, dass für PLR- bzw. IV-Verfahren ebenfalls rekursive Varianten verfügbar sind (siehe [SS89, Kap. 11] und [Lju12, Kap. 11]), welche im englischen Sprachraum ebenfalls mit dem Akronym RLS („*Recursive least squares*“) abgekürzt werden. Zudem existieren auch für parametrische Modelle *korrelationsbasierte LS-Verfahren* (CORLS) (siehe [IM11, Abschn. 9.3.2] und [Nel01, Abschn. 16.5.1]), die auf dem gleichen Ansatz wie die Korrelationsanalyse beruhen, jedoch einen parametrischen Modellansatz verwenden. Auf die ausführliche Beschreibung und die Eigenschaften der auf der polynomialen Darstellung (2.7) aufbauenden Algorithmen wird in den folgenden Abschnitten eingegangen.

3.2 Prädiktionsfehlerminimierung

In vielen Anwendungsfällen wird ein Systemmodell entweder direkt oder indirekt für die Prädiktion des zukünftigen Ausgangssignals eingesetzt. Eine naheliegende Identifikationsidee besteht daher darin, den Prädiktor $\hat{\mathbf{y}}[t|t-k, \boldsymbol{\theta}]$ in der Gleichung (2.25) derart zu parametrieren, dass die gegebenenfalls gefilterte Folge der Prädiktionsfehler

$$\begin{aligned} \boldsymbol{\varepsilon}_L[t|t-k, \boldsymbol{\theta}] &= \mathbf{L}(q^{-1})\boldsymbol{\varepsilon}[t|t-k, \boldsymbol{\theta}] \\ &= \mathbf{L}(q^{-1})(\mathbf{y}[t] - \hat{\mathbf{y}}[t|t-k, \boldsymbol{\theta}]) \end{aligned} \quad (3.3)$$

in einem gewissen Sinne möglichst klein wird, wobei $\boldsymbol{\theta} \in \mathcal{D}_M$ den gesuchten Parametervektor und

$$\mathcal{D}_M = \left\{ \boldsymbol{\theta} \in \mathbb{R}^{n_\theta} \mid \mathbf{H}^{-1}(q^{-1}, \boldsymbol{\theta}) \text{ und } \mathbf{H}^{-1}(q^{-1}, \boldsymbol{\theta})\mathbf{G}(q^{-1}, \boldsymbol{\theta}) \text{ sind asymptotisch stabil} \right\} \quad (3.4)$$

die „erlaubte Parametermenge“ beschreibt (siehe Bem. 2.2). Dabei können durch die Wahl eines geeigneten Vorfilters $\mathbf{L}(q^{-1})$ bestimmte Frequenzen der Prädiktionsfehlerfolge gedämpft bzw. betont werden, sodass eine Anpassung des Modells an vorgegebene Frequenzbereiche ermöglicht wird (siehe Abschnitt 3.2.2).

Bei der PEM erfolgt die Schätzung des Parametervektors $\boldsymbol{\theta}$ durch Minimierung eines Kostenfunktional $V_N(\boldsymbol{\theta})$, welches die Folge der gegebenenfalls gefilterten Prädiktionsfehler (3.3) auf einen skalaren Wert abbildet, d. h.

³Das Akronym PLR ist abgeleitet von der im englischen Sprachraum gebräuchlichen Bezeichnung „*Pseudo-linear regression*“ [SS84; SS89]. Es sei jedoch darauf hingewiesen, dass in der Literatur eine Vielzahl weiterer Bezeichnungen dieser Gruppe von Algorithmen, wie bspw. *Repeated least squares* (RLS) [Nel01, Kap. 16], *Extended least squares* (ELS) [SSA85], *Multistage least squares* (MLS) [Lju12] sowie *Approximate Maximum-Likelihood* (AML) [Tow94] geläufig sind.

⁴engl.: *Instrumental variable methods*

$$\hat{\boldsymbol{\theta}}_N = \arg \min_{\boldsymbol{\theta}} V_N(\hat{\boldsymbol{\Lambda}}_\varepsilon(\boldsymbol{\theta})). \quad (3.5)$$

Formal beschreibt $V_N : \mathbb{R}^{n_y \times n_y} \rightarrow \mathbb{R}$ dabei eine monoton wachsende Funktion, welche auf der Menge der positiv definiten Matrizen definiert ist⁵. Die symmetrische $n_y \times n_y$ -Matrix

$$\hat{\boldsymbol{\Lambda}}_\varepsilon(\boldsymbol{\theta}) = \frac{1}{N} \sum_{t=1}^N \boldsymbol{\varepsilon}_L[t|t-k, \boldsymbol{\theta}] \boldsymbol{\varepsilon}_L^\top[t|t-k, \boldsymbol{\theta}] \quad (3.6)$$

wird als Stichprobenkovarianzmatrix der (gefilterten) k -Schritt Prädiktionsfehler bezeichnet [Lju76; SS89]. Entsprechend wird der Schätzer (3.5) als k -Schritt bzw. *Mehrschritt-PEM* bezeichnet. Weiterhin sei angemerkt, dass unter Prädiktionsfehlerminimierung in der Literatur typischerweise der Spezialfall der 1-Schritt-PEM verstanden wird.

Ein naheliegendes und praktisch häufig eingesetztes Optimierungskriterium ist

$$V_N(\boldsymbol{\theta}) = \text{tr}(\mathbf{W} \hat{\boldsymbol{\Lambda}}_\varepsilon(\boldsymbol{\theta})), \quad (3.7)$$

wobei $\mathbf{W} \in \mathbb{R}^{n_y \times n_y}$ eine symmetrische positiv definite Wichtungsmatrix darstellt. Im oftmals verwendeten Spezialfall einer Diagonalmatrix \mathbf{W} wird dann die Summe der Prädiktionsfehlerquadrate minimiert. Ein weiteres Kostenfunktional, welches insbesondere bei der theoretischen Analyse der PEM Anwendung findet, ist

$$V_N(\boldsymbol{\theta}) = \det(\hat{\boldsymbol{\Lambda}}_\varepsilon(\boldsymbol{\theta})) \quad (3.8)$$

(siehe Unterabschnitt Maximum-Likelihood-Methode).

Bemerkung 3.1 (Startindex der Summe (3.6)): Für die Berechnung der Prädiktionsfehler $\boldsymbol{\varepsilon}[1], \dots, \boldsymbol{\varepsilon}[N]$ werden im Allgemeinen Realisierungen von $\mathbf{u}[t]$ und $\mathbf{y}[t]$ für $t < 1$ benötigt, welche in der Praxis jedoch nicht verfügbar sind. Eine naheliegende Möglichkeit mit dieser Problematik umzugehen, ist es die unbekanntenen Werte als zusätzliche Optimierungsparameter aufzufassen, was aber zu einem unnötig hohen Rechenaufwand bei der Optimierung (3.5) führt. Die praktisch zweckmäßigere Methode besteht darin, die Summe in (3.6) erst ab einem geeignet gewählten Zeitpunkt $t_s \geq 1$ auszuwerten. Dieser wird meist derart gewählt, dass in der Differenzgleichung (2.27) keine Werte $\mathbf{u}[t]$ und $\mathbf{y}[t]$ mit $t < 1$ vorkommen. Wird beispielsweise ein Eingrößensystem der Form (2.36) zugrunde gelegt, dann ist dies genau dann der Fall, wenn

$$t_s = \max(n_A + n_F, n_B) + n_D + k \quad (3.9)$$

gewählt wird. ◀

⁵Konkret bedeutet dies, dass die Ungleichung $f(\mathbf{Q} + \Delta\mathbf{Q}) \geq f(\mathbf{Q})$ für alle positiv definiten \mathbf{Q} und positiv semidefiniten $\Delta\mathbf{Q}$ erfüllt sein muss, wobei Gleichheit lediglich für $\Delta\mathbf{Q} = \mathbf{0}$ gelten darf [SS89].

Maximum-Likelihood-Methode

Es existiert ein wichtiger Zusammenhang der PEM zur Maximum-Likelihood-Methode (ML-Methode): Unter Annahme eines normalverteilten weißen Rauschens ($\mathbf{e}[t] \sim \mathcal{N}(\mathbf{0}, \mathbf{\Lambda}_e) \forall t$) kann gezeigt werden, dass der Schätzer (3.5) für $k = 1$ mit der Kostenfunktion (3.8) ebenfalls die Likelihood-Funktion maximiert (siehe [Lju76, Abschn. 2.4], [Ast79] und [SS89, Abschn. 7.4]), weswegen die Prädiktionsfehlerminimierung mit $k = 1$ oftmals auch als *Maximum-Likelihood-Methode* bezeichnet wird⁶.

Einfluss des Vorfilters $L(q^{-1})$

In der praktischen Anwendung kann es manchmal naheliegend sein, Störungen in den Signalen $\mathbf{u}[t]$ und $\mathbf{y}[t]$ durch geeignete Filterung zu unterdrücken. Dabei ist zu beachten, dass ein wichtiger Zusammenhang zwischen der Vorfilterung der Daten, der Wahl des Residuenfilters $L(q^{-1})$ und dem stochastischen Teilmodell $\mathbf{H}(q^{-1})$ besteht: Im Fall eines skalaren Vorfilters $L(q^{-1})$ lässt sich durch Einsetzen der Beziehung (2.27) in (3.3) leicht verifizieren, dass die Filterung der Prädiktionsfehler $\varepsilon[t|t-k]$ der Vorfilterung der Signale \mathbf{u} und \mathbf{y} entspricht (vgl. (2.27)):

$$\begin{aligned} \varepsilon_L[t|t-k, \boldsymbol{\theta}] &= L(q^{-1})\varepsilon[t|t-k, \boldsymbol{\theta}] \\ &= \mathbf{H}^{[0,k-1]} \mathbf{H}^{-1} \left(L(q^{-1})\mathbf{y}[t] - \mathbf{G}L(q^{-1})\mathbf{u}[t] \right) \end{aligned} \quad (3.10)$$

Weiterhin folgt aus obiger Betrachtung, dass Vorfilterung mit $L(q^{-1})$ dem stochastischen Modell $\bar{\mathbf{H}}(q^{-1}) = \mathbf{H}(q^{-1})/L(q^{-1})$ entspricht. Mit anderen Worten kann das inverse Vorfilter deshalb als fester Anteil des stochastischen Teilsystems betrachtet werden. Der in Abschnitt 2.4 erwähnte Spezialfall eines Modells mit integrierten Rauschen kann daher durch ein Vorfilter $L(q^{-1}) = 1 - q^{-1}$, d. h. durch Differenziation der Ein- und Ausgangssignale erzwungen werden.

3.2.1 Implementierungsaspekte von PEM-Verfahren

In diesem Abschnitt sollen einige für die numerische Implementierung des PEM-Verfahrens besonders wichtige Aspekte diskutiert werden. Dabei wird nachfolgend die polynomiale Parametrierung (2.36) und die Verwendung des Gütekriteriums

$$V_N(\boldsymbol{\theta}) = \frac{1}{N} \sum_{t=1}^N \varepsilon^2[t|t-k, \boldsymbol{\theta}] \quad (3.11)$$

zugrunde gelegt.

⁶Genau genommen wird in diesem Fall aufgrund der als unbekannt anzunehmenden Anfangswerte des Residuenfilters die *bedingte* Likelihood-Funktion maximiert. Für die Maximierung der *exakten* Likelihood-Funktion müssen die Anfangswerte des Residuenfilters als zusätzliche Optimierungsparameter betrachtet werden.

Numerische Lösung des Optimierungsproblems

Die optimale Prädiktion $\hat{y}[t|t-k, \boldsymbol{\theta}]$ ist im Allgemeinen eine nichtlineare Funktion der Modellparameter $\boldsymbol{\theta}$ (siehe Gl. (2.25) und Tabelle 1). Das durch (3.11) beschriebene Optimierungsproblem stellt aufgrund der Forderung $\boldsymbol{\theta} \in \mathcal{D}_{\mathcal{M}}$ (siehe Gl. (3.4)) ein beschränktes nichtlineares kleinste Quadrate Problem (Nonlinear least squares – NLS) dar. Da jedoch der Rand von $\mathcal{D}_{\mathcal{M}}$ dem Stabilitätsgebiet des optimalen Prädiktors entspricht und das Kostenfunktional $V_N(\boldsymbol{\theta})$ außerhalb von $\mathcal{D}_{\mathcal{M}}$ ohnehin stark anwächst, kann bei der Implementierung meist auf die im Allgemeinen nichtlinearen Nebenbedingungen verzichtet werden⁷. Bei der Lösung des Optimierungsproblems ist es daher zumeist ausreichend, Algorithmen aus dem Bereich der unbeschränkten Optimierung anzuwenden (siehe u. a. [BV04, Kap. 9]). Dabei wird der Schätzwert des Parametervektors durch eine Iterationsvorschrift der Form

$$\hat{\boldsymbol{\theta}}_{(i+1)} = \hat{\boldsymbol{\theta}}_{(i)} - \alpha_{(i)} \boldsymbol{\delta}_{(i)} \hat{\boldsymbol{\theta}}_{(i)} \quad (3.12)$$

berechnet, wobei $\boldsymbol{\delta}_{(i)} \in \mathbb{R}^{n_{\theta}}$ die Suchrichtung und $\alpha_{(i)} > 0$ Schrittweite im Iterationsschritt i darstellen. Man beachte dabei auch, dass durch geeignete Wahl der Schrittweite die Bedingung $\hat{\boldsymbol{\theta}}_{(i+1)} \in \mathcal{D}_{\mathcal{M}}$ stets sichergestellt werden kann.

Für die Wahl der Suchrichtung existieren unter anderem die folgenden Möglichkeiten, wobei $V'_{(i)}$ und $V''_{(i)}$ den Gradienten und die HESSE-Matrix des Kostenfunktionals V_N im Punkt $\hat{\boldsymbol{\theta}}_{(i)}$ kennzeichnen:

- (1) *Gradientenabstiegsverfahren*: $\boldsymbol{\delta}_{(i)} = V'_{(i)}{}^{\top}$
- (2) *Konjugierte-Gradienten-Verfahren*: $\boldsymbol{\delta}_{(i)} = V'_{(i)}{}^{\top} - \gamma_{(i)} \boldsymbol{\delta}_{(i-1)}$ mit Faktor $\gamma_{(i)} \in \mathbb{R}_{>0}$
- (3) *NEWTON-Verfahren*: $\boldsymbol{\delta}_{(i)} = (V''_{(i)})^{-1} V'_{(i)}{}^{\top}$
- (4) *Quasi-NEWTON-Verfahren*: $\boldsymbol{\delta}_{(i)} = \mathbf{M}_{(i)} V'_{(i)}{}^{\top}$ mit $\mathbf{M}_{(i)} \in \mathbb{R}^{n_{\theta} \times n_{\theta}}$
 - (a) *GAUSS-NEWTON* (GN)
 - (b) *LEVENBERG-MARQUARDT* (LM)
 - (c) *BROYDEN-FLETCHER-GOLDFARB-SHANNO* (BFGS)

Die bei der PEM am häufigsten angewendeten Optimierungsverfahren sind vom Quasi-Newton-Typ. Diese basieren auf dem Newton-Verfahren, berechnen die Inverse der Hesse-Matrix jedoch nicht direkt, sondern approximieren diese durch eine geeignet gewählte Matrix $\mathbf{M}_{(i)} \in \mathbb{R}^{n_{\theta} \times n_{\theta}}$. Für das quadratische Kostenfunktional (3.11), ergibt

⁷Bei Verwendung der Darstellung (2.36) kann die asymptotische Stabilität der optimalen Prädiktion für Modelle mit Polynomordnungen $n_C, n_F \leq 2$ durch lineare Nebenbedingungen erzwungen werden. Das Stabilitätsgebiet entspricht dann einem Dreiecksgebiet in der $c^{[1]}, c^{[2]}$ - bzw. $f^{[1]}, f^{[2]}$ -Ebene (siehe [BJR08, Kap. 3]). Für Polynome höherer Ordnung erhält man im Allgemeinen (nichtkonvexe) polynomiale Nebenbedingungen [Gro14, Kap. 2].

sich

$$V'_N(\boldsymbol{\theta}) = \frac{2}{N} \sum_{t=1}^N \frac{\partial \varepsilon_k[t, \boldsymbol{\theta}]}{\partial \boldsymbol{\theta}} \varepsilon_k[t, \boldsymbol{\theta}] \quad (3.13a)$$

$$V''_N(\boldsymbol{\theta}) = \frac{2}{N} \sum_{t=1}^N \left(\frac{\partial \varepsilon_k[t, \boldsymbol{\theta}]}{\partial \boldsymbol{\theta}} \right)^\top \frac{\partial \varepsilon_k[t, \boldsymbol{\theta}]}{\partial \boldsymbol{\theta}} + \frac{\partial^2 \varepsilon_k[t, \boldsymbol{\theta}]}{\partial \boldsymbol{\theta}^2} \varepsilon_k[t, \boldsymbol{\theta}], \quad (3.13b)$$

wobei die verkürzte Schreibweise $\varepsilon_k[t, \boldsymbol{\theta}] := \varepsilon[t|t-k, \boldsymbol{\theta}]$ verwendet wurde.

Im globalen Minimum $\boldsymbol{\theta}_0$ ergibt sich für $\varepsilon[t|t-1]$ näherungsweise weißes Rauschen $e[t]$ (siehe Korollar 2.1). Der k -Schritt Prädiktionsfehler ist dann wegen $\varepsilon_k[t, \boldsymbol{\theta}_0] = \mathbf{H}^{[0,k-1]}(\mathbf{q}^{-1}, \boldsymbol{\theta}_0)e[t]$ nur von $e[t], \dots, e[t-k+1]$ abhängig. Die Ableitungen $\partial^2 \varepsilon_k[t]/\partial \boldsymbol{\theta}^2$ hängen hingegen nur von $e[t]$ mit $t \leq t-k$ ab, sodass der zweite Summand von $V''_N(\boldsymbol{\theta})$ in der Umgebung des Optimums $\boldsymbol{\theta}_0$ vernachlässigt werden kann. Wird $V''_N(\boldsymbol{\theta})$ in die Iterationsvorschrift des Newton-Verfahrens eingesetzt, dann ergibt sich

$$\hat{\boldsymbol{\theta}}_{(i+1)} = \hat{\boldsymbol{\theta}}_{(i)} - \left(\sum_{t=1}^N \varepsilon_k'^\top[t, \hat{\boldsymbol{\theta}}_{(i)}] \varepsilon_k'[t, \hat{\boldsymbol{\theta}}_{(i)}] \right)^{-1} \sum_{t=1}^N \varepsilon_k'^\top[t, \hat{\boldsymbol{\theta}}_{(i)}] \varepsilon_k[t, \hat{\boldsymbol{\theta}}_{(i)}], \quad (3.14)$$

wobei $\varepsilon_k'[t]$ den Gradient bezüglich $\boldsymbol{\theta}$ kennzeichnet. Die Rekursion (3.14) entspricht dann dem Gauß-Newton-Verfahren (vgl. [BV04, Kap. 9]). Wird der zu invertierenden Matrix zudem der Regularisierungsterm $\gamma_{(i)} \mathbf{I}$ mit dem Dämpfungsparameter $\gamma_{(i)} \in \mathbb{R}_+$ hinzugefügt, dann ergibt sich der Levenberg-Marquardt-Algorithmus [Kel99, Abschn. 3.2.3]. Muss bei der Optimierung zum Beispiel aus Effizienzgründen auf eine Matrixinversion verzichtet werden, dann ist die Verwendung eines alternativen Quasi-Newton-Verfahrens, wie beispielsweise des BFGS-Verfahrens denkbar [Kel99, Kap. 4].

Gradient

Die genannten Optimierungsverfahren setzen die Kenntnis des Gradienten der Prädiktionsfehlerfolge $\varepsilon[t|t-k]$ bezüglich des Parametervektors $\boldsymbol{\theta}$ voraus. Bei der Berechnung von $\varepsilon'[t|t-k]$ ist es zweckmäßig auszunutzen, dass der k -Schritt Prädiktionsfehler als MA-gefilterte 1-Schritt Prädiktionsfehlerfolge

$$\boldsymbol{\varepsilon}[t|t-k, \boldsymbol{\theta}] = \mathbf{H}^{[0,k-1]}(\mathbf{q}^{-1}, \boldsymbol{\theta}) \boldsymbol{\varepsilon}[t|t-1, \boldsymbol{\theta}] \quad (3.15)$$

dargestellt werden kann (2.32). Die Ableitung bezüglich des Parameters θ_i ($i = 1, \dots, n_\theta$) ergibt sich dann zu

$$\frac{\partial \boldsymbol{\varepsilon}[t|t-k]}{\partial \theta_i} = \mathbf{H}^{[0,k-1]}(\mathbf{q}^{-1}) \frac{\partial \boldsymbol{\varepsilon}[t|t-1]}{\partial \theta_i} + \frac{\partial \mathbf{H}^{[0,k-1]}(\mathbf{q}^{-1})}{\partial \theta_i} \boldsymbol{\varepsilon}[t|t-1], \quad (3.16)$$

wobei die Ableitungen $\partial \boldsymbol{\varepsilon}[t|t-1]/\partial \theta_i$ effizient durch lineare Filterung der Ein- und Ausgangssignale $\mathbf{u}[t]$ bzw. $\mathbf{y}[t]$ gewonnen werden können (siehe Anhang B.1). Dabei ist die Ableitung der Filtermatrix $\mathbf{H}^{[0,k-1]}(\mathbf{q}^{-1})$ als elementweise Ableitung der Filterkoeffizienten bezüglich eines Parameters θ_i zu verstehen, d. h.

$$\frac{\partial \mathbf{H}^{[0,k-1]}(\mathbf{q}^{-1})}{\partial \theta_i} = \frac{\partial \mathbf{h}[1]}{\partial \theta_i} \mathbf{q}^{-1} + \dots + \frac{\partial \mathbf{h}[k-1]}{\partial \theta_i} \mathbf{q}^{-k+1}. \quad (3.17)$$

Für die konkrete Berechnung der Ableitungen $\partial \mathbf{h}[j]/\partial \theta_i$ sei auf Anhang B.1 verwiesen.

Anfangswerte des Residuenfilters

Im Allgemeinen stellen das Residuenfilter (2.27) sowie das Gradientenfilter (3.16) asymptotisch stabile Filter mit unendlicher Impulsantwort (Infinite impulse response - IIR) dar. Deren Anfangswerte sind in der Praxis meistens unbekannt und werden daher oftmals zu $\boldsymbol{\varepsilon}[t] \equiv \mathbf{0}$ bzw. $\hat{\mathbf{y}}[t] \equiv \mathbf{y}[1]$ für $t \leq 0$ initialisiert. Dies kann unter Umständen zu unerwünschten, langsam abklingenden transienten Anteilen der geschätzten Prädiktionsfehlerfolge (2.27) führen, welche wiederum die Genauigkeit der Schätzung $\hat{\boldsymbol{\theta}}_N$ negativ beeinflusst. Dies ist insbesondere dann der Fall, wenn $\mathbf{H}^{-1}(z)$ oder $\mathbf{H}^{-1}(z)\mathbf{G}(z)$ Polstellen in der Nähe des Einheitskreises der komplexen Zahlenebene besitzen [Knu94; Knu96]. Um die unerwünschten transienten Anteile zu vermeiden, ist es sinnvoll die Anfangswerte während der Optimierung geeignet zu wählen. Hierfür existieren die folgenden Ansätze:

- (1) Die Initialwerte $\boldsymbol{\varepsilon}[t, \boldsymbol{\theta}]$ ($t \leq 0$) werden als zusätzliche Optimierungsparameter berücksichtigt
- (2) In jeder Iteration werden die Anfangswerte durch
 - (a) Lösung eines unterlagerten MkQ-Problems
 - (b) Backforecasting (BFC)
 bestimmt (siehe Anhang B.2).

Unter der Annahme eines normalverteilten Rauschsignals $\mathbf{e}[t]$ stellt der erste Ansatz die sogenannte „exakte Maximum-Likelihood“-Methode dar. Diese ist aufgrund der statistischen Effizienz aus theoretischer Sicht vorzuziehen [LB79; GHP80; WR86]. Praktisch jedoch führt dies oftmals zu vielen zusätzlichen Optimierungsparametern und somit zu einem deutlich erhöhten Rechenaufwand, weshalb meist eine der letzteren Methoden bevorzugt wird. Dies ist insbesondere dann der Fall, wenn bei der Identifikation mehrere nicht zusammenhängende Datensätze oder Modelle mit hohem Zählergrad des Rauschfilters $\mathbf{H}(q^{-1})$ verwendet werden (z. B. saisonale Modellansätze).

Implementierung eines Multischrittverfahrens

Die Identifikation mittels (3.5) basiert auf der Stichprobenkovarianz der k -Schritt Prädiktionsfehler und wird im Folgenden als k -Schritt oder als *Mehrschritt-PEM* bezeichnet. Bei der Implementierung eines prädiktiven Reglers wird das Kostenfunktional hingegen meist über den gesamten Prädiktionshorizont $k = 1, \dots, k_{\max}$ ausgewertet [Mac00; CB07; BBM17]. Aus der Sicht der Identifikation ist daher die Nutzung der „gemittelten Kovarianzmatrix“

$$\hat{\mathbf{\Lambda}}_{\boldsymbol{\varepsilon}}(\boldsymbol{\theta}) = \frac{1}{Nk_{\max}} \sum_{k=1}^{k_{\max}} \sum_{t=1}^N \boldsymbol{\varepsilon}[t|t-k, \boldsymbol{\theta}] \boldsymbol{\varepsilon}^{\top}[t|t-k, \boldsymbol{\theta}] \quad (3.18)$$

naheliegender, was in der Literatur unter den Schlagworten *Multischritt-PEM* oder *Model predictive control relevant identification (MRI)* diskutiert wird [GPS02; HMT03; ZPV11].

Auf einen theoretischen Vergleich der 1- und der Mehr- bzw. Multischritt-PEM wird im folgenden Unterabschnitt eingegangen.

Bei der Implementierung eines Multischrittverfahrens kann ausgenutzt werden, dass die k -Schritt Prädiktionsfehlerfolge durch MA-Filterung der 1-Schritt Prädiktionsfehler gewonnen werden kann (siehe (3.15)): Die Folgen $\boldsymbol{\varepsilon}[t|t-k] =: \boldsymbol{\varepsilon}_k[t]$ lassen sich für $k = 2, \dots, k_{\max}$ daher rekursiv mittels

$$\begin{aligned} \boldsymbol{\varepsilon}_k[t] &= \underbrace{\boldsymbol{\varepsilon}_1[t] + \mathbf{h}[1] \mathbf{q}^{-1} \boldsymbol{\varepsilon}_1[t] + \dots + \mathbf{h}[k-1] \mathbf{q}^{-k+1} \boldsymbol{\varepsilon}_1[t]}_{\boldsymbol{\varepsilon}_{k-1}[t]} \\ &= \boldsymbol{\varepsilon}_{k-1}[t] + \mathbf{h}[k-1] \boldsymbol{\varepsilon}_1[t-k+1] \end{aligned} \quad (3.19)$$

berechnen, wobei $\mathbf{h}[i]$ ($i = 1, \dots, k_{\max}$) die Koeffizienten der Impulsantwort des Filters $\mathbf{H}(\mathbf{q}^{-1})$ kennzeichnen. Es sei zudem darauf hingewiesen, dass dieser Ansatz in ähnlicher Weise bei der Berechnung der Gradienten (3.16) verwendet werden kann.

3.2.2 Analyse der Mehrschritt-PEM

In diesem Abschnitt werden einige bekannte theoretische Eigenschaften des 1-Schritt PEM-Verfahrens auf den allgemeineren Fall der Mehrschritt-PEM übertragen. Zuletzt werden die Ergebnisse beispielhaft demonstriert.

Identifizierbarkeit und Konsistenz

Im Bereich der Systemidentifikation wird der Begriff der Identifizierbarkeit nicht nur als Eigenschaft einer Modellstruktur \mathcal{M} betrachtet (siehe Anhang A.2.4.1), sondern wird oftmals wesentlich enger gefasst. Dies liegt darin begründet, dass ein System \mathcal{S} nur unter bestimmten Voraussetzungen an die Eigenschaften der Störung $\mathbf{e}[t]$, die Identifikationsmethode \mathcal{I} und die Versuchsbedingungen \mathcal{B} (insbesondere der Wahl des Eingangssignals $\mathbf{u}[t]$) identifizierbar ist.

Für eine genaue Definition der Identifizierbarkeit von stochastischen LTI-Systemen ist es zunächst sinnvoll die Menge

$$\begin{aligned} \mathcal{D}_T(\mathcal{S}, \mathcal{M}) &= \left\{ \boldsymbol{\theta} \in \mathcal{D}_{\mathcal{M}} \mid \mathbf{G}(e^{-j\omega}, \boldsymbol{\theta}) \equiv \mathbf{G}_0(e^{-j\omega}); \right. \\ &\quad \left. \mathbf{H}(e^{-j\omega}, \boldsymbol{\theta}) \equiv \mathbf{H}_0(e^{-j\omega}); \quad \boldsymbol{\Lambda}(\boldsymbol{\theta}) \equiv \boldsymbol{\Lambda}_0; \quad -\pi \leq \omega \leq \pi \right\} \end{aligned} \quad (3.20)$$

einzuführen, wobei für das wahre System \mathcal{S} im Folgenden die Beschreibung (2.7) zugrunde gelegt sei. Diese beschreibt die Menge der Werte von $\boldsymbol{\theta}$ aus der zulässigen Menge $\mathcal{D}_{\mathcal{M}}$, für welche die Frequenzgänge $\mathbf{G}_0(e^{-j\omega})$ und $\mathbf{H}_0(e^{-j\omega})$ sowie die Kovarianzmatrix $\boldsymbol{\Lambda}_0$ des wahren Systems und des Modells übereinstimmen. Anschaulich gesprochen entspricht der Fall $\mathcal{D}_T \neq \emptyset$ daher der Situation, dass die Modellmenge die wahre Systembeschreibung enthält, was im Folgenden vereinfachend mit $\mathcal{S} \in \mathcal{M}$ notiert wird.

Mit der Definition (3.20) stellt sich daher die Frage, ob der PEM-Schätzer $\hat{\boldsymbol{\theta}}_N$ für $N \rightarrow \infty$ mit Wahrscheinlichkeit 1 in die Menge $\mathcal{D}_T \neq \emptyset$ konvergiert, was in der

Literatur oftmals mit dem Begriff der Systemidentifizierbarkeit beschrieben wird. Ist \mathcal{S} *systemidentifizierbar unter \mathcal{I} , \mathcal{B} und \mathcal{M}* ⁸ und besteht die Menge $\mathcal{D}_T(\mathcal{S}, \mathcal{M})$ aus nur einem Punkt, dann wird das System *parameteridentifizierbar unter \mathcal{I} , \mathcal{B} und \mathcal{M}* genannt, was wiederum eine konsistente Schätzung impliziert [SS89, Abschn. 6.4].

Genau genommen stellt sich daher die Frage, ob und unter welchen Bedingungen \mathcal{B} ein LTI-System (2.7) system- bzw. parameteridentifizierbar mittels der PEM ist. Auf diese Problematik wird unter anderem im folgenden Unterabschnitt eingegangen.

Interpretation der Mehrschritt-PEM im Frequenzbereich

Um einen Einblick in die Konvergenz- und Approximationseigenschaften des PEM-Schätzers zu erhalten, ist es zweckmäßig das Kostenfunktional $V_N(\boldsymbol{\theta})$ im Frequenzbereich zu betrachten. Dabei wird im Folgenden der Fall eines SISO-Systems und die Identifikation mittels eines quadratischen Gütekriteriums vorausgesetzt.

Sei $V_N(\boldsymbol{\theta})$ das auf der Basis von N Datenpunkten berechnete Kostenfunktional (3.7) und $\hat{\boldsymbol{\theta}}_N$ dessen globales Minimum. Durch die Anwendung der inversen DTFT (A.2) ergibt sich für $N \rightarrow \infty$ zunächst der Zusammenhang

$$\begin{aligned} V_\infty(\boldsymbol{\theta}) &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N \varepsilon^2[t|t-k, \boldsymbol{\theta}] \\ &= \text{E} \left\{ \varepsilon^2[t|t-k, \boldsymbol{\theta}] \right\} \\ &= r_{\varepsilon_k}[0, \boldsymbol{\theta}] \\ &\stackrel{(A.2)}{=} \frac{1}{2\pi} \int_{-\pi}^{\pi} \phi_{\varepsilon_k}(\omega, \boldsymbol{\theta}) \, d\omega. \end{aligned} \tag{3.21}$$

Dabei kennzeichnet $r_{\varepsilon_k}[\tau, \boldsymbol{\theta}]$ die Kovarianzfunktion der Residuen $\varepsilon[t|t-k]$ und $\phi_{\varepsilon_k}(\omega, \boldsymbol{\theta})$ das Autoleistungsdichtespektrum, dessen Berechnung in Anhang B.3 demonstriert wird. Wird die Beziehung (B.16) in die Gl. (3.21) eingesetzt, dann ergibt sich letztendlich

$$\begin{aligned} V_\infty(\boldsymbol{\theta}) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \frac{H^{[0,k-1]}(\boldsymbol{\theta})}{H(\boldsymbol{\theta})} \right|^2 \left(\left| G_0 + \Delta H(\boldsymbol{\theta}) \frac{\phi_{eu}}{\phi_u} - G(\boldsymbol{\theta}) \right|^2 \phi_u + \right. \\ &\quad \left. |\Delta H(\boldsymbol{\theta})|^2 \left(\lambda_e - \left| \frac{\phi_{eu}}{\phi_u} \right|^2 \phi_u \right) \right) \, d\omega + \lambda_e \sum_{i=0}^{k-1} h^2[i] \end{aligned} \tag{3.22}$$

mit

$$\Delta H(\boldsymbol{\theta}) = H_0 - H_0^{[0,k-1]} \frac{H(\boldsymbol{\theta})}{H^{[0,k-1]}(\boldsymbol{\theta})}, \tag{3.23}$$

wobei $G_0(\omega)$ und $H_0(\omega)$ die Frequenzgänge des wahren Systems \mathcal{S} und $G(\omega, \boldsymbol{\theta})$ und $H(\omega, \boldsymbol{\theta})$ die des Modells \mathcal{M} darstellen (vgl. [HW99]).

⁸Man beachte, dass Systemidentifizierbarkeit im Allgemeinen nicht die Konvergenz der Schätzung $\hat{\boldsymbol{\theta}}$ impliziert.

Bemerkung 3.2 (Konvergenz der Stichprobenkovarianz): Man beachte, dass in (3.21) die gleichmäßige Konvergenz der Stichprobenkovarianz gegen den entsprechenden Erwartungswert unterstellt wird (mit Wahrscheinlichkeit 1). In [Lju12, App. 2B] wird jedoch gezeigt, dass die quasi-Stationarität von $\varepsilon[t|t-k]$, welche wiederum aus der geforderten Stabilität der Filter $H^{-1}(q^{-1}, \boldsymbol{\theta})$ und $H^{-1}(q^{-1}, \boldsymbol{\theta})G(q^{-1}, \boldsymbol{\theta})$ (vgl. (3.4)) und der quasi-Stationarität von $u[t]$ und $y[t]$ gefolgt werden kann, eine hinreichende Bedingung für die gleichmäßige Konvergenz von $V_N(\boldsymbol{\theta})$ ist. \triangleleft

Durch Analyse von (3.22) können einige wichtige Eigenschaften der PEM-Schätzung abgeleitet werden:

Identifizierbarkeit und Konsistenz: Durch Auswertung der Beziehung (3.22) erkennt man, dass $V_\infty(\boldsymbol{\theta})$ minimal wird, wenn ein $\boldsymbol{\theta}_0 \in \mathcal{D}_M$ existiert, sodass $G(\boldsymbol{\theta}_0) \equiv G_0$ und $H(\boldsymbol{\theta}_0) \equiv H_0$ gilt. Damit die Gleichheit für alle ω erfüllt werden kann, muss daher die naheliegende Bedingung $\mathcal{S} \in \mathcal{M}$ erfüllt sein. Dabei ist allerdings zu beachten, dass dies lediglich eine notwendige Bedingung für die Identifizierbarkeit von \mathcal{S} darstellt, denn abhängig von der Wahl des Eingangssignals $u[t]$ können Fälle auftreten, für welche das Kostenfunktional zwar minimal wird, die Gleichheit der Frequenzgänge für alle ω jedoch nicht gefolgt werden kann. Aus diesem Grund muss zusätzlich die *fortwährende Anregung*⁹ (hinreichend hoher Ordnung) durch den Systemeingang u gefordert werden. Diese Bedingung ist oftmals dann verletzt, wenn „ungünstige Eingangssignale“, wie beispielsweise eine sinusförmige Anregung verwendet werden [SS89, Bsp. 2.6], oder wenn das Eingangssignal durch Ausgangsrückkopplung mittels eines linearen Regelgesetzes geringer dynamischer Ordnung herrührt [SS89, Bsp. 2.7, Bsp. 10.4]. Werden hingegen weißes Rauschen bzw. Pseudo-Rausch-Binär-Signale (PRBS) [Ise92a, Abschn. 5.3] oder ein nichtlineares Regelgesetz verwendet, dann ist die Bedingung der fortwährenden Anregung in der Regel erfüllt.

Zusammenfassend lässt sich daher feststellen, dass das System \mathcal{S} systemidentifizierbar mittels der k -Schritt-PEM ist, wenn die folgenden Identifizierbarkeitsbedingungen erfüllt sind:

- (B1) Die wahre Systembeschreibung \mathcal{S} ist in der Modellmenge enthalten, d. h. $\mathcal{S} \in \mathcal{M}$
- (B2) Die Signale $u[t]$ und $y[t]$ sind Realisierungen quasi-stationärer Prozesse
- (B3) Das Eingangssignal $u[t]$ ist fortwährend anregend
- (B4) Im Falle der Identifikation im geschlossenen Regelkreis ist die Anzahl der Verzögerungsglieder des geöffneten Regelkreises $\geq k$ (siehe Bem. 2.8)

Besteht die in (3.20) definierte Menge \mathcal{D}_T zusätzlich aus nur einem Punkt, dann ist \mathcal{S} sogar parameteridentifizierbar mittels der k -Schritt PEM. Wird zudem normalverteiltes weißes Rauschen $e[t]$ angenommen, dann kann allgemein für den Fall eines MIMO-Systems gezeigt werden, dass die PEM-Schätzung mit der Kostenfunktion (3.8) und Prädiktionshorizont $k = 1$ asymptotisch statistisch effizient ist [SS89, Abschn. 7.5].

⁹Die Übersetzung des im englischen Sprachraums üblichen Begriffs *persistence of excitation* ist an die Werke [Unb11] und [BU16] angelehnt. Für eine formale Definition sei auf [SS89, Abschn. 5.4] verwiesen.

Bemerkung 3.3 (Identifizierbarkeitsbedingungen für die Multischritt-PEM): Bei der Identifikation mittels eines Multischrittkriteriums der Form (3.18) ergibt sich für $V_\infty(\boldsymbol{\theta})$ eine Summe von Ausdrücken der Form (3.22) (über unterschiedliche k – vgl. [HW99]). Bei der Argumentation zur Identifizierbarkeit und Konsistenz kann dann analog vorgegangen werden, sodass die Identifizierbarkeitsbedingungen ebenfalls für die Multischritt-PEM gelten. \triangleleft

Unabhängig parametrisiertes Rauschmodell: Aufgrund nicht messbarer exogener Störungen oder des eventuell unrealistischen Störgrößenmodells $v[t] = H(q^{-1})e[t]$ ist die Bedingung $\mathcal{S} \in \mathcal{M}$ aus praktischer Sicht sehr einschränkend. Basierend auf physikalischem Vorwissen des zu identifizierenden Systems kann die Annahme, dass das deterministische Teilsystem G_0 zumindest näherungsweise in der Modellmenge \mathcal{M} enthalten ist oftmals gerechtfertigt werden, wohingegen dies für das stochastische Modell der Störung eventuell nicht möglich ist. Durch die Analyse von (3.22) kann dann jedoch für den Spezialfall eines unabhängig parametrisierten Rauschmodells, d. h. $y[t] = G(q^{-1}, \boldsymbol{\theta}_1)u[t] + H(q^{-1}, \boldsymbol{\theta}_2)e[t]$, eine weitere interessante Eigenschaft der PEM-Schätzung abgeleitet werden:

Im Folgenden wird die Identifikation im offenen Regelkreis, d. h. $\phi_{eu}(\omega) = 0$, vorausgesetzt. Wird zusätzlich angenommen, dass das deterministische und das stochastische Teilmodell unabhängig parametrisiert sind (z. B. FIR-, OE- oder Box-Jenkins-Modell), dann wird ersichtlich, dass der erste Summand innerhalb des Integrals von (3.22) unabhängig von $H(q^{-1}, \boldsymbol{\theta}_2)$ minimal wird, wenn $G(\boldsymbol{\theta}_1) \equiv G_0$ gilt. Dies bedeutet, dass $G(q^{-1}, \boldsymbol{\theta}_1)$ auch dann gegen die wahre Systembeschreibung $G_0(q^{-1})$ konvergiert, wenn die Beschreibung der Störung nicht in der Modellmenge enthalten ist. In analoger Weise lässt sich schlussfolgern, dass ein unabhängig parametrisiertes LTI-Modell ein nichtlineares System \mathcal{S} im Sinne des MSE des k -Schritt Prädiktionsfehlers bestmöglich approximiert.

Festes Rauschmodell: Entsprechend der Beziehung (3.22) ergibt sich im Fall eines festgelegten Rauschmodells $H(q^{-1}, \boldsymbol{\theta}) := H_*(q^{-1})$ die Schätzung

$$\hat{G}(e^{-j\omega}) = \arg \min_{G(\boldsymbol{\theta})} \int_{-\pi}^{\pi} \left| \left(G_0 + \Delta H_* \frac{\phi_{eu}}{\phi_u} - G(\boldsymbol{\theta}) \right) \frac{H_*^{[0,k-1]}}{H_*} \right|^2 \phi_u \, d\omega, \quad (3.24)$$

wobei $B(e^{-j\omega}) = \Delta H_* \phi_{eu} / \phi_u$ als der Bias der Schätzung im Frequenzbereich betrachtet werden kann. Die Schätzung \hat{G} ist daher nur dann asymptotisch erwartungstreu, wenn $\phi_{eu}(\omega) = 0$ oder $H_*(q^{-1}) \equiv H_0(q^{-1})$ gilt. Ist das deterministische Teilsystem G_0 nicht in der Modellmenge \mathcal{M} enthalten und gilt $B(e^{-j\omega}) = 0$, dann ist für eine spezifische Frequenz ω zu erwarten, dass die Abweichung $|G_0(e^{-j\omega}) - G(e^{-j\omega}, \boldsymbol{\theta})|^2$ klein wird, wenn der Term

$$W_*(\omega) = \left| \frac{H_*^{[0,k-1]}(e^{-j\omega})}{H_*(e^{-j\omega})} \right|^2 \phi_u(\omega) \quad (3.25)$$

groß ist, weshalb (3.25) auch als Wichtung im Frequenzbereich interpretiert werden kann. Wird bei der Identifikation beispielsweise ein tieffrequentes Eingangssignal verwendet, dann ist zu erwarten, dass die Approximation des deterministischen Teilsystems im

unteren Frequenzbereich besonders gut ist. Zudem sei erwähnt, dass im Falle eines freien aber unabhängig parametrisierten Rauschmodells in analoger Weise argumentiert werden kann, wobei jedoch die Wichtungsfunktion (3.25) erst nach der Identifikation berechnet werden kann.

Genauigkeit der Mehrschritt-PEM

In vielen Anwendungsfällen, wie beispielsweise dem Entwurf robuster Regelungen, ist eine quantitative Güteangabe des identifizierten Modells notwendig. Diese kann beispielsweise auf der Basis von Konfidenzintervallen oder der Kovarianzmatrix $\mathbf{\Lambda}_\theta$ der geschätzten Parameter $\hat{\boldsymbol{\theta}}_N$ erfolgen.

Der grundlegende Gedanke bei der Berechnung der Schätzgenauigkeit ist es den Schätzwert $\hat{\boldsymbol{\theta}}_N$ als Zufallsgröße mit gesuchter Wahrscheinlichkeitsdichteverteilung zu betrachten. In Anhang B.4 wird gezeigt, dass der Schätzfehler $\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_0$ asymptotisch normalverteilt ist, d. h.

$$\sqrt{N}(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_0) \xrightarrow{i.V.} \mathcal{N}(\mathbf{0}, \mathbf{\Lambda}_\theta), \quad (3.26)$$

wobei die Schätzfehlerkovarianzmatrix durch

$$\mathbf{\Lambda}_\theta = [V_\infty''(\boldsymbol{\theta}_0)]^{-1} \mathbf{\Lambda}_k(\boldsymbol{\theta}_0) [V_\infty''(\boldsymbol{\theta}_0)]^{-1}, \quad (3.27)$$

mit

$$\begin{aligned} V_\infty''(\boldsymbol{\theta}_0) &= 2 \mathbb{E} \left\{ \varepsilon_k'^T[t, \boldsymbol{\theta}_0] \varepsilon_k'[t, \boldsymbol{\theta}_0] \right\} \quad \text{sowie} \\ \mathbf{\Lambda}_k(\boldsymbol{\theta}_0) &= 4\lambda_e \sum_{i=-k+1}^{k-1} \mathbb{E} \left\{ \varepsilon_k'^T[t, \boldsymbol{\theta}_0] \varepsilon_k'[t+i, \boldsymbol{\theta}_0] \right\} \sum_{j=0}^{k-1-|j|} h[j, \boldsymbol{\theta}_0] h[j+|i|, \boldsymbol{\theta}_0] \end{aligned} \quad (3.28)$$

gegeben ist. Dabei kennzeichnen $\varepsilon_k'[t]$ den Gradienten der Prädiktionsfehlerfolge $\varepsilon_k[t] \equiv \varepsilon[t|t-k]$ bezüglich $\boldsymbol{\theta}$ und $h[j, \boldsymbol{\theta}_0]$ die Koeffizienten der Impulsantwort des Rauschfilters $H(q^{-1}, \boldsymbol{\theta}_0)$ (siehe Anhang B.4). Man beachte, dass mithilfe des Resultates (3.27) eine Schätzung $\hat{\mathbf{\Lambda}}_\theta$ berechnet werden kann, indem die Erwartungswerte in (3.28) durch deren empirische Schätzungen ersetzt werden.

Die Varianzen der Systemparameter $\boldsymbol{\theta}$, d. h. der Koeffizienten der Übertragungsfunktionen $G(q^{-1}, \boldsymbol{\theta})$ und $H(q^{-1}, \boldsymbol{\theta})$, sind oftmals schwer interpretierbar. In der Praxis sind daher Genauigkeitsaussagen bezüglich anschaulicher Systemmerkmale wie beispielsweise der Frequenzgänge, Impulsantwort oder der Pol- und Nullstellen des Modells wünschenswert. Unter Berücksichtigung der folgenden Beziehungen lassen sich diese jedoch leicht aus der Schätzfehlerkovarianzmatrix $\mathbf{\Lambda}_\theta$ ableiten:

Wird $\mathbf{f} : \mathcal{D}_M \mapsto \mathbb{C}^m$ allgemein als (analytische) Abbildung des Parametervektors $\boldsymbol{\theta}$ auf die interessierenden Systemmerkmale betrachtet, dann ist in der Umgebung von $\boldsymbol{\theta}_0$ die Näherung

$$\mathbf{f}(\hat{\boldsymbol{\theta}}_N) \approx \mathbf{f}(\boldsymbol{\theta}_0) + \left. \frac{\partial \mathbf{f}}{\partial \boldsymbol{\theta}} \right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} (\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_0) \quad (3.29)$$

zulässig. Wegen Lemma A.2 ist die Zufallsgröße $\mathbf{f}(\hat{\boldsymbol{\theta}}_N)$ näherungsweise asymptotisch normalverteilt, wobei die Kovarianzmatrix durch

$$\boldsymbol{\Lambda}_f = \text{Cov} \{ \mathbf{f}(\hat{\boldsymbol{\theta}}_N) \} \approx \mathbf{f}'(\boldsymbol{\theta}_0) \boldsymbol{\Lambda}_\theta \mathbf{f}'^\top(\boldsymbol{\theta}_0) \quad (3.30)$$

berechnet werden kann [Lju12, Abschn. 9.4]. Mittels der Näherung (3.29) können aus der Schätzung $\hat{\boldsymbol{\Lambda}}_\theta$ daher die Kovarianzen und Konfidenzintervalle beliebiger Systemmerkmale geschätzt werden.

Diskussion und Beispiele

Wie bereits im Abschnitt 3.2.2 erwähnt, stellt die 1-Schritt-PEM unter den genannten Identifizierbarkeitsbedingungen und der zusätzlichen Annahme eines normalverteilten Rauschsignals eine asymptotisch statistisch effiziente Parameterschätzung dar [SS89, Abschn. 7.5]. Die 1-Schritt-PEM ist daher im Sinne der Schätzfehlerkovarianzmatrix die theoretisch genauestmögliche erwartungstreue Schätzung (siehe Anhang A.2.4.4). Aus regelungstechnischer Sicht stellt sich daher die Frage nach dem Nutzen von Mehr- bzw. Multischrittverfahren, welche in der Literatur unter den Schlagworten *control relevant identification* diskutiert wird [HW99; GPS02].

Die Identifizierbarkeitsbedingung $\mathcal{S} \in \mathcal{M}$ ist zweifelsohne eine theoretische Annahme, welche in der Praxis mehr oder weniger gut erfüllt sein wird. Im Gegensatz zur genauestmöglichen Ermittlung der wahren Prozessbeschreibung ist die Identifikation daher viel mehr als eine Approximation des Prozesses, z. B. in einem festgelegten Frequenzbereich, zu verstehen. Geht man in der Gl. (3.24) bspw. davon aus, dass das System G_0 nicht durch das Modell $G(\boldsymbol{\theta})$ beschrieben werden kann, dann ergibt sich bei der Identifikation im offenen Regelkreis ($\phi_{eu} = 0$) wiederum ein Bias

$$B(e^{-j\omega}) = G_0(e^{-j\omega}) - G((e^{-j\omega}), \boldsymbol{\theta}) \neq 0, \quad (3.31)$$

dessen Verteilung im Sinne des Mehr- bzw. Multischrittkriteriums optimal über den Frequenzbereich verteilt wird. Bei der Verwendung des Modells für den Entwurf eines prädiktiven Reglers kann dies durchaus einen positiven Einfluss auf die Güte des geschlossenen Regelkreises haben, was in den Referenzen [SMS92] und [HW99] anhand zweier Beispiele veranschaulicht wird.

Ist umgekehrt, bspw. aufgrund einer sehr gut an den Prozess angepassten Modellstruktur, mit einem vernachlässigbaren *Biasfehler* zu rechnen, dann wird die Genauigkeit des identifizierten Modells im Wesentlichen durch die Länge des Datensatzes, die Anzahl der zu schätzenden Parameter und die Leistung des Eingangs- und Rauschsignals beeinflusst [Lju12, Abschn. 12.3]. In diesem Falle spricht man von einer Dominanz des *Varianzfehlers*, wobei gezeigt werden kann, dass unter diesen Umständen die Nutzung der 1-Schritt-PEM vorteilhaft ist, wenn das Modell für den Entwurf eines modellprädiktiven Reglers verwendet werden soll [ZZP14].

Ein weiterer Nutzen von Mehr- bzw. Multischrittverfahren beruht auf der Feststellung,

dass die k -Schritt-PEM für $k \rightarrow \infty$ gegen die OE-Methode¹⁰ konvergiert (siehe [FP08] und vergleiche Beispiele 3.1 und 3.2). Es ist jedoch auch bekannt, dass bei der Identifikation von OE- und BJ-Modellen die Konvergenz zu lokalen Optima „nicht unüblich“ ist [Lju12, Abschn. 10.5], was sich auch im Rahmen der in dieser Arbeit praktisch durchgeführten Identifikationsexperimente bestätigt hat. Die Wahl geeigneter Optimierungsstartwerte ist in diesem Fall daher als besonders kritisch zu betrachten [Toh08]. Eine mögliche Herangehensweise könnte bspw. darin bestehen, zunächst ein ARX- oder ARMAX-Modell mit $k = 1$ zu identifizieren und anschließend den Prädiktionshorizont k sukzessiv so lange zu erhöhen, bis ein geeignet gewähltes Abbruchkriterium erfüllt wird (z. B. $k > k_{\max}$ oder $|H^{[0,k-1]}(\omega)/H(\omega)| \approx 1$). Die aus der Iteration resultierenden Parameterschätzwerte für $G(q^{-1})$ können anschließend zur Initialisierung bei der OE-Identifikation genutzt werden.

Die folgenden Beispiele sollen die wesentlichen Eigenschaften der Mehrschritt-PEM veranschaulichen:

Beispiel 3.1 (PEM für einen ARX-Prozess erster Ordnung)

Es wird der ARX-Prozess

$$y[t] + a_0^{[1]} y[t - 1] = b_0^{[1]} u[t - 1] + e[t] \quad (3.32)$$

mit $a_0^{[1]} = -0.9$, $b_0^{[1]} = 1$ und $e \sim \mathcal{N}(0, 1)$ betrachtet. Der Systemeingang $u[t]$ wurde mit gleichverteiltem weißen Rauschen mit $u \in [0, 1]$ angeregt ($N = 1000$, $\phi_u(\omega) \approx \frac{1}{12}$).

Die Abbildung 3.2 veranschaulicht den Verlauf der Höhenlinien des logarithmierten Kostenfunktional $\log_{10} V_N(\boldsymbol{\theta})$ mit $\boldsymbol{\theta} = (a^{[1]}, b^{[1]})^\top$ für ein ARX-Modell mit unterschiedlichen Prädiktionshorizonten $k \in \{1, 5, 25\}$ und für ein OE-Modell mit $\boldsymbol{\theta} = (a^{[1]}, f^{[1]})^\top$. Da $\varepsilon[t|t-k] = y[t] - \hat{y}[t|t-k, \boldsymbol{\theta}]$ für $k = 1$ eine lineare Funktion darstellt (siehe Tabelle 1), ist $V_N(\boldsymbol{\theta})$ eine quadratische Funktion, deren parabolische Form in der Darstellung deutlich zu erkennen ist. Zudem ist ersichtlich, dass die Kostenfunktion nicht exakt um den wahren Parameter $\boldsymbol{\theta}_0$ zentriert ist, was den stochastischen Charakter von V_N verdeutlicht. Außerdem ist der stärker werdende nichtlineare Charakter von $V_N(\boldsymbol{\theta})$ für größer werdende k erkennbar, welcher sich der Form des OE-Kostenfunktional annähert.

Die empirisch geschätzten Kovarianzmatrizen $\hat{\Lambda}_\theta$ für $k \in \{1, 5, 25\}$ des ARX-Modells mit $\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}_0 = (-0.9, 1)^\top$ sind in der Tabelle 3 angegeben, wobei festgestellt werden kann, dass die Schätzfehlerkovarianz für größer werdende k zunimmt. In der Abbildung 3.3 (links) ist zudem der Amplitudenfrequenzgang und die empirisch geschätzten 99%-Konfidenzintervalle für unterschiedliche k dargestellt. Da für größer werdende k der tiefe Frequenzbereich stärker betont wird (siehe Abb. 3.3 (rechts)) wirkt sich die größere Schätzfehlerkovarianz $\hat{\Lambda}_\theta$ im Wesentlichen im oberen Frequenzbereich aus. Man erkennt

¹⁰Im Kontext einer deterministischen Betrachtungsweise wird die PEM-Identifikation einer OE-Modellstruktur oftmals auch als *Output Error* oder als *Simulation Error Method* bezeichnet [FP11]. Im Bereich der Identifikation künstlicher neuronaler Netze sind auch die Begriffe *Parallel* und *Series-Parallel Model* gebräuchlich [NP90].

weiterhin, dass die Wichtungsfunktion $\left|H^{[0,k-1]}(e^{-j\omega})/H(e^{-j\omega})\right|$ für größer werdende k gegen 1 konvergiert (OE-Modell). \triangleleft

$k = 1$	$k = 5$	$k = 25$
$\begin{bmatrix} 9.12 \cdot 10^{-5} & 6.38 \cdot 10^{-4} \\ 6.38 \cdot 10^{-4} & 7.40 \cdot 10^{-3} \end{bmatrix}$	$\begin{bmatrix} 1.66 \cdot 10^{-4} & 1.44 \cdot 10^{-3} \\ 1.44 \cdot 10^{-3} & 1.57 \cdot 10^{-2} \end{bmatrix}$	$\begin{bmatrix} 7.49 \cdot 10^{-4} & 7.18 \cdot 10^{-3} \\ 7.18 \cdot 10^{-3} & 7.12 \cdot 10^{-2} \end{bmatrix}$

Tabelle 3: Basierend auf (3.27) empirisch geschätzte Kovarianzmatrizen $\hat{\Lambda}_\theta$ für das ARX-Modell für $\hat{\theta} = \theta_0 = (-0.9, 1)^\top$

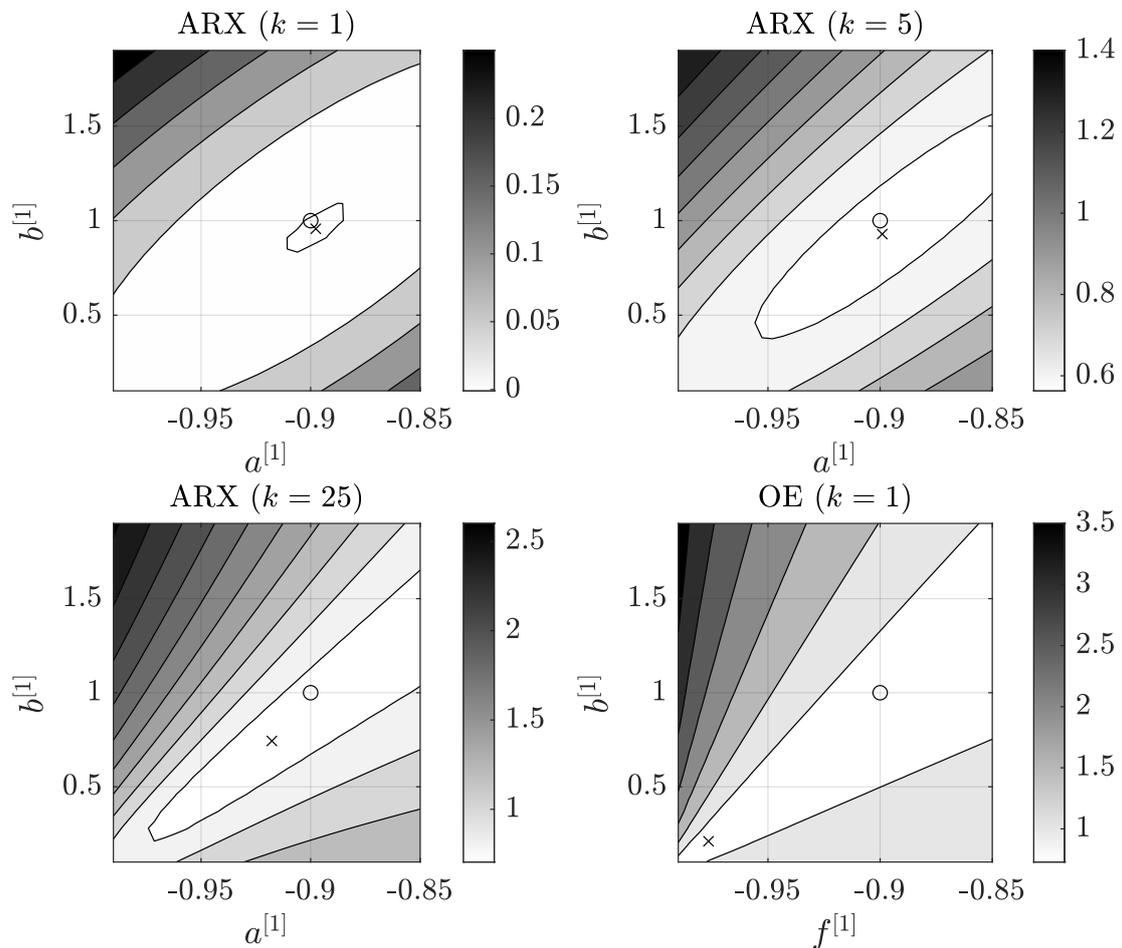


Abbildung 3.2: Höhenlinien $\log_{10} V_N(\theta)$ des ARX-Prozesses (3.32) für unterschiedliche Modellstrukturen (ARX, OE) bzw. Prädiktionshorizonte $k \in \{1, 5, 25\}$. Der Parameter θ_0 und das Argument des Minimums von $V_N(\theta)$ sind mit „o“ bzw. „x“ gekennzeichnet.

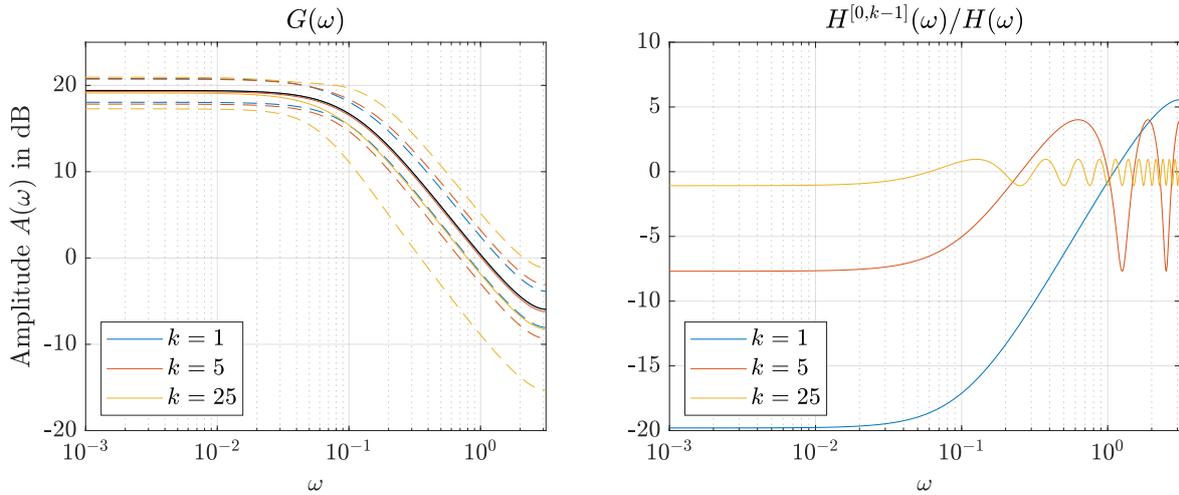


Abbildung 3.3: Amplitudenfrequenzgang (durchgezogene Linie, links) und empirisch geschätzte 99%-Konfidenzintervalle (gestrichelte Linien, links), sowie Wichtungsfunktion (3.25) (rechts), des ARX-Modells (3.32) für $\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}_0 = (-0.9, 1)^\top$ und $k \in \{1, 5, 25\}$.

Das Beispiel 3.1 verdeutlicht, dass die 1-Schritt-PEM der k -Schritt Methode für den Fall $\mathcal{S} \in \mathcal{M}$ im Sinne der Schätzgenauigkeit überlegen ist. Wie bereits erwähnt wurde ist $\mathcal{S} \in \mathcal{M}$ zweifelsohne eine theoretische Annahme, welche in der Praxis nie erfüllt sein wird. Das folgende Beispiel verdeutlicht, dass die k -Schritt-PEM in der praktischen Anwendung durchaus vorteilhaft sein kann:

Beispiel 3.2 (Deterministisches System erster Ordnung mit Ausgangsquantisierung)

Es wird ein deterministisches System erster Ordnung mit einem quantisiertem Ausgangssignal betrachtet, d. h.

$$\begin{aligned} x[t] + a_0^{[1]}x[t-1] &= b_0^{[1]}u[t-1] \\ y[t] &= \lfloor x[t] \rfloor, \end{aligned} \quad (3.33)$$

wobei $\lfloor x \rfloor$ die Rundungsoperation auf eine ganze Zahl kennzeichnet ($b_0^{[1]} = 0.5$ und $a_0^{[1]} = -0.9$). Als Eingangssignal wurde standardisiertes normalverteiltes weißes Rauschen mit $N = 10000$ verwendet, d. h. $u[t] \sim \mathcal{N}(0, 1)$.

Die Abbildung 3.4 veranschaulicht den Verlauf der Höhenlinien des quadratischen Gütefunktional $V_N(\boldsymbol{\theta})$ für das System (3.33) für unterschiedliche Modellstrukturen (ARX, OE), bzw. unterschiedliche Prädiktionshorizonte k . Man erkennt deutlich, dass der ARX-Ansatz keine erwartungstreue Schätzung liefert, dass der Bias für größer werdende k aber kleiner wird. Zudem ist ersichtlich, dass das Minimum des Kostenfunktional der OE-Modellstruktur mit dem wahren Parameter $\boldsymbol{\theta}_0$ übereinstimmt, was durch das unabhängig parametrisierte Rauschmodell begründet werden kann. \triangleleft

Bemerkung 3.4 (Wert des Kostenfunktional $V_N(\hat{\boldsymbol{\theta}})$): Es sei darauf hingewiesen, dass der Wert des Kostenfunktional $V_N(\hat{\boldsymbol{\theta}})$ nicht als ausschließliches Bewertungskriterium

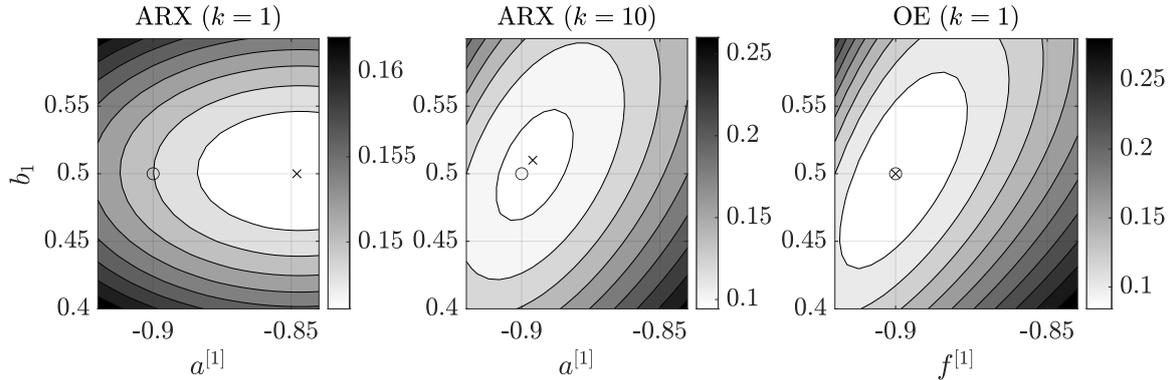


Abbildung 3.4: Höhenlinien des quadratischen Gütefunktions $V_N(\boldsymbol{\theta})$ für das System (3.33) mit $\boldsymbol{\theta}_0 = (-0.9, 0.5)^\top$ für unterschiedliche Modellstrukturen bzw. unterschiedliche Prädiktionshorizonte k . Der wahre Parameter $\boldsymbol{\theta}_0$ ist mit „o“ gekennzeichnet, wohingegen das Argument des Minimums von $V_N(\boldsymbol{\theta})$ mit „x“ markiert wurde.

der Modellgüte herangezogen werden sollte. Für das System (3.33) lassen sich für $k = 1$ beispielsweise Szenarien konstruieren, für welche $V_{N,\text{ARX}} \ll V_{N,\text{OE}}$ gilt. Bei der Modellvalidierung sollte daher stets die Prädiktionsgüte für $k > 1$ überprüft werden. Für eine ausführliche Diskussion der Validierungsproblematik sei auf [Lju12, Kap. 16] und [SS89, Kap. 11] verwiesen. \triangleleft

Bemerkung 3.5 (Wahl einer geeigneten Modellstruktur): Es sei angemerkt, dass im Beispiel 3.2 aufgrund der bekannten Systemstruktur (3.33) die Verwendung eines OE-Ansatzes naheliegend ist. In der Praxis jedoch sind sowohl ARX- als auch OE-Modelle aufgrund des unflexiblen Rauschmodells ($H(q^{-1}) = 1/A(q^{-1})$ bzw. $H(q^{-1}) = 1$) oft zu einschränkend, sodass mit ARMAX- oder BJ-Strukturen oftmals bessere Ergebnisse im Sinne der Prädiktionsgüte erzielt werden können (siehe Kapitel 5). Man beachte allerdings auch, dass gezeigt werden kann, dass ein ARX-Modell jedes stochastische LTI-System beliebig genau approximieren kann, indem die Polynomordnungen n_A und n_B sehr groß gewählt werden (High-Order ARX-Modell) [Lju12, Abschn. 10.4]. \triangleleft

3.3 Methode der kleinsten Fehlerquadrate

In diesem Abschnitt soll der Zusammenhang der PEM mit den auf MkQ-Schätzungen basierenden Verfahren dargelegt werden. Dabei wird insbesondere auf die weit verbreitete Hilfsvariablenmethode und die auf pseudolinearen Regressionen beruhenden mehrstufigen MkQ-Verfahren eingegangen.

3.3.1 Einschnitt-PEM für ARX-Modellstrukturen

In der Tabelle 1 wird darauf hingewiesen, dass der optimale Prädiktor für ARX- und FIR-Modelle eine lineare Funktion des Parametervektors $\boldsymbol{\theta}$ darstellt. Für ein System

mit einer Ausgangsgröße lässt sich für den Prädiktionsfehler daher

$$\begin{aligned}\varepsilon(t|t-1) &= y[t] - \hat{y}[t|t-1, \boldsymbol{\theta}] \\ &= y[t] - \boldsymbol{\varphi}^\top[t] \boldsymbol{\theta}\end{aligned}\tag{3.34}$$

mit

$$\begin{aligned}\boldsymbol{\varphi}^\top[t] &= (-y[t-1] \quad \dots \quad -y[t-n_A] \quad u[t] \quad \dots \quad u[t-n_B]) \\ \boldsymbol{\theta} &= (a^{[1]} \quad \dots \quad a^{[n_A]} \quad b^{[0]} \quad \dots \quad b^{[n_B]})^\top\end{aligned}\tag{3.35}$$

schreiben, wobei $\boldsymbol{\varphi}^\top[t]$ den *Regressionsvektor* darstellt. Die Komponenten von $\boldsymbol{\varphi}[t]$ werden entsprechend als *Regressoren* bezeichnet.

Es kann gezeigt werden, dass das Minimum des quadratischen Optimierungsproblems

$$\hat{\boldsymbol{\theta}}_N = \arg \min_{\boldsymbol{\theta}} \frac{1}{N} \sum_{t=1}^N \varepsilon^2(t|t-1)\tag{3.36}$$

durch die MkQ-Lösung

$$\begin{aligned}\hat{\boldsymbol{\theta}}_N &= \left[\sum_{t=1}^N \boldsymbol{\varphi}[t] \boldsymbol{\varphi}^\top[t] \right]^{-1} \left[\sum_{t=1}^N \boldsymbol{\varphi}[t] y[t] \right] \\ &= \underbrace{\left[\boldsymbol{\Phi}^\top \boldsymbol{\Phi} \right]^{-1}}_{\boldsymbol{\Phi}^\dagger} \boldsymbol{\Phi}^\top \mathbf{Y},\end{aligned}\tag{3.37}$$

mit $\mathbf{Y} = \text{vec} \{y[t]\}$ und $\boldsymbol{\Phi} = \text{vec} \{\boldsymbol{\varphi}^\top[t]\}$ berechnet werden kann (siehe u. a. [Lju12, Kap. 2], [SS89, Kap. 4] und [Ise92a, Kap. 8]), wobei die Regularität der Matrix $\boldsymbol{\Phi}^\top \boldsymbol{\Phi}$ vorausgesetzt wird¹¹. Die in (3.37) vorkommende $n_\theta \times N$ Matrix $\boldsymbol{\Phi}^\dagger$ wird als *Pseudoinverse* bezeichnet. Durch Einsetzen von (3.34) und $\varepsilon'[t|t-1] = -\boldsymbol{\varphi}^\top[t]$ in die Gauß-Newton-Iterationsgleichung (3.14), kann zudem leicht verifiziert werden, dass die MkQ (3.37) der 1-Schritt-PEM für ARX-Modelle bei Verwendung des GN-Verfahrens entspricht.

Der MkQ-Schätzer (3.37) hat die folgenden Eigenschaften¹² [SS89, Abschn. 4.2, Abschn. 7.5]:

- (1) $\hat{\boldsymbol{\theta}}_N$ ist konsistent. Falls $\boldsymbol{\varphi}[t]$ deterministisch ist (FIR-Modell), dann ist $\hat{\boldsymbol{\theta}}_N$ erwartungstreu.
- (2) $\sqrt{N}(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_0) \xrightarrow{i.V.} \mathcal{N}(\mathbf{0}, \boldsymbol{\Lambda}_\theta)$ mit $\boldsymbol{\Lambda}_\theta = \lambda_e E\{\boldsymbol{\varphi}[t] \boldsymbol{\varphi}^\top[t]\}^{-1}$ (Vergleiche (3.26)). Ist die Störung $e[t]$ zudem normalverteilt, dann ist $\hat{\boldsymbol{\theta}}_N$ asymptotisch statistisch effizient. Ist $\boldsymbol{\varphi}[t]$ deterministisch, dann gelten beide Aussagen sogar für endliche

¹¹Es sei angemerkt, dass Spezialfälle existieren, für welche gezeigt werden kann, dass $\boldsymbol{\Phi}^\top \boldsymbol{\Phi}$ singular ist. Diese sind jedoch entweder relativ theoretisch (rauschfreie Daten und Modellordnung zu hoch) oder können durch entsprechende Wahl des Eingangssignals $u[t]$ vermieden werden [SS89, Abschn. 7.1].

¹²Man beachte, dass die angegebenen Resultate nur dann auch im geschlossenen Regelkreis gültig sind, wenn in der Definition von $\boldsymbol{\varphi}^\top[t]$ der Regressor $u[t]$ nicht auftaucht (siehe Bemerkung 2.8 und Gleichung (3.41)).

N . Im Rahmen der Regressionsanalyse wird der Schätzer (3.37) daher als *beste lineare erwartungstreue Schätzfunktion* (Best linear unbiased estimator – BLUE) bezeichnet.

Bemerkung 3.6 (Korrelierte Ausgangsstörung v): Es sei angemerkt, dass die Eigenschaft (1) auch für Modelle der Form

$$y[t] = \boldsymbol{\varphi}^\top[t] \boldsymbol{\theta} + v[t] \quad (3.38)$$

mit deterministischem Regressionsvektor $\boldsymbol{\varphi}^\top[t]$ gilt, wenn für die Ausgangsstörung $v[t]$ mittelwertfreies farbiges Rauschen mit Kovarianzmatrix $\mathbf{E}\{\boldsymbol{v}\boldsymbol{v}^\top\} = \boldsymbol{\Lambda}_v$ angenommen wird ($\boldsymbol{v} = \text{vec}\{v[t]\}$). Die Eigenschaft der Effizienz ist hingegen nur bei der Verwendung der *verallgemeinerten Methode der kleinsten Fehlerquadrate*¹³ (VMkQ)

$$\hat{\boldsymbol{\theta}}_N = [\boldsymbol{\Phi}^\top \boldsymbol{\Lambda}_v^{-1} \boldsymbol{\Phi}]^{-1} \boldsymbol{\Phi}^\top \boldsymbol{\Lambda}_v^{-1} \mathbf{Y} \quad (3.39)$$

gegeben [SS89, Abschn. 4.3]. ◁

3.3.2 Hilfsvariablenverfahren

Der Vorteil der MkQ besteht darin, dass bei der Lösung des Optimierungsproblems (3.36) auf iterative Verfahren der nichtlinearen Optimierung verzichtet werden kann. Der wesentliche Nachteil ist jedoch, dass sich für Prozesse der allgemeinen Form

$$y[t] = \boldsymbol{\varphi}^\top[t] \boldsymbol{\theta}_0 + v[t] \quad (3.40)$$

keine erwartungstreue Schätzung des Parametervektors ergibt. Dies wird ersichtlich, wenn die Systembeschreibung (3.40) in die MkQ-Schätzgleichung (3.37) eingesetzt wird:

$$\begin{aligned} \hat{\boldsymbol{\theta}}_\infty &= \lim_{N \rightarrow \infty} \left[\sum_{t=1}^N \boldsymbol{\varphi}[t] \boldsymbol{\varphi}^\top[t] \right]^{-1} \left[\sum_{t=1}^N \boldsymbol{\varphi}[t] (\boldsymbol{\varphi}^\top[t] \boldsymbol{\theta}_0 + v[t]) \right] \\ &= \boldsymbol{\theta}_0 + \mathbf{E} \left\{ \boldsymbol{\varphi}[t] \boldsymbol{\varphi}^\top[t] \right\}^{-1} \mathbf{E} \left\{ \boldsymbol{\varphi}[t] v[t] \right\} \end{aligned} \quad (3.41)$$

Dabei kennzeichnet der zweite Summand den Bias der Schätzung, der verschwindet, wenn $\boldsymbol{\varphi}[t]$ und $v[t]$ unabhängig sind. Dies ist offensichtlich dann der Fall, wenn der zu identifizierende Prozess tatsächlich der ARX-Modellstruktur genügt ($v[t] \equiv e[t]$).

Eine Möglichkeit zur Vermeidung der Biasproblematik stellt die sogenannte Hilfsvariablenmethode dar. Dabei besteht die wesentliche Idee darin, die MkQ-Lösung derart zu verändern, dass sich in der Betrachtung (3.41) eine (asymptotisch) erwartungstreue Schätzung ergibt. Dieser Gedanke motiviert die sogenannte Hilfsvariablenmethode

$$\hat{\boldsymbol{\theta}}_N = \left[\sum_{t=1}^N \boldsymbol{z}[t] \boldsymbol{\varphi}^\top[t] \right]^{-1} \left[\sum_{t=1}^N \boldsymbol{z}[t] y[t] \right], \quad (3.42)$$

¹³Generalized least squares (GLS)

wobei die Hilfsvariablen $\mathbf{z}[t] \in \mathbb{R}^{n_\theta}$ derart zu wählen sind, dass diese unabhängig vom Störsignal $v[t]$, aber stark korreliert mit den Komponenten des Regressionsvektors $\boldsymbol{\varphi}[t]$ sind. Dies ist beispielsweise dann der Fall, wenn $\mathbf{z}[t]$ die ungestörten Ausgangssignale enthält. Eine typische Wahl für $\mathbf{z}[t]$ erhält man beispielsweise, wenn in der Definition des Regressionsvektors (3.35) die Regressoren $y[t-1], \dots, y[t-n_A]$ durch die gefilterten Signale $\tilde{y}[t] = \hat{G}(q^{-1})u[t]$ ersetzt werden, wobei \hat{G} eine Schätzung der Systemübertragungsfunktion $G(q^{-1})$ darstellt. Für eine ausführliche Diskussion und die theoretische Analyse des Hilfsvariablenverfahrens sei auf die Referenzen [SS83a; SS83b; SS89; SS02] verwiesen.

3.3.3 Pseudolineare-Regression

Eine weitere Klasse von Identifikationsalgorithmen basiert auf der pseudolinearen Darstellung des LTI-Modells (2.36):

$$A(q^{-1})y[t] = \underbrace{\frac{B(q^{-1})}{F(q^{-1})}u[t]}_{=: \tilde{y}[t]} + \underbrace{\frac{C(q^{-1})}{D(q^{-1})}e[t]}_{=: v[t]} \quad \mathbb{E}\{e^2[t]\} = \lambda_e \delta_{t,s} \quad (3.43)$$

Mittels der definierten Hilfssignale $\tilde{y}[t]$ und $v[t]$ lässt sich der Modellausgang $y[t]$ in der pseudolinearen Form

$$y[t] = \boldsymbol{\varphi}^\top[t, \boldsymbol{\theta}] \boldsymbol{\theta} + e[t] \quad (3.44)$$

mit

$$\boldsymbol{\varphi}^\top[t, \boldsymbol{\theta}] = \left(-y[t-1] \dots -y[t-n_A] \quad u[t] \dots u[t-n_B] \quad e[t-1] \dots e[t-n_C] \right. \\ \left. -v[t-1] \dots -v[t-n_D] \quad -\tilde{y}[t-1] \dots -\tilde{y}[t-n_F] \right) \quad \text{und} \quad (3.45)$$

$$\boldsymbol{\theta} = \left(a^{[1]} \dots a^{[n_A]} \quad b^{[0]} \dots b^{[n_B]} \quad c^{[1]} \dots c^{[n_C]} \quad d^{[1]} \dots d^{[n_D]} \quad f^{[1]} \dots f^{[n_F]} \right)^\top$$

darstellen. Durch wiederholte Anwendung des MkQ-Schätzers (3.37) auf die pseudolineare Darstellung (3.44) lässt sich dann die Iterationsvorschrift

$$\hat{\boldsymbol{\theta}}_{(i+1)} = \left[\sum_{t=1}^N \boldsymbol{\varphi}[t, \hat{\boldsymbol{\theta}}_{(i)}] \boldsymbol{\varphi}^\top[t, \hat{\boldsymbol{\theta}}_{(i)}] \right]^{-1} \left[\sum_{t=1}^N \boldsymbol{\varphi}[t, \hat{\boldsymbol{\theta}}_{(i)}] y[t] \right] \\ = \hat{\boldsymbol{\theta}}_{(i)} + \left[\sum_{t=1}^N \boldsymbol{\varphi}[t, \hat{\boldsymbol{\theta}}_{(i)}] \boldsymbol{\varphi}^\top[t, \hat{\boldsymbol{\theta}}_{(i)}] \right]^{-1} \left[\sum_{t=1}^N \boldsymbol{\varphi}[t, \hat{\boldsymbol{\theta}}_{(i)}] \varepsilon[t, \hat{\boldsymbol{\theta}}_{(i)}] \right] \quad (3.46)$$

rechtfertigen, wobei für den Modellausgang $y[t]$ die Beziehung (3.44) mit $e[t] \equiv \varepsilon[t]$ eingesetzt wurde. Die dabei notwendigen Signale $\tilde{y}[t]$, $v[t]$ und $e[t] \equiv \varepsilon[t]$ werden in

jedem Iterationsschritt durch

$$\hat{y}[t, \boldsymbol{\theta}_{(i)}] = \frac{\hat{B}(q^{-1}, \boldsymbol{\theta}_{(i)})}{\hat{F}(q^{-1}, \boldsymbol{\theta}_{(i)})} u[t] \quad (3.47a)$$

$$\hat{v}[t, \boldsymbol{\theta}_{(i)}] = \hat{A}(q^{-1}, \boldsymbol{\theta}_{(i)}) y[t] - \hat{y}[t, \boldsymbol{\theta}_{(i)}] \quad (3.47b)$$

$$\hat{\varepsilon}[t, \boldsymbol{\theta}_{(i)}] = \frac{\hat{D}(q^{-1}, \boldsymbol{\theta}_{(i)})}{\hat{C}(q^{-1}, \boldsymbol{\theta}_{(i)})} \hat{v}[t, \boldsymbol{\theta}_{(i)}] \quad (3.47c)$$

erzeugt, wobei $\hat{A}(q^{-1})$ bis $\hat{F}(q^{-1})$ die im Iterationsschritt i berechneten Schätzungen der Polynome $A(q^{-1})$ bis $F(q^{-1})$ darstellen.

Es besteht ein wichtiger Zusammenhang des PLR- bzw. ELS-Algorithmus (3.46) zum PEM-Verfahren mit $V_N(\boldsymbol{\theta}) = \sum_{t=1}^N \varepsilon^2[t|t-1, \boldsymbol{\theta}]$. Dieser wird ersichtlich, wenn der Gradient der pseudolinearen Darstellung $\varepsilon[t|t-1, \boldsymbol{\theta}] = y[t] - \boldsymbol{\varphi}^\top[t, \boldsymbol{\theta}] \boldsymbol{\theta}$ betrachtet wird:

$$\varepsilon'[t|t-1, \boldsymbol{\theta}] = \frac{\partial \varepsilon[t|t-1, \boldsymbol{\theta}]}{\partial \boldsymbol{\theta}} = -\boldsymbol{\varphi}^\top[t, \boldsymbol{\theta}] - \boldsymbol{\theta}^\top \frac{\partial \boldsymbol{\varphi}[t, \boldsymbol{\theta}]}{\partial \boldsymbol{\theta}} \quad (3.48)$$

Wird in die Iterationsgleichung des Gauß-Newton-Verfahrens (3.14) die Näherung $\varepsilon'[t|t-1] \approx -\boldsymbol{\varphi}^\top[t, \boldsymbol{\theta}]$ eingesetzt, dann ergibt sich die Rekursionsvorschrift des PLR-Algorithmus (3.46). Es lässt sich daher schlussfolgern, dass der PLR-Algorithmus der 1-Schritt-PEM entspricht, wenn die Optimierung mit dem Gauß-Newton-Verfahren und mit der Näherung $\varepsilon'[t|t-1] \approx -\boldsymbol{\varphi}^\top[t, \boldsymbol{\theta}]$ durchgeführt wird.

Es sei jedoch darauf hingewiesen, dass die Konvergenzbedingungen des hier angegebenen einfachen PLR-Algorithmus verhältnismäßig restriktiv sind. Diese können jedoch durch Verwendung einer variablen Schrittweite abgeschwächt werden. Für eine ausführliche Diskussion dieser Problematik sei auf [SSA85] verwiesen.

3.4 PEM-Verfahren für nichtlineare Modellstrukturen

Die PEM-Methode ist ein sehr allgemeines Verfahren, welches prinzipiell auch für die Identifikation nichtlinearer Modellansätze geeignet ist. Die dabei wesentliche Implementierungshürde besteht in der Berechnung der optimalen Prädiktion $\hat{\boldsymbol{y}}[t|t-k, \boldsymbol{\theta}]$ bzw. der entsprechenden Prädiktionsfehlerfolge $\boldsymbol{\varepsilon}[t|t-k, \boldsymbol{\theta}]$. Ist jedoch ein expliziter Ausdruck bekannt, dann kann ein Schätzwert $\hat{\boldsymbol{\theta}}$ analog zu (3.5) durch Optimierung eines geeignet gewählten Kostenfunktionalen erfolgen.

Wie bereits angemerkt wurde, ist es nur in Spezialfällen möglich einen expliziten Ausdruck für die optimale Prädiktion des Systemausgangs anzugeben. Daher wird insbesondere im Rahmen der Identifikation von NARMAX-Modellen, oftmals die implizite Darstellung der 1-Schritt Prädiktion

$$\hat{y}[t|t-1] = F(y[t-1], \dots, y[t-n_A], u[t], \dots, u[t-n_B], e[t-1], \dots, e[t-n_C]) \quad (3.49)$$

verwendet. Werden für die Parametrierung der Nichtlinearität F beispielsweise polynomiale Ansatzfunktionen verwendet, dann lässt sich (3.49) analog zu (3.44) in pseudolinearer Darstellung

$$y[t] = \boldsymbol{\varphi}^\top [t, \boldsymbol{\theta}] \boldsymbol{\theta} \quad (3.50)$$

angeben, wobei der Regressionsvektor $\boldsymbol{\varphi}[t, \boldsymbol{\theta}]$ in diesem Fall aus zeitverzögerten Potenzen von u , y sowie e und gegebenenfalls deren Produkten besteht. Für die Berechnung eines Schätzwertes $\hat{\boldsymbol{\theta}}$ kann daher wiederum ein iteratives Verfahren des Typs (3.46) angewendet werden (siehe [Bil13, Abschn. 3.6]). Ist der Regressionsvektor $\boldsymbol{\varphi}$ nicht von der Störung $e[t]$ abhängig (NFIR- oder NARX-Modell), dann ist keine Iteration notwendig, sodass die im Sinne des MSE optimale Lösung durch die MkQ (3.37) gegeben ist.

Für die in der Abbildung 2.6 dargestellten Modellstrukturen kann die PEM basierend auf den Gleichungen (2.50) bzw. (2.53) direkt implementiert werden. Dabei ist es naheliegend die in Abschnitt 2.4 genannten Parametrierungen der Filter $\mathbf{G}(q^{-1})$ bzw. $\mathbf{H}(q^{-1})$ zu verwenden. Die nichtlinearen Funktionen \mathbf{f} bzw. \mathbf{h} werden in der Praxis unabhängig vom linearen Teilsystem parametrierung, weswegen die entsprechenden Parameter im Folgenden durch $\boldsymbol{\eta}$ gekennzeichnet werden.

Die genannten Modellansätze haben sich aus Sicht der in Kapitel 5 diskutierten Anwendungsbeispiele als besonders geeignet erwiesen. Dies liegt insbesondere daran, dass für das Aufstellen eines geeigneten stochastischen Modellansatzes die Ideen aus der linearen Theorie verwendet werden können. Dies soll anhand des folgenden Beispiels verdeutlicht werden:

Beispiel 3.3 (Bilineares System mit unbekannter periodischer Störung)

Es wird das in der Abbildung 3.5 dargestellte bilineare Beispielsystem betrachtet. Dabei kennzeichnen $u[t]$ und $y[t]$ gemessene Ein- bzw. Ausgangssignale, $v_1[t]$ eine unbekannte periodische Störung mit bekannter Periodendauer p und $v_2[t]$ ein tieffrequentes Störsignal (Drift/Trend). Eine beispielhafte Simulation des Systems mit $a_0^{[1]} = -0.9$, $b_0^{[1]} = 0.5$ und $\eta_{1,0} = 3$ ist in der Grafik 3.6 dargestellt.

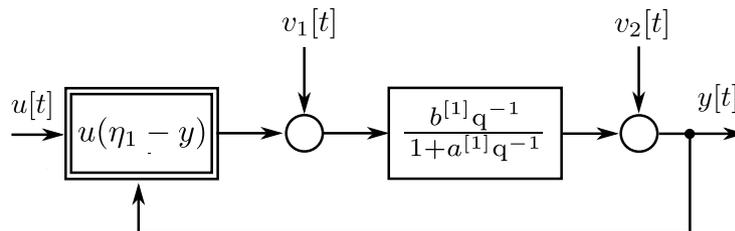


Abbildung 3.5: Bilineares Beispielsystem mit periodischem Störsignal $v_1[t]$ und Drift $v_2[t]$, wobei $a^{[1]}$, $b^{[1]}$ und η_1 zu identifizierende Modellparameter darstellen.

Im Folgenden soll ein geeigneter Modellansatz für das System aufgestellt und identifiziert werden: Da das System ein nichtlinear rückgekoppeltes lineares System erster Ordnung darstellt, ist es naheliegend für das lineare Teilmodell eine Struktur mit autoregressivem Anteil im deterministischen Teilsystem $G(q^{-1})$ zu verwenden (z. B. ARX, ARMAX,

BJ,...). Aufgrund der unbekannt periodischen Störung ist es zudem sinnvoll, das Rauschfilter $H(q^{-1})$ so zu parametrieren, dass die Störung durch das Modell erfasst werden kann. Für den Fall eines driftfreien Systems ($v_2[t] = 0$) kann beispielsweise ein ARARX Modellansatz (siehe Abschnitt 2.4.1)

$$(1 + a^{[1]} q^{-1})y[t] = b^{[1]} q^{-1} f(u[t], y[t], \eta_1) + \frac{1}{1 + d^{[p]} q^{-p}} e[t], \quad (3.51)$$

mit $f(u[t], y[t], \eta_1) = u[t](\eta_1 - y[t])$ verwendet werden. Wird zudem von einem nicht zu vernachlässigendem Trend ausgegangen, d. h. $v_2[t] \neq 0$, dann kann dieser durch einen zusätzlichen Integrator im Rauschfilter $H(q^{-1})$ berücksichtigt werden. Für das zu identifizierende Gesamtmodell ergibt sich daher

$$(1 + a^{[1]} q^{-1})y[t] = b^{[1]} q^{-1} f(u[t], y[t], \eta_1) + \frac{1}{(1 - q^{-1})(1 + d^{[p]} q^{-p})} e[t], \quad (3.52)$$

was einem ARARIX-Modellansatz für das lineare Teilsystem entspricht.

Die Parameter $a^{[1]}$, $b^{[1]}$, η_1 und $d^{[p]}$ mit $p = 50$ (siehe Abb. 3.6) der Modelle (3.51) und (3.52) wurden für den Fall $v_2[t] = 0$ und $v_2[t] = \cos(2\pi \frac{t}{500}) + 0.5$ mittels 1-Schritt PEM-Methode identifiziert ($t = 1, \dots, 500$). Für die Lösung des Optimierungsproblems wurde die durch die MATLAB-Routine `lsqnonlin()` zur Verfügung gestellte Implementierung des Levenberg-Marquardt-Verfahrens verwendet, wobei der Algorithmus mit $\eta_1 = 0$ und $d^{[50]} = 0$ initialisiert wurde (MATLAB 2020b, OPTIMIZATION TOOLBOX 8.2). Die Anfangswerte der Parameter $a^{[1]}$ und $b^{[1]}$ wurden durch MkQ-Schätzung (3.37) berechnet.

Die identifizierten Parameter des Modells für den Fall $v_2[t] = 0$ (3.51) und $v_2[t] \neq 0$ (3.52) sind in der Tabelle 4 angegeben. Dabei lässt sich erkennen, dass für $v_2[t] = 0$ die identifizierten Werte bis auf numerische Fehler den wahren Parametern des Systems entsprechen ($\sigma_e^2 \approx 0$). Hingegen entstehen im Falle eines vorhandenen Drifts $v_2[t] \neq 0$ leichte Modellfehler ($\mathcal{S} \notin \mathcal{M}$).

Der verwendete Datensatz und die beispielhafte 30-Schritt Prädiktion des identifizierten Modells (3.52) kann der Abbildung 3.6 entnommen werden. Es ist zu beachten, dass $\hat{y}[t|t-k]$ eine Simulation und keine im Sinne des MSE optimale Prädiktion darstellt (siehe Bemerkung 2.12). \triangleleft

Parameter	wahrer Wert	$v_2[t] = 0$	$v_2[t] \neq 0$
$a^{[1]}$	-0.9	$-0.9000 \pm 1.24 \cdot 10^{-12}\%$	$-0.9003 \pm 0.016\%$
$b^{[1]}$	0.5	$0.5000 \pm 1.32 \cdot 10^{-12}\%$	$0.5001 \pm 0.016\%$
$d^{[50]}$	–	$-1.000 \pm 1.88 \cdot 10^{-12}\%$	$-0.9999 \pm 0.025\%$
η_1	3	$3.0000 \pm 5.13 \cdot 10^{-13}\%$	$3.0003 \pm 0.008\%$
σ_e^2	–	$1.5085 \cdot 10^{-27}$	$2.7537 \cdot 10^{-7}$

Tabelle 4: Parameterschätzwerte und deren rel. Standardabw. für die Modelle (3.51) und (3.52) ($v_2[t] = 0$ bzw. $v_2[t] = \cos(2\pi \frac{t}{500}) + 0.5$) für den Datensatz in Abb. 3.6.

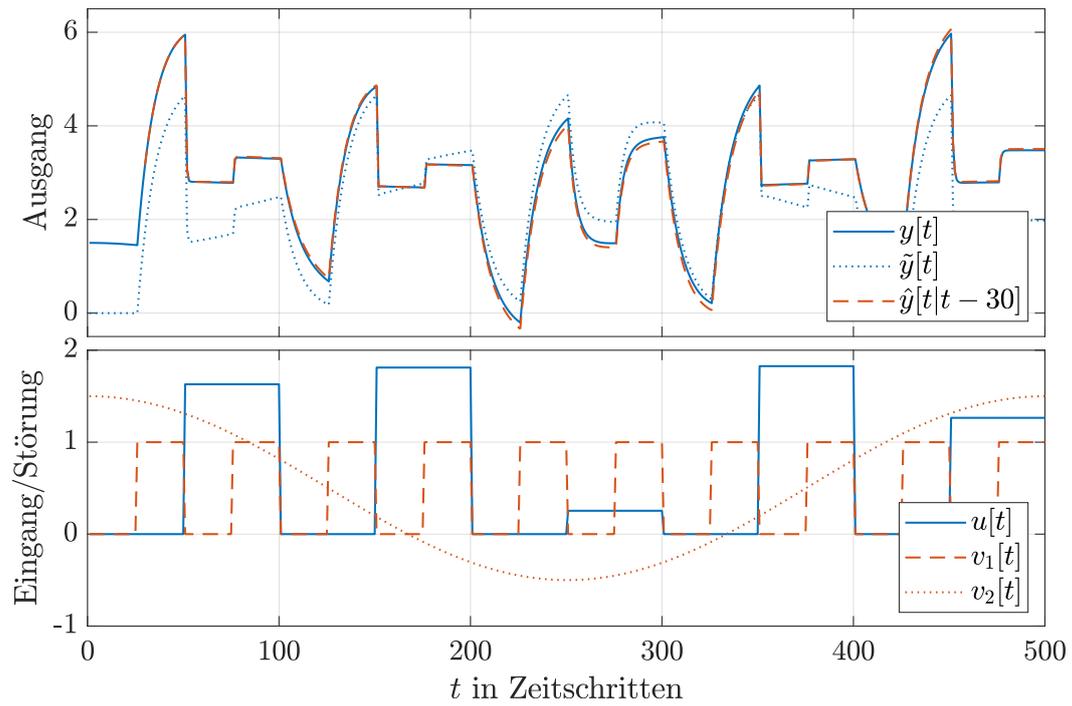


Abbildung 3.6: Simulation des Ausgangssignals des periodisch gestörten bilinearen Beispielsystems mit $a_0^{[1]} = -0.9$, $b_0^{[1]} = 0.5$ und $\eta_{1,0} = 3$ mit ($y[t]$) und ohne Drift ($\tilde{y}[t]$).

Kapitel 4

Regelungstechnik in der Gebäudeautomation

In diesem Kapitel soll dem Leser ein kurzer Einblick in die technische Struktur und in wesentliche Funktionsprinzipien der Gebäudetechnik vermittelt werden. Dabei widmet sich der erste Abschnitt der Diskussion typischer regelungstechnischer Problemstellungen, deren Besonderheiten sowie konventioneller praktischer Lösungsansätze. Im zweiten Abschnitt werden dann, unter Auswertung der wissenschaftlichen Literatur, mögliche Anwendungsfälle der modellprädiktiven Regelung in der Gebäudetechnik dargelegt und klassifiziert.

Generell muss im Folgenden jedoch berücksichtigt werden, dass die meisten Gebäude, deren technische Ausstattung und die Regelung individuell geplant werden, sodass bei konkreten praktischen Implementierungen eine erhebliche Variationsbreite besteht. Es sei daher bereits im Voraus erwähnt, dass bei den Darlegungen kein Anspruch auf Vollständigkeit besteht. Für eine tiefergehende Einführung in die Grundlagen der Regelungstechnik in der Gebäudeautomation sei beispielsweise auf das Buch [APG14] verwiesen. Als grundlegende Nachschlagewerke sind [RSS11] und [Ash17] zu empfehlen.

4.1 Funktionsprinzipien und Regelungsstrategien in der Gebäudetechnik

Die grundlegende Aufgabe der Gebäudeautomation ist die Einhaltung vorgegebener Komfortparameter, typischerweise der Temperatur sowie ggf. der relativen Feuchte und der Qualität der Raumluft (meistens CO₂-Konzentration in parts per million – ppm). Dabei muss stets ein möglichst minimaler Einsatz der vorhandenen Energieträger, bzw. die Minimierung der anfallenden Betriebskosten sichergestellt werden.

Die Funktionen eines Gebäudeenergiesystems können allgemein den vier Ebenen der Energieumwandlung, -speicherung, -verteilung und -übergabe zugeordnet werden (siehe Abbildung 4.1), deren Aufgaben wie folgt beschrieben werden können:

- **Energieumwandlung:** Umwandlung (umgangssprachlich Erzeugung) bzw. Aufbereitung¹ der jeweiligen Energieträger bzw. des -mediums (Wasser, Luft oder elektrischer Strom).
- **Energiespeicherung:** Gegebenenfalls Zwischenspeicherung elektrischer oder thermischer Energie.
- **Energieverteilung:** Transport der Energie zu den jeweiligen Abnehmern.
- **Energieübergabe:** Abnahme der Nutzenergie in den jeweiligen Räumen, bzw. Gebäudezonen.

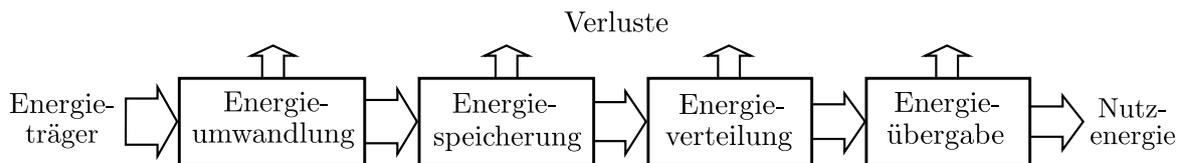


Abbildung 4.1: Schematische Darstellung der Funktionsebenen eines Gebäudeenergiesystems

Den genannten Funktionsebenen können eine Vielzahl unterschiedlicher typischer Regelungsaufgaben zugeordnet werden. Aufgrund der oftmals hohen Komplexität des Gesamtsystems wird dieses, ähnlich wie in anderen ingenieurtechnischen Disziplinen, in geeignete Teilsysteme zerlegt, deren Regelungen weitestgehend unabhängig voneinander entworfen werden. Dabei werden überwiegend klassische Regelungsstrukturen, wie beispielsweise P-, PI- oder PID-Regler sowie unstetige Regler wie Zwei- bzw. Mehrpunktregler verwendet. Diese werden üblicherweise durch eine übergeordnete zeit-, außentemperatur- und/oder zustandsabhängige Steuerlogik ergänzt, welche die Regelkreise der Teilsysteme in - bzw. außer Betrieb nimmt und ggf. deren Sollwerte vorgibt. Dabei erfolgt die Parametrierung der Regelkreise und der Steuerlogik in der Praxis meist basierend auf Erfahrungswissen oder intuitiv auf der Grundlage eines tieferen Anlagenverständnisses.

In Regelkreisen mit kontinuierlichen Steuergrößen werden häufig PI-, manchmal aber auch P- oder PID-Regler² verwendet. Dies lässt sich damit begründen, dass die Regelstrecken in der Gebäudetechnik oftmals näherungsweise PT_n -Charakter (ggf. mit Totzeit) besitzen. Ein typisches Beispiel hierfür ist das dynamische Verhalten der Temperatur in einem Raum mit Radiator- oder Flächenheizsystem, wobei die Raumluft und das Heizsystem als jeweils ein PT_1 -Glied modelliert werden können. Werden die Räume hingegen über eine Lüftungsanlage beheizt, wie es beispielsweise oft in Großraumbüros, Hörsälen, Sport-, Messe- oder Fabrikhallen der Fall ist, dann werden zur Erhöhung

¹Mit diesem Begriff sollen allgemein die Vorgänge der Anpassung von Temperatur-, bzw. Spannungsniveaus sowie alle Luftkonditionierungsmaßnahmen (Befeuchten/Entfeuchten) zusammengefasst werden.

²Das klassische Thermostatventil zur Regelung der Raumtemperatur kann beispielsweise als P-Regler betrachtet werden [APG14, Abschn. 10.3].

der Regelgüte zudem oftmals Zuluft-Abluft-Kaskadenregelungen implementiert (siehe Abbildung 4.2).

Eine Besonderheit bei der Regelung gebäudetechnischer Systeme besteht darin, dass für die Beeinflussung einer Regelgröße oftmals mehrere Stellglieder verwendet werden müssen, weswegen die Aktoren in einer sogenannten *Sequenzansteuerung*³ betrieben werden. Ein typisches Beispiel hierfür ist die Regelung der Zulufttemperatur und ggf. der -feuchte innerhalb einer Klima- bzw. Lüftungsanlage⁴, welche durch die Erhitzer- und Kühlerventile, und ggf. mittels einer Wärmerückgewinnung oder Umluftklappen, sowie über die Drehzahl der Befeuchterpumpe beeinflusst werden. Für eine ausführlichere Diskussion üblicher Regelungsansätze für Lüftungs- und Klimaanlage sei auf [APG14, Kap. 11] und [HC18] verwiesen.

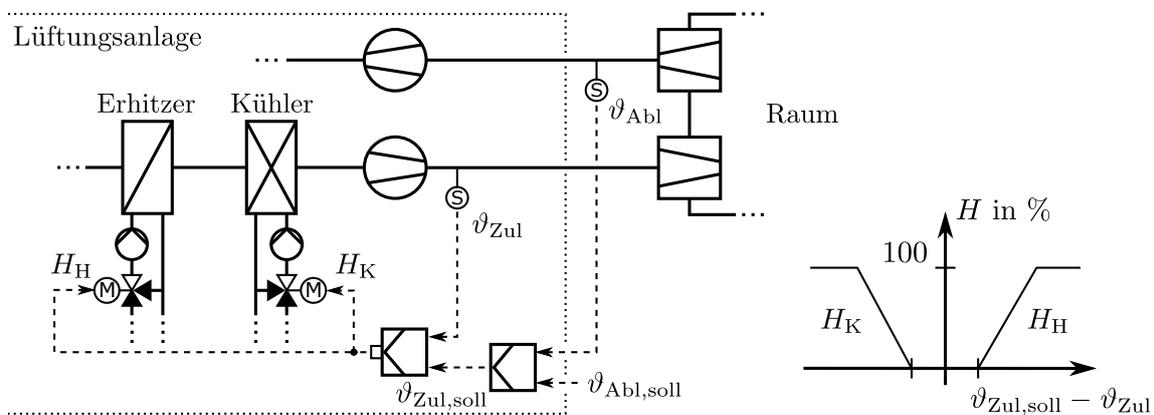


Abbildung 4.2: Zuluft-Abluft-Kaskadenregelung und Sequenzdiagramm für die Ansteuerung des Erhitzer- bzw. Kühlermischventils

Das dynamische Verhalten der Raumtemperatur lässt sich als näherungsweise linear bezüglich der Heiz- bzw. Kühlleistung annehmen (siehe Abschn. 5.1), welche in der Praxis jedoch meistens nicht als Stellgrößen betrachtet werden können. Stattdessen werden diese durch die Veränderung des Massenstroms und/oder der Vorlauf- bzw. Zulufttemperatur des Heizmediums beeinflusst. Die dem Raum zugeführte Heiz- bzw. Kühlleistung verhält sich dann näherungsweise proportional zu dem Produkt aus Massenstrom und der Differenz aus Zuluft- und Raumtemperatur, sodass sich aus systemtheoretischer Sicht

³Zur Veranschaulichung der grundlegenden Funktionsweise der Regelung werden nach VDI 3525 häufig Sequenzdiagramme verwendet (siehe Abbildung 4.2 rechts). Damit sich dabei der Sequenzablauf einfach nachvollziehen lässt, werden die Diagramme in der Praxis grundsätzlich so dargestellt, als würde es sich um P-, oder Mehrpunktregler handeln. Hierbei ist jedoch zu beachten, dass im Falle kontinuierlicher Regler in der Praxis im Allgemeinen nicht P-, sondern wesentlich häufiger auch PI- bzw. PID-Regler, implementiert werden.

⁴Nach DIN EN 12792 wird in Lüftungsanlagen nur die Temperatur geregelt, wohingegen in Klimaanlage auch die Feuchte zu regeln ist.

ein bilineares Systemverhalten ergibt^{5,6} (vgl. bspw. [OPJ10; KB11; OPJ12]).

Die Wärme- bzw. Kälteübergabe in den Räumen kann demnach zum einen durch die Veränderung des Massenstroms und zum anderen durch die Vorlauf- bzw. Zulufttemperatur des Heizmediums beeinflusst werden. Da eine Anpassung jedoch stets mit gerätetechnischem Aufwand verbunden ist, erfolgt die Festlegung einer Größe in der Regel auf einer übergeordneten Ebene meist innerhalb der Heiz- bzw. Lüftungszentrale. Der verbleibende Freiheitsgrad wird dann als Stellgröße für die eigentliche Raumtemperaturregelung verwendet.

In den meisten Gebäuden wird für die Heizung daher eine zentrale Vorlauftemperaturregelung implementiert, wobei mehrere Verbraucher häufig entsprechend ihrer benötigten Temperaturniveaus zu einem Heizkreis zusammengefasst werden. Die Regelung der Vorlauftemperatur ϑ_{V1} eines Heizkreises erfolgt dann mithilfe eines Mischventils und wird oft einem außentemperaturabhängigen Sollwert $\vartheta_{V1,soll}(\vartheta_{Aul})$ nachgeführt (siehe Abbildung 4.3). Der Zusammenhang $\vartheta_{V1,soll} = f(\vartheta_{Aul})$ wird *Heizkurve* genannt und ist ebenfalls beispielhaft in der Abbildung 4.3 dargestellt. Diese kann typischerweise durch einen Steigungsparameter, die Parallelverschiebung und den Heizkörperexponenten parametrisiert werden, wobei letzterer derart eingestellt wird, dass die Überlinearität der Leistungsabgabe der Heizflächen bezüglich der Temperaturdifferenz $\vartheta_{V1} - \vartheta_R$ näherungsweise kompensiert wird (siehe Anhang C.1 und Abb. C.1). Da während der Nacht oftmals reduzierte Komfortanforderungen gelten, wird die Heizkurve häufig durch eine zeitgesteuerte Nachtabsenkung der Vorlauftemperatur ergänzt.

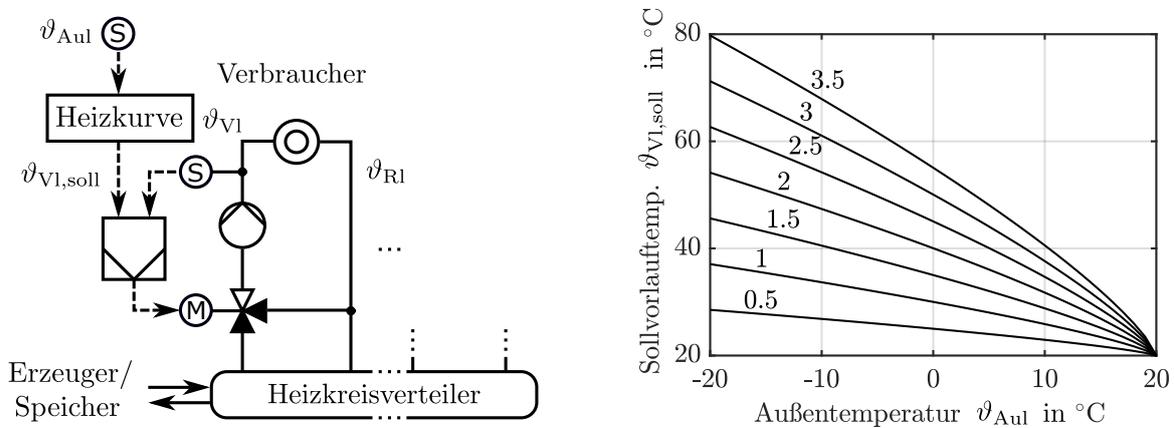


Abbildung 4.3: Vorlauftemperaturregelung (links) und beispielhaft dargestelltes Heizkurvendiagramm mit unterschiedlichen Steigungsparametern $\{0.5, 1, \dots, 3.5\}$ (rechts)

In analoger Weise muss bei der Lüftung mehrerer Räume durch eine zentrale Lüftungsanlage eine übergeordnete Zulufttemperatur vorgegeben werden. Im Unterschied

⁵Bei der Modellierung des dynamischen Verhaltens der absoluten Feuchte und der CO_2 -Konzentration lässt sich ein ähnlicher Ansatz verfolgen, wobei jedoch der Massenstrom und die Zuluftfeuchte bzw. Zuluft- CO_2 Konzentration als unabhängige Größen betrachtet werden müssen.

⁶Im Falle eines Radiator- bzw. Flächenheizsystems müssen weitere Nichtlinearitäten der Leistungsabgabe bezüglich des Massenstroms und der Differenz aus Vorlauf- und Raumtemperatur berücksichtigt werden (siehe Anhang C.1).

zur Regelung der Vorlauftemperatur müssen dabei häufig weitere Randbedingungen wie maximale bzw. minimale Zulufttemperaturen⁷. bzw. -feuchten und/oder minimale Raumlufwechselsraten berücksichtigt werden. Es sei weiterhin erwähnt, dass in vielen hauptsächlich über die Lüftungsanlage beheizten Gebäuden, zudem sogenannte Nachbehandlungszonen existieren, welche die Anpassung der Zulufttemperatur durch zusätzliche Erhitzer oder Kühler ermöglichen.

Weitere regelungstechnische Problemstellungen existieren im Bereich der Energieerzeugung bzw. -umwandlung. Dabei richtet sich deren Komplexität stark nach dem Energiesystem des jeweiligen Gebäudes bzw. Gebäudekomplexes. Wohingegen in kleineren Gebäuden oftmals lediglich ein Erzeuger existiert (monovalentes Energiesystem), werden heutzutage in u. a. moderneren Wohnquartieren, Krankenhäusern oder Freizeitzentren oftmals zentrale multivalente Energiesysteme mit mehreren Erzeugern, wie bspw. Wärmepumpen, Blockheizkraftwerken und/oder Heizkesseln, errichtet⁸. Zudem werden häufig regenerative Energiequellen (Photovoltaik und Solarthermie), sowie zusätzlich elektrische sowie thermische Speicher (Batterien, Pufferspeicher und thermische Bauteilaktivierung) integriert.

Die kosten- bzw. energieoptimale Regelung eines multivalenten Energiesystems kann aufgrund zeitlich eingeschränkt verfügbarer regenerativer Energiequellen, der Nutzung von Speichern und des zeitlich stark veränderlichen Wärme- bzw. Kälte- und Elektroenergiebedarfs eines Gebäudes eine regelungstechnisch anspruchsvolle Problemstellung darstellen. Weiterhin muss berücksichtigt werden, dass die Effizienz vieler Erzeuger von Umgebungsparametern wie der Außentemperatur oder von anderen internen Systemgrößen abhängig ist. Beispielsweise ist die Leistungszahl einer Wärmepumpe bzw. Kältemaschine und die Effizienz eines Gasbrennwertkessels entscheidend von der Rücklauftemperatur abhängig, welche wiederum durch andere Teilsysteme und durch externe Einflussfaktoren beeinflusst werden (Parametrierung der Heizkurven der Heizkreise, Außentemperatur, Raumtemperatursollwerte, ...). Eine regelbasierte Steuerung des Gesamtsystems und eine Optimierung auf der Ebene einzelner Teilsysteme ist für das Sicherstellen der Energie- bzw. Kostenoptimalität des Gesamtsystems daher oft nicht ausreichend.

Die obigen Beispiele verdeutlichen, dass beim Entwurf energie- bzw. kostenoptimaler Regelungsstrategien eine systemübergreifende, modellbasierte und vorausschauende Betrachtungsweise vorteilhaft ist. Aufgrund der Möglichkeit die Systemdynamik sowie Störgrößenprädiktionen explizit zu berücksichtigen, und dabei ein beispielsweise ökonomisches Kostenfunktional zu minimieren, entstand in den vergangenen Jahren verstärkt das Bestreben modellprädiktive Regelungsstrategien im Gebäudebereich einzusetzen.

⁷Diese sind in der Regel von der Raumtemperatur und der Art der Lufteinlässe in den jeweiligen Räumen abhängig.

⁸Praktische Beispiele können u. a. den Broschüren [Geo19] oder [Vie19] entnommen werden.

4.2 Modellprädiktive Regelung in der Gebäudetechnik

4.2.1 Einführung

Beginnend mit Anwendungen innerhalb der chemischen Prozessindustrie Mitte der 1970er Jahre [RRT78; CR80; DP06], hat sich die modellprädiktive Regelung auch in anderen Anwendungsbereichen, u. a. beispielsweise in der Raumfahrt, der Automobiltechnik oder der Landwirtschaft stark verbreitet [HCT12; Har15; DWL18]. Dabei besteht die Besonderheit darin, dass die Stellgrößen durch zyklische Lösung eines Optimalsteuerungsproblems berechnet werden, wobei ein dynamisches Prozessmodell für die Prädiktion der Regelgrößen verwendet wird. Ein wesentlicher Grund für den Erfolg modellprädiktiver Regelungsstrategien besteht u. a. darin, dass MPC insbesondere für die ökonomisch effiziente Lösung komplexer MIMO-Problemstellungen unter Berücksichtigung expliziter Störgrößenprädiktionen sowie Stell- und Regelgrößenbeschränkungen geeignet ist. Für eine ausführliche Einführung in die Theorie der modellprädiktiven Regelung seien dem interessierten Leser u. a. die Lehrbücher [Mac00; CB07; BBM17] und die darin enthaltenen Referenzen nahegelegt.

Aufgrund des zunehmenden ökologischen Bewusstseins, strenger werdender gesetzlicher Vorgaben und der zunehmenden Digitalisierung in der Gebäudetechnik entstand in den vergangenen Jahren das Bestreben MPC auch für den effizienten Betrieb von Gebäuden einzusetzen. Dies wird insbesondere durch die seit etwa 2010 stark zunehmende Anzahl wissenschaftlicher Publikationen deutlich, welche sich der Implementierung und dem Vergleich von MPCs für unterschiedliche Gebäude- und Anlagenkonfigurationen widmen. In der auf der Auswertung von 82 Publikationen beruhenden Überblicksstudie [SFC18] werden die Energieeinsparpotenziale im Mittel mit etwa 20 – 25% (Median: 10 – 15%) bewertet. Dabei muss allerdings berücksichtigt werden, dass abhängig von den konkreten Gebäudecharakteristika und den sonstigen Rand- und Betriebsbedingungen eine erhebliche Streuung existiert. Besonders vorteilhafte Eigenschaften, für welche potenziell hohe Energie- bzw. Kostenersparnisse durch MPC erzielt werden können, sind u. a. in [KH15; HKS16] aufgeführt:

- Langsames dynamisches Verhalten des betrachteten Prozesses („große Zeitkonstanten“), z. B. durch eine große thermische Speicherfähigkeit, gute Isolierung der Gebäudehülle oder durch explizite thermische Speicher (Puffer- oder Latentwärmespeicher⁹ sowie thermische Bauteilaktivierung¹⁰)
- Stark schwankende externe bzw. interne Gewinne durch präzisierbare Störungen (z. B. Außentemperatur, Solarstrahlung bzw. Belegung/Gebäudenutzung)

⁹Diese werden oftmals auch als Phasenwechsel- oder PCM-Speicher bezeichnet (phase change materials – PCM).

¹⁰Bei der thermischen Bauteilaktivierung (auch Betonkern- oder Baukernaktivierung) werden wasserdurchströmte Rohre in die Gebäudemasse eingebettet, welche zur Kühlung bzw. Heizung des Gebäudes verwendet werden können (thermally activated building system – TABS).

- Variable Energiekosten oder Nutzung zeitlich eingeschränkt verfügbarer Energiequellen (erneuerbare Energien)
- Signifikante Abhängigkeiten der Effizienz der Energieerzeuger von veränderlichen internen Systemgrößen (z. B. Rück- oder Vorlauftemperatur), oder von Umgebungsgrößen wie beispielsweise der Außentemperatur
- Großer Behaglichkeitsbereich sowie Tag/Nacht- und/oder Wochenendumschaltungen der Sollraumluftparameter (siehe Abschnitte 6.1 und 6.2)
- Signifikanter Hilfsenergieverbrauch der Gebäudetechnik mit ggf. nichtlinearen Abhängigkeiten zu variablen internen System- oder Stellgrößen (bspw. Elektroenergieaufnahme der Lüfter in einer Lüftungsanlage – siehe Abschnitt 6.2)

In diesem Zusammenhang sei jedoch auch erwähnt, dass der quantitative Vergleich von Regelungsansätzen aufgrund individuell definierter Vergleichsreferenzen und unterschiedlicher Randbedingungen (Wetter, Belegungsprofile, usw.) i. d. R. nicht möglich ist. Neueste Arbeiten zielen daher auf die Entwicklung von Referenzmodellen und definierter Softwareumgebungen zur simulationsbasierten Validierung und dem Vergleich fortschrittlicher Regelungsansätze in der Gebäudetechnik [BJH19]. Weiterhin sollte hervorgehoben werden, dass etwa 80% der Veröffentlichungen rein simulationsbasierte Studien darstellen, und dass sich die praktischen Implementierungen im Wesentlichen auf Demonstratoren im universitären Umfeld beschränken (siehe z. B. [PSF11; MBH12; BKB14; CH15; SGM16; FWM17]). Der Einsatz von MPC in der Gebäudetechnik kann daher nach wie vor als Gegenstand der insbesondere anwendungsorientierten Forschung betrachtet werden.

4.2.2 Literaturüberblick

In der Literatur existieren bereits zusammenfassende Veröffentlichungen, welche sich der Anwendung und dem Entwurf modellprädiktiver Regler im Gebäudebereich widmen [KZ13; AJ14; HKS16; MD16; AJF17; SFC18; DAC20]. Dabei werden die den Untersuchungen zugrundeliegenden Referenzen meist hinsichtlich verschiedener Implementierungsaspekte, wie bspw. dem Typ des verwendeten Modells (linear, bilinear, nichtlinear) sowie des u. a. daraus resultierenden Optimierungsproblems, des Prädiktionshorizonts, der Abtastzeit und der verwendeten Softwareplattform verglichen. Weitere Vergleichsaspekte sind u. a. der betrachtete Gebäudetyp bzw. das zu optimierende Gebäudeteilsystem und die durch MPC erzielten Energie- bzw. Kostenersparnisse.

In diesem Abschnitt soll ein anwendungsorientierter Überblick der publizierten MPC-Ansätze im Gebäudebereich vermittelt werden. Aufgrund der großen Anzahl verfügbarer Veröffentlichungen werden dabei lediglich Beiträge berücksichtigt, die sich der Implementierung von MPCs auf Gebäudeebene widmen. Publikationen die sich dem Entwurf übergeordneter MPC-basierter Regelungsstrategien auf lokaler Netzebene widmen („Microgrids“), werden in der Übersicht daher nicht erfasst (siehe u. a. [PRG14; HVB16; SS16; PWK17]).

Die überwiegende Anzahl der publizierten MPC-Ansätze im Gebäudebereich lassen sich in das Spektrum der modellprädiktiven *Einzelraum-* bzw. *Mehrzonen-* und der *Energiesystemregler* einordnen. Diese können entsprechend der gebäudetechnischen Funktionsebene (siehe Abbildung 4.1) und somit anhand der betrachteten Regel- und Stellgrößen unterschieden werden. Eine entsprechende Übersicht und die jeweilig zugeordneten Referenzen sind in der Tabelle 5 aufgeführt.

	Einzelraum	Mehrzonen	Energiesystem
Funktions- ebene	Übergabe	Übergabe - Erzeugung	Erzeugung und Speicherung
Regel- größen	Temperatur, relative Feuchte, Luftqualität eines Raumes	wie bei Einzelraumreglern jedoch mehrere Räume	Speicher-, Zulufttemp. und/oder Vorlauftemperatur der Heizkreise
Stell- größen	Heiz- bzw. Kühlventilansteuerung, Zuluftvolumenstrom, Verschattungsposition	wie bei Einzelraumreglern jedoch evtl. zusätzlich Zuluft- und/oder Vorlauftemperaturen	Ansteuerung oder Vorlauftemperaturen der Erzeuger, Ventilpositionen sowie Massenströme innerhalb des Energiesystems
Referen- zen	[MBD10; HGP12; LAB12a; PFH16; PFH17]	[YP06; MBD10; KB11; PSF11; MB12; BKB14; SGM16]	[Wim05; KSR08; RH11; MBH12; HBB12; Kub20]

Tabelle 5: *Klassifikation der in der Literatur publizierten MPC-Ansätze für Gebäude*

Ein naheliegendes Anwendungsfeld von MPCs im Gebäudebereich sind modellprädiktive Einzelraumregler, welche in gewisser Weise als „Ersatz“ konventioneller Raumregler betrachtet werden können. Diese zeichnen sich dadurch aus, dass die Stellsignale unabhängig von den übrigen Räumen berechnet werden. Im Gegensatz dazu wird bei Mehrzonenreglern ein, entweder durch den geometrischen Aufbau der Räume und/oder durch die zusätzliche Betrachtung einer übergeordnet festzulegenden Systemgröße (z. B. gemeinsame Zuluft- oder Vorlauftemperatur), verkoppeltes MIMO-Modell verwendet. Die Stellgrößen werden dann durch die Lösung eines für die betrachteten Räume bzw. Zonen gemeinsam formulierten Optimierungsproblems berechnet. Eine beispielhafte schematische Darstellung des in [KB11; MB12] betrachteten Anwendungsfalls ist in Abbildung 4.4 dargestellt. Dabei ist jedoch zu beachten, dass sich der in den genannten Referenzen betrachtete Anlagenaufbau und somit insbesondere die Stellgrößen deutlich unterscheiden können (Vergleiche bspw. [KB11; MB12] und [BKB14] sowie [SGM16]). Im Gegensatz zu Einzelraum- und Mehrzonenreglern operieren Energiesystemregler hauptsächlich auf der Funktionsebene der Energieerzeugung und -speicherung. Für gewöhnlich besteht dabei das wesentliche Ziel in der bedarfsgerechten und kosteneffizienten Umwandlung der Primärenergieträger in Nutzenergie sowie der Minimierung von Speicher- und Verteilverlusten. Ein weiteres charakteristisches Merkmal ist, dass

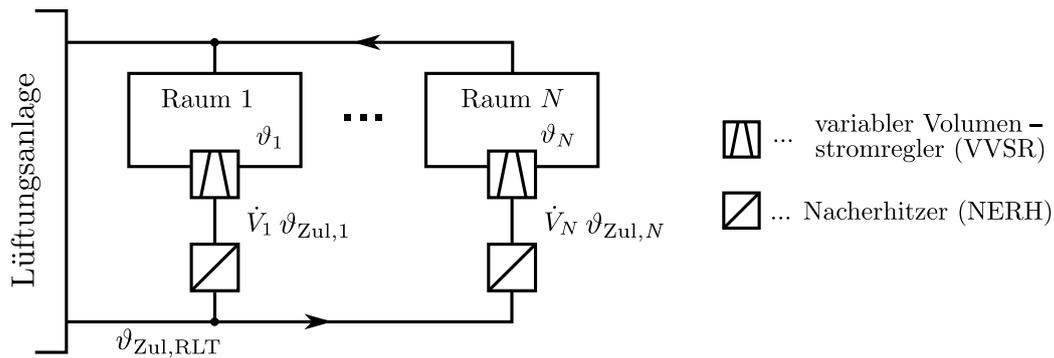


Abbildung 4.4: Schematisch beispielhafter Anlagenaufbau eines Mehrzonensystems (siehe [KB11; MB12])

dabei in der Regel stark vereinfachte, meist basierend auf den realen Messdaten des zu versorgenden Gebäudes bzw. Gebäudekomplexes gewonnene, Prognosemodelle des Wärme- bzw. Kältebedarfs verwendet werden.

Ein typisches Beispiel eines solchen Szenarios ist die in [MBH09; MBH12] diskutierte modellprädiktive Regelstrategie eines durch 2 Kältemaschinen beladenen 7000m^3 Wasserspeichers, der für die Kälteversorgung des Campus der Universität Merced in Kalifornien eingesetzt wird. Im Speziellen besteht dabei die Aufgabe darin den Ladezustand bzw. genauer die Position der Thermokline des Schichtladespeichers derart zu regeln, dass der Kältebedarf des Campus zu jeder Zeit gedeckt werden kann und zum anderen die Betriebskosten für die Kälteerzeugung und -verteilung minimiert werden.

Im Gegensatz zu den bisher erwähnten Klassen werden in der Literatur weiterhin MPC Gesamtgebäuderegler vorgeschlagen, deren Ziel eine möglichst kosten- bzw. energieoptimale ganzheitliche Gebäuderegung ist (siehe u. a. [LAB12a; LAB12b; CH15; FWM17; Jor18; JPS21; HSJ21]). Diese überspannen die Funktionsebenen der Erzeugung bis hin zur Abnahme und können als Kombination von MPC-Mehrzonensystem- und -Energiesystemreglern betrachtet werden. Zudem wird die Optimierung dabei oftmals hierarchisch in über- und untergeordnete Optimierungsprobleme zerlegt, die dann sequentiell gelöst werden (siehe bspw. [LAB12a; LAB12b; FWM17]).

Kapitel 5

Anwendung der Systemidentifikation in der Gebäudetechnik

Die Grundlage für die Auslegung modellprädiktiver Regelungsstrategien ist ein dynamisches Modell des zu regelnden Prozesses. Die dabei wesentlichen Anforderungen sind zum einen die hinreichende Genauigkeit und zum anderen ein möglichst geringer Rechenaufwand für die Prädiktion der Regelgrößen.

Aufgrund der meist individuellen Konstruktion und der technischen Gebäudeausstattung (TGA) ist die Modellierung und Validierung der Gebäude- und TGA-Komponenten in der Praxis ein meist zeitaufwendiger und daher kostspieliger Teil des Entwurfsprozesses. Dies führt dazu, dass der Einsatz von MPC im Gebäudebereich aus ökonomischer Sicht in vielen Fällen noch nicht rentabel ist. Die effiziente und teilautomatisierte Generierung MPC geeigneter Modelle wird daher als eines der wesentlichen Hindernisse für die großflächig-industrielle Verbreitung modellprädiktiver Regelungsstrategien im Gebäudebereich betrachtet [Pri13; PCZ13; SFC18].

Bei der Modellierung der relevanten Gebäude- bzw. TGA-Komponenten werden in der Literatur unterschiedliche Ansätze verfolgt. Erstere Möglichkeit besteht in der Nutzung von detaillierten White-Box-Simulationsmodellen, welche mittels spezieller Software, wie bspw. TRNSYS oder ENERGYPLUS, erzeugt werden können (siehe u. a. [ZAN14; ABS16; ZPY16]). Dabei führt jedoch der vergleichsweise hohe Detailgrad der Modelle zu einer für MPC-Zwecke oftmals zu langen Simulationszeit. Weiterhin muss berücksichtigt werden, dass die Beschaffung der notwendigen Modellparameter in der Praxis oftmals schwierig ist. Die Konfiguration und Validierung des Modells ist daher ein oftmals beschwerlicher und zeitaufwendiger Prozess. Die Verwendung kalibrierter White-Box-Simulationsmodelle für MPC-Zwecke wird in [SFC18] daher als generell wenig vielversprechend betrachtet.

Allerdings sei darauf hingewiesen, dass sich insbesondere im akademischen Bereich die objektorientierte Modellierungssprache MODELICA stark verbreitet hat, was vor allem an der Vielzahl an verfügbaren Modellierungsbibliotheken für den Gebäudebereich ersichtlich wird (BUILDINGS-Library, AIXLIB, BUILDINGSYSTEMS, IDEAS). Unter anderem existieren hierbei auch Bestrebungen die entsprechende Modelle, in Verbindung mit Softwarewerkzeugen für das algorithmische Differenzieren (z. B. CASADI¹), direkt

¹<https://web.casadi.org/>

für den Entwurf prädiktiver Regler einzusetzen [Jor18; JBH19; JPS21].

In den meisten Veröffentlichungen werden jedoch einfachere, durch Bilanzierung gewonnene, konzentriert parametrische Modellansätze bevorzugt (siehe u. a. [OPJ10; LMT12; OPJ12; LGG13; CH15]), was in diesem Zusammenhang meist unter dem Begriff der Grey-Box-Modellierung zusammengefasst wird. Dabei wird insbesondere im Bereich der thermischen Modellierung des Gebäudes oftmals die Analogie zur Elektrotechnik ausgenutzt, indem die Massen der Raumluft und der Wände durch verschaltete Widerstands-Kapazitätsnetzwerke abgebildet werden. Diese können dann mittels der Konstruktionsdaten des Gebäudes bzw. der Datenblätter der TGA parametrisiert werden, wodurch sich letztlich ein lineares bzw. bilineares² Zustandsraummodell gewinnen lässt [Stu14]. Für die vereinfachte Generierung solcher Modelle wurden in der Vergangenheit bereits Softwarewerkzeuge, wie bspw. die frei verfügbare BRCM-TOOLBOX, entwickelt [SGS14; Stu14].

Die Vorteile des genannten Modellierungsansatzes sind insbesondere die gute Interpretierbarkeit des resultierenden Modells sowie die Unabhängigkeit von Messdaten. Diese können daher bereits während der Planungsphase des Gebäudes entwickelt werden. Nachteilig hingegen ist auch hier die oftmals schwierige Beschaffung der notwendigen Parameter und die aufwendige Kalibrierung und Validierung des Modells. Um die genannten Nachteile teilweise zu vermeiden, können die genannten Grey-Box-Modelle deshalb auch mittels aufgezeichneter Messdaten parametrisiert werden³, was u. a. in [BM11; SKM14] verfolgt wird.

Ein weiterhin in der Praxis nicht zu unterschätzender und in der Literatur oftmals vernachlässigter Nachteil obig beschriebener Ansätze besteht in der meist rein deterministisch motivierten Modellbildung. Im Sinne einer stochastischen Betrachtungsweise handelt es sich daher um *Simulationsmodelle* (siehe Abschnitt 2.3). Für den Einsatz modellprädiktiver Regelungsstrategien werden jedoch *Prädiktionen* der relevanten Systemgrößen benötigt. Bei der Verwendung eines Zustandsraummodells bedeutet dies, dass zu jedem Zeitpunkt eine Schätzung des aktuellen Systemzustands benötigt wird, sodass in der Praxis zudem ein Zustandsbeobachter entworfen bzw. ausgelegt werden muss. Aufgrund der Modellkomplexität, der endlichen Auflösung der Sensorik (Quantisierungsrauschen) und oftmals erheblicher nicht messbarer Störeinflüsse, wie bspw. der Wärmeflüsse durch In- bzw. Exfiltration der Luft, menschliche Aktivität oder durch elektrische Geräte, stellt dies eine oftmals nicht triviale Aufgabe dar. Weiterhin sollte beachtet werden, dass aufgrund des nicht berücksichtigten stochastischen Anteils, keinerlei Optimalitätsbedingungen für die Prädiktionen des Modells sichergestellt werden können.

²Dies hängt im Wesentlichen von der Definition der Eingangssignale des Modells ab. Werden beispielsweise Heiz- bzw. Kühlleistungen als Stellsignale verwendet, dann lässt sich das dynamische Verhalten oftmals ausreichend gut durch einen linearen Modellansatz beschreiben [PZ18]. Werden hingegen Zuluft- bzw. Vorlauftemperaturen sowie Ventilpositionen oder Massenströme als Stellsignale betrachtet, dann führt dies in der Regel zu einem bilinearen, möglicherweise sogar nichtlinearen, Zustandsraummodell [Stu14].

³Für die Schätzung der Modellparameter eignet sich beispielsweise die `greyest()` Routine, welche durch die MATLAB SYSTEM IDENTIFICATION TOOLBOX zur Verfügung gestellt wird.

Sofern entsprechende Messdaten aus dem Gebäudebetrieb vorliegen, bietet sich deshalb eine datengetriebene stochastische Betrachtungsweise bei der Modellbildung an. Hierbei wurden in der Vergangenheit sowohl PEM-Verfahren [FS10; MCL10; PZ18; PZR19], als auch die unterraumbasierten Verfahren der Subspace-Identifikation eingesetzt [FS10; PSF11; PCV12; PCZ13]. Letztere zeichnen sich aus Anwendersicht insbesondere durch die verhältnismäßig unkomplizierte Anwendung der Algorithmen aus. Zudem sei erwähnt, dass die Subspace-Identifikation in gewisser Weise auch als eine Optimierung der Mehrschrittprädiktionsfehler betrachtet werden kann [TH05], was in Hinblick auf die Nutzung des Modells für MPC-Zwecke als vorteilhaft erscheint. Nachteilig gegenüber der PEM-Methode hingegen ist, dass die üblichen Implementierungen dieser Algorithmen in der Regel nur sehr begrenzte Möglichkeiten bieten physikalisches Vorwissen in der Modellstruktur zu berücksichtigen. Außerdem sei erwähnt, dass die Anpassung von Subspace-Algorithmen auf nichtlineare Modellstrukturen, wie beispielsweise Hammerstein-Modelle, als aufwendiger betrachtet werden muss (siehe z. B. [GPS05]). Die PEM-Methode hingegen beruht in der Regel auf der Nutzung polynomialer Modellansätze und der numerischen Lösung eines Optimierungsproblems (siehe Abschnitt 3.2). Dabei können bei der Optimierung ausgewählte Parameter gezielt festgelegt werden, wodurch sich der Rechenaufwand in einigen Anwendungen signifikant reduzieren lässt (vgl. Beispiel 3.3).

Aus Gründen der Vollständigkeit seien zudem Identifikationsansätze erwähnt, die auf der Nutzung (rekurrenter) neuronaler Netze basieren [MZ04; MCL11], welche als sehr allgemeine Modellstrukturen betrachtet werden können. Deren wesentlicher Nachteil besteht allerdings in der schweren Interpretierbarkeit des Modells, sodass die Validierung praktisch lediglich datenbasiert erfolgen kann.

Im Rahmen der folgenden Abschnitte soll die Anwendung der Systemidentifikation für einige für den Entwurf von prädiktiven Gebäuderegeln besonders relevant erachtete Beispielprozesse demonstriert werden, wobei für die Identifikation das PEM-Verfahren eingesetzt wird. Die dabei verwendeten Messdaten wurden während der Nutzung unterschiedlicher im Folgenden genauer charakterisierter Gebäude aufgezeichnet und können gemeinsam mit dem Quellcode für die Identifikation aus dem Digitalarchiv

https://github.com/fpaschke/phd_identification_mpc.git

abgerufen werden.

Obwohl wie oben erwähnt die PEM-Methode bereits zur Identifikation im Gebäudebereich eingesetzt wurde, bestehen praktisch relevante Aspekte bzw. Fragestellungen, die in der Literatur bisher nicht oder nur unzureichend genau untersucht wurden:

- Methoden und Ansätze für die Identifikation nicht- bzw. bilinearer Systeme sind im Gebäudebereich bisher nicht untersucht worden.
- Auf den Einfluss und die Modellierung nicht messbarer Störgrößen, durch bspw. Wärmeflüsse aus Nachbarräumen, Gebäudenutzung oder durch elektrische Geräte wird oftmals nicht genauer eingegangen.
- In der Literatur wird bei der Parameteridentifikation im Regelfall lediglich die 1-Schritt-PEM verwendet. Die Sinnhaftigkeit der Nutzung von k -Schritt bzw.

Mehrschrittmethoden ist weitestgehend unerforscht.

- Der Einfluss unterschiedlicher Modellparametrierungen (ARX, ARMAX, BJ, OE) wird oftmals nicht systematisch untersucht.
- In vielen Veröffentlichungen werden bei der Identifikation Messdaten von Demonstratoren benutzt, welche durch vordefinierte Experimente gewonnen wurden [MH95; AMH00; BM11; SKM14]. In dieser Arbeit hingegen werden ausschließlich Messdaten aus der Nutzungszeit realer Gebäude verwendet.

5.1 Thermische Raummodelle

Ein bedeutendes Einsatzgebiet prädiktiver Regelungsansätze im Gebäudebereich sind modellprädiktive Einzelraum- und Mehrzonenregler (siehe Abschn. 4.2.2). Für deren Entwurf werden Modelle für die Vorhersage der Komfortparameter der Luft, insbesondere der Raumtemperatur benötigt. Obwohl bei der Modellierung im Folgenden ein empirischer Ansatz verfolgt wird, ist es zunächst sinnvoll die Wärmebilanz der Luft in einem geschlossenem Raum zu betrachten [OPJ10; Nic12]:

$$C_R \frac{d\vartheta_R(t)}{dt} = \dot{Q}_H(t) + \dot{Q}_K(t) + \dot{Q}_L(t) + \dot{Q}_F(t) + \dot{Q}_S(t) + \dot{Q}_W(t) + \dot{Q}_B(t), \quad (5.1)$$

Dabei kennzeichnen ϑ_R die Temperatur und C_R die als konstant angenommene Wärmekapazität der Raumluft⁴. Weiterhin beschreiben

- \dot{Q}_H bzw. \dot{Q}_K die durch Konvektion und Strahlung durch das Heiz- bzw. Kühlsystem abgegebenen Wärmeströme (Radiator- und/oder Flächenheizung bzw. Umluftkühlgeräte),
- \dot{Q}_L den volumenstromabhängigen Wärmegewinn bzw. -verlust durch erzwungene Lüftung,
- \dot{Q}_F die Transmissionswärmeverluste bzw. -gewinne durch die Fenster,
- \dot{Q}_S die solaren Wärmegewinne,
- \dot{Q}_W die von den Wandinnenoberflächen konvektiv an die Raumluft abgegebene Wärme und
- \dot{Q}_B die durch menschliche Aktivität („Belegung“) oder durch elektrische Geräte an die Raumluft abgegebene Wärme.

Man beachte, dass die Bilanzgleichung (5.1) im Folgenden nicht als allgemeiner Modellansatz für die Identifikation, sondern als Diskussionsgrundlage für u. a. die Wahl der Eingangsgrößen des zu identifizierenden Modells betrachtet wird. Es sei zudem vorweggenommen, dass aufgrund der starken Vereinfachungen nicht notwendigerweise eine

⁴Die Wärmekapazität der Raumluft ist von deren Dichte, und somit u. a. von der Temperatur und der Feuchte abhängig, kann jedoch im in der Gebäudetechnik relevanten Bereich als näherungsweise konstant betrachtet werden (siehe Anhang C.2).

Differentialgleichung erster Ordnung zu erwarten ist. Weiterhin sei erwähnt, dass (5.1) im Allgemeinen zur Identifikation ungeeignet ist, da die darin enthaltenen Wärmeströme teilweise nicht direkt messbar sind.

Die Wahl des Modellansatzes und der Eingangsgrößen ist einer der wesentlichen Aspekte bei der Identifikation dynamischer Systeme. Wie in der Gleichung (5.1) angedeutet, werden dabei im Folgenden 7 Wärmequellen berücksichtigt, von denen jedoch lediglich \dot{Q}_H , \dot{Q}_K , \dot{Q}_L , \dot{Q}_F und \dot{Q}_S explizit, d. h. durch das deterministische Teilsystem des Modells, berücksichtigt werden:

Die Heiz- und Kühlleistung \dot{Q}_H bzw. \dot{Q}_K sowie die Wärmegewinne bzw. -verluste durch erzwungene Lüftung \dot{Q}_L können durch die Aktorik, d. h. durch die Stellgrößen der Raumregelung beeinflusst werden. Im Hinblick auf den Einsatz der Modelle für MPC-Zwecke muss allerdings berücksichtigt werden, dass die Wärmeströme oftmals nicht als Stellgrößen des Systems betrachtet werden können. Stattdessen werden diese durch den Volumenstrom bzw. die Ventilpositionen des Heiz- bzw. Kühlsystems sowie durch die Vorlauf- bzw. Zulufttemperatur des jeweiligen Heizmediums beeinflusst (siehe Abschn. 4.1). In Anlehnung an [OPJ10; KB11; OPJ12] und motiviert durch die Lösungen der Heizkörpergleichungen (C.1) wird im Folgenden daher angenommen, dass sich die Wärmeströme näherungsweise entsprechend der Beziehungen

$$\dot{Q}_H(t) \propto f_H(H_H(t))(\vartheta_{Vl,H}(t) - \vartheta_R(t))^{n_H}, \quad (5.2a)$$

$$\dot{Q}_K(t) \propto f_K(H_K(t))(\vartheta_R(t) - \vartheta_{Vl,K}(t))^{n_K} \quad \text{und} \quad (5.2b)$$

$$\dot{Q}_L(t) \propto \dot{V}_{Zul}(t)(\vartheta_{Zul}(t) - \vartheta_R(t)) \quad (5.2c)$$

verhalten ($\vartheta_{Vl,H} \geq \vartheta_R$ und $\vartheta_{Vl,K} \leq \vartheta_R$), wobei die Bedeutung der verwendeten Symbolik der Tabelle 6 entnommen werden kann. Man beachte, dass dabei die Lösungen der Heizkörpergleichungen (C.1) durch den Ansatz $f_H(H_H)(\vartheta_{Vl} - \vartheta_R)^{n_H}$ angenähert werden, wobei f_H bzw. f_K statische Funktionen darstellen, für welche beispielsweise polynomiale oder exponentielle Ansätze verwendet werden können. Des Weiteren sei darauf hingewiesen, dass im Falle eines Sensoroffsets bei der Temperaturmessung weitere Korrekturparameter in den Ansätzen (5.2) vorgesehen werden können.

Entsprechend der Wärmebilanz (5.1) ist die Veränderung der Raumtemperatur weiterhin an die wetterabhängigen Wärmeverluste bzw. -gewinne \dot{Q}_F und \dot{Q}_S gekoppelt. Für die Transmissionsverluste bzw. -gewinne durch die Fenster wird angenommen, dass sich der Wärmestrom proportional zur Außentemperaturdifferenz

$$\dot{Q}_F(t) \propto \vartheta_{Aul}(t) - \vartheta_R(t) \quad (5.3)$$

verhält.

Für die Berücksichtigung der durch die Globalstrahlung \dot{Q}_{Gs} verursachten solaren Gewinne \dot{Q}_S wurde das insbesondere zur Ertragsberechnung von Photovoltaik- und Solarthermieranlagen verwendete Verfahren in [Qua13, Abschn. 2.6] eingesetzt. Hierbei werden die Messwerte der Globalstrahlung auf der Horizontalen zunächst in einen Diffus-

und einen Direktstrahlungsanteil zerlegt, d. h.

$$\dot{Q}_{\text{Gs}}(t) = \dot{Q}_{\text{dir}}(t) + \dot{Q}_{\text{diff}}(t), \quad (5.4)$$

wobei zur Berechnung des diffusen Strahlungsanteils der empirisch ermittelte Zusammenhang

$$\dot{Q}_{\text{diff}}(t) = \dot{Q}_{\text{Gs}}(t) \begin{cases} (1.020 - 0.254k_T(t) + 0.0123 \sin \alpha_{\text{Sh}}(t)) & \text{für } k_T \leq 0.3, \\ (1.400 - 1.749k_T(t) + 0.177 \sin \alpha_{\text{Sh}}(t)) & \text{für } 0.3 < k_T \leq 0.78 \text{ und} \\ (0.486k_T(t) - 0.182 \sin \alpha_{\text{Sh}}(t)) & \text{für } k_T > 0.78 \end{cases} \quad (5.5)$$

mit dem Faktor

$$k_T(t) = \frac{\dot{Q}_{\text{Gs}}(t)}{E_0 \sin \alpha_{\text{Sh}}(t)}, \quad (5.6)$$

der Solarkonstante $E_0 = 1360.8 \text{ W/m}^2$ und der Sonnenhöhe $\alpha_{\text{Sh}}(t)$ verwendet wird [RBD90; Qua13]. Anschließend wird die auf einer geneigten Ebene, bzw. auf den Fensterflächen, auftreffende Strahlung über

$$\dot{Q}_{\text{gen}}(t) = \dot{Q}_{\text{dir,gen}}(t) + \dot{Q}_{\text{diff,gen}}(t) + \dot{Q}_{\text{refl,gen}}(t), \quad (5.7)$$

mit

$$\dot{Q}_{\text{dir,gen}}(t) = \dot{Q}_{\text{dir}}(t) \frac{\cos \Theta(t)}{\sin \alpha_{\text{Sh}}(t)} \quad (5.8a)$$

$$\dot{Q}_{\text{diff,gen}}(t) = 0.5 \dot{Q}_{\text{diff}}(t) (1 + \cos \alpha_{\text{Fh}}) \quad (5.8b)$$

$$\dot{Q}_{\text{refl,gen}}(t) = 0.5 A \dot{Q}_{\text{Gs}}(t) (1 + \cos \alpha_{\text{Fh}}) \quad (5.8c)$$

berechnet⁵, wobei A den Albedowert des Untergrundes⁶ und $\Theta(t)$ den Winkel zwischen dem Normalenvektor der Fensterfläche und dem Vektor der in den Mittelpunkt der Sonne zeigt kennzeichnet:

$$\cos \Theta(t) = \cos \alpha_{\text{Sh}}(t) \sin \alpha_{\text{Fh}} \cos(\alpha_{\text{Saz}}(t) - \alpha_{\text{Faz}}) + \sin \alpha_{\text{Sh}}(t) \cos \alpha_{\text{Faz}} \quad (5.9)$$

Die standortspezifische Sonnenposition, d. h. der Azimutwinkel $\alpha_{\text{Saz}}(t)$ und die Sonnenhöhe $\alpha_{\text{Sh}}(t)$, wurde mit dem Algorithmus nach der DIN 5034 berechnet (siehe [Qua13, Abschn. 2.5]).

Einen weiteren oftmals signifikanten Einfluss auf die Raumtemperatur hat die Wärmeabgabe \dot{Q}_{B} der Raumbellegung. Dabei wird praktisch meist eine von den Umgebungsfaktoren und der ausgeführten Tätigkeit abhängige personenbezogene Wärmeabgabe

⁵Bei der Verwendung eines Verschattungssystems ist es naheliegend die Lamellenpositionen und ggf. den Anstellwinkel durch einen weiteren multiplikativen Faktor im Ansatz für $\dot{E}_{\text{dir,gen}}$ zu berücksichtigen. In ähnlicher Weise kann die Verdeckung der Sonne durch benachbarte Gebäude, Vegetation oder andere Objekte durch eine Schattenlinie $f_S(\alpha_{\text{Sh}}(t), \alpha_{\text{Saz}}(t))$ einbezogen werden.

⁶Ist der Untergrund bzw. die Umgebung unbekannt, wird meist der Wert $A = 0.2$ verwendet [Qua13].

Symbol	Einheit	Typ	Bedeutung
ϑ_R	$^{\circ}\text{C}$	Ausgangsgr.	Raumtemperatur
\dot{Q}_H, \dot{Q}_K	kW	Eingangsgr.	Heiz- bzw. Kühlleistung
$\vartheta_{V1,H}, \vartheta_{V1,K}$	$^{\circ}\text{C}$	Eingangsgr.	Vorlauftemp. der Heizung bzw. Kühlung
H_H, H_K	-	Eingangsgr.	Stellsignale des Heiz- bzw. Kühlsystems (bspw. Ventilposition oder Ansteuerung eines Umluftkühlgerätes)
$\vartheta_{Zul}, \dot{V}_{Zul}$	$^{\circ}\text{C}, \frac{\text{m}^3}{\text{s}}$	Eingangsgr.	Zulufttemp. bzw. -volumenstrom der Lüftung
ϑ_{Aul}	$^{\circ}\text{C}$	Eingangsgr.	Außenlufttemperatur
\dot{Q}_{Gs}	W/m^2	Eingangsgr.	Globalstrahlung auf der horizontalen Ebene
$\alpha_{Sh}, \alpha_{Saz}$	$\text{rad}/^{\circ}$	Eingangsgr.	Höhe bzw. Azimuth der Sonne
n_H, n_K	-	Parameter	Exponent
$\alpha_{Faz}, \alpha_{Fh}$	$\text{rad}/^{\circ}$	Parameter	Azimuth und Neigungswinkel der Fensterfläche gegenüber der horizontalen Ebene
A	-	Parameter	Albedo

Tabelle 6: Bedeutung der verwendeten Symbolik

von etwa 100 – 300W/Pers. angenommen [RSS11, Abschn. 1.2.2]. Eine naheliegende Möglichkeit zur Berücksichtigung der Wärmegewinne \dot{Q}_B besteht daher darin die Anzahl der Personen im Raum als Eingangsgröße des Modells zu verwenden. In der Praxis ist dies jedoch oftmals nicht möglich, da die entsprechenden Belegungsdaten in der Regel nicht vorliegen. Um diese Problematik bei der Identifikation zu berücksichtigen sind folgende Ansätze denkbar:

- **Datenvorauswahl:** Nutzung binärer Anwesenheitssensorik oder typischer Belegungszeiten (Zeitpläne/Kalender) zur Vorauswahl der Trainingsdaten (siehe Abschnitt 5.1.1).
- **Belegungsschätzung:** Nutzung korrelierter Sensordaten (z. B. CO₂-Gehalt der Raumluft oder relative bzw. absolute Luftfeuchte) zur Schätzung der Raumbelegung (siehe [EBV13; AFG14; EBV15]).
- **Störgrößenansatz:** Implizite Berücksichtigung der Raumbelegung durch Verwendung eines entsprechend parametrisierten Störgrößenmodells (siehe Abschnitt 5.1.2 und [Pas20]).

Eine weitere praktisch nicht messbare Wärmequelle stellt die von den Wandinnenoberflächen konvektiv an die Raumluft abgegebene Wärme \dot{Q}_W dar. Diese hängt u. a. von den jeweiligen Wandoberflächentemperaturen ab, welche wiederum an die Raumtemperaturen der angrenzenden Räume gekoppelt sind. Ein offensichtlicher Ansatz würde daher darin bestehen, die Raumtemperaturen der Nachbarräume als Eingangsgrößen des Modells zu verwenden, wodurch sich durch anschließende Verkopplung aller identifi-

zierten Einzelraummodelle ein thermisches Gesamtgebäudemodell ergibt (vgl. [SGS14; Stu14]). Dies bedeutet allerdings auch, dass für alle Räume des Gebäudes Modelle für die jeweiligen Raumtemperaturen erzeugt werden müssen, was praktisch eine vollständige Ausrüstung des Gebäudes mit entsprechender Sensorik erforderlich macht.

Ein weiteres Problem bei der Implementierung des beschriebenen Identifikationsansatzes besteht darin, dass in den meisten Gebäuden, beispielsweise in Bürogebäuden oder in Schulen, ein regelmäßiger Nutzungsrhythmus der Räume existiert. Dies wiederum führt dazu, dass die Temperaturen in den einzelnen Räumen, d. h. die Eingangs- und Ausgangssignale, untereinander zeitlich stark korreliert sind. Bei der Identifikation kann dies zu ungenauem oder sogar zu physikalisch unplausiblen Übertragungsverhalten führen, sodass u. U. der Entwurf und die Durchführung informativer Identifikationsexperimente⁷ notwendig ist (siehe z. B. [Lju12, Kap. 13]).

Zur Vermeidung der genannten Probleme wird im Folgenden deshalb eine alternative Herangehensweise verfolgt. Dabei wird der Wärmestrom $\dot{Q}_W(t)$ als tieffrequente Störgröße, d. h. als Drift der Ausgangsgröße betrachtet. Im Rahmen der im Folgenden diskutierten Anwendungsbeispiele werden daher Modellstrukturen mit einem Integrator im stochastischen Teilsystem verwendet, was wiederum der Differenziation der Ein- und Ausgangssignale des LTI-Blocks entspricht (siehe Gl. 3.10). Hierbei ist jedoch zu beachten, dass durch die Differenziation insbesondere der hohe Frequenzbereich sehr stark gewichtet wird (vgl. Beziehung (3.25)). Da die Modelle zur Mehrschrittprädiktion eingesetzt werden sollen, ist jedoch ebenfalls eine hinreichend gute Modellgüte im tieferen Frequenzbereich wünschenswert. Insbesondere bei der Verwendung von Modellstrukturen mit einem relativ unflexiblen Rauschmodell (ARX- und OE-Modell), ist aus praktischer Sicht daher die Verwendung eines Hochpassvorfilters

$$L(s) = \frac{s}{1 + T_L s} \quad (5.10)$$

zweckmäßiger, wobei s die komplexe Variable der *LAPLACE*-Transformation kennzeichnet. Die Zeitkonstante des Vorfilters T_L wird dann im Folgenden anwendungsspezifisch basierend auf der näherungsweise zu erwartenden Systemdynamik festgelegt.

5.1.1 Anwendungsbeispiel 1: Besprechungsraum

Der in diesem Abschnitt betrachtete 9.2m×5.7m große Besprechungsraum befindet sich im Erdgeschoss eines in den 1950er Jahren errichteten Massivbaus in Dresden. Der Raum wird sowohl als Demonstrator für Forschungszwecke sowie als gewöhnlicher Konferenzraum genutzt. Das Automationssystem des Raumes besteht aus den folgenden Komponenten:

- Multifunktionsdeckensensor zur Messung der Temperatur, relativen Feuchte und Luftqualität

⁷Anschaulich gesprochen wird ein Datensatz als informativ bezüglich einer gegebenen Modellmenge \mathcal{M} bezeichnet, wenn es möglich ist alle Modelle aus \mathcal{M} mittels des Datensatzes „zu unterscheiden“. Für eine formale Definition des Begriffes sei auf [Lju12, Abschn. 8.2] verwiesen.

- Drei Fenster mit Schaltkontakten und außen liegender elektrischer Verschattung⁸ sowie eine innen liegende Tür mit einem Schaltkontakt
- Binäre Belegungssensorik
- Drei Plattenheizkörper mit gleicher Vorlauftemperatur und jeweils eigenem elektronischen Funk-Heizkörperthermostat mit gemeinsamer Sollwertvorgabe (PID-Regler zur Regulierung der Ventilposition $H_H = 0, \dots, 100\%$)
- Drei Umluftkühlgeräte (Deckenkassetten) mit gleicher Vorlauftemperatur und diskreter Ansteuerung der Lüfterstufe $H_K \in \{0, 1, 2, 3\}$ sowie 4-Punkt Proportionalreglern mit gemeinsamer Sollwertvorgabe
- Zwei getrennte Wärmemengenzähler zur Messung der von den Radiatoren bzw. Umluftkühlgeräten abgegebenen Heiz- bzw. Kühlleistung sowie der jeweiligen Vor- und Rücklauftemperatur
- Zentrale Datenbank zur Speicherung aller verfügbaren Sensor- und Aktorsignale sowie der vorgegebenen Sollwerte

Eine schematische Darstellung des Besprechungsraumes ist in der Abbildung 5.1 dargestellt.

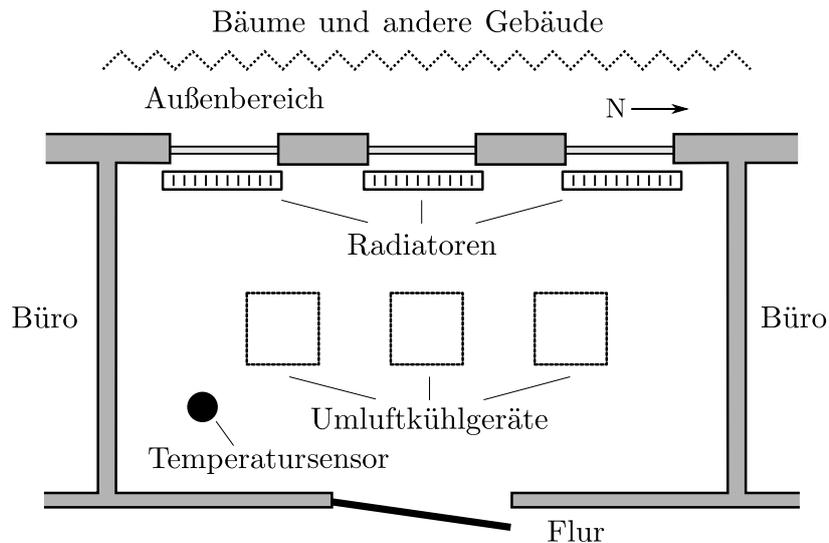


Abbildung 5.1: Schematische Darstellung des betrachteten Besprechungsraumes

Beschreibung des Datensatzes: Die im Folgenden verwendeten Daten wurden im Zeitraum vom 1.9.2015 bis zum 31.8.2016 aufgezeichnet. Aufgrund offensichtlich fehlerhafter Rohdaten, welche insbesondere durch Kommunikationsausfälle entstanden sind, musste eine manuelle Datenvorauswahl durchgeführt werden. Dabei wurden aus dem Jahresdatensatz insgesamt 100 separate Teilstücke mit minimaler Länge von 8h

⁸Die Lamellenpositionen und die Anstellwinkel der Verschattung stehen nicht als Messsignale zur Verfügung, weshalb im Folgenden angenommen wird, dass die Fenster unverschattet sind.

ausgewählt, in welchen der Besprechungsraum nicht genutzt wurde und zudem die benötigten Messsignale keine offensichtlichen Fehler enthielten. Daher wurden für die Identifikation und Validierung des Modells insgesamt lediglich ≈ 79 Tage bzw. $N = 22764$ Datenpunkte genutzt (Abtastzeit $T_A = 5\text{min}$). Für die Messungen der Außentemperatur und der Globalstrahlung wurden die Aufzeichnungen einer vom deutschen Wetterdienst (DWD) betriebenen Wetterstation in Dresden Klotzsche verwendet (Stationsnummer 1048, Abtastzeit 10min).

a) Hammerstein-Modell

Der in diesem Abschnitt betrachtete Konferenzraum wird u. a. als Demonstrator für Forschungszwecke eingesetzt und ist daher mit einer umfangreichen Sensorik ausgestattet, von welcher in der Praxis im Allgemeinen nicht ausgegangen werden kann. Hierzu gehören insbesondere die installierten Wärmemengenzähler, die zur Messung der im Raum abgegebenen Heiz- bzw. Kühlleistung verwendet werden. Im Rahmen der Identifikation können die Wärmeströme $\dot{Q}_H(t)$ und $\dot{Q}_K(t)$ daher zunächst als Eingangsgrößen betrachtet werden (siehe Gleichung (5.1)), was im Folgenden den quantitativen Vergleich eines Hammerstein-Modells mit einem nichtlinear rückgekoppelten Modellansatz der Form (2.51) ermöglicht (siehe Unterabschnitt b)).

Modellansatz: Als Eingangsgrößen des linearen Teilsystems wurden die Messwerte der Wärmemengenzähler $\dot{Q}_H[t]$ und $\dot{Q}_K[t]$ sowie die gemessene Außenlufttemperatur $\vartheta_{\text{Aul}}[t]$ der DWD Wetterstation verwendet. Ferner wurde zur Berücksichtigung solarer Gewinne die Strahlung $\dot{Q}_{\text{gen}}[t]$ als weitere Eingangsgröße des LTI-Blocks definiert (Gl. (5.7)), wobei die Ausrichtung der Fensterflächen zu $\alpha_{\text{Faz}} = 270^\circ$ und $\alpha_{\text{Fh}} = 90^\circ$ festgelegt wurde (vgl. Abb. 5.1).

Für den LTI-Block des Hammerstein-Modells wurden die in Abschnitt 2.4 genannten Parametrierungen verwendet, wobei die Ordnungen n der jeweilig verwendeten Polynome bei der Identifikation schrittweise erhöht wurden ($n = n_A = n_{B_i} = n_C = n_D = n_F$). Um die Modellstruktur des deterministischen Teilsystems $G_i(q^{-1})$ ($i = 1, \dots, 4$) auf die Klasse der nicht schwingungsfähigen Modelle einzuschränken, wurde eine linear faktorisierte Darstellung der Polynome $A(q^{-1})$ bzw. $F(q^{-1})$ und $B_i(q^{-1})$ verwendet (siehe Bem. 2.9). Es sei zudem erwähnt, dass alle Filter $G_i(q^{-1})$ mit einem gemeinsamen Nennerpolynom $A(q^{-1})$ bzw. $F(q^{-1})$ parametrisiert wurden. Außerdem wurde aufgrund des zu erwartenden Drifts der Ausgangsgröße ein Hochpassvorfilter (5.10) mit der Zeitkonstante $T_L = 2\text{h}$ verwendet.

Identifikation: Die Modelle wurden mithilfe der 102 ausgewählten Ausschnitte des Datensatzes mittels 1-Schritt bzw. Multischritt-PEM identifiziert ($k_{\text{max}} = 12$). Für die initiale Parameterschätzung $\hat{\theta}_{(0)}$ einer Modellstruktur der Polynomordnung n wurde der Wert des identifizierten Modells der Ordnung $n - 1$ verwendet, wobei für $n = 1$ der MkQ-Schätzer (3.37) genutzt wurde. Für die Lösung des Optimierungsproblems wurde die durch die MATLAB-Routine `lsqnonlin()` zur Verfügung gestellte Implementierung des Levenberg-Marquardt Verfahrens verwendet. Weiterhin wurde dem Löser der analytisch ermittelte Gradient des Kostenfunktional bereitgestellt (siehe Abschn. 3.2.1). Für die

Berechnung der Anfangswerte des Residuenfilters (2.27) wurde die Backforecasting-Methode eingesetzt (siehe Anhang B.2.2).

Diskussion: Die Abbildung 5.2 veranschaulicht die Prädiktionsgüte unterschiedlich parametrierter Hammerstein-Modelle im Sinne des Kriteriums

$$\text{RMMSE}(\varepsilon[t]) = \sqrt{\frac{1}{Nk_{\max}} \sum_{k=1}^{k_{\max}} \sum_{t=1}^N \varepsilon^2[t|t-k]} \quad \text{für } k_{\max} = 12. \quad (5.11)$$

Durch die Abbildung wird verdeutlicht, dass ARX- und OE-Strukturen⁹ die schlechteste Prädiktionsgüte im Sinne von (5.11) aufweisen. Dies lässt sich beim ARX-Modell darauf zurückführen, dass für das deterministische Teilsystem $\mathbf{G}(q^{-1})$ Tiefpassverhalten zu erwarten ist, wodurch aufgrund des gemeinsamen Nennerpolynoms $A(q^{-1})$ der hochfrequente Anteil der Prädiktionsfehlerfolge betont wird (vgl. Gl. (3.25)). Die ARX-Modelle besitzen daher oftmals eine befriedigende Güte im Sinne der 1-Schritt Prädiktion, sind für längere Prädiktionshorizonte jedoch eher ungeeignet (vgl. Bsp. 3.2 und Abb. 3.3).

Im Gegensatz zur ARX-Modellstruktur haben OE-Modelle eine frequenzunabhängige Wichtung (3.25), weswegen ein besseres Verhalten für längere Prädiktionshorizonte zu erwarten ist. Dennoch wurde in den Untersuchungen eine deutliche Überlegenheit der ARMAX- und BJ-Strukturen festgestellt, was auf die nicht ausreichende Flexibilität des Rauschmodells bei OE-Modellen zurückgeführt wird. An dieser Stelle sei jedoch auch erwähnt, dass die Güte der ARX- und der OE-Modelle im Gegensatz zur ARMAX- und BJ-Parametrierung stark von der Wahl des Vorfilters (5.10) abhängig ist. Weiterhin ist zu erwähnen, dass bei OE- und BJ-Modellen oftmals eine langsame Konvergenz bei der Optimierung und/oder eine Konvergenz zu lokalen Minima¹⁰ festgestellt wurde. Aus praktischer Sicht ist daher die Nutzung der ARMAX-Struktur empfehlenswert, da für diese eine hinreichende Flexibilität und zum anderen weniger Probleme bei der Lösung der Optimierungsaufgabe festgestellt wurden.

Die Abbildung 5.2 verdeutlicht weiterhin den Unterschied, der Multischritt-PEM gegenüber der 1-Schritt-PEM. Die bezogen auf das RMMSE Kriterium zu erwartende Verbesserung liegt typischerweise unter 10%, wobei für ARX-Modelle die tendenziell größten Verbesserungen erzielt werden konnten. Dies lässt sich wiederum damit begründen, dass durch den längeren Prädiktionshorizont dem obig erwähnten Nachteil der Überbetonung hoher Frequenzen entgegengewirkt wird. Für die übrigen Parametrierungen fällt der Unterschied hingegen wesentlich geringer aus, sodass aus praktischer Sicht die Identifikation mit einem Mehrschrittkriterium nicht zwingend erforderlich erscheint, sofern eine ARMAX- oder BJ-Parametrierung verwendet wird.

⁹Die im Folgenden verwendeten Abkürzungen beziehen sich stets auf den frei parametrierbaren Anteil der jeweiligen Modellstruktur. Berücksichtigt man das Vorfilter bei der Bezeichnung des ARX- bzw. OE-Modells, dann muss genau genommen von einer ARIMAX- bzw. BJ-Struktur gesprochen werden.

¹⁰Dies lässt sich bspw. in der Abb. 5.2 links für die BJ-Struktur mit $n = 2$ beobachten (ARMAX- ist ein Spezialfall der BJ-Parametrierung).

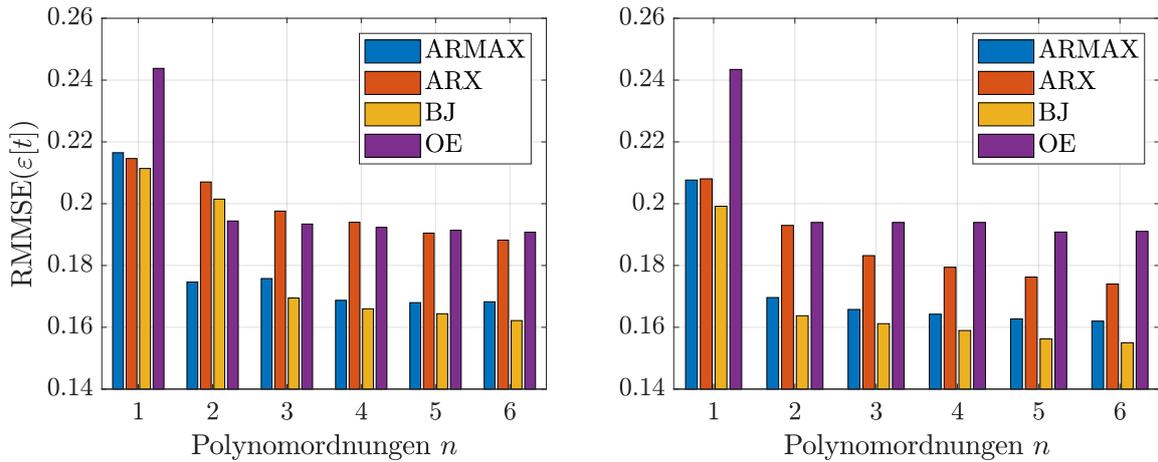


Abbildung 5.2: RMMSE unterschiedlich parametrierter¹⁰ Hammerstein-Modelle in Abhängigkeit der Polynomordnung n ($k_{\max} = 12$). Die Modelle wurden mit 1-Schritt- (links) bzw. 12-Multischritt-PEM (rechts) identifiziert.

Bemerkung 5.1 (Reduktion der Strukturparameter): Aus der Grafik 5.2 kann die notwendige Polynomordnung n einer Modellstruktur abgeschätzt werden. Da jedoch alle Polynome mit der gleichen Ordnung n angesetzt wurden, ist eine Überparametrierung möglich. Zur Ermittlung der minimal notwendigen Polynomordnungen hat es sich als zweckmäßig erwiesen mittels der Kovarianzmatrix (3.27) die relativen Parameterschätzfehler zu berechnen und die Ordnung des jeweiligen Polynoms dann zu reduzieren, falls diese einen vorgegebenen Grenzwert überschreitet. Sofern die Prädiktionsgüte des Modells dadurch nicht signifikant verschlechtert wird, ist es sinnvoll die reduzierte Parametrierung zu verwenden. Weitere Methoden die zur Ermittlung der Strukturparameter eingesetzt werden können sind in [SS89, Kap. 11] angegeben. \triangleleft

Basierend auf der Grafik 5.2 wurde eine ARMAX Parametrierung mit $n = 2$ ausgewählt. Mit dem in der Bemerkung 5.1 beschriebenen Vorgehen konnten die Strukturparameter weiterhin zu $n_A = 2$, $n_C = 1$ und $\mathbf{n}_B = (1, 2, 1, 1)$ reduziert werden, ohne dass eine signifikante Verschlechterung der Prädiktionsgüte festgestellt wurde.

Die Abbildung 5.3 veranschaulicht die Frequenzgänge der entsprechenden Modelle die mittels 1-Schritt bzw. 12-Schritt-PEM identifiziert wurden. Dabei lassen sich nur beim Frequenzgang $G_{12}(\omega)$ (Kühlleistung) nennenswerte Unterschiede feststellen. Auf der rechten Seite der Grafik ist zudem die stärkere Gewichtung niedrigerer Frequenzen bei der 12-Schritt-PEM ersichtlich. Entsprechend veranschaulichen die Box-Whisker-Diagramme in der Grafik 5.4 eine geringfügig bessere Prädiktionsgüte für längere Prädiktionshorizonte bei dem Modell das mit 12-Schritt-PEM identifiziert wurde. Die Zeitkonstanten der Modelle wurden zu 19.4 und 112.5min (1-Schritt-PEM) bzw. zu 21.4 und 98.9min bei der 12-Schritt-PEM bestimmt.

Bei den in der Grafik rechts dargestellten Autokovarianzfolgen $\hat{r}_\epsilon[\tau]$ kann in beiden Fällen eine Schwingung mit etwa 6 – 7 Zeitschritten festgestellt werden, was auf eine nicht vollständig erfasste Systemdynamik hindeutet. Es sei jedoch angemerkt, dass diese durch

Verwendung hoher Polynomordnungen vermieden werden konnte ($n \geq 8$). Aufgrund der jedoch unwesentlich besseren Prädiktionsgüte wurde zugunsten der geringeren Komplexität das Modell mit den obig genannten Strukturparametern bevorzugt.

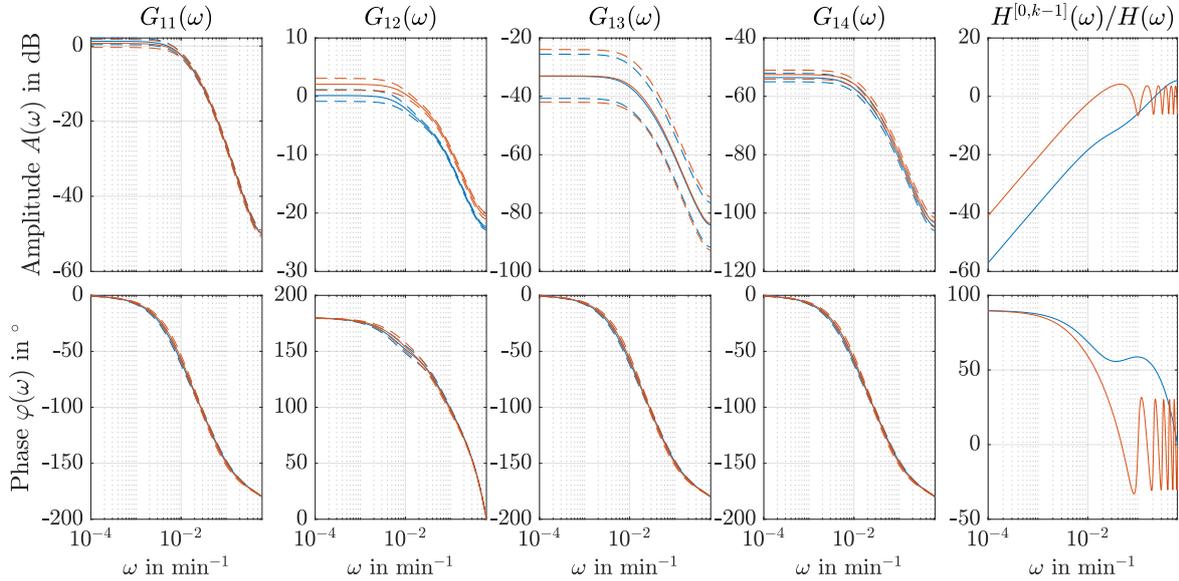


Abbildung 5.3: Bode-Diagramm des mittels 1-Schritt (–) bzw. 12-Schritt (–) PEM identifizierten Raummodells sowie deren 99%-Konfidenzintervalle (– –). Rechts ist zudem die Wichtungsfunktion (3.25) dargestellt.

Bemerkung 5.2 (Kritischer Wert der Autokovarianzfolge \hat{r}_ε): In der Abbildung 5.4 wurde der kritische Wert der Autokovarianzfolge $r_\varepsilon[\tau]$ gekennzeichnet, auf dessen Ursprung folgend kurz eingegangen wird:

Unter der Annahme, dass $\varepsilon[t] := \varepsilon[t|t-1]$ mittelwertfreies weißes Rauschen ist (Nullhypothese), konvergiert

$$\sqrt{N}\hat{r}_\varepsilon[\tau] \xrightarrow{i.V.} \mathcal{N}(0, \lambda_\varepsilon^2) \quad (\forall \tau \neq 0), \quad (5.12)$$

wobei

$$\hat{r}_\varepsilon[\tau] = \frac{1}{N} \sum_{t=1}^N \varepsilon[t]\varepsilon[t-\tau], \quad (5.13)$$

die Schätzung der Autokovarianzfunktion kennzeichnet [SS89, Abschn. 11.1]. Basierend auf (5.12) lässt sich dann eine Teststatistik mit dem kritischen Wert

$$\hat{r}_{\text{krit}} = \frac{1}{\sqrt{N}} \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \quad (5.14)$$

definieren, wobei Φ^{-1} die inverse Verteilungsfunktion der Normalverteilung und α das spezifizierte Signifikanzniveau darstellen. \triangleleft

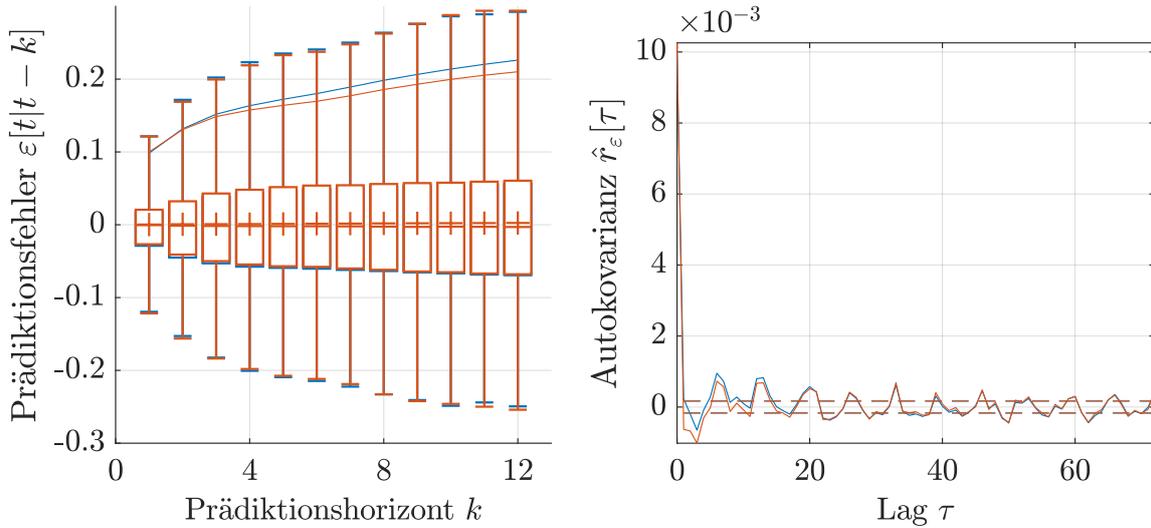


Abbildung 5.4: Box-Whisker-Diagramm der Prädiktionsfehler $\varepsilon[t|t-k]$ (links) und Autokovarianzfolge $\hat{r}_\varepsilon[\tau]$ der Residuen $\varepsilon[t|t-1]$ (rechts) des mittels 1-Schritt (–) bzw. 12-Schritt-PEM (–) identifizierten Raummodells. In der Darstellung links wurden die empirischen 5, 25, 50, 75 und 95% Perzentile sowie der Mittelwert (+) und die Standardabweichung (-) markiert. Die gestrichelte Linie in der Darstellung rechts kennzeichnet den kritischen Bereich zum Signifikanzniveau $\alpha = 0,01$ (siehe Bemerkung 5.2).

b) Nichtlineares Modell mit Ausgangsrückkopplung

Die in diesem Abschnitt diskutierten Prädiktionsmodelle für die Raumtemperatur sollen für den Entwurf von Einzelraum- bzw. Mehrzonen-MPC-Reglern eingesetzt werden. Praktisch können die Wärmeströme \dot{Q}_H und \dot{Q}_K dabei jedoch nicht als Stellgrößen betrachtet werden, sondern werden durch die Vorlauftemperatur des Heiz- bzw. Kühlmediums und durch die Ventilposition bzw. Ansteuerung der Umluftkühlgeräte beeinflusst. Des Weiteren sei nochmals erwähnt, dass in der Praxis aus Kostengründen normalerweise keine Wärmemengenzähler in den Räumen verbaut werden, sodass alternative Modellierungsansätze von Interesse sind.

Modellansatz: Im Rahmen der durchgeführten Identifikationsexperimente wurde daher ein auf der Modellstruktur (2.51) basierender Ansatz untersucht. Analog zum vorherigen Unterabschnitt wurden die Transmissionswärmeverluste \dot{Q}_F und die solaren Gewinne \dot{Q}_S wiederum durch die Eingangsgrößen (5.3) und (5.7) modelliert. Die Heiz- und Kühlleistung \dot{Q}_H bzw. \dot{Q}_K wurden hingegen durch Funktionen des Typs (5.2a) bzw. (5.2b) angenähert¹¹, wobei für $n_H = 1,3$ und $n_K = 1,0$ festgelegt wurde¹² (siehe Anhang

¹¹Im Falle negativer Temperaturdifferenzen $\Delta\vartheta$ in (5.2a) bzw. (5.2b), ergeben sich für nicht ganzzahlige Exponenten komplexe Funktionswerte. Um dies zu vermeiden wurde der Ansatz $\text{sgn}(\Delta\vartheta) |\Delta\vartheta|^n$ implementiert.

¹²Die Betrachtung der Exponenten als freie Optimierungsparameter führte zu keiner signifikanten Verbesserung der Prädiktionsgüte und zu einer relativ hohen Varianz der jeweiligen Parameterschätzwerte (Gl. (3.27)).

C.1). Anstelle der Messwerte der Wärmemengenzähler wurden die Vorlauftemperaturen und Thermostatventilpositionen bzw. die Ansteuerung der Umluftkühlgeräte als Eingangsgrößen des Modells betrachtet. Weiterhin wurde für die Nichtlinearitäten

$$f_H[t] = \bar{H}_H^{\eta_1}[t] \quad \text{bzw.} \quad (5.15a)$$

$$f_K[t] = \bar{H}_K^{\eta_2}[t] \quad (5.15b)$$

angesetzt, wobei \bar{H}_H und \bar{H}_K die gemittelte Ventilposition der Thermostate bzw. die Ansteuerung der Umluftkühlgeräte und η_1 sowie η_2 Optimierungsparameter kennzeichnen. Die Parametrierung des LTI-Blocks wurde entsprechend der im vorherigen Unterabschnitt dargelegten Vorgehensweise durchgeführt, wobei die Verzögerung des offenen Kreises zu $d = 1$ festgelegt wurde. Die Startwerte des Optimierungsparameters $\boldsymbol{\eta}$ wurden zu $\hat{\boldsymbol{\eta}}_{(0)} = (1, 1)^\top$ gewählt. Zudem wurde analog zum vorherigen Unterabschnitt wiederum ein Hochpassvorfilter (5.10) mit der Zeitkonstante $T_L = 2\text{h}$ verwendet.

Identifikation: Die Modelle wurden mittels 1-Schritt-PEM und Backforecasting-Initialisierung des optimalen Prädiktors identifiziert. Das weitere Vorgehen bei der Identifikation und die Einstellungen des Optimierungsalgorithmus entsprechen denen im vorherigen Unterabschnitt.

Diskussion: Die Abb. 5.5 zeigt die Prädiktionsgüte unterschiedlich parametrierter Modelle im Sinne des RMMSE Kriteriums (5.11), wobei die Prädiktionen $\hat{y}[t|t-k]$ für $k > 1$ durch Simulation des deterministischen Teilmodells berechnet wurden (siehe Bem. 2.12). Die Grafik verdeutlicht, dass wiederum die ARMAX- und BJ-Parametrierungen gegenüber ARX- und OE-Strukturen vorzuziehen sind. Vergleicht man weiterhin die Abb. 5.2 und 5.5, dann lässt sich eine $\approx 5 - 10\%$ schlechtere Prädiktionsgüte der Modelle mit Ausgangsrückkopplung feststellen, sofern genügend hohe Polynomordnungen verwendet werden. Dies kann zum einen durch die gewählte Modellstruktur begründet werden, welche die Prozessdynamik möglicherweise schlechter abbildet. Dabei sei insbesondere darauf hingewiesen, dass ein hinreichend guter hydraulischer Abgleich¹³ des Heiz- bzw. Kühlsystems vorausgesetzt wird. Zum anderen sei erwähnt, dass die Ventilpositionen der Funk-Heizkörperthermostate nur im Intervall von 10min an die Datenbank der Raumautomation übertragen werden.

In der Grafik 5.6 und 5.7 lässt sich die Güte des mittels 1-Schritt-PEM identifizierten ARMAX Hammerstein-Modells aus der Abb. 5.4 mit der eines vergleichbaren nichtlinear rückgekoppelten Modells direkt vergleichen. Es wird veranschaulicht, dass sich beide Ansätze qualitativ sehr ähnlich verhalten.

¹³Beim hydraulischen Abgleich werden die verschiedenen hydraulischen Widerstände der Heizstränge, die bspw. aufgrund unterschiedlicher Rohrleitungslängen verursacht werden, durch gezielte Drosselung ausgeglichen. Ist ein Heizkreis nicht ausreichend gut abgeglichen, dann kann dies zur Unterversorgung einzelner Heizkörper führen. Hiermit ist gemeint, dass das Ventil öffnet, aber der entsprechende Heizkörper in bestimmten Situationen nicht durchströmt wird.

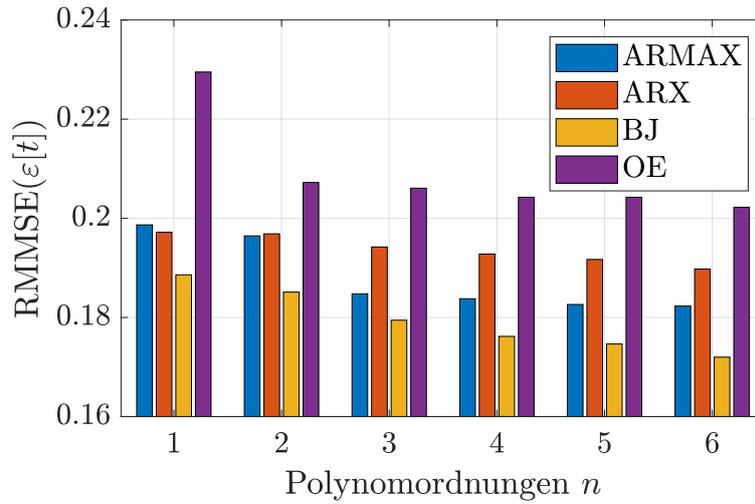


Abbildung 5.5: $RMMSE$ (5.11) unterschiedlich parametrierter nichtlinearer Modelle des Typs (2.51) in Abhängigkeit der Polynomordnungen n .

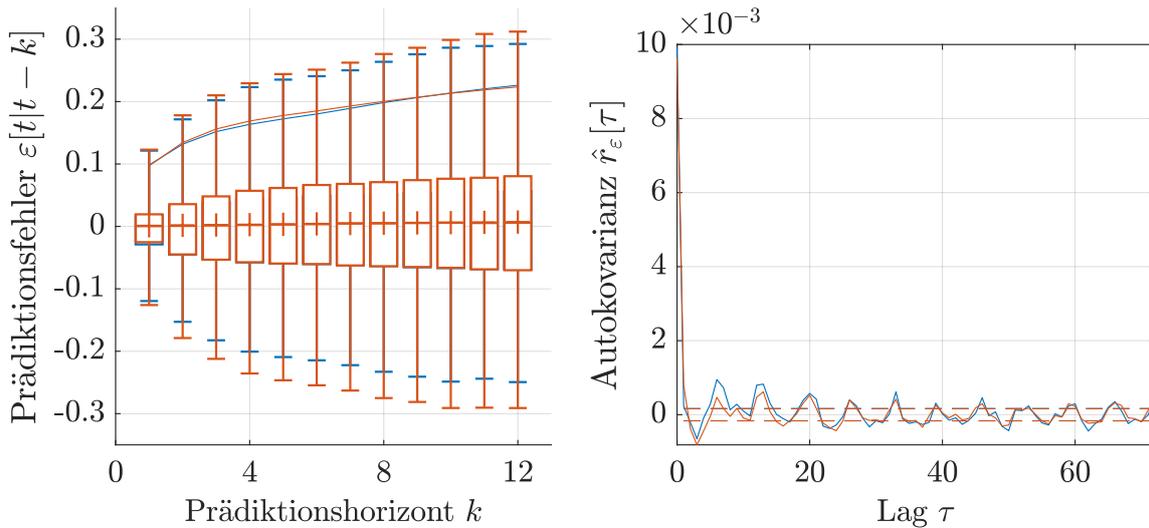


Abbildung 5.6: Box-Whisker-Diagramm der Prädiktionsfehler $\varepsilon[t|t-k]$ (links) und Autokovarianzfolge $\hat{r}_\varepsilon[\tau]$ der Residuen $\varepsilon[t|t-1]$ eines Hammerstein-Modells (—) sowie eines vergleichbaren nichtlinearen Modells mit Ausgangsrückkopplung (---). In der Darstellung links wurden die empirischen 5, 25, 50, 75 und 95% Perzentile sowie der Mittelwert (+) und die Standardabweichung (-) markiert. Die gestrichelte Linie rechts kennzeichnet den kritischen Bereich zum Signifikanzniveau $\alpha = 0.01$ (siehe Bemerkung 5.2).

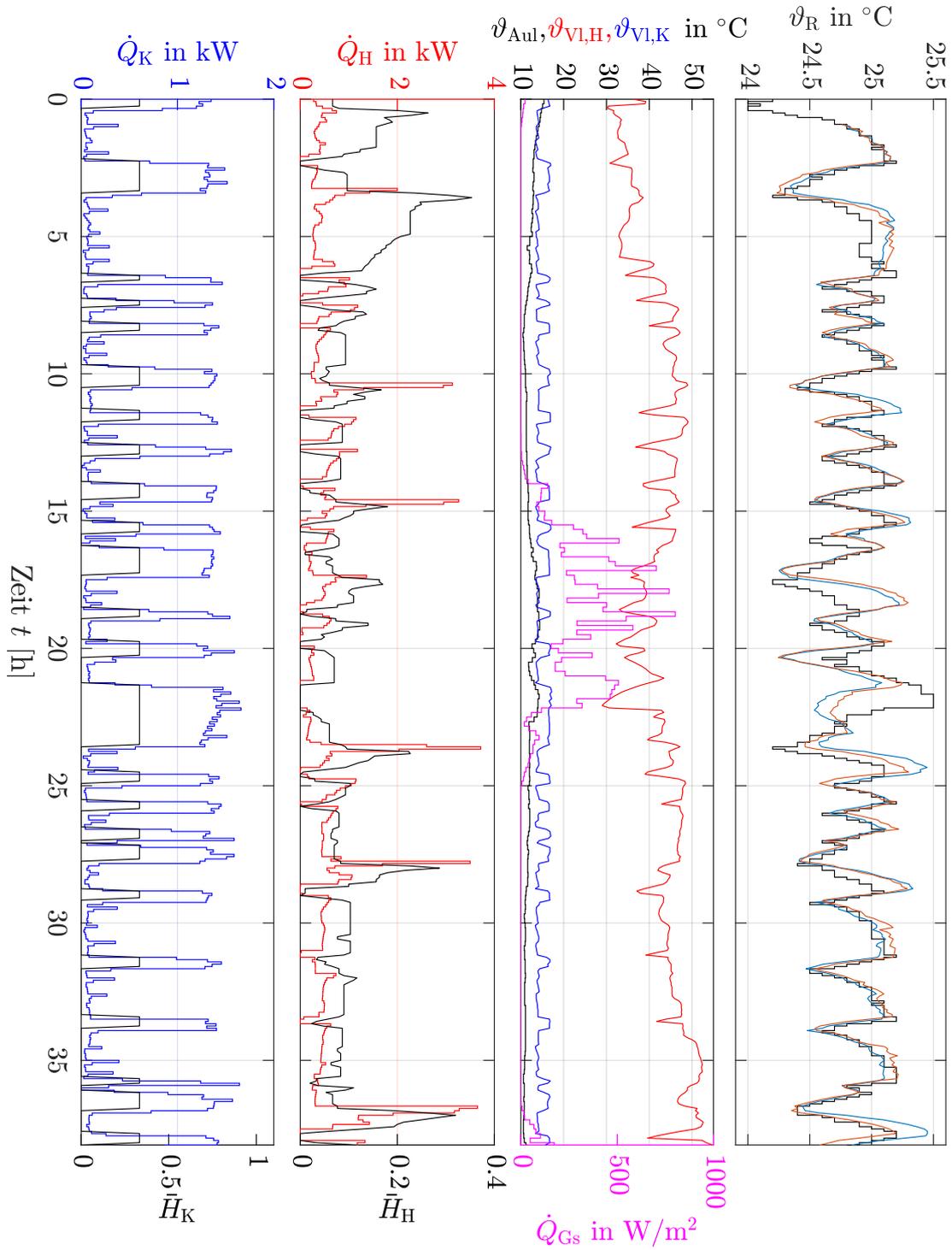


Abbildung 5.7: Vergleich der 12-Schritt Prädiktion (1h) auf einem ausgewählten Testdatensatz eines Hammerstein-Modells (—) sowie eines vergleichbaren nichtlinearen Modells mit Ausgangsrückkopplung (---).

5.1.2 Anwendungsbeispiel 2: Büroraum mit Fußbodenheizung

Der in diesem Anwendungsbeispiel betrachtete 12m² große Büroraum befindet sich im ersten Obergeschoss eines gut gedämmten und in den 2010er Jahren modernisierten Firmengebäudes. Das nach Norden ausgerichtete Büro verfügt lediglich über eine Fußbodenheizung, welche durch einen Zweipunktregler angesteuert wird.

Beschreibung des Datensatzes: Die verwendeten Daten wurden während der Nutzungsphase des Büros vom 1.1.2015 bis zum 30.8.2015 aufgezeichnet. Wegen offensichtlicher Datenfehler musste wiederum eine manuelle Vorauswahl durchgeführt werden, wobei 3 separate Datensätze mit der Länge von etwa 66, 48 und 79 Tagen selektiert wurden. Weiterhin ist wegen der guten Wärmedämmung und der Trägheit der Fußbodenheizung ein relativ langsames dynamisches Verhalten des Systems zu erwarten, weswegen die Abtastzeit zu $T_A = 15\text{min}$ festgelegt wurde. Für die Messungen der Außentemperatur und der Globalstrahlung wurden die im Intervall von 10min aufgezeichneten Daten der vom DWD betriebenen Wetterstation in Chemnitz verwendet (Stationsnummer 0853).

Modellansatz: Im Folgenden wird wiederum ein auf der Modellstruktur (2.51) basierender Identifikationsansatz für die Raumtemperatur $y[t] := \vartheta_R[t]$ vorgestellt. Da der Raum ausschließlich über eine Fußbodenheizung temperiert wird, wurde in diesem Anwendungsfall

$$\mathbf{f}(\mathbf{u}, y) = \begin{pmatrix} H_H[t] \operatorname{sgn}(\vartheta_{V1}[t] - \vartheta_R[t]) |\vartheta_{V1}[t] - \vartheta_R[t]|^{n_H} \\ \vartheta_{Aul}[t] \\ \dot{Q}_{\text{gen}}[t] \end{pmatrix} \quad (5.16)$$

angesetzt ($n_H = 1.1$), wobei $\dot{Q}_{\text{gen}}[t]$ wiederum die solaren Gewinne (Gl. (5.7)), $\vartheta_{Aul}[t]$ und $\vartheta_{V1}[t]$ die Außen- bzw. Vorlauftemperatur sowie $H_H[t] \in \{0, 1\}$ die Heizventilposition beschreiben.

In diesem Beispiel soll der in Abschnitt 5.1 erwähnte Störgrößenansatz zur Modellierung der Wärmeabgabe durch die Belegung des Raumes \dot{Q}_B veranschaulicht werden. Dabei besteht die wesentliche Annahme darin, dass dem Raum eine periodische Nutzungscharakteristik zugrunde liegt, bei welcher $\dot{Q}_B[t] \approx \dot{Q}_B[t - p]$ angenommen werden kann, was insbesondere bei Büros, Klassenräumen oder Hörsälen eine sinnvolle Annahme darstellen kann. Wie bereits in Beispiel 3.3 demonstriert wurde, kann dann ein saisonaler Ansatz zur Modellierung der Wärmeabgabe der Belegung eingesetzt werden.

Für das im Folgenden identifizierte Büromodell, wurde entsprechend der Bemerkung 2.11 daher eine ARARMAX Parametrierung mit dem Rauschmodell

$$H(q^{-1}) = \frac{C(q^{-1})}{A(q^{-1})D(q^{-1})} = \frac{\tilde{C}(q^{-1}) + \bar{C}(q^{-1})}{A(q^{-1})(1 + \bar{D}(q^{-1}))} \quad (5.17)$$

angesetzt (siehe Abschn. 2.4.1), wobei $C(q^{-1})$ und $D(q^{-1})$ durch Polynome der Form

$$\tilde{C}(q^{-1}) = 1 + c^{[1]} q^{-1} + \dots + c^{[n]} q^{-n} \quad (5.18a)$$

$$\bar{C}(q^{-1}) = c^{[p_{\min}]} q^{-p_{\min}} + \dots + c^{[p_{\max}]} q^{-p_{\max}} \quad (5.18b)$$

$$\bar{D}(q^{-1}) = d^{[p_{\min}]} q^{-p_{\min}} + \dots + d^{[p_{\max}]} q^{-p_{\max}} \quad (5.18c)$$

parametriert wurden. Die minimale und maximale Periodendauer wurde aufgrund der sich wöchentlich wiederholenden Nutzungscharakteristik des Büros zu $p_{\min} = 671$ bzw. $p_{\max} = 673$ gewählt (Abtastzeit $T_A = 15\text{min}$). Hierbei kennzeichnet n wiederum den Strukturparameter des Modells, welcher bei der Identifikation analog zum vorherigen Anwendungsbeispiel schrittweise erhöht wurde ($n = n_A = n_{B_i}$, mit $i = 1, \dots, 3$). Für die Parametrierung des deterministischen Teilsystems $\mathbf{G}(q^{-1})$ wurde wiederum ein faktorisiertes Ansatz der Polynome $B_i(q^{-1})$ und $A(q^{-1})$ gewählt (vgl. Abschn. 5.1.1). Aufgrund des bei Fußbodenheizungen deutlich trägeren dynamischen Verhaltens, wurde die Zeitkonstante des Vorfilters (5.10) zu $T_L = 6\text{h}$ festgelegt.

Identifikation: Die Modelle wurden wiederum mittels 1–Schritt-PEM und Backforecasting-Initialisierung des Residuenfilters identifiziert. Für die Lösung des Optimierungsproblems wurde die MATLAB Implementierung des Trust-Region-Verfahrens eingesetzt, wobei für die initiale Parameterschätzung $\hat{\boldsymbol{\theta}}_{(0)}$ der MkQ-Schätzer (3.37) genutzt wurde. Alle weiteren Einstellungen wurden wie im Abschn. 5.1.1 beschrieben gewählt.

Diskussion: Die Abbildung 5.8 veranschaulicht den Vergleich ausgangsrückgekoppelter Modelle des Typs (2.51) ohne ($\bar{C}(q^{-1}) = \bar{D}(q^{-1}) = 0$) sowie mit saisonalem Rauschmodell in Abhängigkeit des Strukturparameters n . Dabei wurden die Prädiktionen $\hat{y}[t|t-k]$ für $k > 1$ wiederum durch Simulation des deterministischen Teilmodells berechnet (siehe Bemerkung 2.12). Die Grafik verdeutlicht, dass durch das saisonale Störgrößenmodell eine $\approx 15 - 20\%$ bessere Genauigkeit im Sinne des RMMSE (5.11) erzielt werden kann.

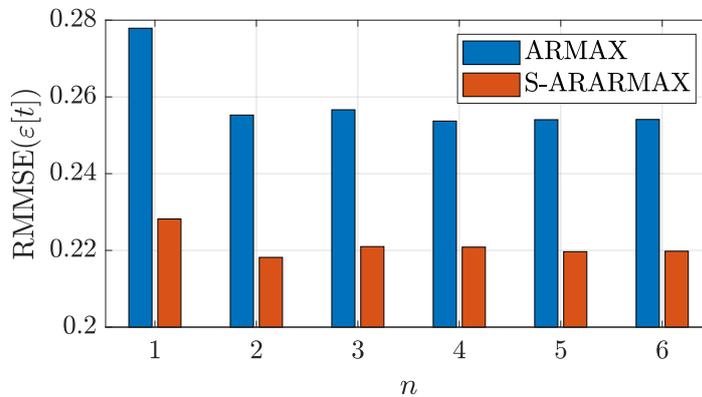


Abbildung 5.8: RMMSE (5.11) der Modelle mit (S-ARARMAX) und ohne saisonalem Rauschmodell (ARMAX) in Abhängigkeit des Strukturparameters n ($k_{\max} = 12$).

Die Parameter eines Modells, welches entsprechend des in Bem. 5.1 geschilderten Vorgehens ausgewählt wurde, sind in Tab. 7 angegeben. In der Abb. 5.9 sind zudem ein Box-Whisker-Diagramm sowie die Autokovarianzfolge der Prädiktionsfehler für das Modell mit und ohne ($\bar{C}(q^{-1}) = \bar{D}(q^{-1}) = 0$) saisonalem Störgrößenansatz dargestellt. Dabei kann für das Modell ohne saisonalen Ansatz eine Asymmetrie der Prädiktionsfehlerverteilung für große k erkannt werden, was auf die nicht modellierte Störung hindeutet. Weiterhin kann in der Darstellung der Autokovarianzfolge, insbesondere für das Modell ohne saisonalen Anteil, die Periodizität der Störung deutlich erkannt werden.

Polynom bzw. Parameter	Rel. Standardabw. [%]
$A(q^{-1}) = (1 - 0.888 q^{-1})(1 - 0.9552 q^{-1})$	1.08; 0.52
$B_1(q^{-1}) = 1.177 \cdot 10^{-3} q^{-1}$	6.17
$B_2(q^{-1}) = 7.768 \cdot 10^{-4} q^{-1}$	9.35
$B_3(q^{-1}) = 1.764 \cdot 10^{-5} q^{-1}$	14.4
$C(q^{-1}) = 1 - 0.8777 q^{-1} - 0.8638 q^{-671} + 0.8089 q^{-672}$	0.69; 0.37; 0.69
$D(q^{-1}) = 1 - 0.7284 q^{-672}$	0.85

Tabelle 7: Parameterschätzwerte und rel. Abweichungen des identifizierten Modells.

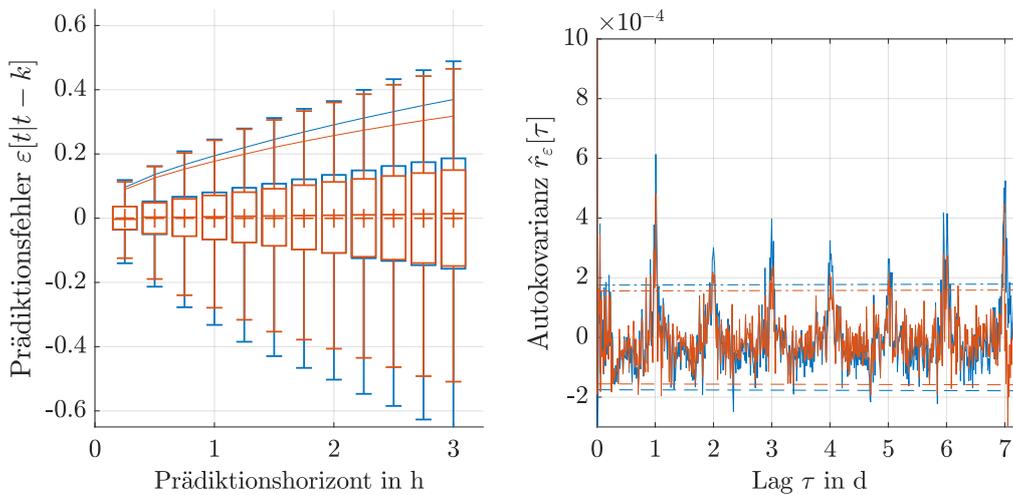


Abbildung 5.9: Box-Whisker-Diagramm der Prädiktionsfehler $\varepsilon[t|t-k]$ (links) und Autokovarianzfolge $\hat{r}_\varepsilon[\tau]$ der Residuen $\varepsilon[t|t-1]$ (rechts) des Modells mit (–) und ohne saisonalem Rauschmodell (–). In der Darstellung links wurden die empirischen 5, 25, 50, 75 und 95% Perzentile sowie der Mittelwert (+) und die Standardabweichung (–) markiert. Die gestrichelte Linie rechts kennzeichnet den kritischen Bereich zum Signifikanzniveau $\alpha = 0.01$ (siehe Bemerkung 5.2).

5.1.3 Anwendungsbeispiel 3: Turnhalle mit Deckenheizsystem

Die in diesem Abschnitt betrachtete Einfeldsporthalle ist Teil eines Schulgebäudes und hat eine Grundfläche von etwa 400m^2 sowie eine Raumhöhe von 6.5m. Die Turnhalle wird über ein Deckenheizsystem beheizt, deren Wärmeabgabe über ein zentrales Mischventil gesteuert und mittels eines Wärmemengenzählers gemessen wird. Eine schematische Darstellung des Aufbaus lässt sich der Abbildung 5.10 entnehmen.

Beschreibung des Datensatzes: Die zur Systemidentifikation verwendeten Daten wurden vom 1.11.2012 bis zum 1.11.2013 aufgezeichnet. Aus dem Rohdatensatz wurden 3 separate Datensätze mit der Länge von etwa 53, 27 und 253 Tagen extrahiert, wobei eine Abtastzeit von 15min verwendet wurde. Für die Messwerte der Außentemperatur wurden die aufgezeichneten Daten der DWD Wetterstation 01048 in Dresden verwendet.

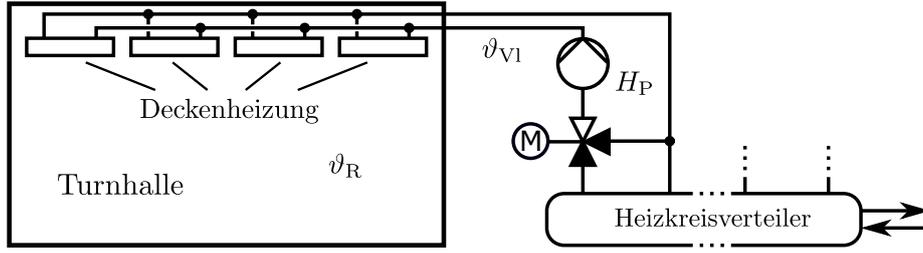


Abbildung 5.10: Schematische Darstellung der Turnhalle mit Deckenheizsystem

Modellansatz: Aufgrund des installierten Wärmemengenzählers kann im Folgenden wiederum der Vergleich einer linearen Modellstruktur mit einem nichtlinear rückgekoppelten Modellansatz der Form (2.51) dargestellt werden. Dabei folgt die Definition der Modellansätze im Wesentlichen der bereits in Abschnitt 5.1.1 geschilderten Methodik:

Für das LTI-Modell wurde der Eingangsvektor zu $\mathbf{u}[t] = (\dot{Q}_H[t], \vartheta_{\text{Aul}}[t])^\top$ definiert, wobei die Heizleistung \dot{Q}_H durch numerische Differenziation der Messwerte des Wärmemengenzählers ermittelt wurde. Für das nichtlineare Modell wurde die statische Nichtlinearität

$$\mathbf{f}(\mathbf{u}, y) = \begin{pmatrix} H_P[t] \operatorname{sgn}(\vartheta_{v1}[t] - \vartheta_R[t]) |\vartheta_{v1}[t] - \vartheta_R[t]|^{n_H} \\ \vartheta_{\text{Aul}}[t] \end{pmatrix}, \quad (5.19)$$

mit $n_H = 1.1$ festgelegt, wobei $H_P[t]$ das binäre Ansteuersignal der Heizkreispumpe kennzeichnet. Die solaren Gewinne wurden aufgrund des großen Volumens und der verhältnismäßig geringen Fensterfläche der Turnhalle vernachlässigt.

Für das LTI- sowie das nichtlineare Modell mit (5.19) wurde eine ARARMAX-Struktur mit einem saisonalen Rauschmodell verwendet. Dabei hat sich herausgestellt, dass in beiden Fällen eine Parametrierung mit $n_A = 1$ und $\mathbf{n}_B = [1, 1]$ ausreichend ist um die wesentliche Systemdynamik abzubilden. Die zu identifizierenden Polynome des Systemmodells wurden aufgrund der wöchentlichen Nutzungscharakteristik zu

$$\begin{aligned} A(q^{-1}) &= 1 + a^{[1]} q^{-1} & B_1(q^{-1}) &= 1 + b_1^{[1]} q^{-1} & B_2(q^{-1}) &= 1 + b_2^{[1]} q^{-1} \\ C(q^{-1}) &= 1 + c^{[672]} q^{-672} & D(q^{-1}) &= 1 + d^{[672]} q^{-672} \end{aligned} \quad (5.20)$$

angesetzt. Die Zeitkonstante des Vorfilters (5.10) wurde zu $T_L = 6\text{h}$ festgelegt.

Identifikation: Beide Modelle wurden wieder mittels 1-Schritt-PEM und Backforecasting Initialisierung des optimalen Prädiktors identifiziert, wobei die gleichen Einstellungen wie im Abschnitt 5.1.1 verwendet wurden. Die identifizierten Polynomparameter sind in der Tabelle 8 angegeben, wobei die Polstelle von $A(q^{-1})$ im kontinuierlichen Zeitbereich einer Zeitkonstante von $\approx 7.4\text{h}$ entspricht. Es sei zudem erwähnt, dass durch die Erhöhung der Polynomordnungen oder durch die Verwendung der Multischritt-PEM keine signifikanten Verbesserungen erzielt werden konnten ($< 3\%$ im Sinne des RMMSE).

Parameter	LTI-Modell	Nichtlin. Modell
$a^{[1]}$	$-0.9662 \pm 0.16\%$	$-0.9680 \pm 0.16\%$
$b_1^{[1]}$	$1.576 \cdot 10^{-3} \pm 32.16\%$	$1.826 \cdot 10^{-3} \pm 15.54\%$
$b_2^{[1]}$	$2.407 \cdot 10^{-3} \pm 9.50\%$	$2.755 \cdot 10^{-3} \pm 8.60\%$
$c^{[672]}$	$-0.8949 \pm 0.30\%$	$-0.8953 \pm 0.30\%$
$d^{[672]}$	$-0.9691 \pm 0.14\%$	$-0.9692 \pm 0.14\%$

Tabelle 8: Parameterschätzwerte und relative Abweichungen des linearen Modells mit $\mathbf{u} = (\dot{Q}_H, \vartheta_{\text{Aul}})^\top$ bzw. des nichtlinearen Modells (2.51) mit (5.19).

Diskussion: Die Abbildungen 5.12 und 5.11 veranschaulichen den direkten Vergleich beider Modelle, wobei erkannt werden kann, dass sich beide Modelle ähnlich verhalten. Des Weiteren sieht man in der Darstellung der Autokorrelationsfolge eine Schwingung mit der Periodendauer von einem Tag, welche zwar durch ein komplexeres Rauschmodell vermieden werden konnte (mehr freie Parameter), jedoch zu keiner signifikanten Verbesserung der Prädiktionsgüte geführt hat, sodass die in Tab. 8 angegebenen Modelle bevorzugt wurden. Die Abbildung 5.12 veranschaulicht den direkten Vergleich der 6h-Prädiktion beider Modelle, wobei sich wiederum feststellen lässt, dass sich beide Modelle nahezu identisch verhalten. Des Weiteren weisen die fehlerhaften Prädiktionen an den Tagen 11 und 28 auf Abweichungen vom wöchentlichen Nutzungsrhythmus der Turnhalle hin.

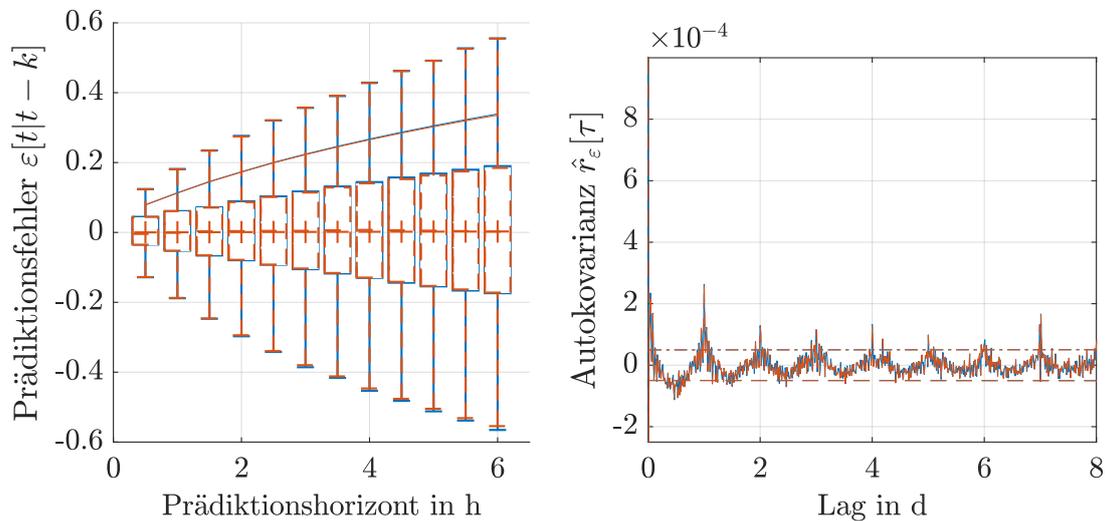


Abbildung 5.11: Box-Whisker-Diagramm der Prädiktionsfehler $\varepsilon[t|t-k]$ (links) und Autokovarianzfolge $\hat{r}_\varepsilon[\tau]$ der Residuen $\varepsilon[t|t-1]$ (rechts) des LTI (—) und des entsprechenden ausgangsrückgekoppelten Modells (---). In der Darstellung links wurden die empirischen 5, 25, 50, 75 und 95% Perzentile sowie der Mittelwert (+) und die Standardabweichung (-) markiert. Die gestrichelte Linie rechts kennzeichnet den kritischen Bereich zum Signifikanzniveau $\alpha = 0.01$ (siehe Bemerkung 5.2).

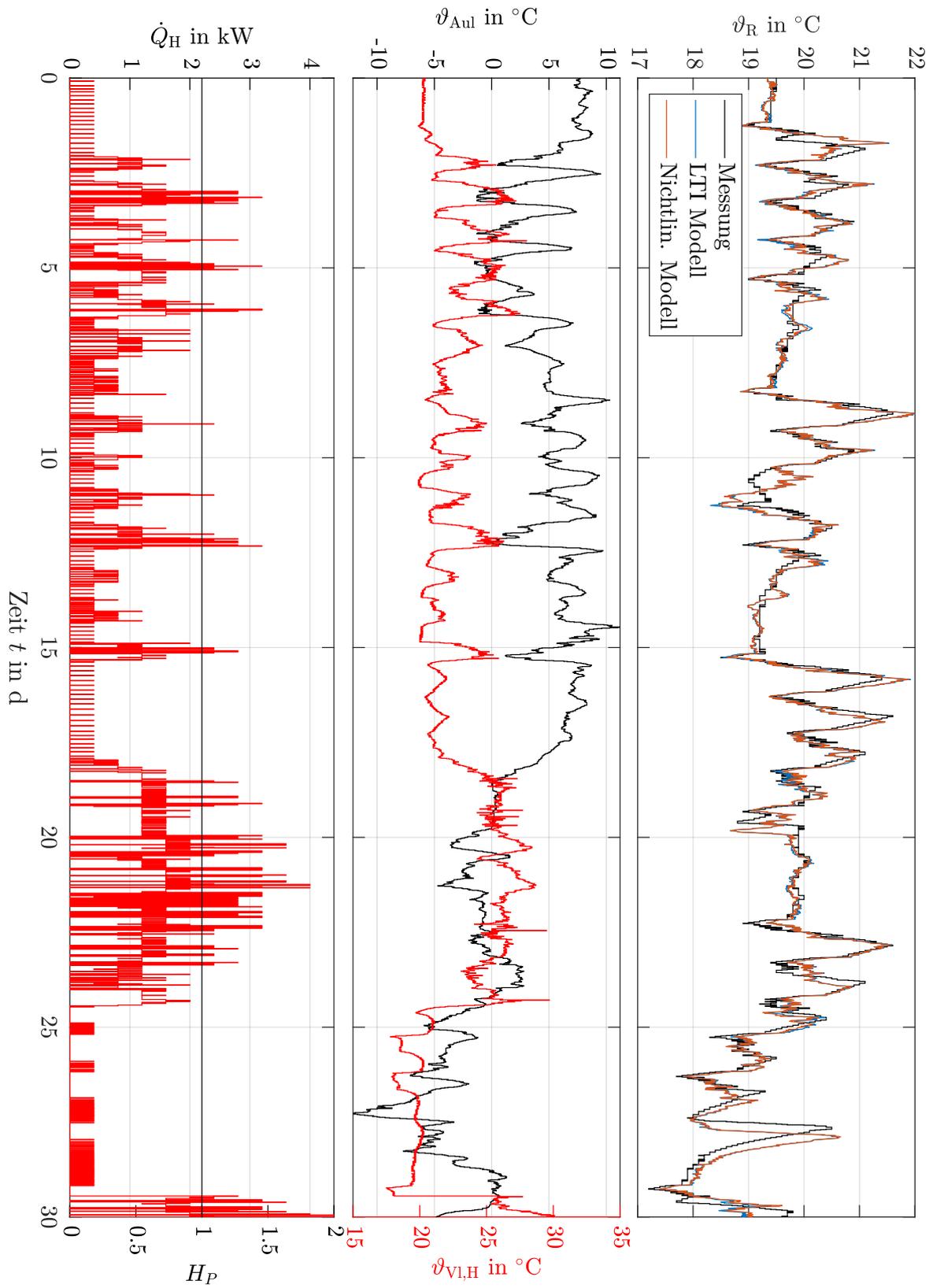


Abbildung 5.12: Vergleich der 24-Schritt Prädiktion (6h) des linearen Modells (—) sowie des entsprechenden nichtlinearen Modells mit Ausgangsrückkopplung (—).

5.1.4 Anwendungsbeispiel 4: Hörsaalgebäude

Die in diesem Abschnitt dokumentierten Erfahrungen bei der Modellidentifikation anhand der Betriebsdaten eines Hörsaalgebäudes sollen typische Probleme bei der Identifikation von Raumtemperaturmodellen verdeutlichen. Die Ausführungen sollen daher als Hinweise bzw. Empfehlungen für zukünftige Versuchsaufbauten und Experimente dienen.

Die im Rahmen der Arbeit untersuchten Hörsäle mit einer Kapazität von etwa 450 bis 1000 Plätzen werden über eine zentrale Lüftungsanlage und lokal installierte variable Volumenstromregler bedarfsgerecht mit Wärme bzw. Kälte und mit Frischluft versorgt. Eine schematische Darstellung des Aufbaus und die ungefähre Sensorplatzierung ist in der Abbildung 5.13 dargestellt.

Basierend auf den Betriebsdaten der Hörsäle konnten mit den in Abschnitt 5.1.1 bis 5.1.3 genutzten Modellansätzen keine zufriedenstellenden Identifikationsergebnisse erzielt werden. Die dabei wesentliche Problematik bestand zum einen in der sehr stark von der unbekanntem Raumauslastung abhängigen Temperaturentwicklung in den Hörsälen. Zum anderen wurde durch qualitativen Vergleich mehrerer Messungen an unterschiedlichen Orten eine stark von der Sensorplatzierung und den lokalen Strömungsverhältnissen abhängige Temperatur sowie ein unterschiedlicher CO_2 -Gehalt der Raumluft festgestellt. Die Annahme einer gleichmäßig vermischten Luft in den Hörsälen musste daher in Frage gestellt werden, sodass ein konzentriert parametrischer Modellansatz aus praktischer Sicht nicht gerechtfertigt war.

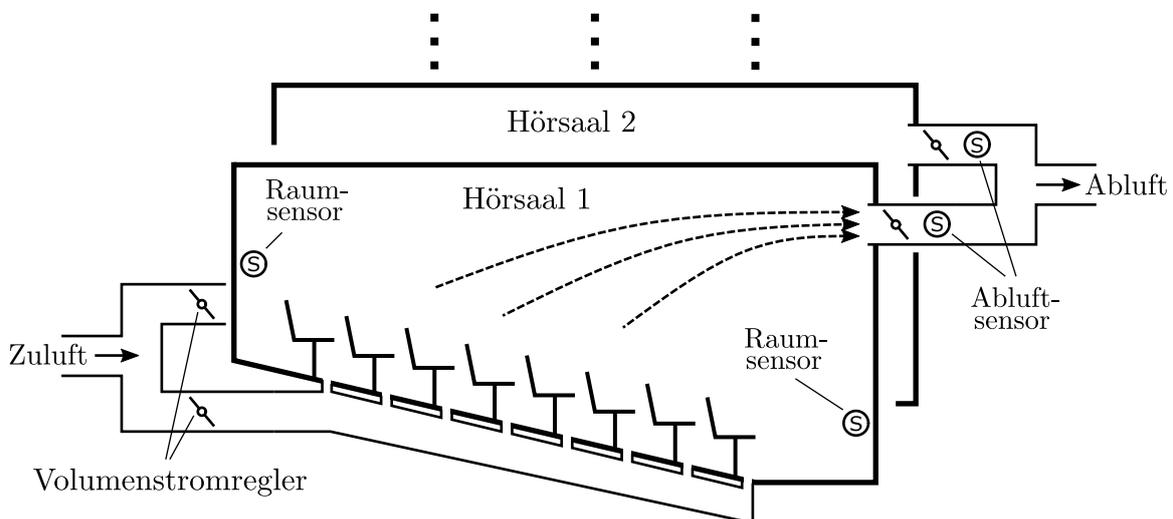


Abbildung 5.13: Schematische Darstellung der Hörsäle und der Lüftung.

Für die *Ablufttemperatur* hingegen erschien dieser aufgrund der anzunehmenden gleichmäßigen Luftvermischung als wesentlich gerechtfertigter. Diese Vermutung konnte durch Vergleich der Ablufttemperatur mit dem zu erwartenden typischen Nutzungsmuster der Hörsäle bekräftigt werden (siehe Tag 2 in Abb. 5.14). Des Weiteren wurden

mitunter vereinzelte manuelle Belegun­gs­zählungen in den Hörsälen durch­ge­führt, wobei ein qualitativ plausibler Zusammen­hang der Belegung mit den Messwerten der Abluft festgestellt wurde. Im Hinblick auf eine modellbasierte Hörsaalregelung erschien die Identifikation eines Modells für die Ablufttemperatur daher als zielführender Ansatz.

Im Rahmen der dabei durch­ge­führten Identifikationsexperimente wurden allerdings weitere Schwierigkeiten festgestellt. Hierbei ist insbesondere die von den Druckverhältnissen im Abluftkanal abhängige Temperaturmessung bei ausgeschalteter Lüftung und zu nennen ($\dot{V}_{\text{Zul}} \approx \dot{V}_{\text{Abl}} \approx 0$). In der Abbildung 5.14 lässt sich dieses Verhalten beispielhaft durch den Anstieg der Ablufttemperatur vor dem Einschaltzeitpunkt der Lüftung erkennen, welcher durch die Analyse der Messdaten auf einen „Rückstau“ der Abluft anderer Hörsäle zurück­ge­führt wurde. Die Messwerte der Ablufttemperatur sind bei ausgeschalteter Hörsaallüftung für die Identifikation daher teilweise unbrauchbar. Da allerdings für saisonale Modellansätze lange zusammenhängende Datensätze benötigt werden (siehe Gl. (3.9)), war eine Modellidentifikation für die Ablufttemperatur nicht möglich.

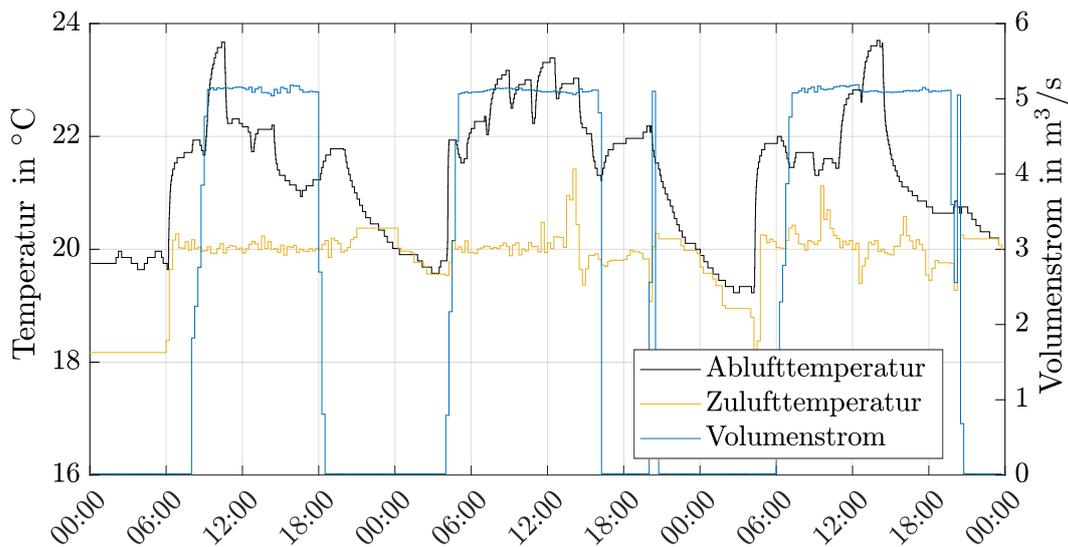


Abbildung 5.14: *Beispielhaft aufgezeichnete Abluft- und Zulufttemperatur sowie der Zuluftvolumenstrom eines Hörsaals.*

Die Ausführungen dieses Abschnittes verdeutlichen, dass die Wahl des Messortes der Raumluftparameter genau bedacht werden sollte. Dabei wird eine Platzierung der Sensorik in unmittelbarer Nähe der Abluftkanäle, jedoch vor den Abluftklappen empfohlen.

5.1.5 Zusammenfassung

Im Rahmen dieser Arbeit wurden eine Vielzahl von praktischen Identifikationsexperimenten durch­ge­führt, deren wesentliche Ergebnisse und Erfahrungen in diesem Abschnitt nochmals zusammengefasst werden sollen.

Generell sei an dieser Stelle zunächst erwähnt, dass die PEM-Identifikation plausibler Modelle mittels während des Gebäudebetriebs aufgezeichneter Messdaten möglich war. Dabei konnten mit ARMAX und BJ-Strukturen vergleichbar gute Ergebnisse erzielt werden, wobei im letzteren Fall eine tendenziell wesentlich langsamere Konvergenz bei der Lösung des Optimierungsproblems beobachtet wurde. Weiterhin sei erwähnt, dass sogar mit ARX sowie mit OE-Modellen eine gute Prädiktionsgüte erreicht werden konnte, sofern das Vorfilter (5.10) sinnvoll angesetzt wurde. Hingegen ist die Wahl des Vorfilters für ARMAX- und für BJ-Modelle oftmals weniger bedeutsam, da diese ohnehin ein ausreichend flexibles Rauschmodell besitzen¹⁴. Weiterhin hat sich während der Arbeit herausgestellt, dass durch die Verwendung eines PEM-Multischrittkriteriums meist nur relativ geringe Gewinne im Sinne der Prädiktionsgenauigkeit zu erwarten sind. Die Wahl der Initialisierungsmethode des optimalen Prädiktors hatte hingegen, insbesondere bei der Verwendung kurzer Datensätze (Abschn. 5.1.2), einen größeren Einfluss auf die Prädiktionsgüte des identifizierten Modells. Des Weiteren sei an dieser Stelle nochmals vermerkt, dass mittels eines nichtlinear rückgekoppelten Modellansatzes (2.51) meist vergleichbar gute Ergebnisse wie mit einem entsprechendem LTI- bzw. Hammerstein-Modell erzielt werden konnten.

Obwohl oftmals eine gute Prädiktionsgüte der Modelle erreicht werden konnte, hat sich während der Arbeit jedoch auch herausgestellt, dass die Identifikation eine praktisch aufwendige, fehleranfällige und herausfordernde Aufgabe darstellt. Daher ist im Hinblick auf eine (teil)automatisierte Generierung der Modelle aus den Gebäudebetriebsdaten noch Forschungs- und Entwicklungsarbeit zu leisten.

Eines der festgestellten Probleme war zum einen die häufig mangelhafte Qualität der Rohdaten, welche oftmals viele Datenlücken enthielten. Durch entsprechende regelbasierte Abfragen, konnten diese Lücken jedoch meist mit guter Genauigkeit gefunden werden, sodass eine automatisierte Beseitigung offensichtlich fehlerhafter Messdaten praktisch machbar erscheint. Ein weiteres Problem der Betriebsdaten ist deren oftmals relativ geringer Informationsgehalt. Damit ist insbesondere gemeint, dass durch den charakteristischen Tagesgang des Wetters, die tägliche Nutzungscharakteristik der Räume und die oftmals tageszeitabhängig fest programmierte Tag/Nacht-Umschaltung zeitliche Korrelation zwischen den Störungen und den Stellsignalen des Prozesses bestehen. Bei den praktisch durchgeführten Identifikationsexperimenten hat dies häufig zu physikalisch unplausiblen Modellen mit Resonanzfrequenzen bei einem Tag^{-1} geführt, weshalb eine Einschränkung der Modellstruktur des deterministischen Teilsystems $\mathbf{G}(q^{-1})$ mittels eines (linear) faktorisierten Ansatzes sinnvoll ist. Es sei deshalb angemerkt, dass eine Veränderung der Tag/Nacht-Umschaltzeitpunkte zu informativeren Datensätzen und somit zu vermutlich genaueren Identifikationsergebnissen führen würde.

Die jedoch größte Herausforderung bei der Identifikation stellen praktisch nicht messbare Störeinflüsse dar. Insbesondere die Wärmeabgabe bzw. sonstige Einflüsse der Raumnutzer sind in vielen Fällen von herausragender Bedeutung. Wie sich während

¹⁴Wurde bei der Identifikation beispielsweise auf die Vorfilterung der Ein- und Ausgangssignale verzichtet ($L(q^{-1}) = 1$), dann wurde der Integrator im stochastischen Teilsystem i. d. R. auch während der Optimierung „gefunden“.

der Arbeit gezeigt hat, kann die Verwendung saisonal parametrierter Modellansätze einen möglichen Lösungsansatz für diese Problematik darstellen. Deren praktische Anwendungsfähigkeit ist jedoch zum einen aufgrund der Annahme einer periodischen Raumnutzung erheblich eingeschränkt. Zum anderen werden bei der Identifikation lange zusammenhängende Trainingsdatensätze benötigt (siehe Gl. 3.9), welche aufgrund von Sensorausfällen häufig nicht verfügbar sind. Für zukünftige Arbeiten ist daher die Untersuchung alternativer Identifikationsansätze sinnvoll. Ein praktisch vielversprechender Ansatz könnte hierbei die Nutzung zusätzlicher CO₂ und Feuchte-Sensorik sein, welche eine Modellidentifikation ohne explizite Kenntnis der Raumbelastung ermöglichen könnte. Hierbei sei insbesondere auf die Methoden der *Blind System Identification* bzw. *Unknown Input Identification* verwiesen [AQH97; ORD14; Lin17; LE17].

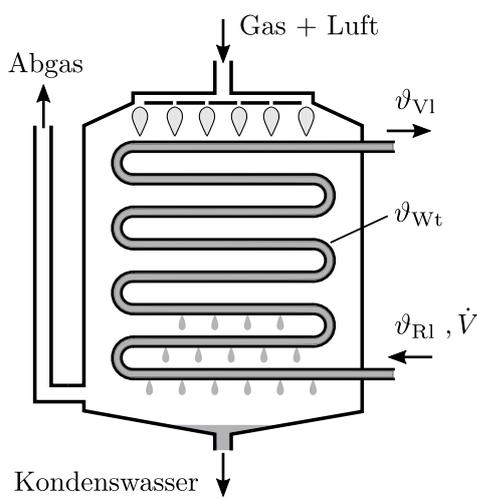
5.2 Wärmeerzeuger

Ein wesentlicher Bestandteil moderner Gebäudeenergiesysteme sind Wärme- bzw. Kälteerzeuger. Hierzu gehören u. a. Gas-, Öl- oder Holzheizkessel, Wärmepumpen, Kältemaschinen sowie Blockheizkraftwerke (BHKWs). Da deren dynamisches Verhalten in der Regel deutlich schneller als das der Pufferspeicher und der Verbraucher ist, werden diese beim Entwurf modellprädiktiver Regler oftmals als statische Systeme modelliert (siehe bspw. [MBH09; MBH12]). Die Identifikation entsprechender Modelle kann daher mit Methoden der Regressionsanalyse durchgeführt werden, wobei jedoch bei der Auswahl der Identifikationsdaten beachtet werden muss, dass sich das System näherungsweise im stationären Zustand befindet. Aufgrund der oftmals veränderlichen Rücklauf- und/oder Quelltemperaturen kann dies während des Anlagenbetriebs möglicherweise nur schwer sichergestellt werden.

In der Praxis ist daher durchaus ein dynamischer Identifikationsansatz wünschenswert, aus welchen ggf. die statischen Modelle extrahiert werden können. Im Folgenden wird daher das entsprechende Vorgehen am Beispiel eines Gas-Brennwertkessels und anhand einer Wärmepumpe veranschaulicht.

5.2.1 Heizkessel

Heizkessel dienen der Bereitstellung von Wärme durch die Verbrennung fester, flüssiger oder gasförmiger Brennstoffe. Dabei wird über einen Wärmetauscher in der Regel Wasser erhitzt, welches mit einer Pumpe dem Heizsystem des Gebäudes zugeführt wird. Im Unterschied zu veralteten Heizkesseln können Brennwert-Heizkessel, durch die Kondensation des im Abgas enthaltenen Wasserdampfes höhere Wirkungsgrade erreichen. Dabei ist der Anteil der nutzbaren Kondensationswärme dann am größten, wenn durch niedrige Rücklauftemperaturen möglichst viel Wasserdampf im Heizkessel kondensiert (siehe Abb. 5.15). Bei der Verbrennung von Erdgas entstehen bezogen auf den Heizwert beispielsweise maximal etwa 11% Kondensationswärme (siehe z. B. [RSS11, Tab. 2.3.1-1]). Eine schematische Darstellung eines Gas-Brennwertkessels ist in der Abbildung 5.15 gezeigt.



Symbol	Bedeutung
$\vartheta_{Vl}, \vartheta_{Rl}, \dot{V}$	Vor-, Rücklauftemp. bzw. Volumenstrom des Wassers
ϑ_{Wt}	Temp. des Wärmetauschers
C_W, C_{Wt}	Wärmekapazität des Wassers bzw. Wärmetausches
c_W, ρ_W	spezifische Wärmekapazität und Dichte des Wassers
\dot{Q}_K, H_{Br}	Heizleistung des Kessels und Brennermodulation
\dot{Q}_{Verl}	Abgas- und Umgebungsverluste
U	Wärmedurchgangskoeffizient

Abbildung 5.15: Schematische Darstellung eines Gas-Brennwertkessels und Bedeutung der verwendeten Symbolik.

Zunächst wird zur Darlegung des Identifikationsansatzes wiederum ein kontinuierliches konzentriert parametrisches Modell herangezogen. Dabei werden das Wasser mit der Temperatur $\vartheta_W =: \vartheta_{Vl}$ und das Metall des Wärmetauschers als konzentrierte Wärmespeicher betrachtet:

$$C_W \frac{d\vartheta_{Vl}(t)}{dt} = c_W \rho_W \dot{V}(t) (\vartheta_{Rl}(t) - \vartheta_{Vl}(t)) + U (\vartheta_{Wt}(t) - \vartheta_{Vl}(t)) \quad (5.21a)$$

$$C_{Wt} \frac{d\vartheta_{Wt}(t)}{dt} = U (\vartheta_{Vl}(t) - \vartheta_{Wt}(t)) + H_{Br}(t) \dot{Q}_K (\vartheta_{Rl}(t)) + \dot{Q}_{Verl}(t) \quad (5.21b)$$

Weiterhin wird angenommen, dass die an den Wärmetauscher abgegebene Heizleistung proportional zur Brennermodulation $H_{Br}(t) \in [0, 1]$, und dass die zusätzlich nutzbare Kondensationswärme abhängig von der Rücklauftemperatur $\vartheta_{Rl}(t)$, ist. Die Bedeutung der übrigen Symbole kann der Abb. 5.15 entnommen werden.

Modellidentifikation für einen Gas-Brennwertkessel

Der in diesem Abschnitt betrachtete Gas-Brennwertkessel mit 285kW Nennleistung wird zur Brauch- und Heizwasseraufbereitung eines Wohnquartiers eingesetzt. Der Brenner des Kessels wird über einen Zweipunktregler angesteuert, sobald die Temperatur des angeschlossenen 3m²-Pufferspeichers unter einen vorgegebenen Wert fällt. Die Vorlauf- und Rücklauftemperatur wird im Zu- bzw. Ablauf des Kessels gemessen.

Beschreibung des Datensatzes: Die in diesem Abschnitt verwendeten Betriebsdaten wurden im Zeitraum vom 1.12.2018 bis zum 30.6.2019 mit einer Abtastzeit von $T_A = 1\text{min}$ aufgezeichnet. Für die Identifikation wurden aus den Rohdaten 171 separate Datensätze ausgewählt, bei denen die Kesselpumpe eingeschaltet war ($\dot{V} > 0$). Aufgrund der hohen Kesselleistung und des vergleichsweise geringen Speichervolumens sind die

Taktzeiten des Kessels relativ kurz, sodass die mittlere Datensatzlänge lediglich 16min beträgt. Die Gesamtanzahl der zur Identifikation verwendeten Datenpunkte beträgt daher nur $N = 2808$.

Modellansatz: Der hier angewendete Modellierungsansatz basiert wiederum auf der Modellstruktur (2.51). Dabei sind die Vorlauftemperatur $y[t] := \vartheta_{V1}$ als Ausgangsgröße und der Volumenstrom \dot{V} , die Rücklauftemperatur ϑ_{R1} sowie die Brennermodulation H_{Br} als die Eingangsgrößen $\mathbf{u}[t]$ des Modells zu betrachten.

Wie in den Gleichungen (5.21) angedeutet, wird die Vorlauftemperatur des Kessels im Wesentlichen durch den konvektiv zugeführten Wärmestrom $c_W \rho_W \dot{V}(t)(\vartheta_{R1}(t) - \vartheta_{V1}(t))$ und durch die Heizleistung $H_{Br}(t)\dot{Q}_K(\vartheta_{R1}(t))$ des Brenners beeinflusst. Die Eingangsnichtlinearität des Modells wurde daher zu

$$\mathbf{f}(\mathbf{u}, y, \boldsymbol{\eta}) = \begin{pmatrix} \dot{V}[t](\vartheta_{R1}[t] - \vartheta_{V1}[t]) \\ \dot{Q}_0 H_{Br}[t] \left(1 - \frac{\dot{q}_{\text{kond}}}{1 + \exp(-\eta_2(\vartheta_{R1}[t] - \eta_1))} \right) \end{pmatrix} \quad (5.22)$$

angesetzt, wobei $\dot{Q}_0 = 285\text{kW}$ die Nennleistung des Kessels und $\dot{q}_{\text{kond}} = 0.11$ den relativen Anteil der Kondensationswärme kennzeichnen. Entsprechend (5.22) wird der Übergang zwischen keiner und vollständiger Nutzung der Kondensationswärme daher durch eine logistische Funktion mit den Optimierungsparametern η_1 und η_2 modelliert (Startwert: $\boldsymbol{\eta}_{(0)} = (50, 0.1)^\top$). Die typischen Umgebungsverluste betragen lediglich $\approx 1\%$ der Heizleistung und wurden daher vernachlässigt [RSS11, Abschn. 2.3.1].

Für das dynamische Teilsystem des Modells wurde eine ARMAX-Parametrierung mit dem linear faktorisierten Polynom $A(z)$ der Ordnung 2 angesetzt. Da integrierendes Verhalten bezüglich der ersten Komponente von \mathbf{f} zu erwarten ist (vgl. Gl. (5.21)), wurde eine Nullstelle des Polynoms $A(z)$ bei $z = 1$ festgelegt und die verbleibende Nullstelle als Optimierungsparameter betrachtet. Die Verwendung einer BJ- oder OE-Parametrierung ist aufgrund der Instabilität des optimalen Prädiktors daher nicht ratsam (siehe Gl. (3.4) und Bem. 2.2). Man beachte zudem, dass durch den Integrator im Rauschmodell $H(z) = C(z)/A(z)$ tieffrequente Störgrößen, z. B. Umgebungsverluste, toleriert werden. Da insbesondere das statische Verhalten des Systems von Interesse ist, wurde weiterhin $C(z) = (z - 0.95)^2$ festgelegt¹⁵.

Identifikation: Die Identifikation des Modells mit der Eingangsnichtlinearität (5.22) wurde mittels 1-Schritt-PEM durchgeführt. Aufgrund der kurzen Datensatzlänge und der Nähe der Nullstellen von $C(z)$ zum Einheitskreis, wurde die MkQ-Methode zur Berechnung der Anfangswerte des optimalen Prädiktors verwendet (siehe Anhang B.2). Die Lösung des Optimierungsproblems wurde mittels des Levenberg-Marquardt-Algorithmus berechnet. Die identifizierten Parameterschätzwerte sind in der Tabelle 9 angegeben. Die entsprechende Kennlinie $\dot{Q}_K(\vartheta_{R1}, \boldsymbol{\eta})$ und das 99%-Konfidenzintervall bei 100% Brennerleistung ist in der Grafik 5.16 veranschaulicht.

¹⁵Wurden die Koeffizienten von $C(z)$ als freie Optimierungsparameter betrachtet, dann wurde $C(z) \approx 1$ identifiziert (ARX-Modell), was entsprechend des Rauschmodells $H(z) = C(z)/A(z)$ und der Beziehung (3.25) einer Betonung des hohen Frequenzbereiches entsprechen würde.

Diskussion: Die in der Tabelle 9 angegebenen Parameterschätzwerte und die in Abb. 5.17 dargestellte 5-Schritt Prädiktionen (5min) veranschaulichen, dass sich mit den Betriebsdaten plausible Modelle mit einer guten Prädiktionsgüte und Schätzgenauigkeit identifizieren lassen.

Polynom bzw. Parameter	Rel. Standardabw. [%]
$A(q^{-1}) = (1 - q^{-1})(1 - 0.9463q^{-1})$	–; 0.95
$B_1(q^{-1}) = 4.311 \cdot 10^{-2} q^{-1} - 3.736 \cdot 10^{-2} q^{-2}$	2.30; 2.60
$B_2(q^{-1}) = 3.436 \cdot 10^{-2} q^{-1} - 2.722 \cdot 10^{-2} q^{-2}$	5.00; 6.72
$C(q^{-1}) = (1 - 0.95q^{-1})^2$	–
$\eta_1 = 47.05$	1.59
$\eta_2 = 0.4794$	26.84

Tabelle 9: Parameterschätzwerte und relative Abweichungen des identifizierten Modells für die Gasbrennwerttherme.

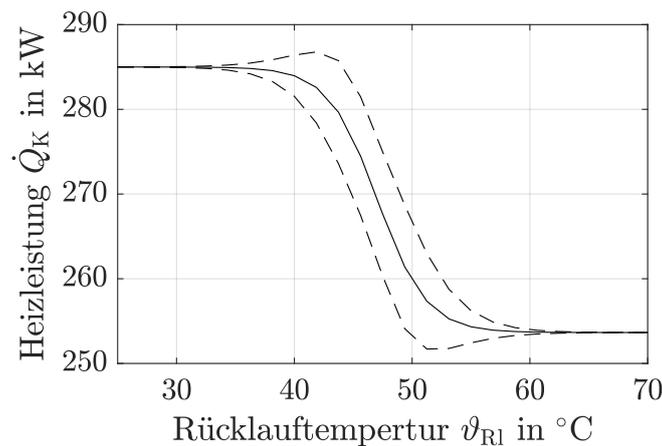


Abbildung 5.16: Identifizierte Heizleistung bei \dot{Q}_K in Abhängigkeit der Kesselrücklauf-temperatur ϑ_{RI} und 99%-Konfidenzintervall bei 100% Brennerleistung ($H_{Br} = 1$).

5.2.2 Wärmepumpe

Wärmepumpen entziehen der Umgebung unter Aufwendung mechanischer Arbeit Wärme und geben diese anschließend auf einem höheren Temperaturniveau wieder ab. Dabei wird die der Quelle entnommene Wärme (z. B. Außenluft, Grundwasser oder Erdreich) durch Verdampfung, Verdichtung und anschließende Kondensation des Kältemittels auf der Sekundärseite wieder abgegeben (siehe Abb. 5.18). Die hierfür notwendige mechanische Arbeit wird durch einen Verdichter erbracht, wobei die der Umgebung entzogene Wärme ein Vielfaches des Wärmeäquivalentes der aufgewendeten Arbeit

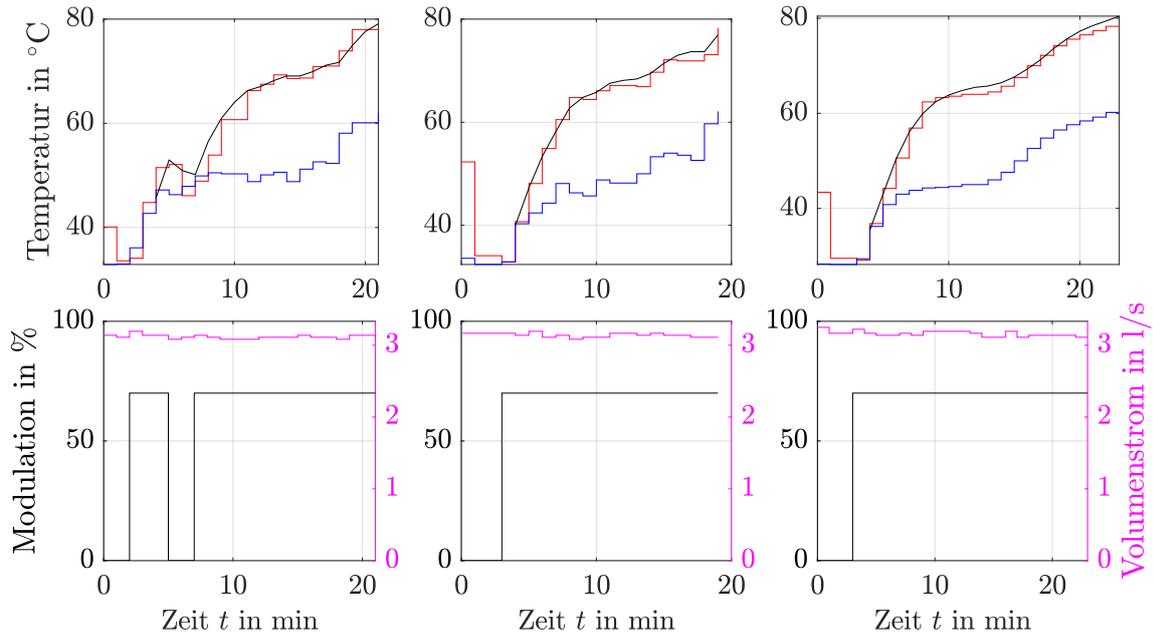


Abbildung 5.17: Oben: Gemessene Vorlauf- (–) und Rücklauf- (–) sowie 5min-Prädiktion (–) des identifizierten Brennwärmtankmodells für drei ausgewählte Datensätze. Unten: Modulation des Brenners (–) und gemessener Volumenstrom (–).

ist. Das Verhältnis der Nutzwärmeabgabe \dot{Q}_K (Kondensatorleistung) zur elektrischen Antriebsleistung P_{el} des Verdichters wird als *Leistungszahl*

$$\varepsilon = \frac{\dot{Q}_K}{P_{el}} = \frac{\dot{Q}_V + aP_{el}}{P_{el}} \quad (5.23)$$

bezeichnet (Coefficient of Performance – COP), wobei die Verdampferleistung \dot{Q}_V den der Quelle entzogenen Wärmestrom und a ¹⁶ den Anteil der Antriebsleistung, der auf der Sekundärseite in Form von Wärme wieder abgegeben wird, beschreiben [RSS11, Abschn. 5.2.1-4]. Die Leistungszahl ε ist zum einen von der Wärmepumpenbauart, und zum anderen von den Temperaturniveaus, insbesondere von der Temperaturspreizung zwischen Quelle und dem sekundären Vorlauf, d. h. $\Delta\vartheta = \vartheta_{sek,V1} - \vartheta_{prim,V1}$, abhängig. Der Zusammenhang $\varepsilon(\Delta\vartheta)$ kann im praktisch relevanten Anwendungsbereich von $\Delta\vartheta \approx 20 \dots 60\text{K}$ näherungsweise mit einer quadratischen Beziehung beschrieben werden [VDL12; SBB12; FWW17]. Für eine anwendungsorientierte Beschreibung der Funktionsweise von Wärmepumpen sei auf [RSS11, Abschn. 5.2] und [Vie11] verwiesen. Detaillierte Ausführungen zur Modellierung können bspw. in [Jin12] nachgeschlagen werden.

¹⁶Der Faktor a ist abhängig von der verwendeten Bauart des Verdichters, der Wärmedämmung und den Betriebstemperaturen der warmen Bauteile. Für überschlagsmäßige Berechnungen kann $a \approx 0.8 \dots 1$ angesetzt werden [RSS11, Abschn. 5.2.1-3].

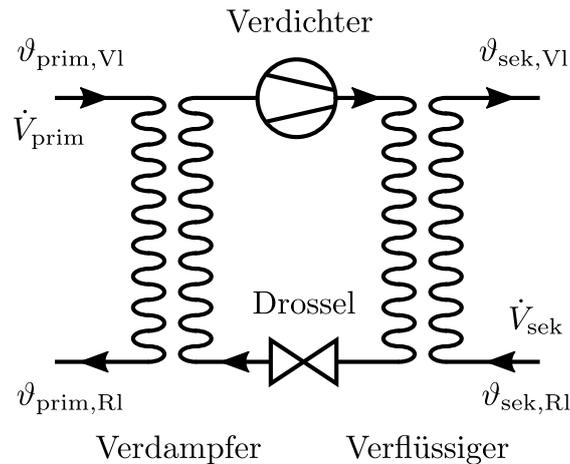


Abbildung 5.18: Schematische Darstellung einer Wärmepumpe.

Folgend wird wiederum beispielhaft ein Modellierungs- und Identifikationsansatz für eine Wärmepumpe vorgestellt. Dabei wird auf die Verwendung interner Systemgrößen (Temperaturen und Drücke des Kältemittels) verzichtet, da aus regelungstechnischer Sicht vor allem das Eingangs-Ausgangs-Verhalten von Interesse ist. Weiterhin wird angenommen, dass sich die internen Systemgrößen stets im jeweilig vorgesehenen Arbeitsbereich befinden, sodass keine Leistungsreduktion oder Abschaltung der Wärmepumpe durch die interne Regelung, bspw. aufgrund des Über- oder Unterschreitens der Verflüssigungs- bzw. Verdampfungstemperatur des Kältemittels, erfolgt.

Modellidentifikation für eine Wärmepumpe

Die hier betrachtete Wärmepumpe mit einer Nennleistung von $2 \times 50\text{kW}$ wird primärseitig an einem Erdwärmesondenfeld betrieben und dient der Versorgung eines Fußbodenheizungssystems eines Wohnquartieres. Die Wärmepumpe besitzt zwei Verdichterstufen, welche jedoch lediglich abwechselnd betrieben werden. Im nachfolgend dargestellten Modellansatz wird daher angenommen, dass lediglich eine der beiden Kompressorstufen eingeschaltet ist, und dass diese die gleiche Leistungsabgabe besitzen.

Beschreibung des Datensatzes: Die zur Identifikation verwendeten Betriebsdaten wurden im Zeitraum vom 6.9.2019 bis zum 11.11.2019 mit einer Abtastzeit von $T_A = 1\text{min}$ aufgezeichnet. Dabei wurden aus den Rohdaten 146 separate Datensätze selektiert, für welche die Pumpen des Primär- und Sekundärkreises eingeschaltet waren ($\dot{V}_{\text{prim}} > 0$ und $\dot{V}_{\text{sek}} > 0$). Die Gesamtanzahl der verwendeten Datenpunkte beträgt $N = 13949$.

Modellansatz: Das prinzipielle Vorgehen orientiert sich an dem im vorherigen Abschnitt dargestellten Modellierungsansatz für Heizkessel, d. h. das Wasservolumen des Verdampfers und des Kondensators werden wiederum als konzentrierte Wärmespeicher betrachtet. Die Vorlauf bzw. Rücklauftemperatur auf der Sekundär- bzw. Primärseite sind daher als die Ausgangsgrößen $\mathbf{y} := (\vartheta_{\text{sek,Vl}}, \vartheta_{\text{prim,Rl}})^\top$ des Modells zu betrachten. Diese werden wiederum durch die entsprechenden konvektiven Wärmeströme sowie durch die

Verdampfer- bzw. Verflüssigerleistung beeinflusst, weshalb die Eingangsnichtlinearität

$$\mathbf{f}(\mathbf{u}, \mathbf{y}, \boldsymbol{\eta}) = \begin{pmatrix} \dot{V}_{\text{sek}}[t] (\vartheta_{\text{sek,Rl}}[t] - \vartheta_{\text{sek,Vl}}[t]) \\ H_{\text{Komp}}[t] \dot{Q}_{\text{K}}(\vartheta_{\text{sek,Vl}}[t], \vartheta_{\text{prim,Vl}}[t], \boldsymbol{\eta}) \\ \dot{V}_{\text{prim}}[t] (\vartheta_{\text{prim,Vl}}[t] - \vartheta_{\text{prim,Rl}}[t]) \\ H_{\text{Komp}}[t] \dot{Q}_{\text{V}}(\vartheta_{\text{sek,Vl}}[t], \vartheta_{\text{prim,Vl}}[t], \boldsymbol{\eta}) \end{pmatrix} \quad (5.24)$$

angesetzt wurde. Dabei gilt aufgrund der abwechselnd betriebenen Verdichterstufen $H_{\text{Komp}}[t] = 1$, sobald einer der beiden Kompressoren eingeschaltet ist.

Entsprechend der Funktion (5.24) werden die Verdampfer- und die Kondensatorleistung als statische Funktionen in Abhängigkeit der Temperaturniveaus $\vartheta_{\text{sek,Vl}}$ und $\vartheta_{\text{prim,Vl}}$ modelliert. Dabei wurde für die Verflüssigerleistung ein linearer Modellansatz

$$\dot{Q}_{\text{K}}[t, \boldsymbol{\eta}] = \dot{Q}_{\text{K},0} + \eta_1(\vartheta_{\text{sek,Vl}}[t] - \vartheta_{\text{sek,Vl},0}) + \eta_2(\vartheta_{\text{prim,Vl}}[t] - \vartheta_{\text{prim,Vl},0}) \quad (5.25)$$

mit den Nenndaten $\dot{Q}_{\text{K},0} = 50\text{kW}$ bei $\vartheta_{\text{sek,Vl},0} = 35^\circ\text{C}$ und $\vartheta_{\text{prim,Vl},0} = 10^\circ\text{C}$ gewählt¹⁷. Die der Wärmequelle entzogene Verdampferleistung \dot{Q}_{V} wurde dann in Abhängigkeit der Verflüssigerleistung und der Leistungszahl ε modelliert (vgl. (5.23))

$$\dot{Q}_{\text{V}}[t, \boldsymbol{\eta}] = \dot{Q}_{\text{K}}[t, \boldsymbol{\eta}] \frac{\varepsilon[t, \boldsymbol{\eta}] - a}{\varepsilon[t, \boldsymbol{\eta}]}, \quad (5.26)$$

wobei für die Leistungszahl ein quadratischer Zusammenhang

$$\varepsilon[t, \boldsymbol{\eta}] = \varepsilon_0 + \eta_3(\Delta\vartheta[t] - \Delta\vartheta_0) + \eta_4(\Delta\vartheta[t] - \Delta\vartheta_0)^2 \quad (5.27)$$

angesetzt wurde ($\Delta\vartheta_0 = \vartheta_{\text{sek,Vl},0} - \vartheta_{\text{prim,Vl},0} = 25\text{K}$). Der COP ε der betrachteten Wärmepumpe unter den Normbedingungen $35^\circ\text{C}/10^\circ\text{C}$ beträgt laut Datenblatt $\varepsilon_0 = 5.6$.

Für das dynamische Teilsystem wurde eine MIMO ARMAX-Parametrierung

$$\mathbf{A}(q^{-1})\mathbf{y}[t] = \mathbf{B}(q^{-1})\mathbf{f}(\mathbf{u}, \mathbf{y}, \boldsymbol{\eta}) + \mathbf{C}(q^{-1})\mathbf{e}[t] \quad (5.28)$$

mit

$$\begin{aligned} \mathbf{A}(q^{-1}) &= \begin{pmatrix} A_1(q^{-1}) & 0 \\ 0 & A_2(q^{-1}) \end{pmatrix}, \quad \mathbf{C}(q^{-1}) = \begin{pmatrix} C_1(q^{-1}) & 0 \\ 0 & C_2(q^{-1}) \end{pmatrix}, \quad \text{und} \\ \mathbf{B}(q^{-1}) &= \begin{pmatrix} B_{11}(q^{-1}) & B_{12}(q^{-1}) & 0 & 0 \\ 0 & 0 & B_{23}(q^{-1}) & B_{24}(q^{-1}) \end{pmatrix} \end{aligned} \quad (5.29)$$

¹⁷Bei der Validierung des beschriebenen Modellansatzes und der Identifikationsmethodik mit einem in der MODELICA-Sprache implementierten Modell einer Wasser/Wasser-Wärmepumpe (BUILDINGS LIBRARY 7.0.0: Buildings.Fluid.HeatPumps.ScrollWaterToWater), konnten mit einem quadratischen Ansatz für die Verflüssigerleistung \dot{Q}_{K} genauere Ergebnisse erzielt werden. Bei der Identifikation mittels der realen Messdaten konnten die entsprechenden Parameter der quadratischen Terme jedoch mit keiner hinreichend hohen Genauigkeit identifiziert werden, sodass die lineare Ansatzfunktion (5.25) favorisiert wurde.

verwendet, wobei für die Ordnungen aller in (5.29) enthaltenen Polynome eins gewählt wurde. Weiterhin wurden die Nullstellen von A_1 und A_2 zu $z_0 = 1$ vorgegeben, weil näherungsweise integrierendes Verhalten bezüglich der konvektiven Wärmeströme zu erwarten ist¹⁸. Da wiederum insbesondere das statische Verhalten der Wärmepumpe von Interesse ist, wurde $C_1(q^{-1}) = C_2(q^{-1}) = 1 - 0.95q^{-1}$ festgelegt (vgl. Abschn. 5.2.1).

Identifikation: Die Parameteridentifikation des Modells (5.28) mit (5.24) wurde mittels 1-Schritt-PEM durchgeführt, wobei das Kostenfunktional (3.7) verwendet wurde. Die Lösung des Optimierungsproblems wurde mittels sequentieller quadratischer Programmierung (MATLAB: `fmincon()`¹⁹, `Algorithm='sqp'`) durchgeführt, wobei durch entsprechende Nebenbedingungen an $\boldsymbol{\eta}$ monoton fallendes Verhalten der Leistungszahl ε im Bereich $\Delta(\vartheta) = 20 \dots 60\text{K}$ erzwungen wurde. Die Parameter der Eingangsnichtlinearität (5.24) wurden zu $\boldsymbol{\eta}_{(0)} = (-0.003 \cdot \dot{Q}_{K,0}, 0.03 \cdot \dot{Q}_{K,0}, -0.15, 7.34 \cdot 10^{-4})^\top$ initialisiert²⁰.

Diskussion: In der Tabelle 10 sind die identifizierten Parameterschätzwerte des Modells und deren relative Abweichungen angegeben. Die entsprechenden Kennlinienfelder für die Verdampfer- und die Verflüssigerleistung sowie deren 99%-Konfidenzintervalle sind in der Grafik 5.19 dargestellt. Weiterhin wird in Abbildung 5.20 die 5-Schritt Prädiktion (5min) des Modells für 3 ausgewählte Datensätze veranschaulicht. Die Ausführungen verdeutlichen qualitativ, dass sich mit den Betriebsdaten plausible Modelle mit einer guten Prädiktionsgüte und Schätzgenauigkeit identifizieren lassen.

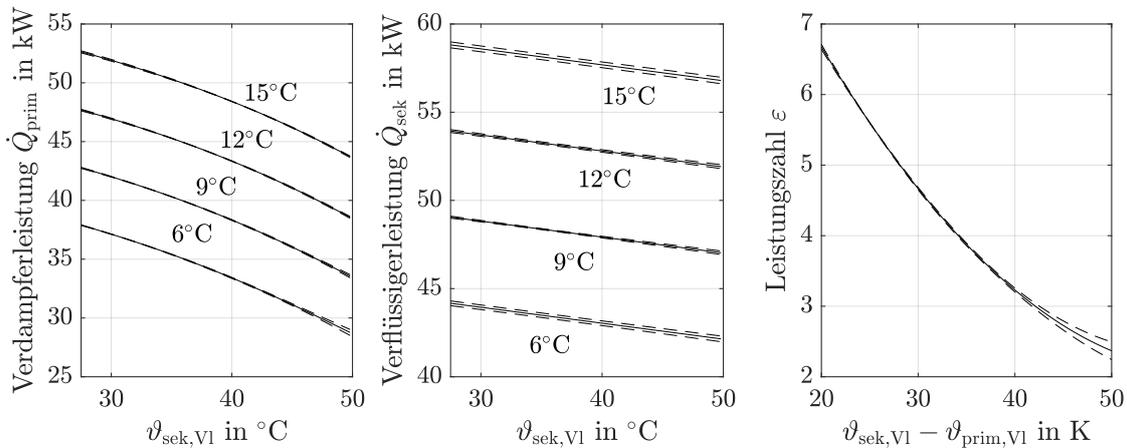


Abbildung 5.19: Kennlinienfelder der Verdampfer- und Verflüssigerleistung für unterschiedliche Quellentemperaturen $\vartheta_{\text{prim,V1}} \in \{6, 9, 12, 15\}^\circ\text{C}$ sowie der Leistungszahl ε der Wärmepumpe und die 99%-Konfidenzintervalle (eine Verdichterstufe).

¹⁸Bei der Wahl des Modellansatzes, insbesondere der Polynomordnungen, wurde das dynamische Verhalten der Wärmetauscher (Verflüssiger und Verdampfer), dass näherungsweise in der Größenordnung der Abtastzeit $T_A = 1\text{min}$ zu erwarten ist und daher ohnehin nicht mit einer sinnvollen Genauigkeit identifiziert werden kann, vernachlässigt.

¹⁹Die Funktion `lsqnonlin()` unterstützt keine allgemeinen Ungleichheitsnebenbedingungen.

²⁰Die Werte $\boldsymbol{\eta}_{(0)}$ wurden bei Experimenten mit einem in der MODELICA-Sprache implementierten Modell einer Wasser/Wasser-Wärmepumpe (BUILDINGS LIBRARY 7.0.0: `Buildings.Fluid.HeatPumps.ScrollWaterToWater`) ermittelt.

Polynom bzw. Parameter	Rel. Standardabw. [%]
$A_1(q^{-1}) = A_2(q^{-1}) = (1 - q^{-1})$	–
$B_{11}(q^{-1}) = 6.496 \cdot 10^{-2} q^{-1}$	1.73
$B_{12}(q^{-1}) = 8.794 \cdot 10^{-2} q^{-1}$	1.66
$B_{23}(q^{-1}) = 1.849 \cdot 10^{-2} q^{-1}$	1.15
$B_{24}(q^{-1}) = -2.510 \cdot 10^{-2} q^{-1}$	1.10
$C_1(q^{-1}) = C_2(q^{-1}) = 1 - 0.95 q^{-1}$	–
$\eta_1 = -9.043 \cdot 10^{-2}$	2.74
$\eta_2 = 1.628$	0.76
$\eta_3 = -0.2011$	0.98
$\eta_4 = 2.872 \cdot 10^{-3}$	5.18

Tabelle 10: Parameterschätzwerte und relative Abweichungen des identifizierten Modells für die Wärmepumpe.

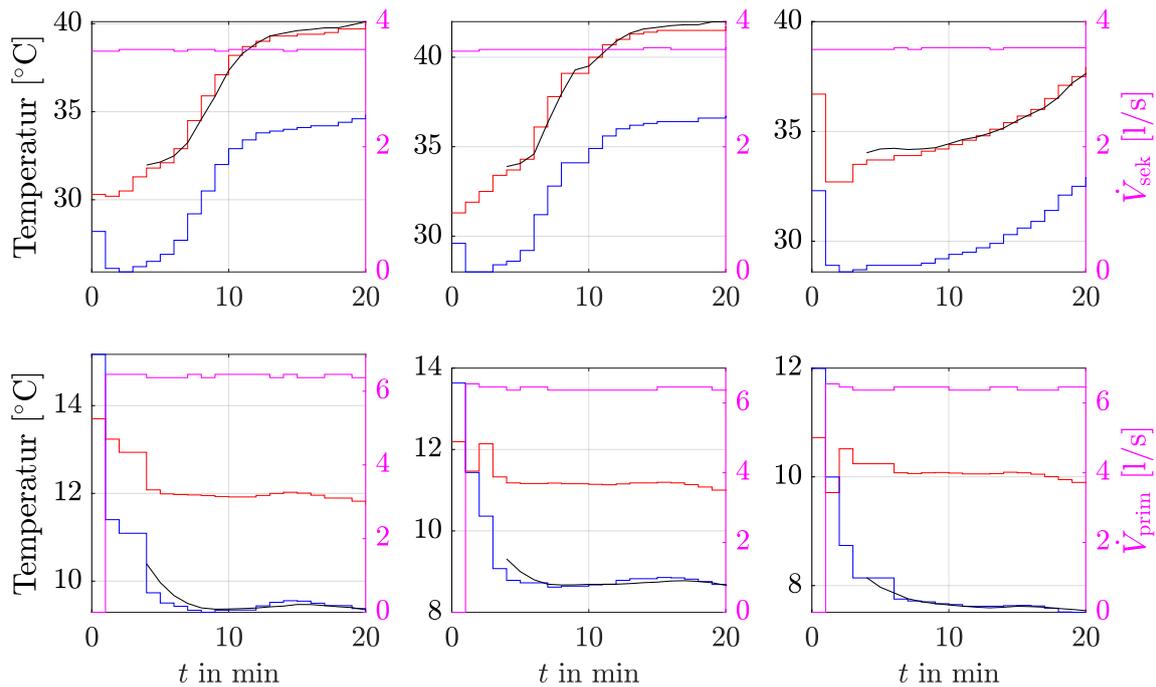


Abbildung 5.20: Oben: Gemessene sekundäre Vorlauf- (---), Rücklauf- (—) und Volumenstrom \dot{V}_{sek} (—) der Wärmepumpe sowie 5min Prädiktion der Vorlauf- (---) für drei ausgewählte Datensätze. Unten: Messwerte der primären Vorlauf- (---), Rücklauf- (—) und Volumenstrom \dot{V}_{prim} (—) sowie 5min Prädiktion der primären Rücklauf- (---).

Kapitel 6

Beispielimplementierungen modellprädiktiver Regler in der Gebäudetechnik

In diesem Kapitel soll die Implementierung modellprädiktiver Regler auf der Basis des in Abschnitt 5.1 vorgestellten Identifikationsansatzes veranschaulicht werden. Dabei soll zum einen gezeigt werden, dass die durch Identifikation gewonnenen Modelle in der Lage sind bereits in der Forschung etablierte und validierte Modelle hinreichend gut zu beschreiben. Des Weiteren sollen anhand zweier praxisnaher Beispiele Implementierungsansätze bzw. -vorschläge für den Entwurf von MPC-Reglern vorgestellt werden. Dabei orientiert sich ersteres an den Anwendungsbeispielen der Abschnitte 5.1.1 bzw. 5.1.2 und veranschaulicht den Entwurf eines Einzelraumreglers für einen Raum mit Fußbodenheizung und -kühlung. In einem weiteren Fallbeispiel wird die Realisierung eines Mehrzonenreglers für die Temperaturregelung eines Hörsaalgebäudes demonstriert.

Die Grundlage für die Identifikation und simulationsbasierte Validierung der entworfenen Regler ist in beiden Fällen ein entsprechendes Regelstreckenmodell, das in der Modellierungssprache MODELICA entworfen wurde. Aufgrund der großen Verbreitung wurde hierfür die frei verfügbare BUILDINGS Bibliothek (Version 7.0.0) verwendet [Wet21], wobei für die Raummodellierung die Klasse `Buildings.ThermalZones.Detailed.MixedAir` genutzt wurde. Diese bildet ein konzentriert parametrisches koppelbares Raummodell mit einer beliebigen Anzahl von Wänden und Fenstern nach, die über Wärmeleitung, Konvektion, Infrarot- und Solarstrahlung an der Wärmebilanz des Raumes teilnehmen. Eine ausführliche Beschreibung der durch das Modell abgebildeten Vorgänge ist der Dokumentation der BUILDINGS Bibliothek und der Referenz [WZN11] zu entnehmen.

Die für beide Beispiele entworfenen MODELICA Modelle wurden zum einen für die Generierung „synthetischer“ Identifikationsdaten verwendet. Zum anderen wurden diese auch als Prozessmodell über die FMI-Schnittstelle (siehe Bemerkung 6.1) in die Simulationsumgebung MATLAB/SIMULINK integriert und für die Validierung und den Vergleich des MPC-Reglers mit einer konventionellen Regelungsstrategie verwendet. Ein schematisches Prinzip des entsprechenden Regelkreises ist in der Abbildung 6.1 dargestellt. Der Quellcode und die Simulationsmodelle beider Beispiele sind im Digitalarchiv

https://github.com/fpaschke/phd_identification_mpc.git
verfügbar.

Bemerkung 6.1 (FMI-Schnittstelle): Das Functional Mock-up Interface (FMI) ist eine standardisierte Schnittstelle zur Kopplung unterschiedlicher Simulationswerkzeuge bzw. Modelle, welche entsprechend als Functional Mock-up Unit (FMU) bezeichnet werden [BOA11]. Hierbei existieren im Wesentlichen zwei unterschiedliche Kopplungsmechanismen, die als „Model-Exchange“ bzw. „Co-Simulation“ bezeichnet werden. Dabei besteht der wesentliche Unterschied darin, dass im ersteren Fall nur das Differential-algebraische Gleichungssystem als C-Code exportiert wird, wohingegen bei Co-Simulation zusätzlich ein Integrationsalgorithmus in die FMU integriert wird. Im Rahmen der Arbeit wurde der Kopplungsmechanismus Model-Exchange in Kombination mit dem MATLAB/SIMULINK Löser ode45 verwendet. ◁

Bemerkung 6.2 (Störgrößen und Störgrößenprädiktionen): Im Folgenden wird eine Aufteilung des Vektors \mathbf{u} , der im Rahmen der Systemidentifikation allgemein als Eingangsvektor des Modells betrachtet wird, in Stellgrößen und (präzifizierbare) Störgrößen vorgenommen. Nach der in der Regelungstheorie gebräuchlichen Konvention wird für die Stellgröße jedoch auch das Symbol $\mathbf{u}[t]$ verwendet [Mac00; Rei06; Lun16a]. Die Unterscheidung der gemeinten Bedeutung muss daher durch den jeweiligen Kontext (Identifikation/Reglerentwurf) erfolgen. Weiterhin sei auf die explizite Unterscheidung präzifizierbarer und nicht-präzifizierbarer Störungen $\mathbf{d}(t)$ (z. B. Außentemperatur) bzw. $\mathbf{v}(t)$ hingewiesen (siehe Abb. 2.1, 2.2 und 6.1). ◁

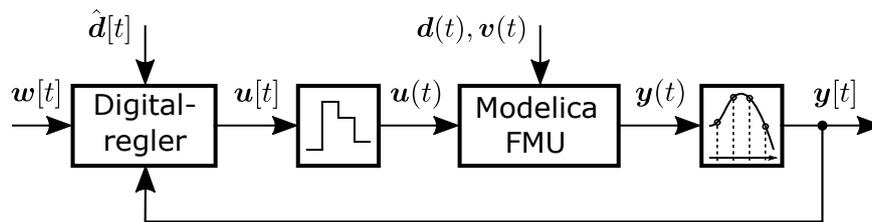


Abbildung 6.1: Schematische Darstellung der implementierten Regelung mit Prozessmodell-FMU sowie Abtast- und Halteglied 0-ter Ordnung.

6.1 Modellprädiktiver Einzelraumregler

Dieser Abschnitt befasst sich mit dem Entwurf eines MPC-Reglers für einen Einzelraum, wobei die Heizventilposition der Fußbodenheizung sowie die Höhe der Außenverschattung als Stellgrößen und die Lufttemperatur als Regelgröße betrachtet werden. Wie in der Einleitung des Kapitels bereits erwähnt wurde, wird für die Generierung des MPC-Modells und die Validierung des Reglers ein in MODELICA entworfenes Prozessmodell verwendet, welches im ersten Unterabschnitt kurz beschrieben wird. Basierend auf dem Simulationsoutput dieses Modells wird nachfolgend ein reduziertes Modell trainiert, welches anschließend für die Formulierung des prädiktiven Regelungsansatzes verwendet wird. Abschließend wird der MPC nach dem in der Abb. 6.1 dargestellten

Prinzip in einer Jahressimulation validiert und die Ergebnisse mit einer konventionellen Regelungsstrategie verglichen.

6.1.1 Prozessmodell

Das Modell des hier betrachteten Raumes mit einer Abmessung von $6 \times 10 \times 3\text{m}$ wird als Bestandteil eines Massivbaus angenommen. Der Raum besitzt eine westwärts orientierte Außenwand mit einer doppelt verglasten Fensterfläche von insgesamt $6 \times 2\text{m}$ (Rahmenanteil 10%). Die äußeren Witterungsbedingungen, werden durch einen Wetterdatensatz des Standortes Berlin nachgebildet (Testreferenzjahr¹ - TRY). Die Temperaturen der angrenzenden Nachbarräume wurden durch Messwerte eines realen Gebäudes vorgegeben, wohingegen die Temperatur des Erdreiches durch einen sinusförmigen Jahresverlauf mit einem Mittelwert von 9°C und einer Amplitude von 6K modelliert wurde (Minimum am 1. März). Der Wand-, Decken- und Fußbodenaufbau des Modells kann der Tabelle 11 entnommen werden, wobei die in der BUILDINGS Bibliothek hinterlegten Standardwerte für die Materialparameter verwendet wurden.

Der Raum verfügt über eine Fußbodenheizung, die in den Wintermonaten zum Heizen und vom 1. Mai bis zum 30. September zum Kühlen verwendet werden kann. Dabei wird die Vorlauftemperatur im Sommer konstant zu 18°C und im Winter außentemperaturgeführt durch eine Heizkurve eingestellt. Zudem wird angenommen, dass die Höhe der Außenverschattung angepasst werden kann. Die thermische Last der extern vorzugebenden Raumbelastung wird zu $100\text{W}/\text{Person}$ angesetzt (75W Strahlung und 25W Konvektion). Eine grafische Darstellung des MODELICA-Modells ist in der Abbildung 6.2 gezeigt, wobei zusätzlich noch erwähnt sei, dass die Raumtemperatur mit einer Auflösung von 0.1K quantisiert wurde (nicht in den Abb. 6.1 und 6.2 dargestellt).

Typ	Aufbau ²	Randbedingung
Innenwand (N,O,S)	1.25cm Gipskarton, 12cm Backstein, 1.25cm Gipskarton	Temperaturmesswerte
Außenwand (W)	1.25cm Gipskarton, 20cm Isolations- material, 11.5cm Backstein	Wetterdatensatz Berlin (Stun- denwerte)
Decke	20cm Beton	Temperaturmesswert
Fußboden	5.5cm Betonestrich, 20cm Isolations- material, 25cm Beton	sinusförmiger Jahrestemperatur- verlauf

Tabelle 11: Wandaufbau und -ausrichtung sowie Randbedingungen des Modelica Einzelraummodells.

¹Verfügbar unter <https://energyplus.net/weather> (zuletzt besucht am 11.06.2021).

²Aufzählung beginnend mit der dem Raum zugewandten Seite.

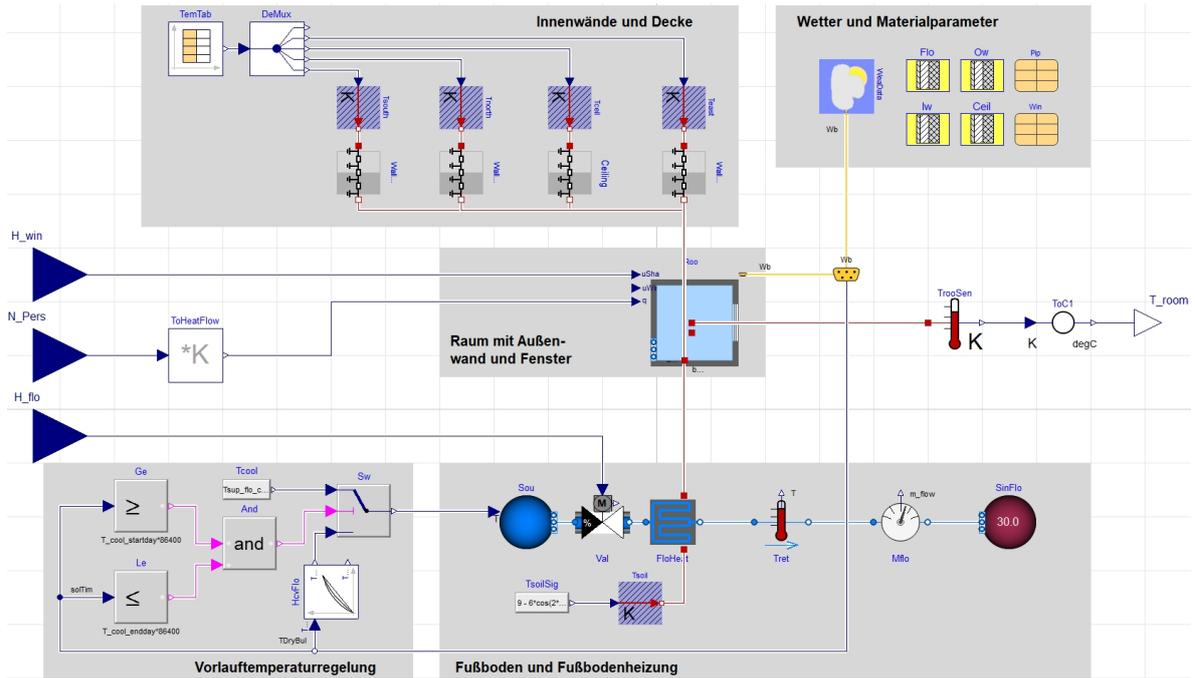


Abbildung 6.2: Grafische Darstellung des Modelica Einzelraummodells mit Fußbodenheizung.

6.1.2 Identifikation

Beschreibung des Datensatzes: Der für die Identifikation verwendete Datensatz wurde durch Simulation des MODELICA Modells über 4 Wochen erzeugt ($N = 2689$), wobei die Stellgrößen durch sprungförmige Signale mit zufällig gewählter Sprunghöhe angeregt wurden (siehe Abb. 6.5). Die Abtastzeit wurde zu $T_A = 15\text{min}$ gewählt.

Modellansatz: Die Identifikation wurde analog zu dem in Abschnitt 5.1 dargestellten Vorgehen durchgeführt. Dabei wurde die Eingangsnichtlinearität des Modellansatzes (2.51) zu

$$\mathbf{f}(\mathbf{u}, \mathbf{y}, \boldsymbol{\eta}) = \begin{pmatrix} \vartheta_{\text{AuI}}[t] \\ H_F[t] \dot{Q}_{\text{gen}}[t] \\ H^n[t] \text{sgn}(\vartheta_{\text{VI}}[t] - \vartheta_{\text{R}}[t]) |\vartheta_{\text{VI}}[t] - \vartheta_{\text{R}}[t]|^n \end{pmatrix} \quad (6.1)$$

gewählt ($n = 1.1$ – vgl. Anhang C.1), wobei $H_F \in [0, 1]$ die Position der Außenverschattung kennzeichnet und weiterhin die Bezeichnungen aus der Tabelle 6 gelten. Es sei zudem darauf hingewiesen, dass die auf der Trennung der Globalstrahlung in einen Diffus- und einen Direktstrahlungsanteil basierende Methode zur Berechnung der solaren Gewinne verwendet wurde (siehe Gl. (5.7)). Ferner wurde für das LTI-Teilsystem die ARMAX-Parametrierung mit einem festen Hochpassvorfilter (5.10) mit $T_L = 6\text{h}$ angesetzt, wobei für die Polynome $A(q^{-1})$ und $B_i(q^{-1})$ ein faktorisiertes Ansatz verwendet wurde.

Identifikation: Die Modelle wurden mittels 1-Schritt-PEM und MkQ-Initialisierung

des optimalen Prädiktors identifiziert. Zur Lösung des Optimierungsproblems wurde das Trust-Region-Verfahren verwendet, wobei $\hat{\eta}_{1,(0)} = 1$ gewählt und die initiale Parameterschätzung $\hat{\theta}_{(0)}$ mittels MkQ berechnet wurde (Gl. (3.37)). Die identifizierten Modellparameter und deren relative Standardabweichungen sind in Tabelle 12 aufgeführt. Die entsprechenden Frequenzgänge sind in der Abbildung 6.3 dargestellt.

Polynom bzw. Parameter	Rel. Standardabw. [%]
$A(q^{-1}) = (1 - 0.5662 q^{-1})(1 - 0.9666 q^{-1})$	4.14; 0.18
$B_1(q^{-1}) = 3.145 \cdot 10^{-2} q^{-1}(1 - 0.9579 q^{-1})$	6.05; 0.57
$B_2(q^{-1}) = 1.132 \cdot 10^{-4} q^{-1}$	5.87
$B_3(q^{-1}) = 4.514 \cdot 10^{-3} q^{-1}$	5.27
$C(q^{-1}) = 1 - 0.7503 q^{-1} + 0.2195 q^{-2}$	3.4; 6.25
$\eta_1 = 1.597$	6.55

Tabelle 12: Parameterschätzwerte des Raummodells und deren relative Abweichungen.

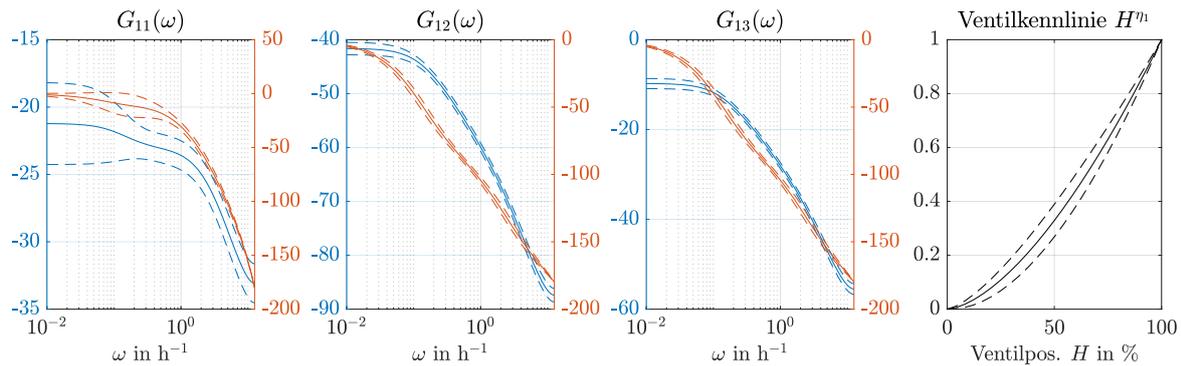


Abbildung 6.3: Amplituden- (—) bzw. Phasenfrequenzgänge (---) und deren 99%-Konfidenzintervalle (---), sowie die identifizierte Ventilkenlinie mit 99% Konfidenzintervall (rechts).

Diskussion: Die Abbildungen 6.3 bis 6.5 veranschaulichen, dass durch Identifikation ein Modell mit angemessener Prädiktionsgüte aus den Simulationsdaten erzeugt werden kann. Insbesondere ist hierbei bemerkenswert, dass das MODELICA Modell mit insgesamt 40 Zustandsgrößen durch das identifizierte Modell mit nur 3 Zuständen (ARMAX-Teilsystem und Vorfilter) approximiert wird.

Die Abbildung 6.4 zeigt, dass der Fehler $\varepsilon[t|t - k]$ für lange Prädiktionshorizonte k deutlich ansteigt, was durch den lediglich als Störgröße modellierten Einfluss der Wärmeabgabe der Wände zu erklären ist. Die Darstellung der Autokovarianzfunktion zeigt weiterhin, dass die Residuen näherungsweise als weißes Rauschen betrachtet werden können, was verdeutlicht, dass die Modellstruktur eine hinreichende Komplexität besitzt. Durch Transformation des LTI-Teilsystems in den kontinuierlichen Zeitbereich, können

die Zeitkonstanten zudem zu $T_1 \approx 7.36\text{h}$ und $T_2 \approx 26\text{min}$ ermittelt werden, was in etwa den praktischen Erwartungen³ entspricht.

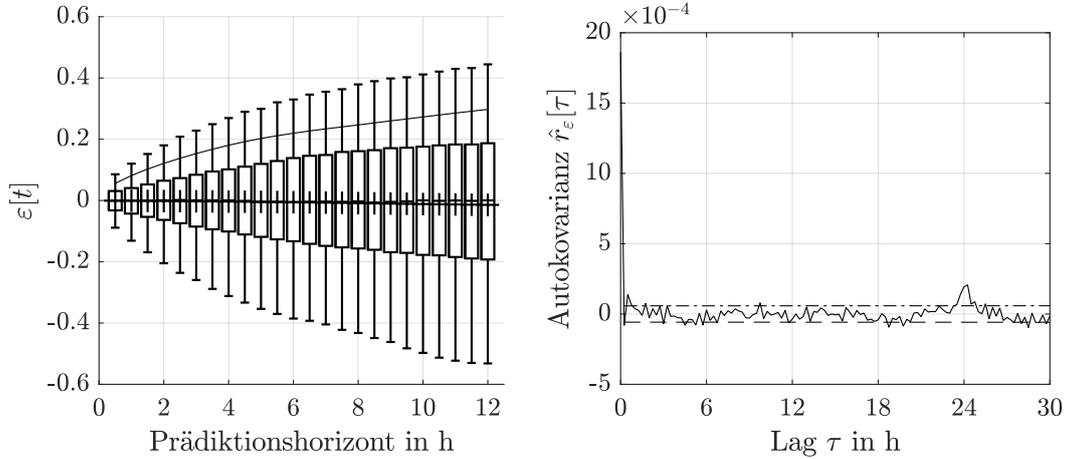


Abbildung 6.4: Box-Whisker-Diagramm der Prädiktionsfehler $\varepsilon[t|t-k]$ (links) und Autokovarianzfolge $\hat{r}_\varepsilon[\tau]$ der Residuen $\varepsilon[t|t-1]$ (rechts). In der Darstellung links wurden die empirischen 5, 25, 50, 75 und 95% Perzentile sowie der Mittelwert (+) und die Standardabweichung (-) markiert. Die gestrichelte Linie rechts kennzeichnet den kritischen Bereich zum Signifikanzniveau $\alpha = 0.01$ (siehe Bemerkung 5.2).

6.1.3 Reglerentwurf

Dieser Abschnitt veranschaulicht die Implementierung eines MPC-Reglers für einen Einzelraum. Das Ziel der Regelung besteht im Allgemeinen darin, unter Aufwendung von möglichst wenig Heiz- bzw. Kühlenergie, die Raumtemperatur in einem vorgegebenen Komfortbereich, d. h. zwischen einem Minimal- und einem Maximalwert, zu halten. Dabei werden die Heizventilposition der Fußbodenheizung sowie die Höhe der Außenverschattung als Stellgrößen und die Außentemperatur sowie die Solarstrahlung als prädizierbare Störgrößen betrachtet:

$$\mathbf{u}[t] := \begin{pmatrix} H[t] \\ H_F[t] \end{pmatrix} \quad \mathbf{d}[t] := \begin{pmatrix} \vartheta_{\text{Aul}}[t] \\ \dot{Q}_{\text{gen}}[t] \end{pmatrix} \quad \mathbf{y}[t] := \vartheta_{\text{R}}[t] \quad (6.2)$$

Vereinfachend wird im Folgenden angenommen, dass die minimale Beleuchtungsstärke innerhalb des Raumes durch Anpassung des Lamellenwinkels der Verschattung oder durch die elektrische Beleuchtung gewährleistet wird, sodass der Regler die Höhe der Außenverschattung und somit die solaren Erträge unabhängig von der Beleuchtungsstärke wählen kann.

³Aufgrund der hohen Wärmekapazität des Betonstrichs und der vergleichsweise geringen Kapazität der Raumluft sind 2 Zeitkonstanten im Stunden- bzw. Minutenbereich zu erwarten.

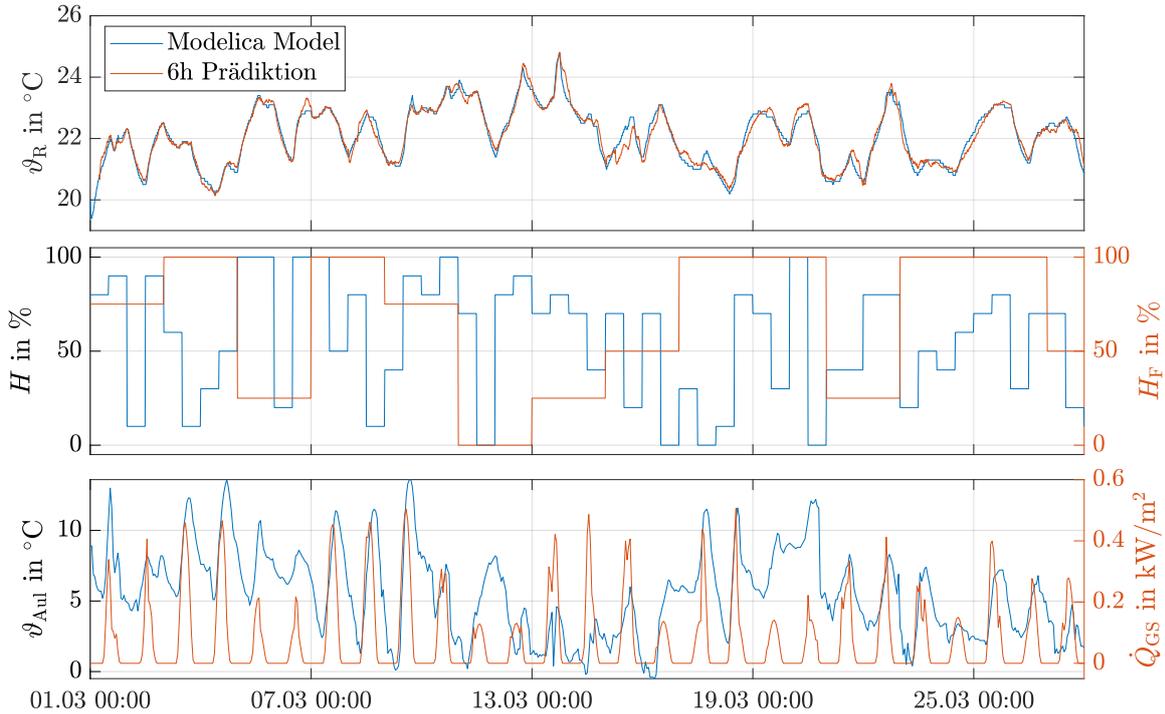


Abbildung 6.5: Vergleich der 24-Schritt Prädiktion (6h) des identifizierten Raummodells mit der Simulation des Modelica Modells auf dem Identifikationsdatensatz.

Die Formulierung des Optimierungsproblems basiert auf der Zustandsraumdarstellung des deterministischen Teilsystems des Modells (2.51)

$$\begin{aligned} \mathbf{x}[t+1] &= \mathbf{A}\mathbf{x}[t] + \mathbf{B}_u \mathbf{f}_u(\mathbf{u}[t], \mathbf{d}[t], \mathbf{y}[t]) + \mathbf{B}_d \mathbf{f}_d(\mathbf{d}[t], \mathbf{y}[t]) \\ \mathbf{y}[t] &= \mathbf{C}\mathbf{x}[t], \end{aligned} \quad (6.3)$$

wobei eine Aufteilung der Nichtlinearität (6.1) in einen von den Stell- bzw. Störgrößen abhängigen Teil $\mathbf{f}_u = (f_2, f_3)^\top$ bzw. $\mathbf{f}_d = f_1$ vorgenommen wird⁴. Man beachte, dass das Modell aufgrund von $d = 1$ keinen Durchgriff besitzt, sodass $\mathbf{D}_u = \mathbf{D}_d = \mathbf{0}$ gilt (siehe Bem. (2.12) und vgl. Tabelle 12).

Bei dem im Folgenden dargestellten Implementierungsansatz für den MPC wird die virtuelle Stellgröße $\tilde{\mathbf{u}}[t] := (H_F[t], f_3)^\top$ als Optimierungsvariable betrachtet. Dies ist vorteilhaft, da zum einen die Komponente f_3 von (6.1) als näherungsweise proportional zur Heiz- bzw. Kühlleistung angenommen wird (siehe Gln. (5.2)), wodurch die entsprechende Energie bei der Formulierung des Optimierungsproblems berücksichtigt werden kann. Weiterhin kann aufgrund der Linearität des Ersatzmodells die Prädiktion des Systemausgangs durch ein einfaches Matrix-Vektor-Produkt

$$\mathbf{y}[t] = \mathbf{\Psi}\mathbf{x}[t] + \mathbf{\Omega}\mathbf{D}[t] + \mathbf{\Theta}\mathbf{M}[t]\tilde{\mathbf{u}}[t] \quad (6.4)$$

⁴Es sei darauf hingewiesen, dass im Folgenden aufgrund einer übersichtlicheren Notation auf die explizite Kennzeichnung prädizierter bzw. geschätzter Größen durch $(\hat{\cdot})$ verzichtet wird.

berechnet werden (vgl. [Mac00, Kap. 2.6]), wobei $\mathbf{M}[t] = \text{diag}(\dot{Q}_s[t], 1, \dot{Q}_s[t+1], 1, \dots, \dot{Q}_s[t+k-1], 1)$. Weiterhin kennzeichnen $\mathbf{x}[t]$ den Systemzustand und $\mathbf{y}[t] = \text{vec}\{\mathbf{y}[t+1], \dots, \mathbf{y}[t+k]\}$ sowie $\mathbf{D}[t] = \text{vec}\{\mathbf{d}[t], \dots, \mathbf{d}[t+k-1]\}$ die vektorisierten Folgen der Ausgangs- bzw. Störgrößenprädiktionen. Die Matrizen Ψ , Θ und Ω sind durch

$$\Psi = \begin{pmatrix} \mathbf{CA} \\ \mathbf{CA}^2 \\ \vdots \\ \mathbf{CA}^k \end{pmatrix}, \quad \Omega = \begin{pmatrix} \mathbf{CB}_d & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{CAB}_d & \mathbf{CB}_d & \ddots & \vdots \\ \vdots & \vdots & \ddots & \mathbf{0} \\ \mathbf{CA}^{k-1}\mathbf{B}_d & \mathbf{CA}^{k-2}\mathbf{B}_d & \cdots & \mathbf{CB}_d \end{pmatrix} \text{ und} \quad (6.5)$$

$$\Theta = \begin{pmatrix} \mathbf{CB}_u & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{CAB}_u & \mathbf{CB}_u & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{CA}^{k-1}\mathbf{B}_u & \mathbf{CA}^{k-2}\mathbf{B}_u & \cdots & \mathbf{CB}_u \end{pmatrix}$$

definiert. Weiterhin ist es bei der Implementierung von MPC-Reglern üblich die Stellgrößenfolge $\tilde{\mathbf{u}}[t]$ durch deren Differenzen, d. h. mittels

$$\underbrace{\begin{pmatrix} \tilde{\mathbf{u}}[t] \\ \vdots \\ \tilde{\mathbf{u}}[t+k-1] \end{pmatrix}}_{\tilde{\mathbf{u}}[t]} = \underbrace{\begin{pmatrix} \mathbf{1} & \mathbf{0} & \mathbf{0} \\ \vdots & \ddots & \mathbf{0} \\ \mathbf{1} & \vdots & \mathbf{1} \end{pmatrix}}_{\mathbf{I}_\Delta} \underbrace{\begin{pmatrix} \Delta\tilde{\mathbf{u}}[t] \\ \vdots \\ \Delta\tilde{\mathbf{u}}[t+k-1] \end{pmatrix}}_{\Delta\tilde{\mathbf{u}}[t]} + \underbrace{\begin{pmatrix} \mathbf{1} \\ \vdots \\ \mathbf{1} \end{pmatrix}}_{\mathbf{1}} \otimes \tilde{\mathbf{u}}[t-1] \quad (6.6)$$

auszudrücken, wodurch sich für die Prädiktion des Systemausgangs⁵ letztlich

$$\mathbf{y}[t] = \underbrace{\Psi\mathbf{x}[t] + \Omega\mathbf{D}[t] + \Theta\mathbf{M}[t](\mathbf{1} \otimes \tilde{\mathbf{u}}[t-1])}_{\mathbf{w}[t]} + \Theta[t]\mathbf{M}[t]\mathbf{I}_\Delta\Delta\tilde{\mathbf{u}}[t] \quad (6.7)$$

ergibt (\otimes kennzeichnet das *KRONECKER*-Produkt). Dabei ist der Vektor $\mathbf{w}[t] = \text{vec}\{w[t+1], \dots, w[t+k]\}$ in (6.7) als freie Bewegung des Systems zu interpretieren ($\Delta\tilde{\mathbf{u}} = \mathbf{0}$).

Basierend auf der eingeführten Notation wird im Folgenden auf die Formulierung des Optimierungsproblems eingegangen. Dabei ist zu beachten, dass die Raumtemperatur zwischen einem vorgegebenen Minimal- und Maximalwert \mathbf{y}_{\min} bzw. \mathbf{y}_{\max} gehalten werden soll, sodass eine Folge- bzw. Festwertregelung aus praktischer Sicht nicht sinnvoll erscheint. Es ist daher naheliegend den Komfort über die Nebenbedingungen

$$\mathbf{y}_{\min}[t] - \mathbf{1}\gamma \leq \underbrace{\mathbf{w}[t] + \Theta\mathbf{I}_\Delta\Delta\tilde{\mathbf{u}}[t]}_{\stackrel{(6.7)}{=} \mathbf{y}[t]} \leq \mathbf{y}_{\max}[t] + \mathbf{1}\gamma \quad (6.8)$$

⁵Aus Gründen der einfacheren Notation wird auf die übliche Unterscheidung des Prädiktions- und des Steuerhorizonts verzichtet. Es sei jedoch angemerkt, dass diese durch einfache Anpassung der Matrix \mathbf{I}_Δ berücksichtigt werden kann.

sicherzustellen, wobei es notwendig ist die Nebenbedingungen mit einer Schlupfvariable $\gamma \geq 0$ „aufzuweichen“ um eventuell auftretende unlösbare Optimierungsprobleme zu vermeiden (vgl. [Mac00, Abschn. 3.4]). Um weiterhin zu garantieren, dass die Ventil- und Verschattungsposition stets im zulässigen Bereich liegen, sind weitere Nebenbedingungen an die Optimierungsvariablen $\Delta\tilde{\mathbf{u}}$ notwendig:

$$\tilde{\mathbf{u}}_{\min}[t] \leq \underbrace{\mathbf{1} \otimes \tilde{\mathbf{u}}[t-1] + \mathbf{I}_{\Delta}\Delta\tilde{\mathbf{u}}[t]}_{\stackrel{(6.6)}{=} \tilde{\mathbf{u}}[t]} \leq \tilde{\mathbf{u}}_{\max}[t], \quad (6.9)$$

wobei für die Grenzwerte der Verschattungsposition

$$0 \leq \tilde{u}_1[j] \leq 1 \quad (j = 0, \dots, t-k-1) \quad (6.10)$$

gilt. Weiterhin ergibt sich aus (6.1) für $H[t] = 1$ die maximale Heiz- bzw. Kühlleistung, welche durch Einsetzen von (6.7) zu

$$\begin{aligned} \tilde{u}_{2,\max}[j] &= (\vartheta_{\text{V1}}[j] - y[j])^n \\ &= (\vartheta_{\text{V1}}[j] - w[j] - \Theta_{j,:}\mathbf{I}_{\Delta}\Delta\tilde{\mathbf{u}}[t])^n \end{aligned} \quad (6.11)$$

ausgedrückt werden kann, wobei $\Theta_{j,:}$ die j -te Zeile von Θ beschreibt. Die prädiizierte Vorlauftemperatur ϑ_{V1} kann mit der Heizkurve und der vorhergesagten Außentemperatur berechnet werden.

Die Gleichungen (6.11) sind nichtlinear in $\Delta\tilde{\mathbf{u}}[t]$, wodurch sich nichtlineare Nebenbedingungen ergeben. Um nachfolgend auf Algorithmen der linearen bzw. quadratischen Programmierung zurückgreifen zu können, sind jedoch lineare Nebenbedingungen notwendig. Da die Beziehung (6.11) aufgrund von $n = 1.1$ näherungsweise linear ist, ist eine Linearisierung entlang der optimalen Stellgrößenfolge des vorherigen Zeitschrittes $\Delta\tilde{\mathbf{u}}_-$ zweckmäßig⁶:

$$\begin{aligned} \tilde{u}_{2,\max}[j] &= (\vartheta_{\text{V1}}[j] - w[j] - \Theta_{j,:}\mathbf{I}_{\Delta}\Delta\tilde{\mathbf{u}}[t])^n \\ &\approx \tilde{u}_{2,\max}[j] \Big|_{\Delta\tilde{\mathbf{u}}_-} - n(\vartheta_{\text{V1}}[j] - w[j] - \Theta_{j,:}\mathbf{I}_{\Delta}\Delta\tilde{\mathbf{u}}_-)^{n-1} \Theta_{j,:}\mathbf{I}_{\Delta}(\Delta\tilde{\mathbf{u}} - \Delta\tilde{\mathbf{u}}_-) \end{aligned} \quad (6.12)$$

Das Kostenfunktional der Optimierung wird zu

$$V(\Delta\tilde{\mathbf{u}}, \gamma) = w_{\gamma}\gamma^2 + \Delta\tilde{\mathbf{u}}^{\top} \mathbf{W}_{\Delta\tilde{\mathbf{u}}}\Delta\tilde{\mathbf{u}} + \mathbf{w}_{\tilde{\mathbf{u}}}^{\top}(\mathbf{1} \otimes \tilde{\mathbf{u}}[t-1] + \mathbf{I}_{\Delta}\tilde{\mathbf{u}}) \quad (6.13)$$

gewählt, wobei w_{γ} , $\mathbf{W}_{\Delta\tilde{\mathbf{u}}}$ und $\mathbf{w}_{\tilde{\mathbf{u}}}$ geeignet zu wählende Wichtungparameter kennzeichnen. Dabei berücksichtigt der erste Summand von V die Abweichung von den geforderten Komfortparametern (siehe Abb. 6.6 links), wohingegen letzterer die aufgewendete Energie repräsentiert. Mit der positiv semidefiniten Wichtungsmatrix $\mathbf{W}_{\Delta\tilde{\mathbf{u}}}$, welche typischerweise als Diagonalmatrix gewählt wird, lässt sich die Reaktionsgeschwindigkeit des MPCs auf Störungen bzw. Modelabweichungen beeinflussen.

⁶Genaugenommen kennzeichnet $\Delta\tilde{\mathbf{u}}_- = \text{vec}\{\Delta\tilde{\mathbf{u}}[t], \dots, \Delta\tilde{\mathbf{u}}[t-k-2], \mathbf{0}\}$, wobei die Folgenglieder $\Delta\tilde{\mathbf{u}}$ den entsprechenden Elementen des Vektors $\Delta\tilde{\mathbf{u}}[t-1]$ entsprechen.

Das Kostenfunktional (6.13) und die (linearisierten) Nebenbedingungen (6.8) sowie (6.9) beschreiben ein quadratisches Optimierungsproblem, welches auch für relativ viele Optimierungsvariablen effizient mit Methoden der quadratischen Programmierung gelöst werden kann (siehe [Mac00, Abschn. 3.3] sowie [GMW82; BV04]). Basierend auf dem berechneten Optimum können durch Inversion von (6.1) anschließend die Ventil- und die Verschattungsposition $H[t]$ bzw. $H_F[t]$ berechnet werden.

6.1.4 Simulationsergebnisse

Der im vorigen Abschnitt beschriebene MPC-Regler wurde entsprechend der Abbildung 6.1 in einer Ein-Jahres Simulation validiert und mit einer konventionellen Regelstrategie verglichen. Dabei wurden zwei unterschiedliche Nutzungsszenarien des Raumes angenommen:

- **a) Szenario Konferenzraum:** Der Raum ist Montags bis Freitags zwischen 9:00 und 17:00 mit einer zufälligen Anzahl von 0 – 8 Personen belegt (vgl. Abb. 6.7). Die Raumbelegung wird daher als eine dem prädiktiven Regler unbekanntes, d. h. nicht prädizierbare Störgröße angenommen. Dies ist beispielsweise eine sinnvolle Annahme, wenn der Raum als Konferenzraum genutzt wird.
- **b) Szenario Büro:** Der Raum unterliegt einer näherungsweise periodischen Nutzungscharakteristik. Dabei wird der Raum Montags bis Freitags zwischen 9:00 bis 13:00 und 14:00 bis 18:00 von durchschnittlich 6 Personen genutzt, wobei jedoch die tatsächliche Belegungszahl als (näherungsweise⁷) normalverteilte Zufallsgröße mit einem Erwartungswert von 6 und einer Standardabweichung von 1.2 Personen modelliert wurde. Um die Periodizität der Störgröße zu berücksichtigen, wurde daher analog zu dem in Abschn. 6.1.2 beschriebenen vorgehen ein MPC-Modell mit saisonalen Störgrößenansatz trainiert (vgl. Abschnitt 5.1.2).

Zum Vergleich des prädiktiven Regelungsansatzes wurde zudem eine konventionelle Regelungsstrategie implementiert. Dabei wurde für die Fußbodenheizung mittels des um $\vartheta_{VI} - \vartheta_R = 15K$ und $H = 0.5$ linearisierten identifizierten Modells ein PI-Regler nach dem Betragsoptimum ausgelegt [LW19, Abschn. 10.4.4.1]. Die Abtastzeit wurde sowohl für den MPC als auch für den PI-Regler zu $T_A = 15min$ gewählt. Weiterhin wurde im Fall der konventionellen Regelstrategie die Verschattungsposition mit einem Zweipunktregler angesteuert, wobei für die Schaltpunkte $21^\circ C$ bzw. $23^\circ C$ gewählt wurde. Um dabei jedoch eine zu häufige Anpassung der Verschattungsposition zu vermeiden, wurden für beide Regelungsansätze lediglich alle 3h Stelleingriffe zugelassen⁸.

Um weiterhin die Robustheit des MPC gegenüber fehlerhaften Wettervorhersagen einschätzen zu können, wurde zum einen mit einer perfekten Wettervorhersage und zum anderen mit einem einfachen Prognosemodell für die Außentemperatur $\vartheta_{Au\ell}$ und die Globalstrahlung \dot{Q}_{Gs} simuliert. Das Prädiktionsmodell wurde mittels der im Wetterdatensatz

⁷Die Ausgabe des Zufallsgenerators wurde auf eine ganze Personenanzahl gerundet.

⁸Im Falle des prädiktiven Reglers kann dies sehr einfach durch entsprechende Wahl der Matrix \mathbf{I}_Δ realisiert werden (siehe Gl. (6.6)).

enthaltenen Zeitreihen trainiert (siehe Abschnitt 6.1.1), wobei ein einfacher S-ARMA Modellansatz mit zwei Ausgangsgrößen und einer Saisonalität von 24h verwendet wurde⁹. Weiterhin wurden für die Simulation folgende Parameter festgelegt:

- **Raumtemperaturgrenzwerte:** Montags bis Freitags zwischen 7:00 und 19:00 entsprechend der Abb. 6.6 (rechts) und $\vartheta_{R,\min} = 16^\circ\text{C}$ bzw. $\vartheta_{R,\max} = 30^\circ\text{C}$ sonst
- **Raumbelegung:** Entsprechend des obig beschriebenen Szenarien
- **Konventionelle Regelstrategie:** Zeitdiskreter PI-Regler für die Fußbodenheizung mit unterschiedlich gewähltem Aufheizzeitpunkt T_{vor} ¹⁰: Proportionalverstärkung $K_P \approx 1.294$ und Nachstellzeit $T_N \approx 7.36\text{h}$ (Betragsoptimum); Schaltpunkte des Zweipunktreglers für die Verschattung: 21°C bzw. 23°C
- **MPC-Regler:** Prädiktions- und Steuerhorizont $k = 48$ (12h); $w_\gamma = k \cdot 1000$; $\mathbf{W}_{\Delta\tilde{\mathbf{u}}} = \text{diag}(\mathbf{1} \otimes (0.001, 0.1)^\top)$; $\mathbf{w}_{\tilde{\mathbf{u}}} = \mathbf{1} \otimes (0, 1)^\top$; Löser: MATLAB `quadprog()`, `Algorithm='active-set'`

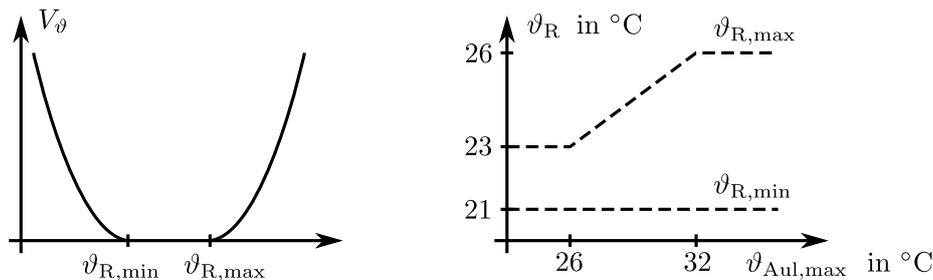


Abbildung 6.6: Komfortabhängiger Teil der Kostenfunktion (6.13) (links) und min. bzw. max. Raumtemperaturgrenzwert in Abh. der max. Tagesaußentemperatur (rechts).

Die Abbildung 6.7 zeigt beispielhaft den simulativen Vergleich der konventionellen mit der prädiktiven Regelstrategie in der Übergangszeit für das Konferenzraum-Szenario. In der Grafik lässt sich zum einen die bessere Wahl des Aufheizzeitpunktes am Morgen durch den MPC erkennen. Zum anderen ist ersichtlich, dass insbesondere durch verbesserte Regelung der Verschattungsposition, weniger Komfortverletzungen entstehen.

Die Grafik 6.8 veranschaulicht die Gesamtmenge der aufgewendeten Energie (Heiz- und Kühlenergie) sowie der Komfortverletzungen des MPC und der konventionellen Regelstrategie bei unterschiedlichen Raum-Vorhaltezeiten T_{vor} . Dabei ist für beide Szenarien, sowohl bei perfekter Wettervorhersage als auch bei der Verwendung der Wetterprognosemodelle, eine deutliche Überlegenheit des prädiktiven Reglers hinsichtlich

⁹Die Prognosequalität der Wettermodelle ist als relativ schlecht zu beurteilen, sodass die in Abb. 6.8 dargestellten Ergebnisse als tendenziell ungünstiger Fall zu beurteilen sind. Insbesondere sollte in der Realität, beispielsweise durch Verwendung von Online-Vorhersagediensten, eine tendenziell bessere Prognosequalität zu erwarten sein.

¹⁰Da die Fußbodenheizung eine große Trägheit besitzt, ist es für einen sinnvollen Vergleich notwendig den Sollwert des PI-Reglers am Morgen um eine durch T_{vor} bestimmte Zeit „vorzuziehen“ (Vorhaltezeit), damit der Raumtemperatursollwert um 7:00 erreicht werden kann.

des Raumkomforts erkennbar. Bei einem vergleichbaren Energieverbrauch werden durch den MPC für das Büro- etwa 33% und für das Konferenzraum-Szenario $\approx 27\%$ weniger Komfortverletzungen verursacht, wobei die Verbesserung im Szenario b) auf die Berücksichtigung der Raumbelegung durch das saisonale Störgrößenmodell zurückzuführen ist. Der Komfortverlust bei der Nutzung der Wetterprognosemodelle lässt sich in beiden Fällen auf etwa 6% beziffern.

Bemerkung 6.3: Die mittlere Dauer zur Lösung des QP mit 53 Optimierungsvariablen und 201 Nebenbedingungen betrug auf einem Windows Laptop mit Intel Core i7-3520M CPU mit 2x2.9GHz und 8GB Arbeitsspeicher etwa 12ms. \triangleleft

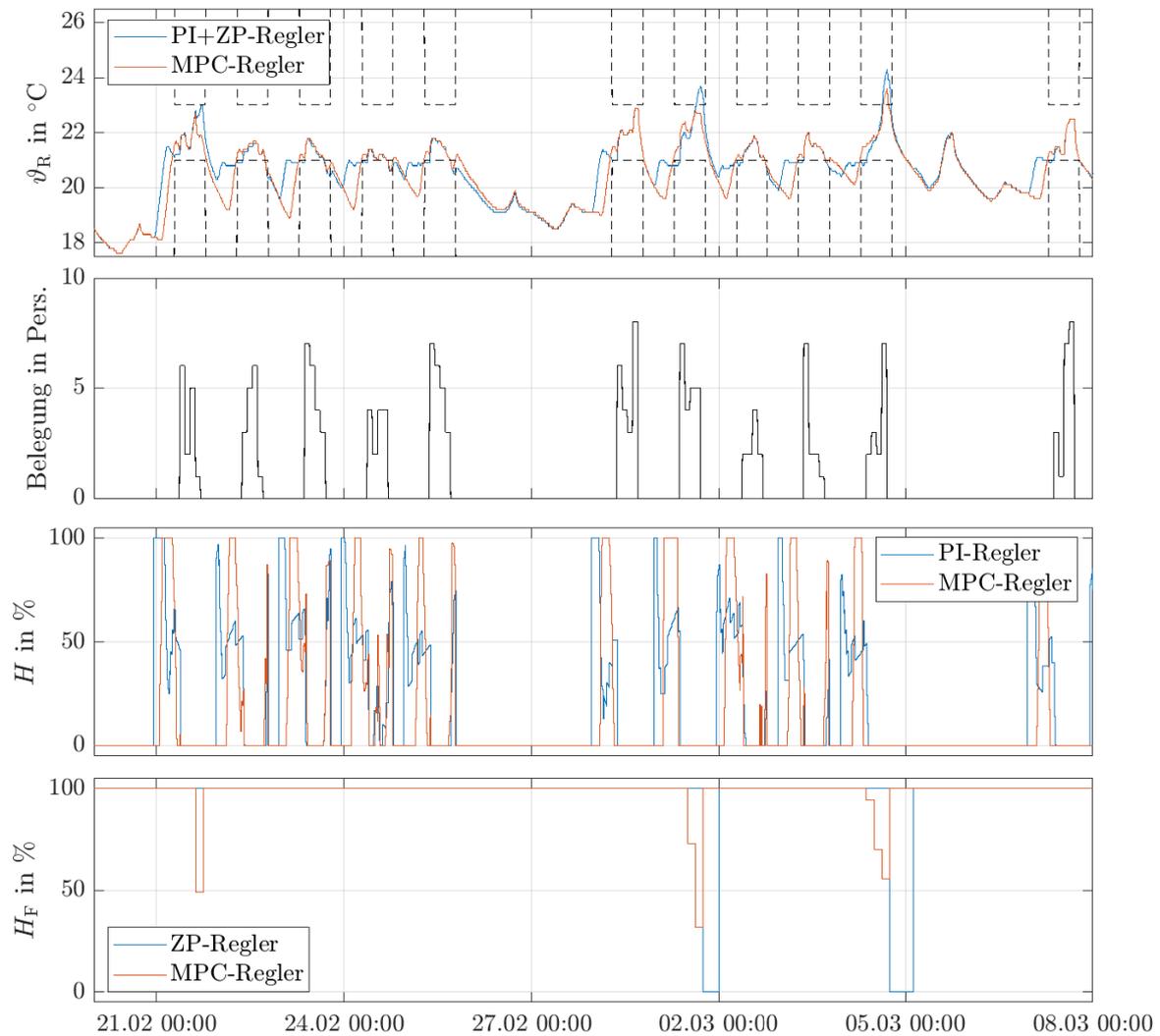


Abbildung 6.7: Vergleich des Regelverhaltens der prädiktiven mit der konventionellen Regelstrategie mit der Vorhaltezeit $T_{\text{vor}} = 8h$ „vorgezogen“ in der Übergangszeit für das Konferenzraum-Szenario.

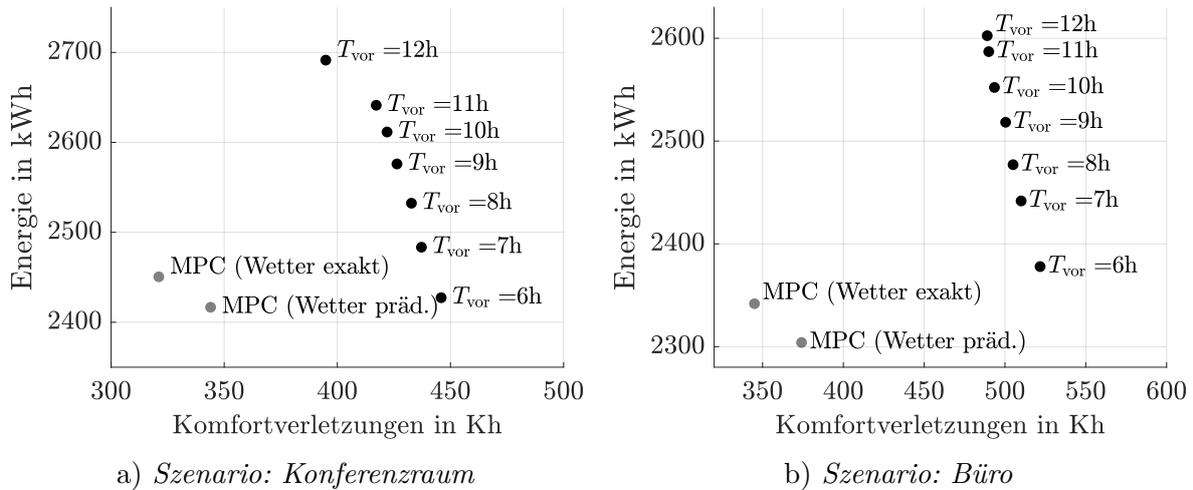


Abbildung 6.8: Jahresvergleich der aufgewendeten Energie und der Komfortverletzungen des MPC (•) mit der konventionellen Regelstrategie (●) bei unterschiedlichen Raum-Vorhaltezeiten T_{vor} .

6.2 Modellprädiktiver Mehrzonenregler

Der Betrieb von Lüftungs- bzw. Klimaanlage stellt bei der Betriebsführung von Gebäuden einen oftmals erheblichen Kostenfaktor dar. Hierbei verursacht die elektrische Leistungsaufnahme der Lüfter, die bei konstanter Druckdifferenz theoretisch kubisch¹¹ vom geförderten Massenstrom abhängig ist [RSS11, Abschn. 3.3.1], einen nicht zu vernachlässigenden, oftmals sogar dominierenden, Betriebskostenanteil.

Wird beim Betrieb des Gebäudes von einer Tag-/Nacht- und/oder Wochenendabsenkung der Sollwerte wie bspw. im Abschnitt 6.1 ausgegangen, dann wird in der Praxis oft der Ansatz verfolgt, das Gebäude in der Nacht und am Wochenende auskühlen zu lassen und kurz vor den Betriebszeiten mittels eines zügigen Luftaustausches wieder aufzuheizen. Aufgrund der dadurch verhältnismäßig großen Massenströme läuft die Lüftungsanlage dann in einem gewissen Zeitraum unter hoher Last, was insbesondere aufgrund der erwähnten nichtlinearen Abhängigkeit der Leistungsaufnahme des Ventilators eine nicht optimale Betriebsführung darstellen kann.

Im Folgenden wird daher beispielhaft die Implementierung eines prädiktiven Mehrzonenreglers für ein Hörsaalgebäude beschrieben, welches prinzipiell an das Anwendungsbeispiel des Abschnittes 5.1.4 angelehnt ist. Dabei wird davon ausgegangen, dass die 4 Hörsäle des Gebäudes durch eine zentrale druckgeregelte¹² Lüftungsanlage versorgt werden, sodass durch die Regelung nur eine gemeinsame Zulufttemperatur für

¹¹Bei der Untersuchung praktisch aufgenommener Messdaten hat sich jedoch gezeigt, dass für die Charakterisierung der Kennlinie ein quadratischer Zusammenhang ausreichend ist (vgl. auch [KB11]).

¹²In einer druckgeregelten Lüftungsanlage wird durch entsprechende Ansteuerung der Frequenzumrichter der Zuluft- und Abluftventilatoren in der Lüftungsanlage ein i. d. R. konstanter Zuluft- und Abluftdruck im Kanalnetz aufrechterhalten. Die Versorgung der Zonen wird dann über lokal installierte variable Volumenstromregler sichergestellt.

alle Hörsäle festgelegt werden kann. Der Raumkomfort in den jeweiligen Zonen wird dann durch lokal installierte variable Volumenstromregler sichergestellt. Als Stellgrößen sind daher zum einen die Zulufttemperatur der Lüftungsanlage und zum anderen die Massen- bzw. Volumenströme in die Hörsäle zu betrachten¹³. Ziel der Regelung ist wiederum die Einhaltung der Raumtemperaturgrenzwerte in den jeweiligen Hörsälen, wobei unter Berücksichtigung der nichtlinearen Ventilator Kennlinie und durch geschickte Vortemperierung der thermischen Gebäudemasse im Absenkbetrieb eine Minderung der Gesamtkosten für die Betriebsführung erzielt werden soll. Eine schematische Darstellung des betrachteten Szenarios ist in Abbildung 6.9 veranschaulicht.

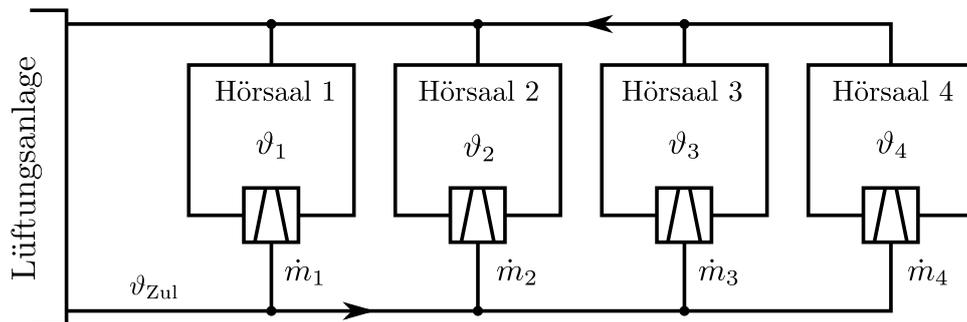


Abbildung 6.9: Schematische Darstellung des Hörsaalgebäudes mit zentraler Lüftungsanlage

6.2.1 Prozessmodell

Das Modell des Gebäudes wurde wiederum mittels der MODELICA BUILDINGS Bibliothek entworfen, wobei als Wetterdatensatz wieder das Testreferenzjahr des Standortes Berlin verwendet wurde. Der Aufbau der Wände und der Grundriss des Gebäudes sowie die schematische Darstellung des MODELICA Modells sind in der Abbildung 6.10 bzw. 6.11 dargestellt. Für die Materialparameter wurden die in der BUILDINGS Bibliothek hinterlegten Standardwerte verwendet. Die Raum- bzw. Fensterhöhe der Hörsäle beträgt 8m bzw. 3m. Für die Temperatur des Erdreiches, die Quantisierung der Temperatursensoren (0.1K) sowie für die thermische Last der extern vorzugebenden Raumbelegung (100W/Person) gelten die in Abschnitt 6.1.1 genannten Annahmen. Die variablen Volumenstromregler der Hörsäle wurden als idealisierte Massenstromquellen mit vorzugebender Temperatur modelliert (siehe Abb. 6.11). Der Luftaustausch durch bspw. Undichtigkeiten oder Türen wurde nicht modelliert. Das Gesamtmodell besitzt 137 Zustandsgrößen.

¹³Da der durch die Lüfter der Lüftungsanlage aufgebaute Druck in der Regel sehr gering ist (typischerweise im Bereich einiger Hundert Pascal) und der relevante Temperaturbereich relativ gering ist, kann die Dichte der Luft als näherungsweise konstant betrachtet werden kann (vgl. Anhang C.2).

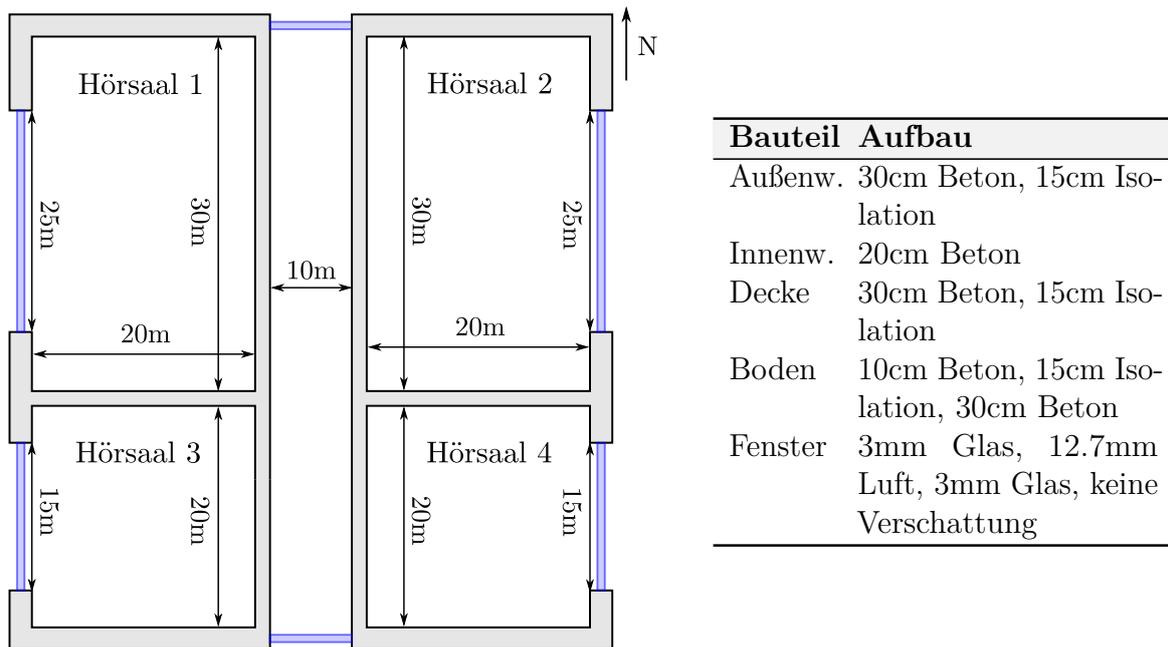


Abbildung 6.10: Grundriss und Materialparameter des Hörsaalgebäudes.

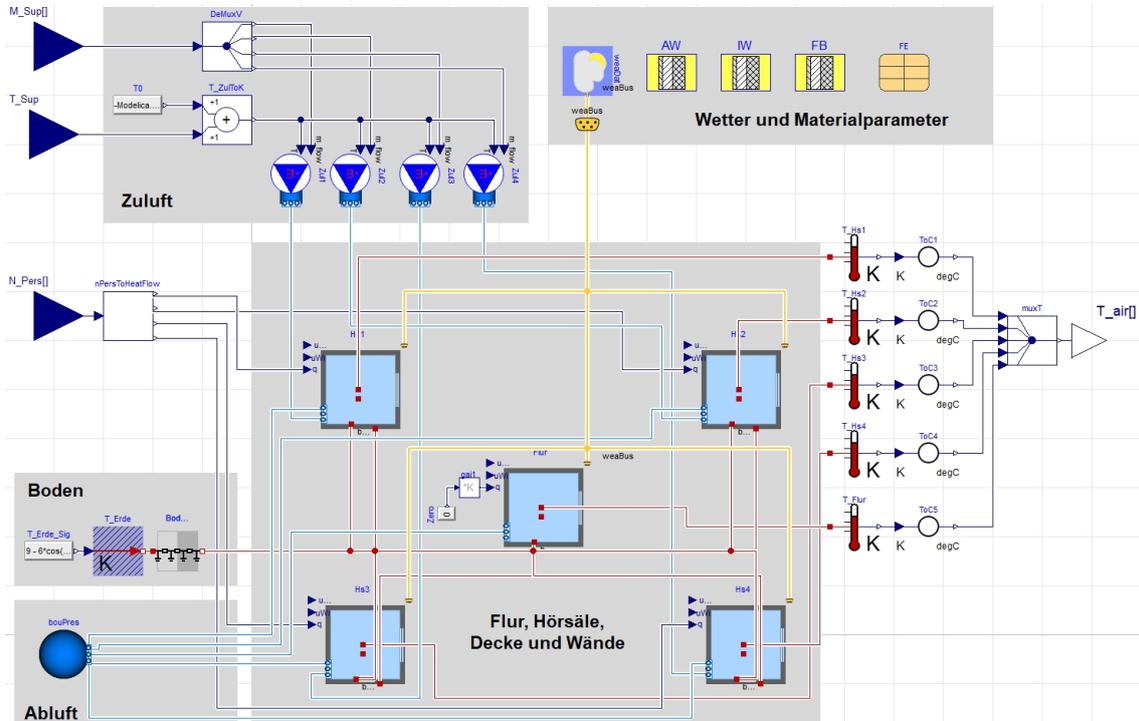


Abbildung 6.11: Grafische Darstellung des Modelica Prozessmodells des Gebäudes mit vier Hörsälen und einem Flur.

6.2.2 Identifikation

Beschreibung des Datensatzes: Der für die Identifikation verwendete Datensatz wurde durch Simulation des Prozessmodells über 12 Wochen erzeugt, wobei die Stellgrößen (Zuluftmassenströme und Zulufttemperatur) durch sprunghöhenartige Signale mit zufällig gewählter Sprunghöhe angeregt wurden (siehe Abb. 6.13). Die Abtastzeit wurde zu $T_A = 60\text{min}$ gewählt ($N = 2017$).

Bemerkung 6.4 (Wahl der Abtastzeit): Es wurde eine Abtastzeit von $T_A = 60\text{min}$ gewählt, weil der prädiktive Regler zur Vortemperierung der thermischen Gebäudemasse im Absenkbetrieb eingesetzt werden soll und daher insbesondere das träge Systemverhalten von Interesse ist. In der Praxis sind im Nutzungsbetrieb aufgrund von Belegungswechseln jedoch schnell veränderliche interne Gewinne in den Zonen und daher aufgrund der vergleichsweise geringen Wärmekapazität der Raumluft relativ schnelle Änderungen der Regelgrößen zu erwarten. Im Nutzungsbetrieb muss daher eine deutlich geringere Abtastzeit, bspw. 5min, verwendet werden. \triangleleft

Modellansatz: Der folgend dargestellte prädiktive Regelungsansatz wurde basierend auf zwei unterschiedliche Modellansätzen des Typs (2.51) getestet. Ersterer entspricht wiederum der in Kapitel 5 beschriebenen Methode, d. h. es wird ein separates MISO-Einzelraummodell mit

$$\mathbf{f}_i(\mathbf{u}, y_i) = \begin{pmatrix} \vartheta_{\text{Aul}}[t] \\ \dot{Q}_{\text{gen},i}[t] \\ \dot{m}_i[t](\vartheta_{\text{Zul}}[t] - \vartheta_i[t]) \end{pmatrix} \quad (i = 1, \dots, 4) \quad (6.14)$$

für jeden der vier Hörsäle identifiziert ($y_i[t] := \vartheta_i[t]$), wobei die solaren Gewinne $\dot{Q}_{\text{gen},i}[t]$ mittels (5.7) berechnet wurden. Die Wärmeströme aus den Nachbarräumen wurden wiederum als nicht prädizierbare tieffrequente Störung berücksichtigt. Um insbesondere die thermische Trägheit des Gebäudes zu erfassen, wurde die Zeitkonstante des Hochpassvorfilters (5.10) mit $T_L = 7\text{d}$ angesetzt. Für das dynamische LTI-Teilsystem wurde für alle Hörsaalmodelle die ARMAX-Parametrierung mit den Polynomordnungen $n_A = n_{B_j} = 2$ und $n_C = 1$ verwendet ($j = 1, \dots, 3$).

Der zweite Modellansatz hingegen beruht auf einem verkoppelten Mehrzonenmodell, wobei für den LTI-Block eine MIMO ARMAX-Parametrierung

$$\mathbf{A}(q^{-1})\mathbf{y}[t] = \mathbf{B}(q^{-1})\mathbf{f}(\mathbf{u}, \mathbf{y}) + \mathbf{C}(q^{-1})\mathbf{e}[t] \quad (6.15)$$

mit $\mathbf{y}[t] = (\vartheta_1[t], \dots, \vartheta_5[t])^\top$ und $\mathbf{f}(\mathbf{u}, \mathbf{y}) = \text{vec}\{\mathbf{f}_1(\mathbf{u}, y_1), \dots, \mathbf{f}_4(\mathbf{u}, y_4), \mathbf{f}_5(\mathbf{u})\}$ verwendet wurde¹⁴. Aufgrund der Gebäudestruktur wurde für die Polynommatrizen

$$\mathbf{A}(q^{-1}) = \begin{pmatrix} A_{11}(q^{-1}) & 0 & A_{13}(q^{-1}) & 0 & A_{15}(q^{-1}) \\ 0 & A_{22}(q^{-1}) & 0 & A_{24}(q^{-1}) & A_{25}(q^{-1}) \\ A_{31}(q^{-1}) & 0 & A_{33}(q^{-1}) & 0 & A_{35}(q^{-1}) \\ 0 & A_{42}(q^{-1}) & 0 & A_{44}(q^{-1}) & A_{45}(q^{-1}) \\ A_{51}(q^{-1}) & A_{52}(q^{-1}) & A_{53}(q^{-1}) & A_{54}(q^{-1}) & A_{55}(q^{-1}) \end{pmatrix} \quad (6.16)$$

¹⁴Die Indizes 1 bis 4 beziehen sich auf die Hörsäle 1 bis 4, wohingegen der Index 5 den Flur des Gebäudes kennzeichnet (siehe Abbildungen 6.9 bis 6.11). Da angenommen wird, dass der Flur nicht an die Lüftungsanlage angeschlossen ist, gilt $\mathbf{f}_5(\mathbf{u}) = (\vartheta_{\text{Aul}}[t], \dot{Q}_{\text{gen},5}[t])^\top$.

und

$$\begin{aligned} \mathbf{B}(q^{-1}) &= \text{blkdiag}(\mathbf{B}_1(q^{-1}), \dots, \mathbf{B}_5(q^{-1})) \quad \text{und} \\ \mathbf{C}(q^{-1}) &= \text{diag}(C_1(q^{-1}), \dots, C_5(q^{-1})) \end{aligned} \quad (6.17)$$

angesetzt, wobei die Symmetrie von $\mathbf{A}(q^{-1})$ ($A_{ij}(q^{-1}) = A_{ji}(q^{-1})$) durch Gleichheitsnebenbedingungen für die entsprechenden Parameter erzwungen wurde (vgl. Abb. 6.10). Für die Hauptdiagonalpolynome der Matrix $\mathbf{A}(q^{-1})$ sowie für alle Polynome von $\mathbf{B}(q^{-1})$ wurde die Ordnung 2 verwendet, wohingegen für die übrigen Polynome die Ordnung 1 verwendet wurde.

Identifikation: Sowohl der MISO- als auch der MIMO-Modellansatz wurde mittels 1-Schritt-PEM mit dem Gütekriterium (3.7) identifiziert. Dabei wurden beim MISO-Modellansatz für jeden Hörsaal jeweils eine schnelle Zeitkonstante für das dynamische Verhalten der Raumluft ($\approx 1.5 - 1.7\text{h}$) und eine langsame Zeitkonstante für die Dynamik der trägen Gebäudemasse ($\approx 6 - 8$ Tage) identifiziert. Zur Lösung der Optimierungsaufgabe wurde das Levenberg-Marquardt- (MISO-Modelle) bzw. das SQP-Verfahren (MIMO-Modell) eingesetzt.

Diskussion: Die Grafiken 6.12 veranschaulichen die Box-Whisker-Diagramme der Prädiktionsfehler und die Autokovarianzfolgen der Residuen für den MISO- sowie den MIMO-Modellansatz. In den Darstellungen lässt sich erkennen, dass sich die Prädiktionsgüte der beiden Modellierungsansätze für kurze Vorhersagehorizonte, und entsprechend auch die Autokovarianzfolgen, kaum unterscheiden. Weiterhin erkennt man, dass das MIMO-Modell für längere Prädiktionshorizonte eine bessere Prädiktionsgüte besitzt. Dies kann damit begründet werden, dass durch die verkoppelte Struktur des MIMO-Modellansatzes eine genauere Beschreibung des dynamischen Verhaltens der trägen Gebäudemasse möglich ist.

6.2.3 Reglerentwurf

Basierend auf den beiden Modellansätzen wird im Folgenden wiederum die Implementierung eines MPC-Reglers für das Anwendungsbeispiel beschrieben. Dabei werden die Stell-, Regel- und (präzifizierbaren) Störgrößen zu

$$\mathbf{u}[t] := \begin{pmatrix} \vartheta_{\text{Zul}}[t] \\ \dot{m}_1[t] \\ \vdots \\ \dot{m}_4[t] \end{pmatrix}, \quad \mathbf{y}[t] := \begin{pmatrix} \vartheta_1[t] \\ \vdots \\ \vartheta_4[t] \end{pmatrix} \quad \text{bzw.} \quad \mathbf{d}[t] := \begin{pmatrix} \vartheta_{\text{Aul}}[t] \\ \dot{Q}_{\text{Gs}}[t] \\ \alpha_{\text{Saz}}[t] \\ \alpha_{\text{Sh}}[t] \end{pmatrix} \quad (6.18)$$

festgelegt¹⁵. Die Belegung der Hörsäle wurde als nicht präzifizierbare Störung zu betrachten. Die Vektoren \mathbf{u} , \mathbf{y} und \mathbf{d} kennzeichnen wiederum die vektorisierten Folgen der Stell-, Regel- bzw. Störgrößen.

¹⁵Mittels der Globalstrahlung \dot{Q}_{Gs} und der Sonnenposition ($\alpha_{\text{Saz}}, \alpha_{\text{Sh}}$) sowie der Ausrichtung der Fensterflächen wurden die solaren Erträge $\dot{Q}_{\text{gen},i}[t]$ in jedem Hörsaal prognostiziert (siehe Abschn. 5.1).

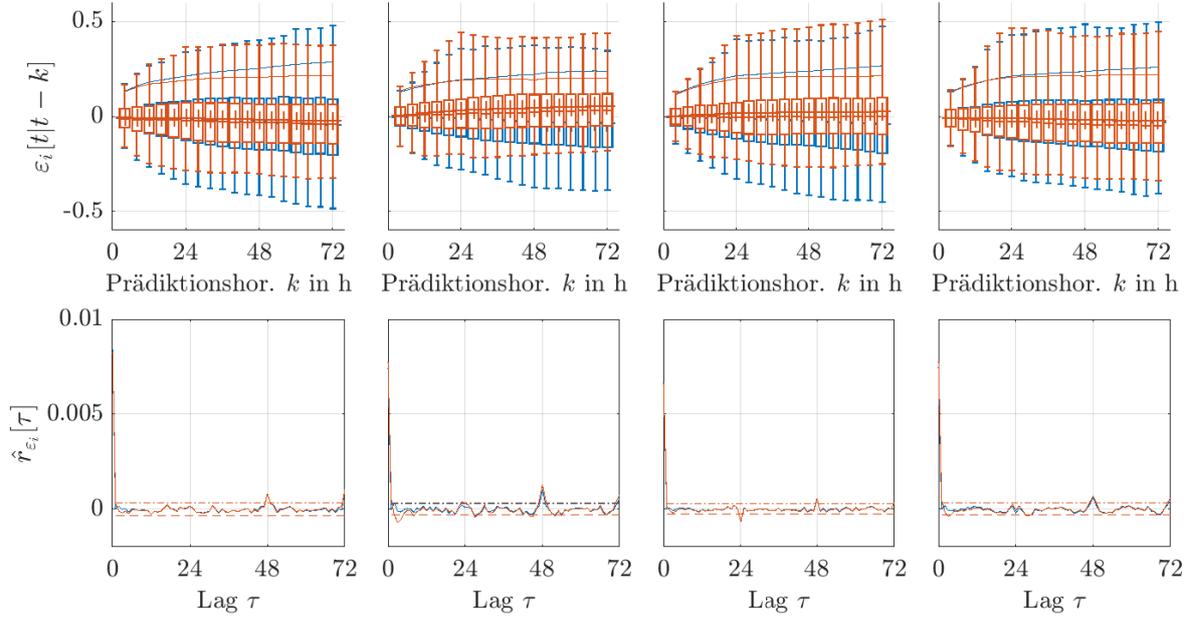


Abbildung 6.12: Box-Whisker-Diagramme der Prädiktionsfehler $\varepsilon_i[t|t-k]$ (oben) und Autokovarianzfolgen $\hat{r}_{\varepsilon_i}[\tau]$ der Residuen $\varepsilon_i[t|t-1]$ (unten) für die Hörsäle $i = 1, \dots, 4$ (von links nach rechts). Die Diagramme sind für den MISO- (–) und den MIMO-Modellansatz (–) dargestellt. In den Grafiken oben sind die empirischen 5, 25, 50, 75 und 95% Perzentile sowie der Mittelwert (+) und die Standardabweichung (–) markiert. Die gestrichelten Linien in den unteren Grafiken kennzeichnen den kritischen Bereich zum Signifikanzniveau $\alpha = 0.01$ (siehe Bemerkung 5.2).

Wie bereits in der Einleitung des Abschnittes beschrieben wurde, besteht das Ziel des MPC-Reglers darin durch geschickte Vortemperierung der thermischen Gebäudemasse die Betriebskosten, d. h. die Summe der Heiz-, Kühl- und Elektroenergiekosten, zu senken. Das Kostenfunktional der Optimierung zum Zeitpunkt τ wird daher zu

$$\begin{aligned}
 V(\mathbf{u}, \gamma, \tau) &= V_{\text{HK}}(\mathbf{u}) + V_{\text{El}}(\mathbf{u}) + V_{\vartheta}(\gamma) \\
 &= T_A \sum_{t=\tau}^{\tau+k-1} w_{\text{HK}c_L} (1 - \eta_{\text{Wrg}}) \dot{m}_{\text{Zul}}[t] \left| \vartheta_{\text{Zul}}[t] - \vartheta_{\text{Aul}}[t] \right| + w_{\text{El}} \kappa_{\text{El}} \dot{m}_{\text{Zul}}^2[t] \\
 &\quad + \sum_{i=1}^4 w_{\gamma} \gamma_i^2
 \end{aligned} \tag{6.19}$$

angesetzt, wobei V_{HK} und V_{El} den Betriebskosten (Heiz-, Kühl- und Elektroenergie) und V_{ϑ} den vom Komfort abhängigen Teil des Kriteriums darstellen. Weiterhin kennzeichnen c_L [kJ/kgK] die spezifische Wärmekapazität der Luft, w_{HK} und w_{El} [€/kWh] die Heiz-, Kühl- und Elektroenergiekosten, η_{Wrg} die Effizienz der Wärmerückgewinnung der Lüftungsanlage, κ_{El} [kW/(kg/s)²] den Anlagenkoeffizienten des Lüfters und $\dot{m}_{\text{Zul}}[t] = \sum_{i=1}^4 \dot{m}_i[t]$ den durch die Lüftungsanlage geförderten Gesamtmassenstrom. Die Komfortbedingungen in jedem Hörsaal werden wiederum durch Nebenbedingungen

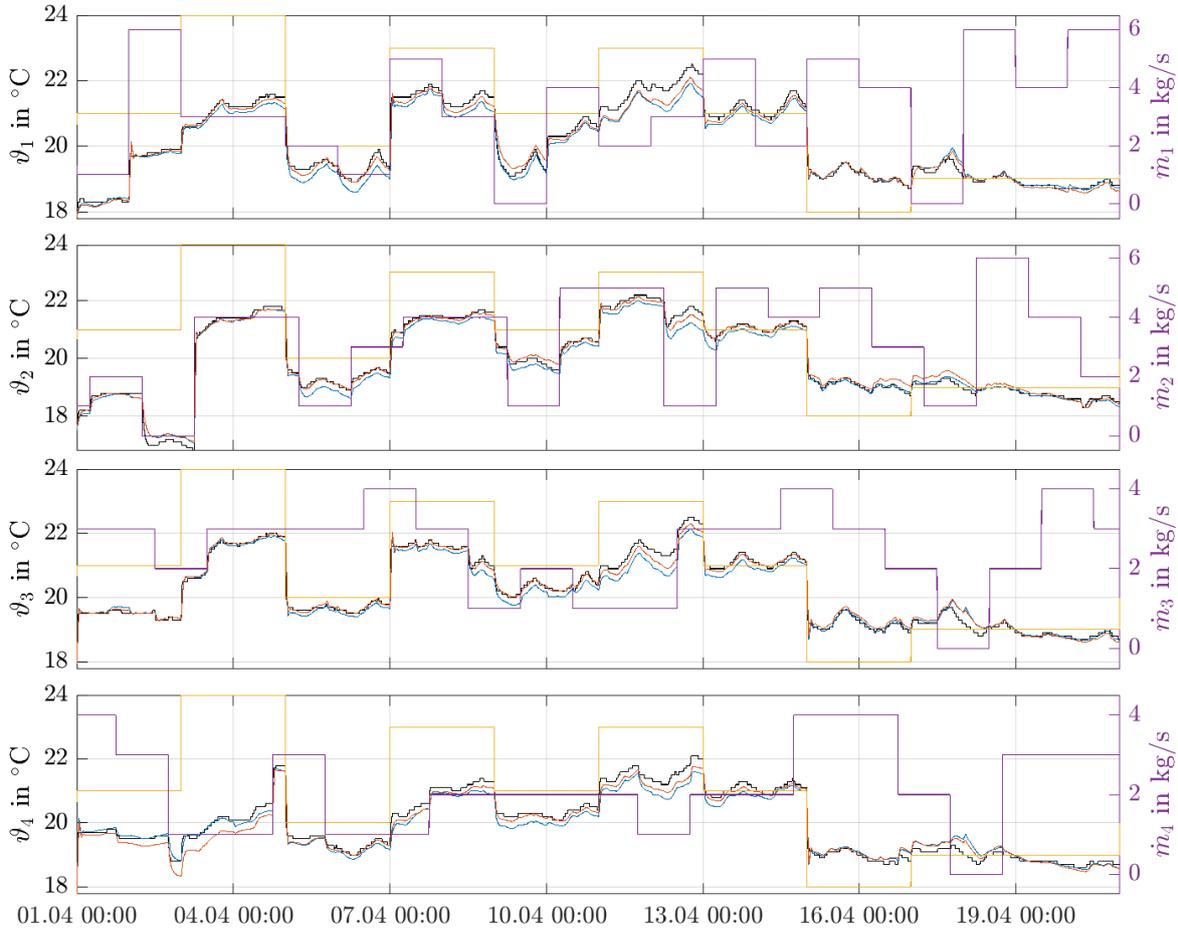


Abbildung 6.13: Beispielhafter Vergleich der 72-Schritt Prädiktion (72h) der identifizierten MISO- (–) bzw. des MIMO-Modells (–) mit der Simulation des Modelica Prozessmodells (–) für die 4 Hörsäle (von oben nach unten). In den Grafiken sind zudem die Zulufttemperatur ϑ_{Zul} (–) und die Zuluftmassenströme (–) dargestellt.

$$\mathbf{y}_{\min,i}[\tau] - \gamma_i \leq \mathbf{y}_i[\tau] \leq \mathbf{y}_{\max,i}[\tau] + \gamma_i \quad (i = 1, \dots, 4) \quad (6.20)$$

sichergestellt, wobei die Elemente des Vektors $\boldsymbol{\gamma}$ Schlupfvariablen, die zur Gewährleistung der Lösbarkeit des Optimierungsproblems benötigt werden, darstellen. Die Stellgrößenbeschränkungen können durch lineare Nebenbedingungen

$$\mathbf{u}_{\min}[\tau] \leq \mathbf{u}[\tau] \leq \mathbf{u}_{\max}[\tau] \quad (6.21)$$

ausgedrückt werden, wobei sich die Zeitabhängigkeit von \mathbf{u}_{\min} bzw. \mathbf{u}_{\max} durch unterschiedliche Zulufttemperaturgrenzwerte während und außerhalb der Nutzungszeiten des Gebäudes ergibt¹⁶.

¹⁶In der Praxis sind während der Nichtnutzungszeit typischerweise höhere (Heizen) bzw. geringere (Kühlen) Zulufttemperaturen möglich, wohingegen während der Belegungszeit aufgrund möglicher Zugserscheinungen strengere Vorgaben gelten. Konkrete Grenzwerte sind u. a. von der gewünschten Raumtemperatur, sowie von der Ausführung der Luftauslässe abhängig.

Vergleicht man die Struktur des obig beschriebenen Optimierungsproblems mit der des im vorigen Abschnitt beschriebenen prädiktiven Einzelraumreglers, dann sind insbesondere zwei Unterschiede hervorzuheben: Zum einen wird in diesem Beispiel die Zulufttemperatur (vgl. Vorlauftemperatur) nicht als übergeordnet festgelegte Systemgröße, sondern als eine zusätzliche Optimierungsvariable, betrachtet. Zum anderen ist zu berücksichtigen, dass die Hilfsenergie für den Stofftransport des Heizmediums (V_{El}) explizit im Kostenfunktional berücksichtigt wird. Eine Minimierung der Heiz- bzw. Kühlenergie, basierend auf den Eingangsgrößen des LTI-Blocks des MPC-Modells wie im vorigen Beispiel, ist insbesondere aufgrund der Abhängigkeit der Elektroenergiekosten vom geförderten Gesamtmassenstrom \dot{m}_{Zul} nicht möglich (vgl. Abschn. 6.1). Eine Darstellung als quadratisches Programm ist daher nicht möglich, sodass das Optimierungsproblem (6.19) mit den Nebenbedingungen (6.20) und (6.21) im Folgenden als nichtlineares Programm (NLP) betrachtet wird.

Bei der Formulierung des Optimierungsproblems wurde wiederum von der Zustandsraumdarstellung

$$\begin{aligned} \mathbf{x}[t+1] &= \mathbf{A}\mathbf{x}[t] + \mathbf{B}_u \mathbf{f}_u(\mathbf{u}[t], \mathbf{y}[t]) + \mathbf{B}_d \mathbf{f}_d(\mathbf{d}[t]) \\ \mathbf{y}[t] &= \mathbf{C}\mathbf{x}[t], \end{aligned} \quad (6.22)$$

des MPC-Modells ausgegangen. Wie bereits oben beschrieben, ergibt sich aufgrund der nichtlinearen Modellstruktur ein NLP, sodass auf Algorithmen der nichtlinearen Programmierung zurückgegriffen werden muss (siehe bspw. [GMW82, Kap. 6]). Für eine effiziente Lösung des Optimierungsproblems ist es sinnvoll dem Löser die Gradienten des Kostenfunktionals (6.19) sowie der nichtlinearen Nebenbedingungen (6.20) bereitzustellen. Hierfür werden insbesondere die Ableitungen der Ausgangsgrößen \mathbf{y} nach den Stellgrößen \mathbf{u} benötigt, welche durch Differentiation von (6.22) gewonnen werden können:

$$\begin{aligned} \frac{\partial \mathbf{x}[t+1]}{\partial \mathbf{u}[s]} &= \begin{cases} \mathbf{A} \frac{\partial \mathbf{x}[t]}{\partial \mathbf{u}[s]} + \mathbf{B}_u \left(\frac{\partial \mathbf{f}_u(\mathbf{u}[t], \mathbf{y}[t])}{\partial \mathbf{y}[t]} \frac{\partial \mathbf{y}[t]}{\partial \mathbf{u}[s]} + \underbrace{\frac{\partial \mathbf{f}_u(\mathbf{u}[t], \mathbf{y}[t])}{\partial \mathbf{u}[s]}}_{\text{falls } s=t, \text{ sonst } \mathbf{0}} \right) & \forall s \leq t \\ \mathbf{0} & \text{sonst} \end{cases} \\ \frac{\partial \mathbf{y}[t]}{\partial \mathbf{u}[s]} &= \mathbf{C} \frac{\partial \mathbf{x}[t]}{\partial \mathbf{u}[s]} \end{aligned} \quad (6.23)$$

6.2.4 Simulationsergebnisse

Der obig beschriebene MPC-Regler wurde entsprechend der Abbildung 6.1 in der Simulationsumgebung MATLAB/SIMULINK implementiert und in einer Ein-Jahres-Simulation validiert. Zum Vergleich wurde eine konventionelle Regelstrategie herangezogen, die auf der Regelung der Hörsaaltemperaturen mittels der Zuluftmassenströme durch lokale PI-Regler basiert. Dabei wurde die Zulufttemperatur der Lüftungsanlage übergeordnet

durch einen Zweipunktregler

$$\vartheta_{\text{Zul}}[t] = \begin{cases} \vartheta_{\text{Zul,min}}[t], & \text{falls RLT-Kühlen} = 1 \text{ und sonst} \\ \vartheta_{\text{Zul,max}}[t] \end{cases} \quad (6.24)$$

bestimmt, wobei der Betriebsmodus (RLT-Kühlen) entsprechend des in der Abbildung 6.14 dargestellten Regelschemas festgelegt wurde. Die schematische Darstellung der konventionellen Regelstrategie ist in der Grafik 6.14 veranschaulicht, wobei für die Abtastzeit $T_A = 5\text{min}$ gewählt wurde.

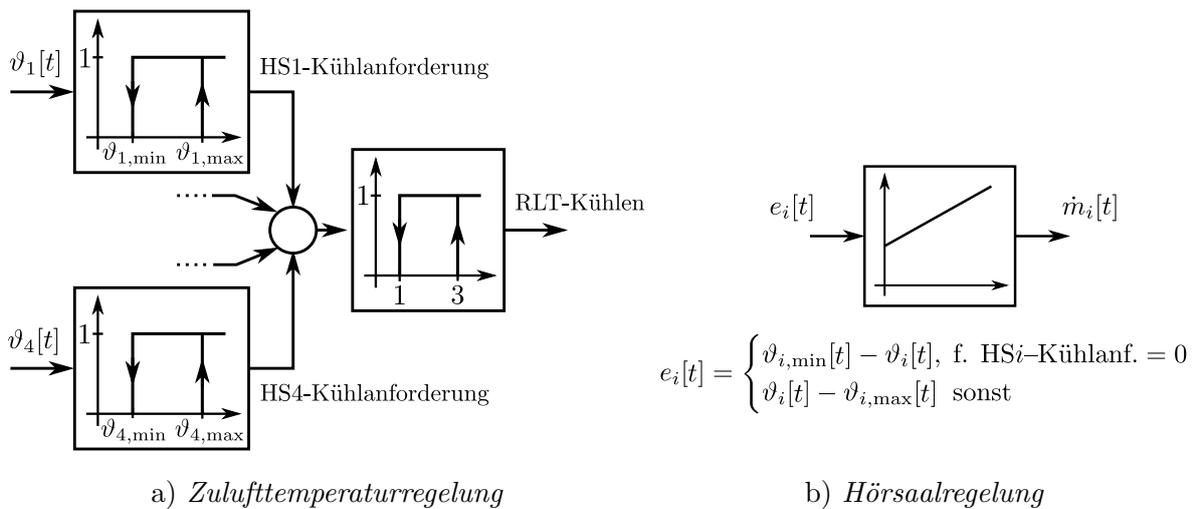


Abbildung 6.14: Referenzregelstrategie für das Hörsaalgebäude.

Ergänzend zur beschriebenen Referenzregelstrategie wurde zum Zweck der Kostenreduktion ein MPC-Regler mit der Abtastzeit $T_A = 60\text{min}$ und einem Prädiktions- und Steuerhorizont von 4 Tagen implementiert ($k = 96$). Die Optimierung wurde jedoch nur während der Absenkezeit zur Vortemperierung des Gebäudes durchgeführt, wohingegen in den Nutzungszeiten auf die Referenzregelstrategie umgeschaltet wurde¹⁷. Für die Simulation wurden weiterhin folgende Parameter verwendet ($i = 1, \dots, 4$):

- **Temperaturgrenzwerte der Hörsäle:** Montags bis Freitags zwischen 8:00 und 20:00 (Nutzungszeit) entsprechend der Abb. 6.6 (rechts) sowie $\vartheta_{i,\text{min}} = 16^\circ\text{C}$ und $\vartheta_{i,\text{max}} = 30^\circ\text{C}$ sonst (Nichtnutzungszeit)
- **Stellgrößenbeschränkungen:** $\dot{m}_{i,\text{max}} = 6$ bzw. 4kg/s (Hörsäle 1/2 bzw. 3/4);

¹⁷Wie bereits in der Bemerkung 6.4 angedeutet, sind im Nutzungsbetrieb schnell veränderliche interne Gewinne in den Zonen und daher relativ schnelle Änderungen der Regelgrößen zu erwarten, weswegen in der Nutzungszeit eine deutlich kleinere Abtastzeit verwendet werden muss. Da der MPC-Regler während dieser Zeit i. d. R. ohnehin versucht die Hörsaaltemperaturen auf deren minimalen bzw. maximalen Grenzwerten zu halten, was im Wesentlichen der konventionellen Regelungsstrategie entspricht (vgl. Abb. 6.17 und 6.18), wurde zum Zweck der Verkürzung der Simulationszeit auf die Referenzregelstrategie mit $T_A = 5\text{min}$ umgeschaltet.

$\vartheta_{\text{Zul,min}}[t] = 18^\circ\text{C}/\min(14^\circ\text{C}, \vartheta_{\text{Aul}}[t])$ ¹⁸ und $\vartheta_{\text{Zul,max}}[t] = 26/30^\circ\text{C}$ (Nutzungszeit/Nichtnutzungszeit)

- **Raumbelegung:** Innerhalb einer Woche wurde in jedem Vorlesungsblock von 90min und für jeden Hörsaal eine zufällige Anzahl an Personen vorgegeben, d. h. die Personenanzahl $N_i(t)$ wurde als gleichverteilte Zufallsgröße im Intervall $[0, N_{i,\text{max}}]$ angenommen ($N_{i,\text{max}} = 300/350/250/200$). Das Belegungsmuster wurde dann für jeden Hörsaal wöchentlich wiederholt.
- **Referenzregelstrategie:** Empirisch parametrisierte diskrete PI-Regler für die Temperaturregelung in den Hörsälen mit Proportionalverstärkung $K_P = 1\text{kgs}^{-1}/\text{K}$ und Integralverstärkung $K_I = 0.1\text{kgs}^{-1}/\text{Kmin}$. Weiterhin wurde der Sollwert für ein rechtzeitiges Aufheizen/Abkühlen der Hörsäle um die Vorhaltezeit $T_{\text{vor}} = 30\text{min}$ ¹⁹ „vorgezogen“.
- **MPC:** Prädiktions- und Steuerhorizont $k = 96$ (4d); $\eta_{\text{Wrg}} = 0.75$; $w_{\text{HK}} = 0.1\text{€}/\text{kWh}$; $w_{\text{El}} = 0.3\text{€}/\text{kWh}$; $w_\gamma = k \cdot 1000\text{€}/\text{K}^2$; $c_L = 1.005\text{kJ}/\text{kg}$; $\kappa_{\text{El}} = 0.2\text{kW}/(\text{kg}/\text{s})^2$

Der direkte Vergleich der aufgewendeten Jahresenergie und der Kosten für die Regelung mit und ohne MPC-basierter Vortemperierung der Hörsäle ist in den Grafiken 6.15 und 6.16 zusammengefasst, wobei die Simulation mit MPC für 4 verschiedene Fälle durchgeführt wurde:

- (a) MISO-Modell mit idealer Wettervorhersage, d. h. $\hat{\vartheta}_{\text{Aul}}[t] \equiv \vartheta_{\text{Aul}}[t]$ und $\hat{Q}_{\text{Gs}}[t] \equiv \dot{Q}_{\text{Gs}}[t]$
- (b) MIMO-Modellansatz (6.15) mit idealer Wettervorhersage
- (c) MISO-Modellansatz mit S-ARMA Wettermodell (siehe Abschnitt 6.1.4)
- (d) MISO-Modellansatz mit saisonalem Rauschmodell zur Berücksichtigung interner Gewinne durch die periodische Hörsaalbelegung (siehe Abschn. 5.1.2) und ideale Wettervorhersage

In der Abbildung 6.15 lässt sich erkennen, dass durch den MPC eine Kostenreduktion von etwa 25% erreicht werden kann, wobei die wesentlichen Ersparnisse auf eine Verringerung der Elektroenergieaufnahme der Lüfter zurückzuführen sind. Weiterhin lässt sich feststellen, dass der verwendete Modellierungsansatz und die Art der Wetterprognose einen vergleichsweise geringen Einfluss auf die Gesamtenergiekosten besitzen. Dabei ist insbesondere bemerkenswert, dass der verkoppelte MIMO-Modellansatz (MPC (b)) trotz etwas besserer Prädiktionsgüte des Modells zu höheren Gesamtenergiekosten bei der Jahressimulation geführt hat. Die Gründe hierfür wurden jedoch nicht weiter untersucht, werden aufgrund der festgestellten deutlich langsameren Konvergenzgeschwindigkeit bei einem „erhöhten Schwierigkeitsgrad“ der resultierenden Optimierungsaufgabe vermutet.

¹⁸In den Sommer- und Übergangsmonaten kann während der Nichtnutzungszeit des Gebäudes eine „freie Kühlung“ durch die Außenluft genutzt werden.

¹⁹Aufgrund der hohen maximalen Zuluftmassenströme $\dot{m}_{i,\text{max}}$ kann ein zügiger Luftaustausch in den Hörsälen erfolgen, sodass eine relativ kurze Vorhaltezeit ausreichend ist.

Der zeitliche Verlauf der Jahresenergiekosten für die Regelung mit und ohne MPC-basierter Vortemperierung ist in der Abbildung 6.16 dargestellt. Man erkennt, dass die größten Ersparnisse während der Kühlperioden im Sommer erzielt werden.

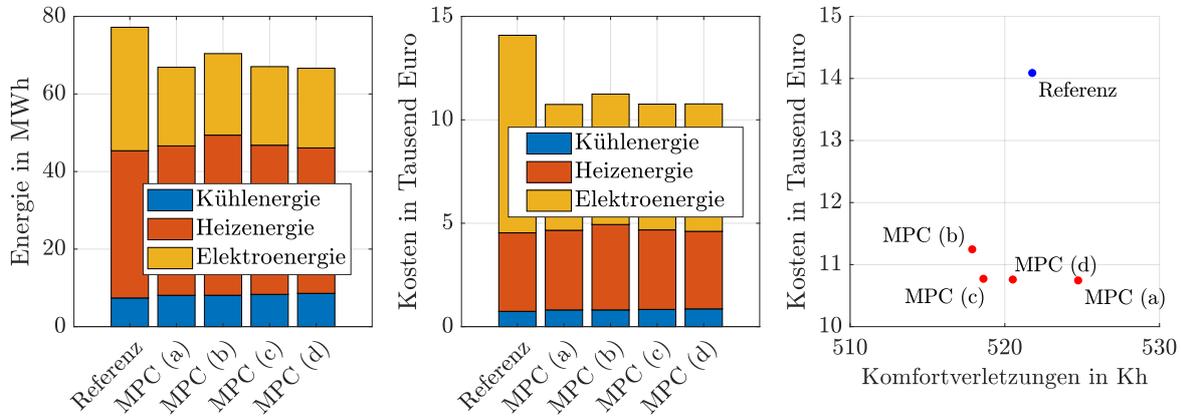


Abbildung 6.15: Kosten- und Jahresenergievergleich für die Regelung ohne (Referenz) und mit MPC-basierter Vortemperierung der Hörsäle.

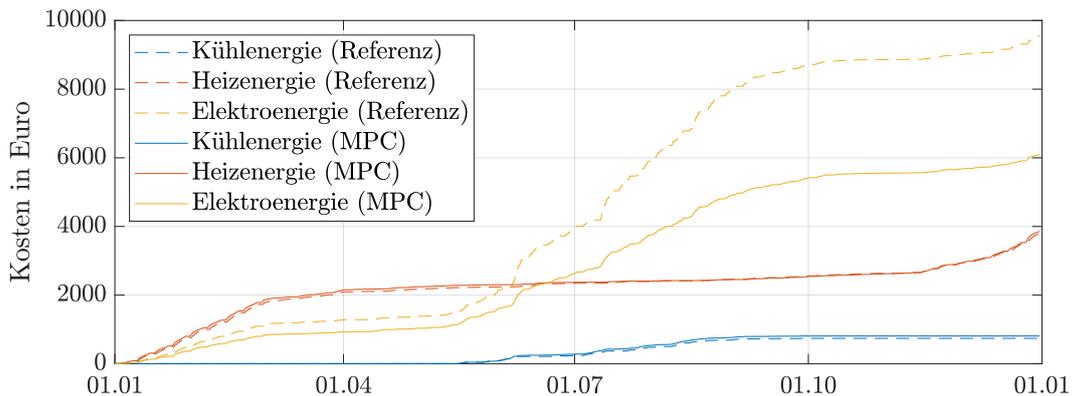


Abbildung 6.16: Verlauf der Jahresenergiekosten für die Regelung ohne (Referenz) und mit MPC-basierter Vortemperierung (MISO-Modell mit idealer Wettervorhersage).

Die Grafik 6.17 veranschaulicht den Vergleich der Regel- und Stellgrößenverläufe mit und ohne MPC-basierter Vortemperierung der Hörsäle. Dabei erkennt man zunächst, dass sich beide Regelstrategien bezüglich der gewählten Zulufttemperatur ϑ_{Zul} ähnlich verhalten, was u. a. die Sinnhaftigkeit der Implementierung der in Abb. 6.14 dargestellten Referenzregelstrategie rechtfertigt. Des Weiteren sieht man, dass sich aufgrund der Vortemperierung durch den MPC geringere Massenströme während des Tages ergeben, was sich aufgrund der quadratischen Abhängigkeit der Leistungsaufnahme des Lüfters vom Gesamtmassenstrom positiv auf die Elektroenergiekosten auswirkt.

Die Abbildung 6.18 zeigt beispielhaft eine durch den MPC berechnete optimierte Stell- bzw. Regelgrößenfolge im Sommer. Dabei lässt sich zum einen ein zeitlich abwechselnd

versetztes Kühlen der Hörsäle während der Nichtnutzungszeit beobachten, was wiederum zu einem geringeren Gesamtmassenstrom und somit zu einer reduzierten Leistungsaufnahme der Lüfter führt. Zudem kann erkannt werden, dass insbesondere während der Nacht vorgekühlt wird, was sich wiederum positiv auf die Kühlenergiekosten auswirkt (vgl. Gl. (6.19)). Weiterhin wird aus der Grafik auch ersichtlich, dass die optimierten Regelgrößenfolgen während der Nutzungszeit der Hörsäle entlang der oberen Temperaturgrenzwerte verlaufen, was eine im Sommer intuitiv sinnvolle Betriebsführung darstellt. In analoger Weise verlaufen die optimierten Regelgrößenfolgen im Winter typischerweise entlang der Minimalbegrenzung, was während der Nutzungszeit der Hörsäle eine Umschaltung auf die Referenzregelstrategie rechtfertigt (vgl. Bem. 6.4).

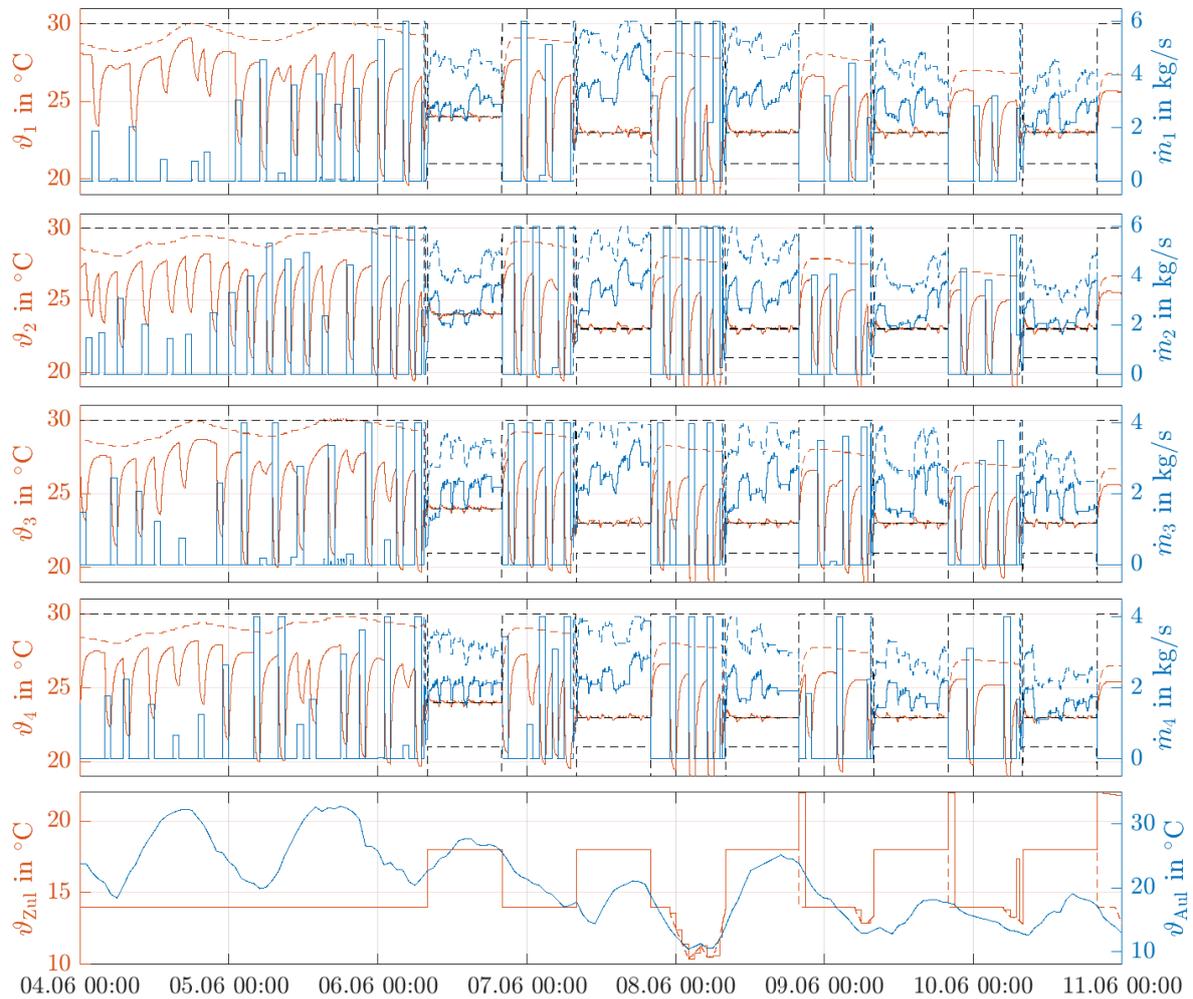


Abbildung 6.17: Vergleich der Verläufe der Hörsaaltemperaturen ϑ_1 bis ϑ_4 , der Zulufttemperatur ϑ_{Zul} (untere Grafik) sowie der zugehörigen Massenströme \dot{m}_1 bis \dot{m}_4 mit (durchgezogene Linien) und ohne MPC-basierte Vortemperierung (gestrichelte Linien) für eine Beispielwoche im Sommer. Die schwarzen gestrichelten Linien kennzeichnen die unteren und oberen Temperaturgrenzwerte $\vartheta_{i,\min}[t]$ bzw. $\vartheta_{i,\max}[t]$.

Bemerkung 6.5: Die Lösung des nichtlinearen Programms mit 484 Optimierungsvariablen, 768 nichtlinearen und 964 linearen Nebenbedingungen wurde mittels der MATLAB Routine `fmincon()` unter Verwendung des SQP-Algorithmus berechnet. Die Dauer für das Auffinden eines Minimums betrug auf einem Windows 10 PC mit Intel Core i7-3520M (2x2.9GHz) und 8GB Arbeitsspeicher im Mittel etwa 36 Sekunden. \triangleleft

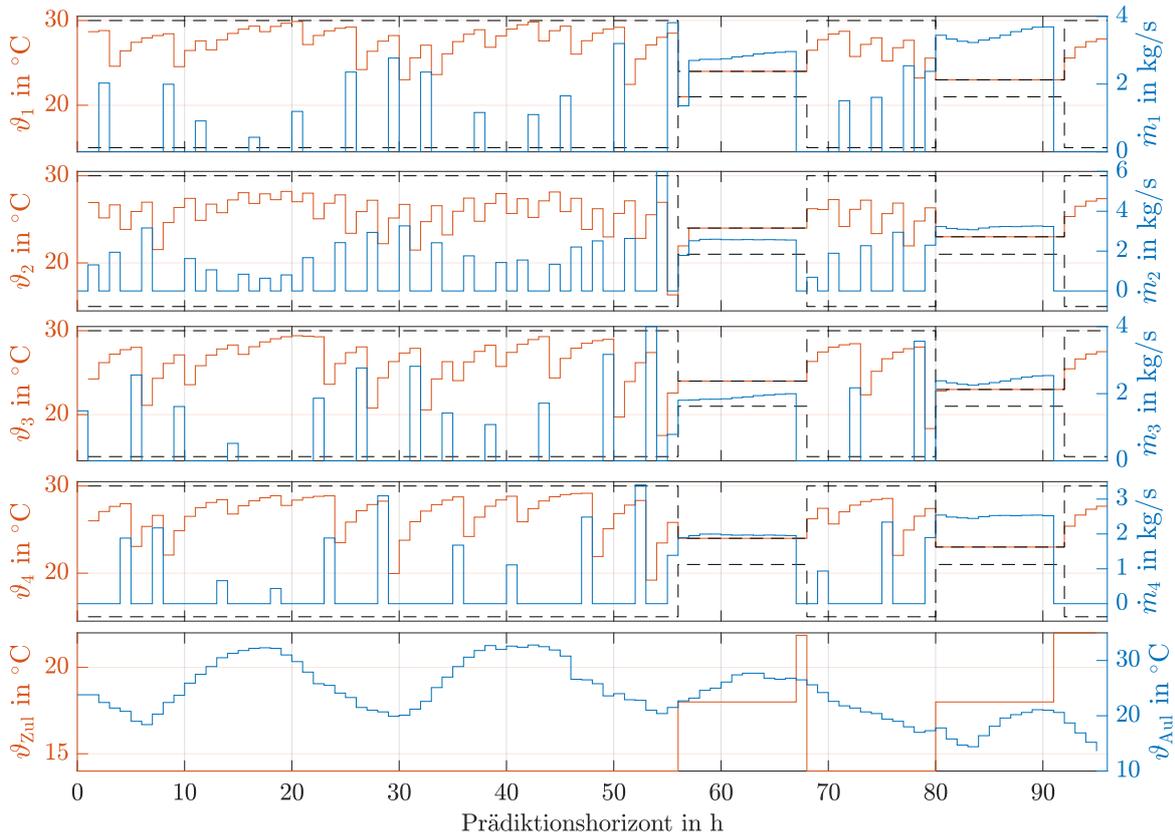


Abbildung 6.18: Beispielfhafte durch den MPC berechnete optimierte Stell- bzw. Regelgrößenfolge im Sommer.

Kapitel 7

Zusammenfassung und Ausblick

7.1 Zusammenfassung

Die vorliegende Arbeit befasst sich mit Aspekten der datengetriebenen Modellierung in der Gebäudetechnik zum Zweck des Entwurfs modellprädiktiver Regler und ist in einen Theorie- und in einen Anwendungsteil gegliedert. Dabei vermittelt der erste Teil der Arbeit (Kapitel 2 und 3) wichtige theoretische Grundlagen der Systemidentifikation, wohingegen sich der übrige Teil der praktischen Anwendung der Systemidentifikation und dem Entwurf von prädiktiven Reglern in der Gebäudetechnik widmet.

Das zweite Kapitel befasst sich mit Modellen zeitdiskreter stochastischer Systeme, welche die Grundlage für die Systemidentifikation darstellen. Dabei wird ein Ausdruck für die optimale Mehrschrittprädiktion eines allgemeinen MIMO-LTI-Systems hergeleitet. Des Weiteren werden im Abschnitt 2.5 nichtlineare Modellstrukturen vorgeschlagen, für welche die Prädiktionsfehlerfolge in analoger Weise einfach berechnet werden kann. Dabei besitzt insbesondere das Modell (2.51) eine große Relevanz in der Gebäudetechnik und stellt somit die Grundlage für die PEM-Identifikation in den Kapiteln 5 bzw. 6 dar.

Das Kapitel 3 widmet sich dann der Identifikation dynamischer Systeme, insbesondere mittels der Methode der Prädiktionsfehlerminimierung. Dabei wird auf die wesentlichen Implementierungsaspekte und die theoretische Analyse der Mehrschritt-PEM eingegangen (Konsistenz und Schätzgenauigkeit) und die wesentlichen Erkenntnisse anhand zweier Beispiele verdeutlicht. Zudem wird anhand des Beispiels 3.3 verdeutlicht, wie sich wesentliche Ideen bei der Parametrierung und Identifikation linearer Systeme auf das nichtlineare Modell (2.51) übertragen lassen.

Das darauf folgende Kapitel 4 dient als Einleitung in den Anwendungsteil der Arbeit. Dabei werden die wesentlichen Funktionsprinzipien in der Gebäudetechnik vermittelt und mögliche Anwendungsfälle prädiktiver Regelungen genannt und klassifiziert.

Im Kapitel 5 wird die Anwendung der Systemidentifikation anhand während des Gebäudebetriebs aufgenommenen Messdaten demonstriert. Dabei wurde gezeigt, dass mittels des Modellansatzes (2.51) und der Nutzung der PEM sinnvolle Identifikationsergebnisse erzielt werden können. Weiterhin wurde eine Methode zur Berücksichtigung einer periodischen Raumnutzung vorgeschlagen, die auf einer geeigneten Parametrierung des Rauschmodells basiert. Zudem wurde u. a. auf den praktischen Vergleich der 1-Schritt- mit der Mehrschritt-PEM eingegangen, wobei sich durch die Verwendung eines

Mehrschrittverfahren keine signifikanten Vorteile im Sinne der Prädiktionsgenauigkeit ergeben haben. Zuletzt wurde in der Zusammenfassung des Abschnittes 5.1 auf die praktischen Probleme bei der Identifikation eingegangen, aus welcher sich zukünftige Forschungsfragen ableiten lassen.

Das letzte Kapitel befasst sich mit der Implementierung modellprädiktiver Regelungsansätze für zwei konkrete Anwendungsbeispiele. Die dabei verwendeten Prozessmodelle wurden unter Nutzung der Sprache MODELICA und der weit verbreiteten BUILDINGS Bibliothek erstellt und zur Generierung synthetischer Messdaten verwendet. Diese wurden anschließend für die Identifikation der für den MPC-Entwurf notwendigen Modelle verwendet, wodurch die Sinnhaftigkeit der in Kapitel 5 vorgeschlagenen Modell- und Identifikationsansätze zusätzlich bekräftigt werden konnte. Der Vergleich des basierend auf den identifizierten Modellen entworfenen MPC-Reglers mit einer konventionellen Regelungsstrategie veranschaulicht beispielhaft die Vorteile der prädiktiven Regelung.

7.2 Ausblick und offene Fragen

Im Rahmen dieser Arbeit haben sich offene Fragestellungen und mögliche neue Forschungsansätze ergeben, die im Folgenden zusammengefasst werden:

Einschritt- und Mehr- bzw. Multischritt-PEM

In Abschnitt 3.2 wurden Implementierungsaspekte und eine theoretische Analyse der Mehr- bzw. Multischritt-PEM vorgestellt. Dabei wurde u. a. ein Ausdruck zur Berechnung der Schätzfehlerkovarianz (3.27) der Mehrschritt-PEM hergeleitet. Dabei ist jedoch ein direkter Vergleich mit der Kovarianzmatrix der 1-Schritt-PEM (B.30) sowie die Herleitung von (3.27) für den allgemeinen MIMO-Fall offengeblieben.

Des Weiteren sei erwähnt, dass für rückgekoppelte Systeme des Typs (2.51) mittels der entwickelten Identifikationssoftware (siehe Anhang D) lediglich eine 1-Schritt PEM-Identifikation möglich ist. Bei der Implementierung von prädiktiven Reglern werden jedoch Mehrschrittprädiktionen benötigt, welche im Anwendungsteil der Arbeit entsprechend der Bemerkung 2.12 durch Simulation des Modells berechnet wurden. Praktisch wäre es daher naheliegend, bereits bei der Identifikation den entsprechenden Simulationsfehler zu minimieren, wodurch evtl. eine bessere Regelgüte bei der MPC-Implementierung erzielt werden könnte.

Rekursive Identifikationsalgorithmen

Ein weiterer in der Arbeit nicht betrachteter Aspekt sind rekursive Identifikationsalgorithmen, die insbesondere bei der Schätzung zeitvarianter Modelle und beim Entwurf adaptiver Regelungsverfahren eine wichtige Rolle spielen [LLM98; NA05]. Insbesondere die Implementierung und Analyse von rekursiven Prädiktionsfehlerverfahren (RPEM) auf Basis von Mehr- bzw. Multischrittprädiktionen wären naheliegende Erweiterungsmöglichkeiten der Ausführungen des Abschnittes 3.2. Direkte Anwendungsfälle von

RPEM-Verfahren in der Gebäude- und Energietechnik könnten sich aufgrund von zeitlichen Veränderungen der Nutzungscharakteristik bei der Lastprognose [Mes11; BMA13; GDJ18], d. h. bei der Identifikation von Prognosemodellen für den Wärme-, Kälte- bzw. Elektroenergiebedarf, ergeben.

Weitere Anwendungen der Modellstruktur (2.51)

Insbesondere aufgrund der Analogie zum linearen Fall hat sich das blockorientierte Modell (2.51) als zielführender Modellierungsansatz für viele Problemstellungen aus der Gebäudetechnik erwiesen. Aus praktischer Sicht sind jedoch eventuell weitere technische Anwendungen und der Vergleich mit typischen im Bereich der nichtlinearen Systemidentifikation verwendeten Ansätzen, wie bspw. NARMAX-Modellen, denkbar. Weitere Anknüpfungspunkte könnten u. a. Ansätze aus dem Bereich des maschinellen Lernens, bspw. die Verwendung eines künstlichen neuronalen Netzes (KNN) als statische Nichtlinearität $f(\mathbf{u}, \mathbf{y})$, darstellen.

Identifikation mit unbekanntem Eingangsgrößen

Bereits in der Zusammenfassung des Abschnittes 5.1 wurde erwähnt, dass nicht messbare Störeinflüsse, wie bspw. der Einfluss der Raumnutzer, eine besondere Herausforderung bei der Identifikation von Prädiktionsmodellen für die Komfortparameter der Raumluft darstellen. In weiterführenden Arbeiten könnte daher die Verwendung zusätzlicher CO₂ und Feuchte-Sensorik und der Methoden der Blind System Identification bzw. Unknown Input Identification in Erwägung gezogen werden [AQH97; ORD14; Lin17; LE17]. Weiterhin könnte die Untersuchung möglicherweise bestehender Zusammenhänge zum Beobachterentwurf für Systeme mit unbekanntem Eingangsgrößen (*Unknown Input Observer*) [CP98; Irl16] ein sowohl aus praktischer als auch aus theoretischer Sicht lohnenswertes Forschungsfeld darstellen.

Modellprädiktive Energiesystemregler

Die in Kapitel 6 vorgeschlagenen Beispielimplementierungen stellen lediglich zwei typische und naheliegende Anwendungsfälle von modellprädiktiven Regelungen in der Gebäudetechnik dar. In der Praxis sind jedoch eine Vielzahl weiterer Szenarien denkbar. Insbesondere im Bereich der modellprädiktiven Regelung von Energiesystemen (siehe Abschnitt 4.2.2) könnten sich aufgrund der zunehmenden Verbreitung von regenerativen Energiequellen, elektrischen und thermischen Speichertechnologien in Kombination mit der Elektrifizierung des Personenkraftverkehrs (zeitlich beschränkte Nutzung der PKW-Batterie im Hausnetz) weitere interessante MPC-Anwendungen ergeben.

Anwendungsspezifische Software

Wie in der Einleitung der Arbeit beschrieben wurde, ist die Modellierung der Systemkomponenten ein kritischer Aspekt beim Entwurf modellprädiktiver Regler, dessen

Zeitaufwand durch die Systemidentifikation reduziert werden kann. Die innerhalb der Arbeit vorgeschlagenen Modell- und Identifikationsansätze sind jedoch nur für Ingenieure mit tiefgehenden Kenntnissen aus dem Bereich der Regelungstheorie und der Systemidentifikation anwendbar. Eine Möglichkeit die Nutzbarkeit der Ergebnisse dieser Arbeit zu verbessern, wäre der Entwurf von anwendungsspezifischer Software, welche die Identifikation und den MPC-Entwurf erleichtert. Hierfür könnte im ersten Schritt eine möglicherweise objektorientierte Komponentenbibliothek, mit den typischen für den MPC-Entwurf notwendigen Komponenten (Räume, Pufferspeicher, Wärmepumpen, Heizkessel, BHKWs, etc.) entworfen werden, die dem Anwender eine nutzerfreundliche Schnittstelle zur Verfügung stellt. Die notwendigen Informationen zur Konfiguration der zu identifizierenden Modellstruktur, sowie sinnvolle Standardeinstellungen für die Identifikation, könnten dann aus den Objektinstanzen extrahiert werden. Die Informationen, welche wiederum zur Instanziierung der Objekte benötigt werden (z. B. Anzahl und Orientierung der Fenster, Informationen zum Heizsystem eines Raumes, Nennleistung eines Wärmeerzeugers), könnten zukünftig bspw. aus BIM-Daten (Building Information Modeling¹) gewonnen werden.

¹Der Begriff Building Information Modeling beschreibt den ganzheitlichen Prozess zum Erstellen und Verwalten von Informationen eines Bauobjekts und umfasst Methoden, Software und Datenaustauschformate für die gewerkeübergreifende Planung, Bewirtschaftung und den Bau von Gebäuden (siehe z. B. [BKB15]). BIM-Daten bzw. BIM-Modelle sind Dateien, welche ausgetauscht und maschinell verarbeitet werden können.

Anhang A

Begriffe und Definitionen

A.1 Diskrete Fourier-Transformation

Gegeben sei ein reelles zeitdiskretes Signal $x[t]$ ($t \in \mathbb{Z}$). Die komplexe Folge

$$X(\omega) = \mathcal{F}\{x[t]\} = \sum_{t=-\infty}^{\infty} x[t]e^{-j\omega t} \quad \omega \in \{\omega \in \mathbb{R} \mid -\pi < \omega < \pi\} \quad (\text{A.1})$$

heißt *Fourier-Transformation des zeitdiskreten Signals* $x[t]$ (Discrete-time Fourier transform – DTFT). Die *inverse DTFT* wird mittels

$$x[t] = \mathcal{F}^{-1}\{X(\omega)\} = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(\omega)e^{j\omega t} d\omega \quad (\text{A.2})$$

berechnet [Foe11].

Man beachte, dass in der Regel endliche Signale $x_N[t]$ mit $t = 1, \dots, N$ vorliegen, weswegen die DTFT meist nur im Bereich der theoretischen Signalanalyse Anwendung findet. In der Praxis wird daher die *diskrete Fourier-Transformation* (Discrete Fourier transform – DFT)

$$X_N[\omega] = \sum_{t=1}^N x[t]e^{-j\omega t} \quad \text{mit } \omega = \frac{2\pi k}{N} \quad \text{und } k = 1, \dots, N, \quad (\text{A.3})$$

bzw. die *inverse DFT*

$$x_N[t] = \frac{1}{N} \sum_{k=1}^N X_N[\omega]e^{-j\frac{2\pi k}{N}t} \quad (\text{A.4})$$

verwendet [Lju12, Abschn. 2.2].

A.2 Begriffe der Stochastik

In diesem Abschnitt werden einige innerhalb dieser Arbeit verwendeten Begriffe der Stochastik eingeführt. Dabei wird davon ausgegangen, dass dem Leser die grundlegenden Begriffe der Wahrscheinlichkeitstheorie geläufig sind. Für eine umfangreichere Einführung sei unter anderem auf [Jaz70; AM79; Soe02; VV07] verwiesen.

A.2.1 Unabhängigkeit und Unkorreliertheit stochastischer Variablen

Seien \mathbf{X} und \mathbf{Y} Zufallsgrößen mit Erwartungswert $\boldsymbol{\mu}_X$ bzw. $\boldsymbol{\mu}_Y$ und den Wahrscheinlichkeitsdichten $f(\mathbf{x})$ bzw. $f(\mathbf{y})$. Die Vektoren \mathbf{X} und \mathbf{Y} werden als *unabhängig* bezeichnet, wenn deren gemeinsame Wahrscheinlichkeitsdichte $f(\mathbf{x}, \mathbf{y})$ in das Produkt der Dichten von \mathbf{X} und \mathbf{Y} zerlegt werden kann, d. h.:

$$f(\mathbf{x}, \mathbf{y}) = f(\mathbf{x})f(\mathbf{y}) \quad (\text{A.5})$$

Gilt hingegen lediglich

$$\mathbb{E} \{ (\mathbf{X} - \boldsymbol{\mu}_X)(\mathbf{Y} - \boldsymbol{\mu}_Y)^\top \} = \mathbf{0}, \quad (\text{A.6})$$

dann werden \mathbf{X} und \mathbf{Y} als *unkorreliert* bezeichnet. Sind die Zufallsgrößen \mathbf{X} und \mathbf{Y} unabhängig, dann sind sie auch unkorreliert, wohingegen die Umkehrung im Allgemeinen nicht gilt [Soe02].

A.2.2 Konvergenz stochastischer Variablen

Es wird eine Folge von Zufallsgrößen \mathbf{X}_t mit der Indexvariable $t = 1, 2, \dots$ betrachtet. Bei der Analyse solcher Folgen werden verschiedene Konvergenzbegriffe unterschieden [Ast70; SS89; Ise92a]: Sei \mathbf{X}^* eine Zufallsgröße, dann konvergiert \mathbf{X}_t

- *in Verteilung*, wenn $\lim_{t \rightarrow \infty} f_{X_t}(\mathbf{x}) = f_{X^*}(\mathbf{x})$, wobei f_{X_t} und f_{X^*} die Wahrscheinlichkeitsdichte von \mathbf{X}_t bzw. \mathbf{X}^* darstellen.
- *in Wahrscheinlichkeit*, wenn für jedes $\epsilon > 0$ $\lim_{t \rightarrow \infty} \mathbb{P}(|\mathbf{X}_t - \mathbf{X}^*| > \epsilon) = 0$ gilt.
- *mit Wahrscheinlichkeit 1*, wenn $\mathbb{P}(\lim_{t \rightarrow \infty} \mathbf{X}_t = \mathbf{X}^*) = 1$ gilt.
- *im quadratischen Mittel*, wenn $\lim_{t \rightarrow \infty} \mathbb{E} \{ \mathbf{X}_t - \mathbf{X}^* \} = \mathbf{0}$ gilt.

A.2.3 Stochastische Prozesse

Dieser Abschnitt befasst sich mit den grundlegenden Eigenschaften und Kenngrößen stochastischer Prozesse. Ein stochastischer Prozess \mathbf{X}_t mit $t \in \mathcal{T}$ wird dabei als eine Menge von Zufallsgrößen $\mathbf{X}_{t_1}, \mathbf{X}_{t_2}, \dots$ verstanden, welche durch die multivariate Verteilungsfunktion

$$F_X(\mathbf{x}_1, \mathbf{x}_2, \dots; t_1, t_2, \dots) = \mathbb{P}(\mathbf{X}_{t_1} \leq \mathbf{x}_1, \mathbf{X}_{t_2} \leq \mathbf{x}_2, \dots) \quad (\text{A.7})$$

charakterisiert ist, wobei $\mathbf{x}(t) \in \mathcal{X} \subseteq \mathbb{R}^n$ allgemein *Realisierung* oder *Pfad* des Prozesses genannt wird. Entspricht t wie oftmals üblich der Zeit, dann wird $\mathbf{x}(t)$ insbesondere im Rahmen der Signalverarbeitung als *Signal* bezeichnet. Falls $\mathcal{T} = \mathbb{R}$ wird \mathbf{X}_t *kontinuierlicher stochastischer Prozess* genannt, wohingegen $\mathcal{T} = \mathbb{Z}$ einen *diskreten stochastischen Prozess* kennzeichnet. Für eine formale Definition stochastischer Prozesse sei unter anderem auf [Jaz70; Ast70] verwiesen.

A.2.3.1 Stationarität und Quasistationarität

Ein stochastischer Prozess \mathbf{X}_t heißt *stationär*, *stark stationär* oder *stationär im strengen Sinn*, wenn die Verteilungsfunktion von $\mathbf{X}_{t_1}, \mathbf{X}_{t_2}, \dots$ für alle $t_i, t_i + \tau \in \mathcal{T}$ translationsinvariant gegenüber der Zeit ist, d. h.

$$F_X(\mathbf{x}_1, \mathbf{x}_2, \dots; t_1, t_2, \dots) = F_X(\mathbf{x}_1, \mathbf{x}_2, \dots; t_1 + \tau, t_2 + \tau, \dots). \quad (\text{A.8})$$

Wird die Zeitinvarianz lediglich für den Erwartungswert und die Kovarianz des Prozesses gefordert, d. h. es gilt

$$\begin{aligned} (1) \quad & \mathbb{E} \{ \mathbf{X}_t \} = \boldsymbol{\mu}_X = \text{const.} \\ (2) \quad & \text{Cov} \{ \mathbf{X}_t, \mathbf{X}_s \} = \text{Cov} \{ \mathbf{X}_{t+\tau}, \mathbf{X}_{s+\tau} \} \\ (3) \quad & \text{Var} \{ \mathbf{X}_t \} < \infty, \end{aligned} \quad (\text{A.9})$$

dann heißt \mathbf{X}_t *schwach stationär*, *stationär im weiteren Sinne* oder *kovarianzstationär* (siehe [Ast70, Abschn. 2.3], [Ise92a, Abschn. 2.2] und [HW14, Abschn. 2.3.1]). Man beachte, dass die Bedingung (2) besagt, dass die Kovarianz des Prozesses nicht von den Absolutwerten der betrachteten Zeitpunkte t und s abhängig ist, sondern lediglich von deren zeitlicher Differenz $\tau = t - s$.

Quasistationarität

Bei der Identifikation dynamischer Systeme ist die Annahme stationärer stochastischer Prozesse oftmals zu einschränkend. Dies liegt u. a. darin begründet, dass im Rahmen der Systemidentifikation Prozesse mit exogenen Eingangsgrößen betrachtet werden, die vorzugsweise als deterministisch betrachtet werden. Für die Beschreibung der Störungen wird hingegen ein stochastischer Ansatz favorisiert, sodass der Systemausgang als stochastischer Prozess mit deterministischem Anteil betrachtet werden kann. Für ein SISO LTI-System der Form $y[t] = G(q^{-1})u[t] + H(q^{-1})e[t]$ (siehe Abschnitt 2.2.1 und Gl. (2.7)) gilt beispielsweise $\mathbb{E} \{ y[t] \} = G(q^{-1})u[t] \neq \text{const.}$, sodass der Systemausgang im Allgemeinen nicht als stationärer Prozess betrachtet werden kann. Um diese Problematik zu berücksichtigen, ist es zweckmäßig die Definition der quasi-Stationarität einzuführen:

Ein stochastischer Prozess \mathbf{X}_t ist *quasi-stationär*, wenn in der Definition der schwachen Stationarität (A.9) an Stelle von (1) die Bedingung $\mathbb{E} \{ \mathbf{X}_t \} = \boldsymbol{\mu}_X[t]$, mit $|\boldsymbol{\mu}_X[t]| \leq \mathbf{C}$ gilt [Lju12, Abschn. 2.2]. Ist \mathbf{X}_t deterministisch, dann bedeutet quasi-Stationarität demnach, dass \mathbf{X}_t absolut beschränkt ist.

A.2.3.2 Kovarianzfunktion

Seien \mathbf{X}_t und \mathbf{Y}_t stochastische Prozesse mit endlicher Varianz¹. Dann nennt man

$$\begin{aligned} \mathbf{r}_{XY}[t, s] &= \text{Cov} \{ \mathbf{X}_t, \mathbf{Y}_s \} \\ &= \mathbb{E} \left\{ \left(\mathbf{X}_t - \mathbb{E} \{ \mathbf{X}_t \} \right) \left(\mathbf{Y}_s - \mathbb{E} \{ \mathbf{Y}_s \} \right)^\top \right\} \end{aligned} \quad (\text{A.10})$$

¹Diese werden oftmals auch als Prozesse zweiter Ordnung bezeichnet [Ast70, Abschn. 2.3]

die *Kreuzkovarianzfunktion* bzw. *Kreuzkovarianzfolge* ($t, s \in \mathcal{T}$). Basierend auf der Definition der Kreuzkovarianzfunktion lässt sich mit

$$\mathbf{r}_X[t, s] = \mathbf{r}_{XX}[t, s] \quad (\text{A.11})$$

die *Autokovarianzfunktion* bzw. *Autokovarianzfolge* des stochastischen Prozesses \mathbf{X}_t definieren. Wird wie häufig der Spezialfall eines (schwach) stationären Prozesses unterstellt, dann ist die Autokovarianzfunktion lediglich von der Differenz der betrachteten Zeitpunkte $\tau = t - s$ abhängig, sodass dann $\mathbf{r}_X[\tau]$ geschrieben werden kann.

Bemerkung A.1: Ein weiterer eng verwandter Begriff ist der Begriff der Auto- bzw. Kreuzkorrelation: Bei der Betrachtung skalarer Prozesse X_t und Y_t wird die normierte Größe

$$\rho_{XY}[t, s] = \frac{r_{XY}[t, s]}{\sqrt{r_{XX}[t, t]r_{YY}[s, s]}} \quad (\text{A.12})$$

als die *Kreuzkorrelationsfunktion* bzw. *Kreuzkorrelationsfolge* bezeichnet [Ast70, Abschn. 2.4]. In analoger Weise lässt sich mit $\rho_{XX}[t, s] = \rho_X[t, s]$ die *Autokorrelationsfunktion* bzw. *Autokorrelationsfolge* definieren. ◁

A.2.3.3 Spektraldichte

Seien \mathbf{X}_t und \mathbf{Y}_t (schwach) stationäre stochastische Prozesse. Die DTFT (A.1) der Kreuz- bzw. Autokovarianzfunktion, d. h.

$$\phi_{XY}(\omega) = \mathcal{F}\{\mathbf{r}_{XY}[\tau]\} \quad \text{mit } \omega \in (-\pi, \pi) \quad (\text{A.13})$$

wird als das *Kreuz-*, bzw. $\phi_{XX}(\omega) = \phi_X(\omega)$ als *Autoleistungsdichtespektrum* oder *Spektraldichte*, bezeichnet. Weiterhin nennt man die auf der DFT (A.3) basierende Schätzung der Spektraldichte

$$\hat{\phi}_x[\omega] = \frac{1}{N} \mathbf{X}_N[\omega] \bar{\mathbf{X}}_N^\top[\omega] \quad (\text{A.14})$$

das *Periodogramm* des Signals $\mathbf{x}_N[t]$ ($t = 1, \dots, N$), wobei $\bar{\mathbf{X}}[\omega]$ die komplex Konjugierte von $\mathbf{X}[\omega]$ beschreibt [SS89, Abschn. 3.5].

Man beachte, dass das Periodogramm zwar eine asymptotisch erwartungstreue, allerdings inkonsistente Schätzung der Spektraldichte darstellt [Bri81; And84]. In der Praxis werden daher oftmals zusätzliche Fensterfunktionen zur Schätzung der Leistungsdichtespektren verwendet, deren Weite den Kompromiss zwischen Bias und Varianz der Schätzung beeinflusst (siehe zum Beispiel [JW68, Kap. 6], [Bri81, Kap. 5] sowie [SM05, Kap. 2]). Für eine ausführliche Diskussion der Schätzproblematik von Leistungsdichtespektren sei unter anderem auf [SM05] verwiesen.

Transformation von Spektraldichten durch LTI-Systeme

Sei $\mathbf{G}(q^{-1})$ ein $(n_y \times n_u)$ -dimensionales asymptotisch stabiles Filter mit Eingangssignal $\mathbf{u}[t]$ und Ausgangssignal $\mathbf{y}[t]$. Für das Auto- bzw. Kreuzleistungsdichtespektrum des

Filterausgangs \mathbf{y} gelten die Beziehungen [SS89, App. 3.1]:

$$\begin{aligned}\phi_Y(\omega) &= \mathbf{G}(e^{-j\omega})\phi_U(\omega)\mathbf{G}^\top(e^{j\omega}) \\ \phi_{YU}(\omega) &= \mathbf{G}(e^{-j\omega})\phi_U(\omega)\end{aligned}\tag{A.15}$$

A.2.3.4 Weißes Rauschen

Mittels der Definition der Spektraldichte wird üblicherweise der Begriff des weißen Rauschens definiert: Ein stochastischer Prozess \mathbf{X}_t mit $\phi_X(\omega) = \mathbf{C} = \text{const.}$ heißt *weißes Rauschen*. Im Gegensatz dazu wird ein Rauschprozess mit einer nicht konstanten spektralen Leistungsdichte *farbiges Rauschen* genannt.

Entsprechend der Rücktransformationsvorschrift (A.2) ergibt sich für die Autokovarianzfunktion eines diskreten weißen Rauschprozesses

$$\mathbf{r}_X[\tau] = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathbf{C} e^{j\omega\tau} d\omega = \mathbf{C} \frac{\sin(\pi\tau)}{\pi\tau} = \begin{cases} \mathbf{C} & \text{für } \tau = 0 \\ \mathbf{0} & \text{sonst} \end{cases}, \tag{A.16}$$

woraus geschlussfolgert werden kann, dass \mathbf{X}_{t_1} und \mathbf{X}_{t_2} für alle t_1 und t_2 unkorreliert sind.

Bemerkung A.2 (kontinuierliches weißes Rauschen): Es sei angemerkt, dass im Falle eines kontinuierlichen weißen Rauschprozesses \mathbf{X}_t bei der Rücktransformation über $\omega \in (-\infty, \infty)$ integriert werden muss, wodurch sich für die Autokovarianzfunktion $\mathbf{r}_X(\tau) = \mathbf{C}\delta(\tau)$ ergibt, wobei $\delta(\tau)$ den kontinuierlichen *DIRAC*-Impuls kennzeichnet. Kontinuierliches weißes Rauschen besitzt daher eine unendliche Varianz und ist deshalb kein stationärer stochastischer Prozess (siehe Definition (A.9)). \triangleleft

A.2.3.5 Ergodizität

Die bei der Definition der schwachen Stationarität und der Kovarianzfunktion verwendeten Erwartungswerte sind als „Ensemble-Mittelung“, d. h. als Mittelung über mehrere Realisierungen des Prozesses \mathbf{X}_t , zu verstehen. Ein stationärer stochastischer Prozess heißt *ergodisch*, wenn durch zeitliche Mittelung einer einzelnen unendlich langen Realisierung $\mathbf{x}[t]$ dieselben wahrscheinlichkeitstheoretischen Kenngrößen des Prozess ermittelt werden können [Ise92a, Abschn. 2.2]. Gilt dies hingegen nur für den Erwartungswert und die Kovarianz des Prozesses, d. h.

$$\begin{aligned}(1) \quad & \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N \mathbf{x}[t] \rightarrow \mathbb{E}\{\mathbf{X}_t\} \quad \text{und} \\ (2) \quad & \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N \mathbf{x}[t]\mathbf{x}_{t+\tau}^\top \rightarrow \mathbb{E}\{\mathbf{X}_t\mathbf{X}_{t+\tau}^\top\}\end{aligned}\tag{A.17}$$

mit Wahrscheinlichkeit 1, dann nennt man den Prozess \mathbf{X}_t *ergodisch bezüglich der ersten beiden Momente* [SS89, App. B.1]. Die Voraussetzung für die Ergodizität eines Prozesses

ist seine Stationarität, wohingegen die Umkehrung der Aussage nicht gültig ist². Der Zusammenhang der hier genannten Eigenschaften stochastischer Prozesse ist in der Abbildung A.1 dargestellt.

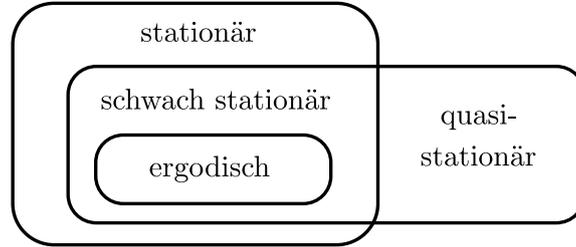


Abbildung A.1: Dendrogramm spezieller stochastischer Prozesse

A.2.4 Begriffe der Schätztheorie

In diesem Abschnitt wird davon ausgegangen, dass eine „wahre Systemstruktur“ \mathcal{S} existiert, welche die in der Wirklichkeit ablaufenden physikalischen, chemischen oder ökonomischen Vorgänge beschreibt, und dass diese durch ein Schätzverfahren theoretisch aufgedeckt werden kann. Der Modellparameter $\boldsymbol{\theta} \in \mathcal{D}_{\mathcal{M}} \subseteq \mathbb{R}^{n_{\theta}}$ für welches das Modell \mathcal{M} das System exakt beschreibt, sei im Folgenden durch $\boldsymbol{\theta}_0$ gekennzeichnet.

A.2.4.1 Identifizierbarkeit

Ein Modell \mathcal{M} wird als *identifizierbar* bezeichnet, wenn es theoretisch möglich ist die „wahre“ Struktur durch ein geeignetes Identifikationsverfahren aufzudecken, indem unendlich viele Beobachtungen gemacht werden [EKR04, Abschn. 3.2]. Mathematisch bedeutet dies, dass unterschiedliche Werte des Parametervektors $\boldsymbol{\theta} \in \mathcal{D}_{\mathcal{M}}$ verschiedene Verteilungsfunktionen $\mathbf{F}_{\boldsymbol{\theta}}(\mathbf{x})$ der beobachteten Variablen \mathbf{x} erzeugen, d. h.:

$$\mathbf{F}_{\boldsymbol{\theta}_1}(\mathbf{x}) = \mathbf{F}_{\boldsymbol{\theta}_2}(\mathbf{x}) \Rightarrow \boldsymbol{\theta}_1 = \boldsymbol{\theta}_2 \quad \forall \boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \in \mathcal{D}_{\mathcal{M}} \quad (\text{A.18})$$

Umgekehrt bezeichnet man eine Modellstruktur als *nicht identifizierbar*, wenn mindestens ein $\boldsymbol{\theta}_1 \neq \boldsymbol{\theta}_2$ existiert ($\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \in \mathcal{D}_{\mathcal{M}}$), für welche $\mathbf{F}_{\boldsymbol{\theta}_1}(\mathbf{x}) = \mathbf{F}_{\boldsymbol{\theta}_2}(\mathbf{x})$ gilt [LC98, Abschn. 1.5]. Dabei ist zu beachten, dass ein Modell typischerweise nur unter bestimmten technischen Voraussetzungen identifiziert werden kann, welche als *Identifizierbarkeitsbedingungen* bezeichnet werden.

Angelehnt an die in [Lju12, Abschn. 4.6] eingeführte Definition kann die durch $\boldsymbol{\theta}$ parametrisierte polynomiale Modellstruktur (2.7) als identifizierbar bezeichnet werden, wenn

$$\begin{aligned} \mathbf{G}(z, \boldsymbol{\theta}_1) = \mathbf{G}(z, \boldsymbol{\theta}_2) \quad \text{und} \quad \mathbf{H}(z, \boldsymbol{\theta}_1) = \mathbf{H}(z, \boldsymbol{\theta}_2) \quad \text{für fast alle } z \\ \Rightarrow \boldsymbol{\theta}_1 = \boldsymbol{\theta}_2 \quad \forall \boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \in \mathcal{D}_{\mathcal{M}}. \end{aligned} \quad (\text{A.19})$$

²Ein Prozess, dessen Realisierungen konstante Signale sind, deren Wert jedoch zufällig ausgewählt wird, ist stationär aber nicht ergodisch [HW14, Abschn. 2.3].

Ein offensichtliches Beispiel für eine nicht identifizierbare Modellstruktur ist das OE-Modell $y[t] = \frac{B(z)(z-a)}{F(z)(z-a)}u[t] + e[t]$, wobei $a \in \mathbb{R}$ einen Modellparameter kennzeichnet. Für eine ausführliche Diskussion der Identifizierbarkeitsproblematik von LTI-Systemen sei auf [Lju12, Kap. 4] verwiesen.

Bemerkung A.3 (Überparametrisierung): Umgangssprachlich wird oftmals auch der nicht genau spezifizierte Begriff der „Überparametrisierung“ verwendet. Dabei sollte jedoch beachtet werden, dass mit diesem Begriff eventuell auch ausgedrückt wird, dass ein Modell \mathcal{M} mehr Parameter enthält, als für die Beschreibung des tatsächlichen Systems \mathcal{S} notwendig wären. Wird für die Identifikation eines ARX-Prozesses beispielsweise ein ARMAX-Modell verwendet, dann ist \mathcal{M} entsprechend des genannten Verständnisses zwar überparametrisiert, jedoch identifizierbar. \triangleleft

A.2.4.2 Erwartungstreue und Bias

Eine Schätzung $\hat{\boldsymbol{\theta}}_N$ heißt *erwartungstreu* bzw. *biasfrei* (unbiased), falls für eine beliebige Anzahl an Messwerten N

$$\mathbb{E}\{\hat{\boldsymbol{\theta}}_N\} = \boldsymbol{\theta}_0 \quad (\text{A.20})$$

gilt. Entsprechend wird die Schätzung als *verzerrt* (biased) bezeichnet, wenn $\mathbb{E}\{\hat{\boldsymbol{\theta}}_N\} \neq \boldsymbol{\theta}_0$ ist, wobei der systematische Fehler $\mathbf{B} = \mathbb{E}\{\hat{\boldsymbol{\theta}}_N\} - \boldsymbol{\theta}_0 \neq \mathbf{0}$ den *Bias* der Schätzung kennzeichnet. Gilt erwartungstreue hingegen lediglich für eine unendliche Anzahl an Messwerten, d. h.

$$\lim_{N \rightarrow \infty} \mathbb{E}\{\hat{\boldsymbol{\theta}}_N\} = \boldsymbol{\theta}_0, \quad (\text{A.21})$$

dann wird $\hat{\boldsymbol{\theta}}_N$ als *asymptotisch erwartungstreu* bezeichnet [Ise92a, App. 4.2].

A.2.4.3 Konsistenz

Eine Schätzung $\hat{\boldsymbol{\theta}}_N$ wird *konsistent* genannt, falls der Schätzwert für eine unendliche Anzahl an Messwerten N mit Wahrscheinlichkeit 1 gegen den wahren Wert $\boldsymbol{\theta}_0$ konvergiert [Ise92a, App. 4.2]:

$$\lim_{N \rightarrow \infty} \mathbb{P}(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_0 = \mathbf{0}) = 1 \quad (\text{A.22})$$

Man beachte, dass entsprechend der Definition A.2.4.1 lediglich identifizierbare Modellstrukturen konsistent geschätzt werden können.

A.2.4.4 Likelihoodfunktion und statistische Effizienz

Sei \mathbf{X} eine Zufallsgröße, deren Wahrscheinlichkeitsdichte $f_{\boldsymbol{\theta}}(\mathbf{x})$ von einem Parametervektor $\boldsymbol{\theta}$ abhängig ist. Die Funktion $L_x : \mathcal{D}_{\mathcal{M}} \rightarrow \mathbb{R}$, mit

$$L_x(\boldsymbol{\theta}) = f_{\boldsymbol{\theta}}(\mathbf{x}), \quad (\text{A.23})$$

heißt *Likelihoodfunktion* von \mathbf{X} .

Die Likelihoodfunktion kann als Maß für die Plausibilität der Beobachtung \mathbf{x} bei gegebenen Parameter $\boldsymbol{\theta}$ interpretiert werden [SS89, App. B.3]. Diese ist zentraler Bestandteil der *Maximum-Likelihoodmethode* (ML-Methode), welche auf der Maximierung der (logarithmierten) Likelihoodfunktion basiert. Dabei ist zu beachten, dass $\boldsymbol{\theta}$ als fester Parameter interpretiert wird. Im Gegensatz dazu wird dieser im Rahmen der *Maximum a posteriori*-Schätzung (MAP-Schätzung) als Zufallsgröße interpretiert, weswegen für die Likelihood-Funktion in diesem Zusammenhang auch $f(\mathbf{x}|\boldsymbol{\theta})$ geschrieben wird [SS89; SLD15].

Eine erwartungstreue Schätzung $\hat{\boldsymbol{\theta}}_N$ wird als *statistisch effizient* bezeichnet, falls sie die kleinstmögliche Varianz besitzt [Ise92a], wobei der untere Grenzwert durch die CRAMÉR-RAO-Ungleichung

$$\begin{aligned} \text{Cov}\{\hat{\boldsymbol{\theta}}_N\} &\geq \left(\mathbb{E} \left\{ \left(\frac{\partial \log L_x(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right)^\top \frac{\partial \log L_x(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \Bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \right\} \right)^{-1} \\ &= - \left(\mathbb{E} \left\{ \frac{\partial^2 \log L_x(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^2} \Bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \right\} \right)^{-1} \end{aligned} \quad (\text{A.24})$$

gegeben ist³ [Lju12, Abschn. 7.4]. Wird der untere Grenzwert hingegen erst für $N \rightarrow \infty$ erreicht, dann ist die Schätzung *asymptotisch statistisch effizient*.

A.3 Lemmas

Lemma A.1

Seien \mathbf{X}_t $n \times n$ - und \mathbf{Y}_t n -dimensionale mittelwertfreie, stationäre (\mathbf{X}_t sogar quasi-stationär), möglicherweise korrelierte multivariate ARMA-Prozesse und

$$\mathbf{Z}_N = \frac{1}{\sqrt{N}} \sum_{t=1}^N \mathbf{X}_t \mathbf{Y}_t. \quad (\text{A.25})$$

Falls der Grenzwert $\boldsymbol{\Lambda}_Z = \lim_{N \rightarrow \infty} \mathbb{E} \{ \mathbf{Z}_N \mathbf{Z}_N^\top \}$ existiert, ist die Zufallsgröße \mathbf{Z}_N asymptotisch normalverteilt [SS89, App. B.1], d. h.

$$\mathbf{Z}_N \xrightarrow{i.V.} \mathcal{N}(0, \boldsymbol{\Lambda}_Z). \quad (\text{A.26})$$

◁

³Es gilt $\mathbf{A} \geq \mathbf{B}$, wenn $\mathbf{A} - \mathbf{B}$ positiv semidefinit ist.

Lemma A.2

Sei \mathbf{X}_N eine Folge n -dimensionaler Zufallsgrößen, die asymptotisch gegen eine Normalverteilung mit $\mathcal{N}(0, \mathbf{\Lambda}_X)$ konvergiert, sowie \mathbf{A}_N und \mathbf{b}_N Folgen zufälliger Matrizen bzw. Vektoren, welche in Wahrscheinlichkeit gegen $\mathbf{A} \in \mathbb{R}^{n \times n}$ bzw. $\mathbf{b} \in \mathbb{R}^n$ konvergieren. Dann konvergiert die Zufallsgröße

$$\mathbf{Y}_N = \mathbf{A}_N \mathbf{X}_N + \mathbf{b}_N \tag{A.27}$$

in Verteilung gegen $\mathcal{N}(\mathbf{b}, \mathbf{A}\mathbf{\Lambda}_X\mathbf{A}^\top)$ [SS89, App. B.1]. ◁

Anhang B

Ergänzungen zur PEM

B.1 Gradient des Einschnitt Prädiktionsfehlers

Für die numerische Lösung der Optimierungsaufgabe (3.5) wird der Gradient der Prädiktionsfehlerfolge $\varepsilon[t|t-k]$ benötigt. Dabei ist es sinnvoll, zunächst die Ableitungen $\partial\varepsilon[t|t-1]/\partial\theta_i$ zu erzeugen.

Im folgenden wird von der allgemeinen Modellstruktur (2.36) ausgegangen, welche der Differenzgleichung

$$CF\varepsilon[t|t-1] = ADFy[t] - BDu[t] \quad (\text{B.1})$$

entspricht, wobei die Argumente q^{-1} und $\boldsymbol{\theta}$ unterdrückt wurden. Durch Differentiation von (B.1) ergeben sich die Ableitungen daher zu (Vergleiche [Lju12, Abschn. 10.3]):

$$C \frac{\partial\varepsilon[t, \boldsymbol{\theta}]}{\partial a^{[i]}} = Dy[t-i] \quad (\text{B.2a})$$

$$CF \frac{\partial\varepsilon[t, \boldsymbol{\theta}]}{\partial b^{[i]}} = -Du[t-i] \quad (\text{B.2b})$$

$$C \frac{\partial\varepsilon[t, \boldsymbol{\theta}]}{\partial c^{[i]}} = -\varepsilon[t-i, \boldsymbol{\theta}] \quad (\text{B.2c})$$

$$CF \frac{\partial\varepsilon[t, \boldsymbol{\theta}]}{\partial d^{[i]}} = AFy[t-i] - Bu[t-i] \quad (\text{B.2d})$$

$$CF \frac{\partial\varepsilon[t, \boldsymbol{\theta}]}{\partial f^{[i]}} = ADy[t-i] - C\varepsilon[t-i] = \frac{BD}{F}u[t-i] \quad (\text{B.2e})$$

Im nächsten Schritt können mittels (3.16) die Gradienten $\partial\varepsilon[t|t-k]/\partial\theta_i$ berechnet werden. Dabei werden die Ableitungen der Impulsantwort $\partial h[j]/\partial\theta_i$ des Rauschfilters $H(q^{-1}) = C/AD$ benötigt. Diese wiederum ergeben sich durch Differentiation der Beziehung (2.15) zu

$$\frac{\partial h[j]}{\partial a^{[l]}} = - \sum_{i=1}^{\min(j, n_R)} d^{[i-l]} h[j-i] + \frac{\partial h[j-i]}{\partial a^{[l]}} r^{[i]}, \quad (\text{B.3a})$$

$$\frac{\partial h[j]}{\partial d^{[l]}} = - \sum_{i=1}^{\min(j, n_R)} a^{[i-l]} h[j-i] + \frac{\partial h[j-i]}{\partial d^{[l]}} r^{[i]} \quad \text{und} \quad (\text{B.3b})$$

$$\frac{\partial h[j]}{\partial c^{[l]}} = \underbrace{1}_{\substack{\text{wenn } j=l, \\ 0 \text{ sonst}}} - \sum_{i=1}^{\min(j, n_R)} \frac{\partial h[j-i]}{\partial c^{[l]}} r^{[i]}, \quad (\text{B.3c})$$

wobei $r^{[i]}$ die Koeffizienten des Polynoms $R(q^{-1}) := A(q^{-1})D(q^{-1})$ darstellen.

Bemerkung B.1: Für die Implementierung der PEM für die in Abbildung 2.6 (b) dargestellte Modellstruktur werden zusätzlich die Gradienten von (2.53) bezüglich der Parameter $\boldsymbol{\eta}$ benötigt, wobei $\boldsymbol{\eta}$ den Parametervektor der Nichtlinearität $f(u[t], y[t], \boldsymbol{\eta})$ kennzeichnet. Diese können analog zu (B.2) durch

$$CF \frac{\partial \varepsilon[t, \boldsymbol{\theta}, \boldsymbol{\eta}]}{\partial \eta_i} = -BD \frac{\partial f(u[t], y[t], \boldsymbol{\eta})}{\partial \eta_i} \quad (\text{B.4})$$

berechnet werden, wobei wiederum die Parametrisierung des LTI-Blocks (2.38) zugrunde gelegt wurde. \triangleleft

B.2 Schätzung von Anfangswerten

Nachfolgend werden die in Abschnitt 3.2.1 erwähnten Initialisierungsmethoden für das Residuenfilter (2.27) und das Gradientenfilter (3.16) erläutert. Dabei wird von der Parametrierung (2.36) ausgegangen, sodass sich aus (2.27) mit $G = B/AF$ und $H = C/AD$ die Filtergleichung

$$\underbrace{CF}_{L(q^{-1})} \varepsilon[t|t-k] = \underbrace{H^{[0, k-1]} D (AFy[t] - Bu[t])}_{=: w[t]} \quad (\text{B.5})$$

mit den in (2.37) definierten Polynomen ergibt. Da die Folge $w[t]$ erst ab dem Zeitpunkt $t_s = \max(n_a + n_f, n_b) + n_d + k$ exakt berechnet werden kann, besteht die Aufgabe daher darin die Anfangswerte $\varepsilon[t|t-k]$ und deren Gradienten $\partial \varepsilon[t|t-k] / \partial \boldsymbol{\theta}$ für $t < t_s$ zu schätzen.

Bemerkung B.2: Es sei angemerkt, dass die Anfangswerte asymptotisch stabiler LTI-Systeme, aufgrund des exponentiell abklingenden Charakters des transienten Vorgangs, im Allgemeinen nicht konsistent geschätzt werden können [SS89, Abschn. 12.7]. \triangleleft

B.2.1 Methode der kleinsten Fehlerquadrate

Eine mögliche Herangehensweise basiert auf der Berechnung der Anfangswerte mittels der Methode der kleinsten Fehlerquadrate. Zu diesem Zweck ist es sinnvoll, die

Filterdarstellung (B.5) in die Regelungsnormalform

$$\begin{aligned} \mathbf{x}[t+1] &= \underbrace{\begin{pmatrix} -l^{[1]} & -l^{[2]} & \dots & -l^{[n_L]} \\ 1 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & 1 & 0 \end{pmatrix}}_A \mathbf{x}[t] + \underbrace{\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}}_B w[t] \\ \varepsilon[t|t-k] &= \underbrace{\begin{pmatrix} 1 & 0 & \dots & 0 \end{pmatrix}}_C \mathbf{x}[t] + \underbrace{1}_D w[t]. \end{aligned} \quad (\text{B.6})$$

zu transformieren, wobei $l^{[i]}$ die Koeffizienten des Polynoms $L(q^{-1}) = C(q^{-1})F(q^{-1})$ darstellen. Durch wiederholtes ineinander einsetzen von (B.6) lässt sich ein überbestimmtes lineares Gleichungssystem der Form

$$\boldsymbol{\epsilon} = \underbrace{\begin{pmatrix} \mathbf{CA} \\ \mathbf{CA}^2 \\ \vdots \\ \mathbf{CA}^{N-t_s+1} \end{pmatrix}}_{\Psi} \boldsymbol{\epsilon}_0 + \underbrace{\begin{pmatrix} \mathbf{D} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{CB} & \mathbf{D} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \mathbf{0} \\ \mathbf{CA}^{N-t_s-1}\mathbf{B} & \dots & \mathbf{CB} & \mathbf{D} \end{pmatrix}}_{\Omega} \mathbf{w}, \quad (\text{B.7})$$

mit $\boldsymbol{\epsilon} = \text{vec} \{\varepsilon[t]\}$, $\mathbf{w} = \text{vec} \{w[t]\}$ ($t = t_s, \dots, N$) und $\boldsymbol{\epsilon}_0 = (\varepsilon[t_s - 1], \dots, \varepsilon[t_s - n_L])^\top$ gewinnen. Da $\varepsilon[t]$ mittelwertfreies weißes Rauschen darstellt, kann zur Schätzung die MkQ-Lösung

$$\boldsymbol{\epsilon}_0 = -(\Psi\Psi^\top)^{-1}\Psi^\top(\Omega\mathbf{w}) \quad (\text{B.8})$$

verwendet werden.

Für die Berechnung der Anfangswerte des Gradientenfilters (3.16) wird in ähnlicher Weise vorgegangen. Hierfür differenziert man (B.5) zunächst bezüglich des Parameters θ_i , wodurch sich

$$L(q^{-1})\frac{\partial\varepsilon[t]}{\partial\theta_i} = \underbrace{\frac{\partial w[t]}{\partial\theta_i} - \frac{\partial L(q^{-1})}{\partial\theta_i}\varepsilon[t]}_{\delta_{\theta_i}[t]} \quad (\text{B.9})$$

ergibt, wobei $\partial L(q^{-1})/\partial\theta_i$ als elementweise Ableitung der Filterkoeffizienten verstanden wird (vergleiche (3.17)). Transformiert man (B.9) wiederum in die Regelungsnormalform, dann ergibt sich in analoger Weise ein überbestimmtes Gleichungssystem, dass mit $\boldsymbol{\delta}_{\theta_i} = (\delta_{\theta_i}[t_s], \dots, \delta_{\theta_i}[N])^\top$ im Sinne der kleinsten Fehlerquadrate gelöst werden kann:

$$\frac{\partial\boldsymbol{\epsilon}_0}{\partial\theta_i} = -(\Psi\Psi^\top)^{-1}\Psi^\top(\Omega\boldsymbol{\delta}_{\theta_i}) \quad (\text{B.10})$$

Bemerkung B.3 (Berechnung der Folgen $\Omega\mathbf{w}$ und $\Omega\boldsymbol{\delta}_{\theta_i}$): Es sei darauf hingewiesen, dass auf die Berechnung von $\Omega\mathbf{w}$ bzw. $\Omega\boldsymbol{\delta}_{\theta_i}$ durch Matrix-Vektormultiplikation aus Effizienzgründen verzichtet werden sollte. Stattdessen ist es meist effizienter, die entsprechenden Folgen durch lineare Filterung, d. h. durch $1/L(q^{-1})w[t]$ bzw. $1/L(q^{-1})\delta_{\theta_i}[t]$, zu erzeugen. \triangleleft

B.2.2 Backforecasting

Die in (B.5) angegebene Filtergleichung

$$w[t] = L(q^{-1})\varepsilon[t] \quad (\text{B.11})$$

beschreibt einen MA-Prozess und ist daher ein Spezialfall eines allgemeinen ARMA-Prozesses. Eine Besonderheit dieser Prozessklasse ist die Eigenschaft, dass deren Kovarianzfunktion eine gerade Funktion ist. Dies bedeutet, dass die Kovarianz zwischen den Beobachtungen $w[t]$ und $w[t - 1]$ der Kovarianz zwischen $w[t]$ und $w[t + 1]$ entspricht, d. h. $\text{Cov}\{w[t], w[t - 1]\} = \text{Cov}\{w[t], w[t + 1]\}$. Der zeitlich umgekehrte Prozess

$$w[t] = L(q)\varepsilon[t] \quad (\text{B.12})$$

erzeugt daher die gleiche Kovarianzfunktion wie (B.11). Diese Feststellung kann für die Schätzung geeigneter Anfangswerte von (B.11) verwendet werden [Knu94; Knu96]:

- (1) Berechne $\varepsilon_b[t] = 1/L(q)w[t]$ durch Rückwärtsfilterung der Folge $w[t]$. Die Anfangswerte $\varepsilon_b[N + 1], \dots, \varepsilon_b[N + n_L]$ werden zu 0 initialisiert und es wird angenommen, dass der transiente Anteil abgeklungen ist, bevor der Zeitpunkt t_s erreicht wurde.
- (2) Prädiziere $w[t]$ für $t < t_s$ mittels

$$w[t_s - i] = \varepsilon_b[t_s - i] + l^{[1]}\varepsilon_b[t_s - i + 1] + \dots + l^{[n_L]}\varepsilon_b[t_s - i + n_L], \quad (\text{B.13})$$

wobei $\varepsilon_b[t] = 0$ für alle $t < t_s$ gilt. ($i = 1, \dots, n_L$).

- (3) Schätze die Anfangswerte $\varepsilon[t]$ für $t < t_s$ mittels

$$w[t_s - i] = \varepsilon[t_s - i] + l^{[1]}\varepsilon[t_s - i - 1] + \dots + l^{[n_L]}\varepsilon[t_s - i - n_L], \quad (\text{B.14})$$

wobei $\varepsilon[t] = 0$ für alle $t < t_s - n_L$ gilt. ($i = 1, \dots, n_L$).

Die Schätzung der Initialisierung des Gradientenfilters (3.16) erfolgt analog, wobei jedoch das Vorwärts- bzw. Rückwärtsmodell von (B.9) zu verwenden sind.

Bemerkung B.4: Es sei darauf hingewiesen, dass die Berechnung der Anfangswerte bei der PEM in jedem Iterationsschritt der Optimierung durchgeführt werden muss. Da deren Berechnung mittels MkQ vergleichsweise rechenaufwendig ist, sollte insbesondere im Falle einer hohen Polynomordnung von $L(q^{-1})$ die Backforecasting Methode verwendet werden. Diese Problematik entsteht oftmals bei der Verwendung saisonaler Modellansätze (siehe Abschn. 2.4.1 und 5.1.2).

Der beispielhafte Vergleich der mittleren Rechenzeit für die Berechnung der Anfangswerte des MA-Filters (B.5) ist in der Abbildung B.1 (links) dargestellt. Dabei lässt sich erkennen, dass die Rechenzeit des MkQ-Verfahrens den Erwartungen entsprechend für hohe Polynomordnungen n_L wesentlich höher als bei der Backforecasting-Methode ist. In der Darstellung rechts erkennt man, dass die MkQ-Methode insbesondere für kurze Datensatzlängen N und Nullstellen von L in der Nähe des Einheitskreises genauere Ergebnisse liefert. Dies lässt sich dadurch begründen, dass bei der Backforecasting-Methode der transiente Anteil von $\varepsilon_b[t]$ bei der Rückwärtsfilterung von $w[t]$ noch nicht ausreichend abgeklungen ist, bevor der Start- bzw. Endzeitpunkt t_s erreicht wurde.

◁

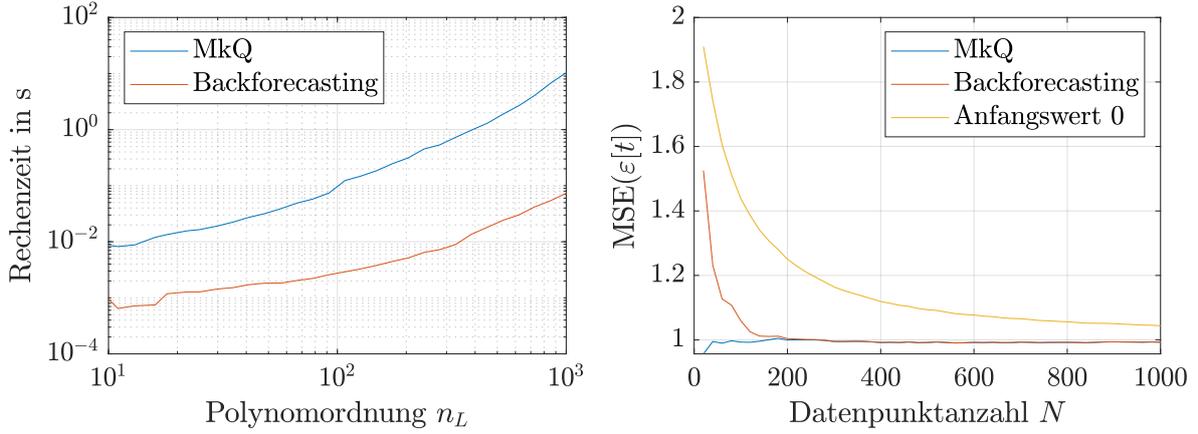


Abbildung B.1: Links: Vergleich der über 100-Realisierungen von $e[t]$ der Länge $N = 10000$ gemittelten Rechenzeit für die Bestimmung der Anfangswerte des MA-Filters $w[t] = (1 - 0.99q^{-n_L})e[t]$ in Abhängigkeit der Polynomordnung n_L mit $\lambda_e = 1$ und $e[t] = 0 \forall t < 1$ mittels MkQ- bzw. Backforecasting-Methode (Core i7-3520M@2.9GHz, 8GB RAM, Windows 10, Matlab 2018b). Rechts: Vergleich des über 100-Realisierungen gemittelten MSE des geschätzten Residuums $\varepsilon[t]$ des MA-Filters $w[t] = (1 - 0.99q^{-1})e[t]$ in Abhängigkeit der Datenpunktzahl N mit $\lambda_e = 1$ und $e[0] = 10$.

B.3 Herleitung des Autoleistungsdichtespektrums der Mehrschritt-PEM

Im Folgenden wird die Berechnung des in Abschnitt 3.2.2 benötigten Autoleistungsdichtespektrums $\phi_{\varepsilon_k}(\omega, \boldsymbol{\theta})$ demonstriert (vgl. [Lju12, Abschn. 8.5]):

Unter der Annahme, dass das zu identifizierende System der Beschreibung (2.7) genügt, kann die Prädiktionsfehlerfolge durch

$$\begin{aligned}
 \varepsilon[t|t-k, \boldsymbol{\theta}] &= \frac{H^{[0,k-1]}(\boldsymbol{\theta})}{H(\boldsymbol{\theta})} \left(y[t] - G(\boldsymbol{\theta})u[t] \right) \\
 &= \frac{H^{[0,k-1]}(\boldsymbol{\theta})}{H(\boldsymbol{\theta})} \left((G_0 - G(\boldsymbol{\theta}))u[t] + H_0e[t] \right) \\
 &= \frac{H^{[0,k-1]}(\boldsymbol{\theta})}{H(\boldsymbol{\theta})} \left(\underbrace{(G_0 - G(\boldsymbol{\theta}))}_{\Delta G(\boldsymbol{\theta})} u[t] + \right. \\
 &\quad \left. \underbrace{\left(H_0 - H_0^{[0,k-1]} \frac{H(\boldsymbol{\theta})}{H^{[0,k-1]}(\boldsymbol{\theta})} \right)}_{\Delta H(\boldsymbol{\theta})} e[t] \right) + H_0^{[0,k-1]} e[t] \\
 &= \frac{H^{[0,k-1]}(\boldsymbol{\theta})}{H(\boldsymbol{\theta})} \begin{bmatrix} \Delta G(\boldsymbol{\theta}) & \Delta H(\boldsymbol{\theta}) \end{bmatrix} \begin{bmatrix} u[t] \\ e[t] \end{bmatrix} + H_0^{[0,k-1]} e[t]
 \end{aligned} \tag{B.15}$$

ausgedrückt werden, wobei $G_0 := G(\boldsymbol{\theta}_0)$ und $H_0 := H(\boldsymbol{\theta}_0)$ die Übertragungsfunktionen

des „wahren Systems“ kennzeichnen. Durch Anwenden der Transformationsgleichung für Spektraldichten (A.15) und unter Berücksichtigung von $\phi_{eu}(\omega) = \overline{\phi_{ue}}(\omega)$ ergibt sich dann

$$\begin{aligned}
 \phi_{\varepsilon_k}(\omega, \boldsymbol{\theta}) &= \left| \frac{H^{[0,k-1]}}{H} \right|^2 [\Delta G \quad \Delta H] \begin{bmatrix} \phi_u & \phi_{ue} \\ \phi_{eu} & \lambda_e \end{bmatrix} \begin{bmatrix} \overline{\Delta G} \\ \overline{\Delta H} \end{bmatrix} + |H_0^{[0,k-1]}|^2 \lambda_e \\
 &= \left| \frac{H^{[0,k-1]}}{H} \right|^2 \left(\left(\Delta G \overline{\Delta G} + \Delta H \overline{\Delta G} \frac{\phi_{eu}}{\phi_u} + \right. \right. \\
 &\quad \left. \left. \Delta G \overline{\Delta H} \frac{\overline{\phi_{eu}}}{\phi_u} \right) \phi_u + \Delta H \overline{\Delta H} \lambda_e \right) + |H_0^{[0,k-1]}|^2 \lambda_e \\
 &= \left| \frac{H^{[0,k-1]}}{H} \right|^2 \left(\left| \Delta G + \Delta H \frac{\phi_{eu}}{\phi_u} \right|^2 \phi_u + \right. \\
 &\quad \left. |\Delta H|^2 \left(\lambda_e - \left| \frac{\phi_{eu}}{\phi_u} \right|^2 \phi_u \right) \right) + |H_0^{[0,k-1]}|^2 \lambda_e,
 \end{aligned} \tag{B.16}$$

wobei die Argumente $\boldsymbol{\theta}$, ω und $e^{-j\omega}$ aus Gründen der Übersicht unterdrückt wurden. Durch Einsetzen der letzten Gleichung in (3.21) ergibt sich schließlich die Beziehung (3.22).

B.4 Berechnung der Schätzfehlerverteilung

In diesem Abschnitt soll auf den theoretischen Vergleich der k - und der 1-Schritt PEM eingegangen werden. Der dabei grundlegende Gedanke ist es den Schätzwert $\hat{\boldsymbol{\theta}}_N$ als Zufallsgröße mit gesuchter Wahrscheinlichkeitsdichteverteilung zu betrachten, wobei insbesondere die Kovarianzmatrix $\boldsymbol{\Lambda}_\theta$ der Schätzung $\hat{\boldsymbol{\theta}}_N$ von Interesse ist. Die Berechnung dieser kann in Analogie zu den in [SS89, Abschn. 7.5] und [Lju12, Abschn. 9.2] dargestellten Ausführungen erfolgen:

Sei $\hat{\boldsymbol{\theta}}_N$ das Minimum des Kostenfunktional

$$V_N(\boldsymbol{\theta}) = \frac{1}{N} \sum_{t=1}^N \varepsilon^2[t|t-k, \boldsymbol{\theta}] \tag{B.17}$$

d. h.

$$V_N'^T(\hat{\boldsymbol{\theta}}_N) = \mathbf{0}. \tag{B.18}$$

In der Umgebung von $\boldsymbol{\theta}_0$ gilt

$$\mathbf{0} = V_N'^T(\hat{\boldsymbol{\theta}}_N) \approx V_N'^T(\boldsymbol{\theta}_0) + V_N''(\boldsymbol{\theta}_0)(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_0), \tag{B.19}$$

wobei V_N' den Gradienten und V_N'' die Hesse-Matrix von V bezüglich $\boldsymbol{\theta}$ darstellen. Für $N \rightarrow \infty$ konvergiert $V_N''(\boldsymbol{\theta}_0)$ mit Wahrscheinlichkeit 1 gegen $V_\infty''(\boldsymbol{\theta}_0)$, sodass sich unter der Annahme der Regularität die Beziehung

$$\sqrt{N}(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_0) \approx -[V_\infty''(\boldsymbol{\theta}_0)]^{-1}[\sqrt{N}V_N'^T(\boldsymbol{\theta}_0)] \tag{B.20}$$

gewinnen lässt. Wegen Lemma A.1 stellt $\sqrt{N}V_N^{\prime\top}(\boldsymbol{\theta}_0)$ eine n_θ -dimensionale asymptotisch normalverteilte Zufallsgröße mit Erwartungswert $\mathbf{0}$ und Kovarianzmatrix $\boldsymbol{\Lambda}_k \in \mathbb{R}^{n_\theta \times n_\theta}$ dar. Durch Anwendung von Lemma A.2 ergibt sich daher

$$\sqrt{N}(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_0) \xrightarrow{i.V.} \mathcal{N}(\mathbf{0}, [V_\infty''(\boldsymbol{\theta}_0)]^{-1} \boldsymbol{\Lambda}_k(\boldsymbol{\theta}_0) [V_\infty''(\boldsymbol{\theta}_0)]^{-1}). \quad (\text{B.21})$$

Für die Berechnung der Schätzfehlerkovarianz müssen die Matrizen $V_\infty''(\boldsymbol{\theta}_0)$ und $\boldsymbol{\Lambda}_k$ berechnet werden. Wegen (3.15) und $\varepsilon_1[t, \boldsymbol{\theta}_0] \equiv e[t]$ ergeben sich für die Ableitungen des Kostenfunktional die Beziehungen

$$\begin{aligned} V_N'(\boldsymbol{\theta}_0) &= \frac{2}{N} \sum_{t=1}^N \varepsilon_k'[t, \boldsymbol{\theta}_0] H^{[0, k-1]}(q^{-1}, \boldsymbol{\theta}_0) e[t] \\ V_N''(\boldsymbol{\theta}_0) &= \frac{2}{N} \sum_{t=1}^N \varepsilon_k'^{\top}[t, \boldsymbol{\theta}_0] \varepsilon_k'[t, \boldsymbol{\theta}_0] + \varepsilon_k''[t, \boldsymbol{\theta}_0] H^{[0, k-1]}(q^{-1}, \boldsymbol{\theta}_0) e[t], \end{aligned} \quad (\text{B.22})$$

wobei die k -Schritt Prädiktionsfehlerfolge und deren Gradient abkürzend durch $\varepsilon_k[t]$ bzw. $\varepsilon_k'[t]$ gekennzeichnet seien. Der Term $H^{[0, k-1]}(q^{-1}, \boldsymbol{\theta}_0) e[t]$ ist lediglich von $e[t], \dots, e[t - k + 1]$ abhängig, wohingegen die Ableitungen $\varepsilon_k'[t]$ und $\varepsilon_k''[t]$ für $\boldsymbol{\theta} = \boldsymbol{\theta}_0$ nur von $e[t]$ mit $t \leq t - k$ abhängen, sodass sich für

$$V_\infty''(\boldsymbol{\theta}_0) = 2 \mathbf{E} \left\{ \varepsilon_k'^{\top}[t, \boldsymbol{\theta}_0] \varepsilon_k'[t, \boldsymbol{\theta}_0] \right\} \quad (\text{B.23})$$

ergibt.

Für die Kovarianzmatrix der Zufallsgröße $\sqrt{N}V_N^{\prime\top}(\boldsymbol{\theta}_0)$ ergibt sich wegen Lemma A.1 zunächst

$$\begin{aligned} \boldsymbol{\Lambda}_k &= \lim_{N \rightarrow \infty} N \mathbf{E} \left\{ V_N^{\prime\top}(\boldsymbol{\theta}_0) V_N'(\boldsymbol{\theta}_0) \right\} \\ &= \lim_{N \rightarrow \infty} \frac{4}{N} \sum_{t=1}^N \sum_{s=1}^N \mathbf{E} \left\{ \varepsilon_k[t, \boldsymbol{\theta}_0] \varepsilon_k'^{\top}[t, \boldsymbol{\theta}_0] \varepsilon_k'[s, \boldsymbol{\theta}_0] \varepsilon_k[s, \boldsymbol{\theta}_0] \right\} \\ &= \lim_{N \rightarrow \infty} \frac{4}{N} \sum_{t=1}^N \sum_{s=1}^{t-k} \mathbf{E} \left\{ \varepsilon_k[t, \boldsymbol{\theta}_0] \right\} \mathbf{E} \left\{ \varepsilon_k'^{\top}[t, \boldsymbol{\theta}_0] \varepsilon_k'[s, \boldsymbol{\theta}_0] \varepsilon_k[s, \boldsymbol{\theta}_0] \right\} \\ &\quad + \lim_{N \rightarrow \infty} \frac{4}{N} \sum_{t=1}^N \sum_{s=t-k+1}^{t+k-1} \mathbf{E} \left\{ \varepsilon_k[t, \boldsymbol{\theta}_0] \varepsilon_k'^{\top}[t, \boldsymbol{\theta}_0] \varepsilon_k'[s, \boldsymbol{\theta}_0] \varepsilon_k[s, \boldsymbol{\theta}_0] \right\} \\ &\quad + \lim_{N \rightarrow \infty} \frac{4}{N} \sum_{t=1}^N \sum_{s=t+k}^N \mathbf{E} \left\{ \varepsilon_k[t, \boldsymbol{\theta}_0] \varepsilon_k'^{\top}[t, \boldsymbol{\theta}_0] \varepsilon_k'[s, \boldsymbol{\theta}_0] \right\} \mathbf{E} \left\{ \varepsilon_k[s, \boldsymbol{\theta}_0] \right\} \\ &= \lim_{N \rightarrow \infty} \frac{4}{N} \sum_{t=1}^N \sum_{s=t-k+1}^{t+k-1} \mathbf{E} \left\{ \varepsilon_k[t, \boldsymbol{\theta}_0] \varepsilon_k'^{\top}[t, \boldsymbol{\theta}_0] \varepsilon_k'[s, \boldsymbol{\theta}_0] \varepsilon_k[s, \boldsymbol{\theta}_0] \right\}, \end{aligned} \quad (\text{B.24})$$

wobei wiederum die Unabhängigkeit von $\varepsilon_k[t, \boldsymbol{\theta}_0]$ und $\varepsilon_k[s, \boldsymbol{\theta}_0]$ für alle $s \neq -k+1, \dots, k-1$ sowie $\varepsilon_k'[s, \boldsymbol{\theta}_0]$ für $s \leq t$ berücksichtigt wurde. Die Auswertung des verbleibenden Terms

erfolgt elementweise:

$$\begin{aligned}
 s = t : \lim_{N \rightarrow \infty} \frac{4}{N} \sum_{t=1}^N \mathbb{E} \left\{ \underbrace{\varepsilon_k^2[t, \boldsymbol{\theta}_0] \varepsilon_k'^{\top}[t, \boldsymbol{\theta}_0] \varepsilon_k'[t, \boldsymbol{\theta}_0]}_{= \mathbb{E}\{\varepsilon_k^2[t, \boldsymbol{\theta}_0]\} \mathbb{E}\{\varepsilon_k'^{\top}[t, \boldsymbol{\theta}_0] \varepsilon_k'[t, \boldsymbol{\theta}_0]\}} \right\} \\
 \stackrel{(2.28)}{=} 4\lambda_e \mathbb{E} \left\{ \varepsilon_k'^{\top}[t, \boldsymbol{\theta}_0] \varepsilon_k'[t, \boldsymbol{\theta}_0] \right\} \sum_{j=0}^{k-1} h^2[j]
 \end{aligned} \tag{B.25}$$

$$\begin{aligned}
 s = t \pm 1 : \lim_{N \rightarrow \infty} \frac{4}{N} \sum_{t=1}^N \mathbb{E} \left\{ (e[t] + h[1]e[t-1] + \dots + h[k-1]e[t-k+1]) \right. \\
 \left. (e[t \pm 1] + \dots + h[k-1]e[t-k+1 \pm 1]) \varepsilon_k'^{\top}[t, \boldsymbol{\theta}_0] \varepsilon_k'[t \pm 1, \boldsymbol{\theta}_0] \right\} \\
 = 4\lambda_e \mathbb{E} \left\{ \varepsilon_k'^{\top}[t, \boldsymbol{\theta}_0] \varepsilon_k'[t \pm 1, \boldsymbol{\theta}_0] \right\} \sum_{j=0}^{k-2} h[j]h[j+1]
 \end{aligned} \tag{B.26}$$

$$\begin{aligned}
 s = t \pm 2 : \lim_{N \rightarrow \infty} \frac{4}{N} \sum_{t=1}^N \mathbb{E} \left\{ (e[t] + h[1]e[t-1] + \dots + h[k-1]e[t-k+1]) \right. \\
 \left. (e[t \pm 2] + \dots + h[k-1]e[t-k+1 \pm 2]) \varepsilon_k'^{\top}[t, \boldsymbol{\theta}_0] \varepsilon_k'[t \pm 2, \boldsymbol{\theta}_0] \right\} \\
 = 4\lambda_e \mathbb{E} \left\{ \varepsilon_k'^{\top}[t, \boldsymbol{\theta}_0] \varepsilon_k'[t \pm 2, \boldsymbol{\theta}_0] \right\} \sum_{j=0}^{k-3} h[j]h[j+2]
 \end{aligned} \tag{B.27}$$

Zusammengefasst ergibt sich für die Schätzfehlerkovarianzmatrix der k -Schritt PEM daher

$$\boldsymbol{\Lambda}_{\theta} = [V_{\infty}''(\boldsymbol{\theta}_0)]^{-1} \boldsymbol{\Lambda}_k [V_{\infty}''(\boldsymbol{\theta}_0)]^{-1}, \tag{B.28}$$

wobei die Matrix $V_{\infty}''(\boldsymbol{\theta}_0)$ durch (B.23) und die Matrix $\boldsymbol{\Lambda}_k$ durch

$$\boldsymbol{\Lambda}_k = 4\lambda_e \sum_{i=-k+1}^{k-1} \mathbb{E} \left\{ \varepsilon_k'^{\top}[t, \boldsymbol{\theta}_0] \varepsilon_k'[t+i, \boldsymbol{\theta}_0] \right\} \sum_{j=0}^{k-1-|i|} h[j, \boldsymbol{\theta}_0] h[j+|i|, \boldsymbol{\theta}_0] \tag{B.29}$$

gegeben ist. Für den Spezialfall $k = 1$ ergibt sich entsprechend

$$\boldsymbol{\Lambda}_{\theta} = \lambda_e \mathbb{E} \left\{ \varepsilon_1'^{\top}[t, \boldsymbol{\theta}_0] \varepsilon_1'[t, \boldsymbol{\theta}_0] \right\}^{-1} \tag{B.30}$$

(Vergleiche [SS89, Abschn. 7.5]).

Anhang C

Thermodynamische Zusammenhänge in der Gebäudetechnik

C.1 Statische Wärmeabgabe von Heizflächen

Die statische Wärmeabgabe von Heizflächen kann im Wesentlichen mittels der Vorlauftemperatur ϑ_{V1} und durch den Volumen- bzw. Massenstrom des Heizmediums beeinflusst werden. Mathematisch wird der Zusammenhang durch die zweite Heizkörpergleichung

$$\frac{\dot{Q}}{\dot{Q}_N} = \left(\frac{\Delta\vartheta_{\text{In}}}{\Delta\vartheta_{\text{In},N}} \right)^n \quad (\text{C.1})$$

beschrieben [RSS11, Abschn. 2.4.4], wobei \dot{Q}_N die Nennleistung und

$$\Delta\vartheta_{\text{In}} = \frac{\vartheta_{V1} - \vartheta_{R1}}{\ln \left(\frac{\vartheta_{V1} - \vartheta_L}{\vartheta_{R1} - \vartheta_L} \right)} \quad (\text{C.2})$$

die logarithmische Übertemperatur der Heizfläche im Arbeitspunkt bzw. im Nennbetriebspunkt kennzeichnen. Weiterhin charakterisiert n den Heizkörperexponenten, welcher den Einfluss gegenüber den Normwerten geänderter Temperaturdifferenzen eines bestimmten Heizkörpertyps auf dessen Wärmeleistung beschreibt und in der Praxis Werte im Bereich 1.0 bis 1.5 annimmt [RSS11, Abschn. 2.4.4]. Dabei ist beispielsweise dann ein großer Wert zu erwarten, wenn die Wärmeabgabe hauptsächlich durch Wärmestrahlung bei relativ hoher Temperatur erfolgt, oder wenn die natürliche Konvektion der Luft eine wichtige Rolle bei der Wärmeabgabe spielt. Ist hingegen die natürliche Konvektion von geringer Bedeutung, oder wird die Wärmeabgabe hauptsächlich durch Wärmeleitung im festen Körper limitiert, dann werden relativ kleine Exponenten n ermittelt. In der Praxis wird meist von den folgenden Werten ausgegangen [RSS11, Abschn. 2.4.4]:

- Fußbodenheizung $n \approx 1.1$
- Plattenheizkörper $n = 1.2 \dots 1.3$
- Konvektoren $n = 1.25 \dots 1.45$

Die beispielhaften Lösungen der transzendenten Gleichung (C.1) für $\Delta\vartheta_{\text{In},N} = 49.8\text{K}$ ¹ sind in der Abbildung C.1 dargestellt. Dabei lässt sich erkennen, dass ein stark nichtlinearer Zusammenhang zwischen dem Volumenstrom und der abgegebenen Heizleistung

¹Dies entspricht den Nennbedingungen $\vartheta_{V1,N} = 75^\circ\text{C}$, $\vartheta_{R1,N} = 65^\circ\text{C}$ und $\vartheta_{L,N} = 20^\circ\text{C}$

\dot{Q} besteht. Die Nichtlinearität bezüglich der Temperaturdifferenz zwischen Vorlauf- und Lufttemperatur wird hingegen nur bei relativ großen Heizkörperexponenten deutlich.

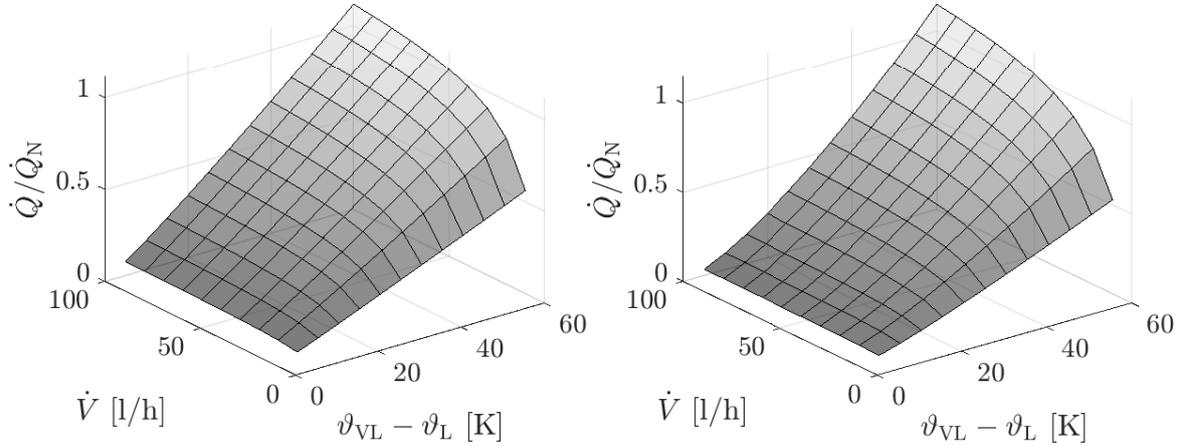


Abbildung C.1: Lösungen der Heizkörpergleichung (C.1) für $\vartheta_{\text{In},N} = 49.8\text{K}$ und $n = 1.1$ (links) sowie $n = 1.45$ (rechts).

C.2 Dichte feuchter Luft

Bei näherungsweise konstantem Druck p_L kann die Dichte feuchter Luft mittels

$$\rho_L = \frac{p_L}{R_f T_L} \quad (\text{C.3})$$

berechnet werden, wobei T_L die absolute Temperatur und

$$R_f = \frac{R_t}{1 - \varphi_L \frac{p_D}{p_L} \left(1 - \frac{R_t}{R_D}\right)} \quad (\text{C.4})$$

die Gaskonstante feuchter und $R_t = 287.06 \text{ J/kgK}$ sowie $R_D = 461.52 \text{ J/kgK}$ die Gaskonstanten trockener Luft bzw. von Wasserdampf bezeichnen. Der Sättigungsdampfdruck p_D von Wasser kann mittels der MAGNUS-Formel

$$p_D = 611.2\text{Pa} \exp\left(\frac{17.62\vartheta_L}{243.12 \text{ }^\circ\text{C} + \vartheta_L}\right), \text{ für } -45^\circ\text{C} \leq \vartheta_L \leq 60^\circ\text{C} \quad (\text{C.5})$$

ermittelt werden [VDI19]. Die auf das Volumen bezogene Wärmekapazität kann dann durch Multiplikation mit der spezifischen Wärmekapazität c_L ermittelt werden, welche in dem in der Gebäudetechnik interessierenden Temperaturbereich als näherungsweise konstant betrachtet werden kann. Die grafische Darstellung der Dichte feuchter Luft ist in Abb. C.2 veranschaulicht.

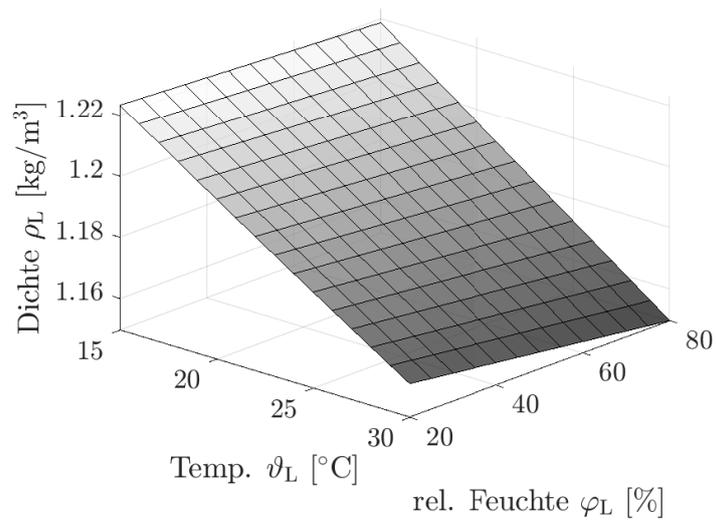


Abbildung C.2: Dichte feuchter Luft in Abhängigkeit der Temperatur und der relativen Feuchte.

Anhang D

Software

Im Rahmen dieser Arbeit ist eine in der Programmiersprache MATLAB implementierte Softwarebibliothek¹ zur Identifikation von MIMO-Modellen des Typs

$$\mathbf{A}(q^{-1})\mathbf{y}[t] = \mathbf{E}^{-1}(q^{-1})\mathbf{B}(q^{-1})\mathbf{f}(\mathbf{u}[t], \mathbf{y}[t], \boldsymbol{\eta}) + \mathbf{C}(q^{-1})\mathbf{D}^{-1}(q^{-1})\mathbf{e}[t], \quad (\text{D.1})$$

entstanden (vgl. Gl. (2.51)), wobei $\mathbf{u}[t] \in \mathbb{R}^{n_u}$ und $\mathbf{y}[t] \in \mathbb{R}^{n_y}$ die zeitdiskreten Ein- bzw. Ausgangssignale und $\mathbf{e}[t] \in \mathbb{R}^{n_y}$ mittelwertfreies² weißes Rauschen mit Kovarianzmatrix $\boldsymbol{\Lambda}_e \in \mathbb{R}^{n_y \times n_y}$ beschreiben. Weiterhin kennzeichnen $\mathbf{f} : \mathbb{R}^{n_u} \times \mathbb{R}^{n_y} \mapsto \mathbb{R}^{n_{\tilde{u}}}$ eine statische Nichtlinearität mit Parametervektor $\boldsymbol{\eta}$ und $\mathbf{A}(q^{-1})$ bis $\mathbf{E}(q^{-1})$ Polynommatrizen der in Tab. 13 angegebenen Dimensionen. Der Spezialfall eines LTI- bzw. Hammerstein-Modells ergibt sich für $\mathbf{f}(\mathbf{u}[t], \mathbf{y}[t], \boldsymbol{\eta}) = \mathbf{u}[t]$ bzw. $\mathbf{f}(\mathbf{u}[t], \mathbf{y}[t], \boldsymbol{\eta}) = \mathbf{f}(\mathbf{u}[t], \boldsymbol{\eta})$.

Matrix	Dimension	Bedingung
\mathbf{A}	$n_y \times n_y$	$\mathbf{A}(0) = \mathbf{I}$
\mathbf{B}	$n_y \times n_{\tilde{u}}$	$B_{ij}(0) = 0$, falls $f_j(\mathbf{u}, \mathbf{y}) \neq f_j(\mathbf{u})$ ³
$\mathbf{C}, \mathbf{D}, \mathbf{E}$	$n_y \times n_y$	Diagonalmatrizen mit $\mathbf{C}(0) = \mathbf{D}(0) = \mathbf{E}(0) = \mathbf{I}$

Tabelle 13: Polynommatrizen des Modells (D.1)

D.1 Kurzbeschreibung

Das durch Gleichung (D.1) repräsentierte Modell wurde objektorientiert in der Klasse `NsfPolyModel` innerhalb des Paketes `idModels` implementiert. Eine Objektinstanz wird mittels des Konstruktoraufrufs

```
M=idModels.NsfPolyModel(nA,nB,nK,nC,nD,nE,'Prop1',Val1,...,'PropN',ValN);
```

erzeugt, wobei die Matrizen bzw. Vektoren `nA` bis `nE` die gewünschten Polynomordnungen von (D.1) und `nK` die Anzahl der Delays von $\mathbf{B}(q^{-1})$ beschreiben (vgl. Tabelle 13). Die durch `'Prop1', Val1, ...` beispielhaft angedeuteten optionalen Argumente (Name-Wert Paare) können zum Setzen beliebiger Attribute aus der Tabelle 14 verwendet werden.

¹https://github.com/fpaschke/phd_identification_mpc

²Es sei angemerkt, dass sich aus dieser Forderung keine Einschränkung der Allgemeingültigkeit des Modellansatzes ergibt (siehe Bemerkung 2.3).

³Damit ist gemeint, dass die Polynome $B_{ij}(q^{-1})$ für alle $i = 1, \dots, n_y$ mindestens einen Delay besitzen müssen, wenn die Nichtlinearität $f_j(\cdot)$ eine Ausgangsrückkopplung besitzt.

Attribut	Beschreibung
A,B,C,D,E	Polynommatrizen des Modells
PreFilter	Vorfilter $\mathbf{L}(q^{-1})$ bzw. fester Teil des Rauschmodells (siehe Abschn. 3.2)
InputNonlinearity	Nichtlinearität $\mathbf{f}(\mathbf{u}, \mathbf{y}, \boldsymbol{\eta})$
NoiseVariance	Kovarianzmatrix $\boldsymbol{\Lambda}_e$ des Rauschsignals $\mathbf{e}[t]$
Ts	Abtastzeit
TimeUnit	Zeiteinheit
Name	Modellname
Info	Struktur für beliebige Objektinformationen
InputName,InputUnit	Namen bzw. Einheiten der Eingangssignale
OutputName,OutputUnit	Namen bzw. Einheiten der Ausgangssignale
InputMin,InputMax	Minima bzw. Maxima der Eingangssignale
OutputMin,OutputMax	Minima bzw. Maxima der Ausgangssignale
OutputOffset	Offset der Ausgangssignale

Tabelle 14: Attribute der Klasse `idModels.NsfPolyModel`

Die Identifikation des Modells mittels PEM-Methode wird dann mithilfe des Aufrufs

```
Opt = M.identifyOptions('Opt1',Val1,...,'OptN',ValN);
M.identify(y,u,Opt);
```

durchgeführt, wobei \mathbf{u} ($N \times n_u$) und \mathbf{y} ($N \times n_y$) die Ein- bzw. Ausgangssignale sowie 'Opt1' bis 'OptN' definierte Optionsnamen kennzeichnen (siehe Tabelle 15)⁴. Alternativ zur oben angegebenen Syntax kann die Identifikation auch in kompakter Schreibweise mittels

```
M.identify(y,u,'Opt1',Val1,...,'OptN',ValN);
```

ausgeführt werden, wobei für nicht spezifizierte Optionen Standardwerte verwendet werden. Weitere wichtige Methoden der Klasse, welche u. a. zur Modellvalidierung verwendet werden können, sind in der Tabelle 16 aufgeführt⁵.

Bemerkung D.1 (Optimierer): In der Voreinstellung werden `lsqnonlin()` bzw. `fmincon()` aus der MATLAB OPTIMIZATION TOOLBOX verwendet (Optionsargument: 'matlab_lsqnonlin' bzw. 'matlab_fmincon'). Alternativ können auch die Optimierungsalgorithmen aus der frei verfügbaren OPTI TOOLBOX⁶ verwendet werden (Optionsargument: 'opti_levmar' oder 'opti_ipopt'). ◀

⁴Eine vollständige Aufzählung der verfügbaren Optionen und der möglichen Werte kann dem Quellcode der Funktion `identifyModelOptions` der Klasse `NsfPolyModel` entnommen werden.

⁵Eine vollständige Aufzählung der Funktionsargumente und der -ausgaben sowie eine kurze Beschreibung zur Funktionalität kann dem Quellcode der jeweiligen Funktion entnommen werden.

⁶<https://github.com/jonathancurrie/OPTI>

Bemerkung D.2 (Mehr- bzw. Multischritt-PEM): In [HW99] wird ein vereinfachter Ansatz für die Implementierung eines Mehr- bzw. Multischrittverfahrens vorgeschlagen. Dieser basiert auf der Identifikation von \mathbf{G} und \mathbf{H} mittels 1-Schritt PEM, anschließender Vorfilterung der Eingangs- und Ausgangsdaten durch $\mathbf{H}^{[0,k-1]}(q^{-1})$ und wiederholte Bestimmung der Parameter von \mathbf{G} bei festgelegtem Rauschmodell \mathbf{H} . Da dabei jedoch die Abhängigkeit von $\mathbf{H}^{[0,k-1]}(q^{-1})$ vom Parametervektor $\boldsymbol{\theta}$ nicht explizit berücksichtigt wird (vgl. Gl. (3.15)), ist der vorgeschlagene Algorithmus im Allgemeinen suboptimal. Hingegen wird bei dem in der Softwarebibliothek implementierten Verfahren bei der Berechnung des Kostenfunktional und des Gradienten die Parameterabhängigkeit von $\mathbf{H}^{[0,k-1]}(q^{-1}, \boldsymbol{\theta})$ in jedem Iterationsschritt der Optimierung explizit berücksichtigt. \triangleleft

Optionsname	Beschreibung
'InitMethod'	Methode, die zur Schätzung der initialen Parameter $\boldsymbol{\theta}_{(0)}$ verwendet wird
'IntegrateNoise'	Erzwingt zusätzlichen Integrator im Rauschmodell (siehe Abschn. 2.4.1)
'EstimateOutputOffset'	Schätzt Offset des Modellausgangs \mathbf{y}
'Hp'	Prädiktionshorizont (Mehrschritt-PEM)
'MultiStep'	Verwendet Multischritt-PEM (Hp-Schritte)
'Ic'	Methode zur Schätzung der Anfangswerte des Residuenfilters (siehe Abschn. B.2)
'Solver'	Löser der Optimierungsaufg. (3.5) (siehe Bem. D.1)

Tabelle 15: Wichtige Optionen der Methode `identify()`

Methode	Beschreibung
<code>simulate()</code>	Simuliert das Modell
<code>identify()</code>	Identifiziert die Parameter des Modells
<code>printParameters()</code>	Gibt Übersicht identifizierter Parameter und deren relative Standardabweichungen auf Kommandozeile aus
<code>show()</code>	Gibt Grafiken zur Modellvalidierung aus (u. a. Box-Whisker-Diagramm und Autokorrelationsfolge der Residuen)
<code>showBode()</code>	Gibt Bode-Diagramme und Konfidenzintervalle des linearen Teilsystems aus
<code>factorize()</code>	Faktorisiert Polynome in Linearfaktoren (siehe Bem. 2.9)
<code>defactorize()</code>	Wandelt Polynome in Summendarstellung um

Tabelle 16: Wichtige Methoden der Klasse `idModels.NsfPolyModel`

D.2 Beispiele

Im Folgenden wird die Nutzung der Software anhand von drei Beispielen demonstriert:

D.2.1 LTI-System zweiter Ordnung

Es wird ein nicht-schwingungsfähiges lineares SISO System zweiter Ordnung mit Ausgangsquantisierung und einem unbekanntem Ausgangsoffset y_0 betrachtet:

$$y[t] = \left\lfloor \frac{b^{[1]} q^{-1}}{1 + a^{[1]} q^{-1} + a^{[2]} q^{-2}} u[t] \right\rfloor + y_0 \quad (\text{D.2})$$

Das System soll durch ein ARMAX-Modell zweiter Ordnung mit zwei reellen Polstellen angenähert werden:

```
% Eingangsdaten erzeugen und System simulieren
rng(0); % für Reproduzierbarkeit
u = kron(rand(20,1),ones(50,1)); % generiere Eingangssignal
B = [0 1e-1]; % Zählerkoeffizienten
A = conv([1 -.95],[1 -.85]); % Nennerkoeffizienten
x = filter(B,A,u) + 3; % simuliere System
y = round(x); % quantisierter Ausgang

% Modell M instanziiieren, identifizieren und validieren
M = idModels.NsfPolyModel(2,1,1,2); % ARMAX Model erzeugen
M.factorize('A'); % faktorisiere A(q^-1)
M.identify(y,u,'EstimateOutputOffset',1); % PEM Identifikation
M.printParameters; % Param. auf Konsole ausgeben
M.show('Prediction',y,u,'Hp',100); % Plot Messung vs. Prädiktion
```

Die Ausgabe des `M.show(...)` Kommandos ist in der Abbildung D.1 dargestellt.

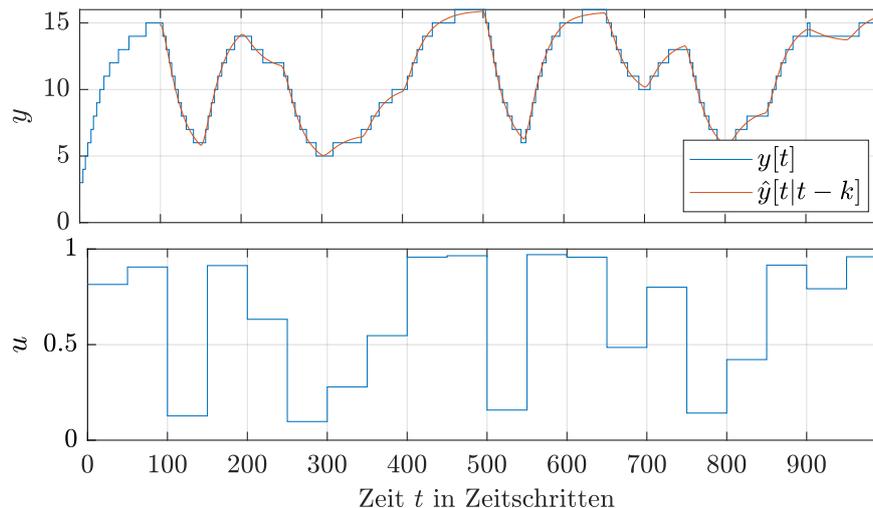


Abbildung D.1: Vergleich des Systemausgangs mit $A(q^{-1}) = (1-0.95q^{-1})(1-0.85q^{-1})$, $B(q^{-1}) = 0.1q^{-1}$ und $y_0 = 3$ mit der 100-Schritt Prädiktion des Modells.

D.2.2 Bilineares System mit periodischer Störung

Es wird das in Abschnitt 3.4 diskutierte bilineare Beispielsystem mit periodischer Störung betrachtet (Beispiel 3.3). Zur Simulation und Identifikation wird zunächst eine Funktion zur Berechnung des Funktionswertes $f(u, y, \eta)$ und der Ableitung $\frac{\partial f}{\partial \eta}(u, y, \eta)$ definiert. Anschließend kann die Simulation mithilfe des folgenden Programmcodes durchgeführt werden:

```
% Parameter und Signale definieren
rng(0); % für Reproduzierbarkeit
T_d = 50; % Periodendauer d_1
N = 500; % Anz. Samples
u = 2*kron(rand(N/T_d/2,1),... % Eingangssignal u
    [zeros(T_d,1); ones(T_d,1)]);
d1 = 1*kron(ones(N/T_d,1), ... % periodische Störung d_1
    [zeros(T_d/2,1);ones(T_d/2,1)]);
d2 = cos(2*pi*linspace(0,1,N)')+0.5; % Störung d_2 ("Drift")
A = [1 -.9]; nA = length(A)-1; % Polynomkoeff. A
B = [0 .5]; nB = length(B)-1; % Polynomkoeff. B
f = @(u,eta,y) deal(u.*(eta-y),... % f(u,eta,y)
    u); % df/deta(u,eta,y)
eta = 3; % Parameter f(u,y,eta)

% System Simulieren
x = NaN(N,1); y = x; x(1) = 0; y(1) = d2(1); % Initialisierung
for t = 2:N
    [fuy,~] = f(u(t-1:-1:t-nB),... % f(u,eta,y) auswerten
        eta,y(t-1:-1:t-nB));
    x(t) = -A(2:end)*x(t-1:-1:t-nA) + ... % Ausgang lin. Teilsystem
        B(2:end)*(fuy + d1(t-1:-1:t-nB));
    y(t) = x(t) + d2(t); % Systemausgang mit Drift
end
```

Wie im Abschnitt 3.4 beschrieben, wird das System mit einem nichtlinearen Modell (2.51) mit ARARIX-Parametrierung des LTI-Teilsystems modelliert. Hierfür wird zunächst ein ARARX-Modell mit der Struktur F instanziiert und anschließend die Identifikation mit einem zusätzlichen Differenzierer als Vorfilter ausgeführt ($L(q^{-1}) = 1 - q^{-1}$):

```
% Struktur für Eingangsnichtlin. definieren
F(1).fun = f; % Funktionshandle f(u,eta,y)
F(1).parameters = 0; % Optimiererstartwert für eta
F(1).free = 1; % als freien Param. festlegen
F(1).input_idx = 1; % Ein- bzw. Ausgänge die bei
F(1).output_idx = 1; % Berechnung von f benötigt werden

% Modell definieren, identifizieren und Parameter ausgeben
M = idModels.NsfPolyModel(... % ARARX Modellansatz M definieren
    nA,1,nB,0,T_d,'InputNonlinearity',F);
```

```
M.D.val(2:end-1) = 0;           % d_1 bis d_(T_d-1) Null setzen
M.D.free(2:end-1) = false;     % und als feste Param. festlegen
M.identify(y,u,'IntegrateNoise',1); % M mit Vorfilter L=1-q^-1 ident.
M.printParameters;             % Parameter ausgeben
```

D.2.3 Nichtlineares MIMO System

Das nichtlineare MIMO-Beispielsystem

$$\begin{pmatrix} 1 - 0.9q^{-1} & 0.5q^{-1} \\ 0.1q^{-1} & 1 - 0.5q^{-1} \end{pmatrix} \mathbf{x}[t] = \begin{pmatrix} 0.1q^{-1} & 0 \\ 0 & 0.05q^{-1} \end{pmatrix} \mathbf{f}(\mathbf{y}[t], \mathbf{u}[t], \boldsymbol{\eta}) \quad (\text{D.3})$$

$$\mathbf{y}[t] = \mathbf{x}[t] + \mathbf{e}[t]$$

mit

$$\mathbf{f}(\mathbf{u}, \mathbf{y}, \boldsymbol{\eta}) = \begin{pmatrix} (u_1 + u_2)(\eta_1 - y_1) \\ u_2^2 y_2 + \eta_2 u_2^3 \end{pmatrix} \quad (\text{D.4})$$

und $\boldsymbol{\eta} = (2, 5)^\top$ soll durch das Modell

$$\mathbf{A}(q^{-1})\mathbf{y}[t] = \mathbf{B}(q^{-1})\mathbf{f}(\mathbf{u}[t], \mathbf{y}[t], \boldsymbol{\eta}) + \mathbf{C}(q^{-1})\mathbf{e}[t] \quad (\text{D.5})$$

angenähert werden, wobei $\mathbf{e}[t]$ normalverteiltes weißes Rauschen darstellt. Für die Identifikation und Simulation werden zunächst zwei Funktionen zur Berechnung der Nichtlinearität $\mathbf{f}(\mathbf{y}, \mathbf{u}, \boldsymbol{\eta})$ definiert⁷:

```
function [f,df_deta,df_du,df_dy] = f_1(u,eta,y)
% Berechnet Funktionswert f_1 und deren Ableitungen
f = sum(u,2).*(eta - y);           % f(u,y,eta)
df_deta = sum(u,2);               % df_deta(u,y,eta)
df_du = ones(1,2).*(eta - y);     % df_du(u,y,eta)
df_dy = -sum(u,2);               % df_dy(u,y,eta)
end
```

```
function [f,df_deta,df_du,df_dy] = f_2(u,eta,y)
% Berechnet Funktionswert f_2 und deren Ableitungen
f = u.^2.*y + eta*u.^3;           % f(u,y,eta)
df_deta = u.^3;                   % df_deta(u,y,eta)
df_du = 2*y.*u+3*eta.*u.^2;      % df_du(u,y,eta)
df_dy = u.^2;                     % df_dy(u,y,eta)
end
```

Der folgende Programmcode simuliert das System (D.3):

⁷Die beiden Ausgabeparameter `df_du` und `df_dy` stellen die partiellen Ableitungen bezüglich der Ein- bzw. Ausgangsgrößen des Modells dar. Diese werden lediglich zur Berechnung der Ableitungen $\frac{\partial \mathbf{y}[t]}{\partial \mathbf{u}[s]}$ benötigt (vgl. Gl. (6.23) und Bemerkung D.3).

```

% Struktur für f(u,y,eta) definieren
F(1).fun = 'f_1'; % Funktionsname von f_1
F(1).parameters = 2; % eta_1
F(1).input_idx = [1 2]; % u1 und u2 an f_1
F(1).output_idx = 1; % y1 and f_1 übergeben
F(2).fun = 'f_2'; % Funktionsname von f_2
F(2).parameters = 5; % eta_2
F(2).input_idx = 2; % u2 an f_2 übergeben
F(2).output_idx = 2; % y2 an f_2 übergeben

% System S instanziiieren und Polynomparameter setzen
S = NsfPolyModel(ones(2),eye(2),eye(2),... % System instanziiieren
    'InputNonlinearity',F,'NoiseVariance',zeros(2));
S.A(1,1).val = [1 -.9]; S.A(1,2).val = [0 .5]; % A Parameter setzen
S.A(2,1).val = [0 .1]; S.A(2,2).val = [1 -.5];
S.B(1,1).val = [0 .1]; S.B(2,2).val = [0 .05]; % B Parameter setzen

% System S Simulieren
rng(0); % für Reproduzierbarkeit
u = kron(rand(10,2),ones(50,1)); % Inputs erzeugen
x = S.simulate(u); % Simulieren
y = x + [.05 .02].*randn(size(u)); % Ausgangsrauschen add.

```

Die Identifikation und Simulation des Modells wird mittels

```

% Modell M instanziiieren, identifizieren, simulieren und Param. ausgeben
F(1).parameters = 1; F(2).parameters = 1; % Startwerte Optimierung
M = NsfPolyModel(ones(2),eye(2),... % Modell M instanziiieren
    eye(2),ones(2,1),'InputNonlinearity',F);
M.identify(y,u); % M identifizieren
M.printParameters; % Parameter ausgeben
ys = M.simulate(u); % M simulieren

```

durchgeführt. Der Vergleich der Simulation des Systems (D.3) und des identifizierten Modells ist in der Abbildung D.2 dargestellt.

Bemerkung D.3 (Ableitungen des Modellausgangs): Zur Berechnung der Ausgangsableitungen $\frac{\partial y[t]}{\partial u[s]}$ müssen bei der Definition der Komponenten von \mathbf{f} die Ableitungen $\frac{\partial \mathbf{f}(\mathbf{u},\mathbf{y})}{\partial \mathbf{y}}$ und $\frac{\partial \mathbf{f}(\mathbf{u},\mathbf{y})}{\partial \mathbf{u}}$ implementiert werden (siehe oben). Der *simulierte* (vgl. Abschn. 2.3) Systemausgang und dessen Ableitungen können dann mittels

```
[ys,x,dy_du] = M.simulate(u);
```

berechnet werden, wobei $\mathbf{dy_du}$ die $N \times n_y \times n_u \times N$ -Matrix der Ausgangsableitungen darstellt. ◀

Bemerkung D.4 (Beschränkungen der Parameter $\boldsymbol{\eta}$): Ist bspw. aufgrund physikalischer Betrachtungen bekannt, dass der Parameter $\boldsymbol{\eta}$ in einem definierten Bereich

liegt, dann kann dies bei der Optimierung durch entsprechende Nebenbedingungen berücksichtigt werden. Hierfür muss für die entsprechende(n) Komponente(n) $f_j(\mathbf{u}, \mathbf{y}, \boldsymbol{\eta}_j)$ ($j \in \{1, \dots, n_{\bar{u}}\}$) der gewünschte Bereich in der Standardform linearer Ungleichungsbedingungen

$$\mathbf{A}_j \boldsymbol{\eta}_j \leq \mathbf{b}_j \quad (\text{D.6})$$

formuliert und bei der Definition der Struktur $\mathbf{F}(j)$ angegeben werden. Soll im obigen Beispiel $0 \leq \eta_1 \leq 5$ gelten, dann muss z. B. zusätzlich folgender Code verwendet werden:

```
F(1).A = [-1; 1];
F(1).b = [0; 5];
```

◁

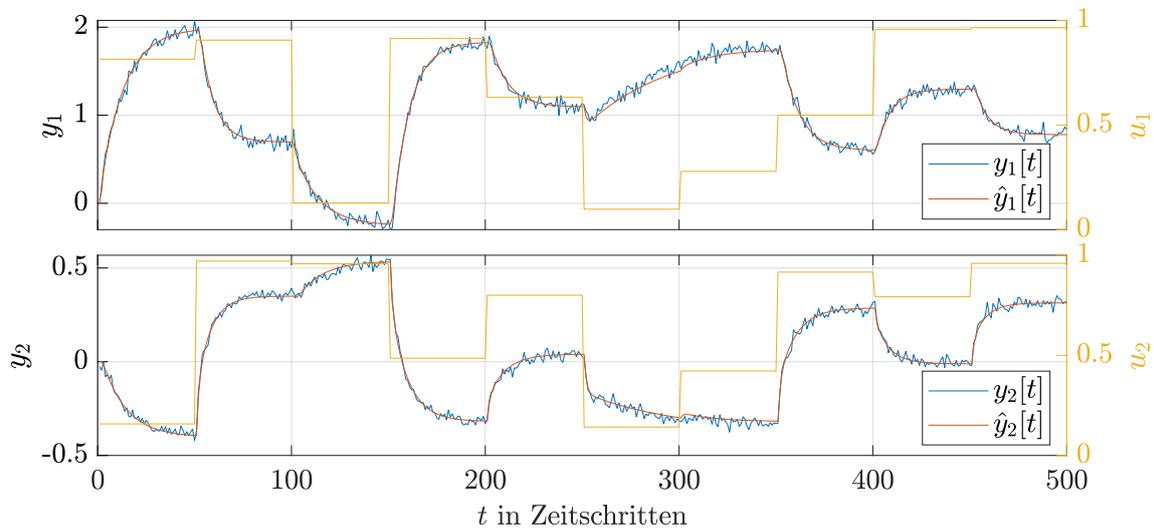


Abbildung D.2: Vergleich des simulierten System- (D.3) und Modellausgangs (D.5).

Abkürzungsverzeichnis

Akronym	Bedeutung
Abl	Abluft
Aul	Außenluft
AR	Autoregressive
ARIMAX	Autoregressive integrated noise moving average exogenous input
BFC	Backforecasting
BHKW	Blockheizkraftwerk
BJ	Box-Jenkins
DFT	Diskrete Fouriertransformation
DTFT	Fouriertransformation für zeitdiskrete Signale
DWD	Deutscher Wetterdienst
ELS	Extended least squares
FIR	Finite impulse response
FMI	Functional Mock-up Interface
FMU	Functional Mock-up Unit
GLS	Generalized least squares
HW	Hammerstein-Wiener
IIR	Infinite impulse response
LTI	Linear time-invariant
LS	Least squares
MA	Moving average
MAP	Maximum a posteriori
MIMO	Multiple input multiple output
MkQ	Methode der kleinsten Fehlerquadrate
ML	Maximum-Likelihood
MPC	Model predictive control
MSE	Mean squared error
NARMAX	Nonlinear autoregressive moving average exogenous input
OE	Output error
PEM	Prediction error minimization
PLR	Pseudo linear regression
ppm	Parts per million
Rl	Rücklauf
R(M)MSE	Root (mean) mean square error
SISO	Single input single output

SQP	Sequenzielle quadratische Programmierung
TGA	Technische Gebäudeausstattung
TRY	Testreferenzjahr
u. i. v.	unabhängig identisch verteilt
VI	Vorlauf
VMkQ	Verallgemeinerte Methode der kleinsten Fehlerquadrate
VSR	Volumenstromregler
WH	Wiener-Hammerstein
ZRM	Zustandsraummodell
Zul	Zuluft

Symbolverzeichnis

Notation	Bedeutung
t	Zeit
q^{-1} bzw. q	Vorwärts- bzw. Rückwärtszeitverschiebungsoperator
z	komplexe Variable der \mathcal{Z} -Transformation
$\boldsymbol{\theta}, \boldsymbol{\eta}$	Parametervektor $\boldsymbol{\theta}$ bzw. $\boldsymbol{\eta}$
N	Datenpunktanzahl
$\mathbf{y}(t)$	Regel- bzw. Ausgangsgrößen
$\mathbf{u}(t)$	Stell- bzw. Eingangsgrößen
$\mathbf{d}(t)$	präzifizierbare Störgrößen
$\mathbf{v}(t)$	nicht präzifizierbare Störgrößen
$\mathbf{e}(t)$	weißes Rauschen
$\mathbf{w}(t)$	Führungsgrößen
$\mathbf{f}(\mathbf{x})$ bzw. $\mathbf{f}[\mathbf{x}]$	Funktion einer kontinuierlichen bzw. diskreten Variable \mathbf{x}
$\mathbf{f}'(\mathbf{x}) = \partial \mathbf{f} / \partial \mathbf{x}$	<i>JACOBI</i> -Matrix bzw. Gradient von \mathbf{f}
$\mathbf{f}''(\mathbf{x}) = \partial^2 \mathbf{f} / \partial \mathbf{x}^2$	<i>HESSE</i> -Matrix von \mathbf{f}
$\mathbb{R}, \mathbb{C}, \mathbb{N}, \mathbb{Z}$	Menge der reellen, komplexen, natürlichen bzw. ganzen Zahlen
$\mathbb{R}_{>0}, \mathbb{Z}_{>0}$	Menge der positiven reellen bzw. ganzen Zahlen
\in, \notin	ist, bzw. ist nicht Element von
\exists, \forall	es existiert, für alle
$\mathbf{0}$	Nullmatrix geeigneter Dimension
\mathbf{I}	Einheitsmatrix geeigneter Dimension
\mathbf{I}_{Δ}	untere Dreiecksmatrix geeigneter Dimension
\mathbf{A}^{\top}	transponierte Matrix von \mathbf{A}
\mathbf{A}^{-1}	inverse Matrix von \mathbf{A}
$\text{tr}(\mathbf{A})$	Spur von \mathbf{A}
$\det(\mathbf{A})$	Determinante von \mathbf{A}
$\delta_{t,s}$	Kronecker-Delta
\mathbf{X} bzw. \mathbf{X}_t	Zufallsgröße bzw. diskreter stochastischer Prozess
\mathbf{x} bzw. $\mathbf{x}[t]$	Realisierung von \mathbf{X} bzw. \mathbf{X}_t
$E\{\mathbf{X}\}$ bzw. $E\{\mathbf{X}[t]\}$	Erwartungswert von \mathbf{X} bzw. \mathbf{X}_t
$\text{Cov}\{\mathbf{X}\}$ bzw. $\text{Cov}\{\mathbf{X}[t]\}$	Kovarianzmatrix bzw. -funktion von \mathbf{X}
$\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Lambda})$	Normalverteilung mit Erwartungswert $\boldsymbol{\mu}$ und Kovarianzmatrix $\boldsymbol{\Lambda}$

μ_x, λ_x bzw. Λ_x	Erwartungswert, Varianz bzw. Kovarianzmatrix der Zufallsgröße \mathbf{X}
$\text{vec}\{\mathbf{x}[t]\}$	(zeilenweise) vektorisierte Folge $\mathbf{x}[t]$
$\mathcal{Z}\{\mathbf{x}[t]\}$	\mathcal{Z} -Transformation der Folge $\mathbf{x}[t]$
$\hat{\mathbf{x}}[t+k t]$	Prädiktion von $\mathbf{x}[t+k]$ zum Zeitpunkt t (bzw. mit Messinformation bis zur Zeit t)
$\mathbf{H}^{[n,m]}(q^{-1})$, mit $n \leq m$	Ausschnitt der Impulsantwort des Filters $\mathbf{H}(q^{-1})$, d. h. $\mathbf{H}^{[n,m]}(q^{-1}) = \mathbf{h}[n]q^{-n} + \dots + \mathbf{h}[m]q^{-m}$
$A_{i,j}$	i -te Zeile und j -te Spalte der Matrix \mathbf{A}
$\mathbf{A}_{i:}$ bzw. $\mathbf{A}_{:,i}$	i -te Zeile bzw. Spalte der Matrix \mathbf{A}
$\boldsymbol{\theta}_{(i)}$	Wert von $\boldsymbol{\theta}$ im Iterationsschritt i
$\hat{\boldsymbol{\theta}}$	Schätzwert von $\boldsymbol{\theta}$

Literatur

- [AB65] K. J. Aström und T. Bohlin. „Numerical identification of linear dynamic systems from normal operating records“. In: *Proceedings of the Second IFAC Symposium on the Theory of Self-Adaptive Control Systems 2* (1965), S. 96–111.
- [ABS16] F. Ascione, N. Bianco, C. De Stasio u. a. „Simulation-based model predictive control by the multi-objective optimization of building energy performance and thermal comfort“. In: *Energy and Buildings* 111 (2016), S. 131–144.
- [ADS04] C. Andrieu, A. Doucet, S. S. Sumeetpal u. a. „Particle Methods for Change Detection, System Identification, and Control“. In: *Proceedings of the IEEE* 92.3 (2004), S. 423–438.
- [AFG14] B. Ai, Z. Fan und R. X. Gao. „Occupancy Estimation for Smart Buildings by an Auto-Regressive Hidden Markov Model“. In: *Proceedings of the American Control Conference* (2014), S. 2234–2239.
- [AJ14] A. Afram und F. Janabi-Sharif. „Theory and applications of HVAC control systems: A review of model predictive control (MPC)“. In: *Building and Environment* 72 (2014), S. 343–355.
- [AJF17] A. Afram, F. Janabi-Sharifi, A. S. Fung u. a. „Artificial neural network (ANN) based model predictive control (MPC) and optimization of HVAC systems: A state of the art review and case study of a residential HVAC system“. In: *Energy and Buildings* 141 (2017), S. 96–113.
- [AM79] B. D. O. Anderson und J. B. Moore. *Optimal Filtering*. Prentice Hall, 1979.
- [AMH00] K. K. Andersen, H. Madsen und L. H. Hansen. „Modelling the heat dynamics of a building using stochastic differential equations“. In: *Energy and Buildings* 31 (2000), S. 13–24.
- [And84] J. Andél. *Statistische Analyse von Zeitreihen*. Akademie-Verlag, Berlin, 1984.
- [APG14] Arbeitskreis der Professoren für Gebäudeautomation und Energiesysteme. *Regelungs- und Steuerungstechnik in der Versorgungstechnik*. Bd. 7. VDE-Verlag, 2014.
- [AQH97] K. Abed-Meraim, W. Qiu und Y. Hua. „Blind system identification“. In: *Proceedings of the IEEE* 85.8 (1997), S. 1310–1322.
- [Ash17] *ASHRAE Handbook - Fundamentals*. American Society of Heating, Refrigerating und Air-Conditioning Engineers (ASHRAE). 2017.

- [Ast64] K. J. Aström. „Control Problems in Paper Making“. In: *IBM Scientific Computing Symposium: Control Theory and Applications* (1964).
- [Ast70] K. J. Aström. *Introduction to Stochastic Control Theory, Volume 70*. Academic Press, 1970.
- [Ast79] K. J. Aström. „Maximum Likelihood and Prediction Error Methods“. In: *Automatica* 16 (1979), S. 551–574.
- [Bai02] E.-W. Bai. „Identification of linear systems with hard input nonlinearities of known structure“. In: *Automatica* 38.5 (2002), S. 853–860.
- [BB63] E. Blandhol und J. G. Balchen. „Determination of System Dynamics by Use of Adjustable Models“. In: *IFAC Proceedings Volumes 1.2* (1963), S. 602–613.
- [BBM17] F. Borelli, A. Bemporad und M. Morari. *Predictive Control for Linear and Hybrid Systems*. 1. Aufl. Cambridge University Press, 2017.
- [BF85] S. A. Billings und M. B. Fadzil. „The Practical Identification of Systems with Nonlinearities“. In: *IFAC Proceedings Volumes 18.5* (1985), S. 155–160.
- [BFS88] S. A. Billings, M. B. Fadzil, J. L. Sulley u. a. „Identification of a non-linear difference equation model of an industrial diesel generator“. In: *Mechanical Systems and Signal Processing 2.1* (1988), S. 59–76.
- [Bil13] S. A. Billings. *Nonlinear System Identification: NARMAX Methods in the Time, Frequency and Spatio-Temporal Domains*. John Wiley & Sons, 2013.
- [Bil85] S. A. Billings. *Introduction to Nonlinear Systems Analysis and Identification*. Techn. Ber. Dept. of Automatic Control und System Engineering - University of Sheffield, 1985.
- [BJH19] D. Blum, F. Jorissen, S. Huang u. a. „Prototyping the BOPTTEST Framework for Simulation-Based Testing of Advanced Control Strategies in Buildings“. In: *Proceedings of the IBPSA Building Simulation 2019* (2019).
- [BJR08] G. E. P. Box, G. M. Jenkins und G. C. Reinsel. *Time Series Analysis: Forecasting and Control*. 4. Aufl. John Wiley & Sons, 2008.
- [BKB14] S. C. Bengea, A. D. Kelman, F. Borrelli u. a. „Implementation of model predictive control for an HVAC system in a mid-size commercial building“. In: *HVAC&R Research* 20 (2014), S. 121–135.
- [BKB15] A. Borrmann, M. König, J. Beetz u. a., Hrsg. *Building Information Modeling: Technologische Grundlagen und industrielle Praxis*. Springer Vieweg, 2015.
- [BM11] P. Bacher und H. Madsen. „Identifying suitable models for the heat dynamics of buildings“. In: *Energy and Buildings* 43 (7 2011), S. 1511–1522.
- [BMA13] P. Bacher, H. Madsen, H. A. Nielsen u. a. „Short-term heat load forecasting for single family houses“. In: *Energy and Buildings* 65 (2013), S. 101–112.

- [BMU21] Bundesministerium für Umwelt, Naturschutz und nukleare Sicherheit. *Lesefassung des Bundes-Klimaschutzgesetzes 2021 mit markierten Änderungen zur Fassung von 2019*. 2021. URL: https://www.bmu.de/fileadmin/Daten_BMU/Download_PDF/Klimaschutz/ksg_aendg_2021_3_bf.pdf (besucht am 28.09.2021).
- [BMWi20] *Energieeffizienz in Zahlen*. Techn. Ber. Bundesministerium für Wirtschaft und Energie (BMWi), 2020. URL: https://www.bmwi.de/Redaktion/DE/Publikationen/Energie/energieeffizienz-in-zahlen-2020.pdf?__blob=publicationFile&v=20.
- [BOA11] T. Blochwitz, M. Otter, M. Arnold u. a. „The Functional Mockup Interface for Tool independent Exchange of Simulation Models“. In: *Proceedings of the 8th International Modelica Conference* (2011).
- [Bri81] D. R. Brillinger. *Time Series Data Analysis and Theory*. Holden-Day, San Francisco, 1981.
- [BU16] C. Bohn und H. Unbehauen. *Identifikation dynamischer Systeme*. 1. Aufl. Springer Vieweg, 2016.
- [Buc10] M. Buchholz. „Subspace-Identification zur Modellierung von PEM-Brennstoffzellen-Stacks“. Diss. Karlsruher Institut für Technologie, Institut für Regelungs- und Steuerungssysteme, 2010.
- [BV04] S. Boyd und L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.
- [BW05] S. A. Billings und H. L. Wei. „The wavelet-NARMAX representation: A hybrid model structure combining polynomial models with multiresolution wavelet decompositions“. In: *International Journal of Systems Science* 36.3 (2005), S. 137–152.
- [BZ94] S. A. Billings und Q. M. Zhu. „Structure Detection Algorithm for Nonlinear Rational Models“. In: *International Journal of Control* 59.6 (1994), S. 1439–1463.
- [CB07] E. F. Camacho und C. Bordons. *Model Predictive control*. 2. Aufl. Springer-Verlag London, 2007.
- [CER01] N. Chiras, C. Evans und D. Rees. „Nonlinear gas turbine modeling using NARMAX structures“. In: *IEEE Transactions on Instrumentation and Measurement* 50.4 (2001), S. 893–898.
- [CG84] S. W. Chan und G. C. Goodwin. „Convergence Properties of the Riccati Difference Equation in Optimal Filtering of Nonstabilizable Systems“. In: *IEEE Transactions on Automatic Control* AC-29.2 (1984), S. 110–118.
- [CH15] R. D. Coninck und L. Helsen. „Practical implementation and evaluation of model predictive control for an office building in Brussels“. In: *Energy and Buildings* 111 (2015), S. 290–298.

- [CP98] J. Chen und R. Patton. *Robust Model-Based Fault Diagnosis for Dynamic Systems*. Kluwer Academic Publishers, 1998.
- [CR80] C. R. Cutler und B. L. Ramaker. „Dynamic Matrix Control - a computer control algorithm“. In: *Proceedings of the Joint Automatic Control Conference* (1980).
- [DAC20] J. Drgoňa, J. Arronyo, I. Cupeiro Figueroa u. a. „All you need to know about model predictive control for buildings“. In: *Annual Reviews in Control* 50 (2020), S. 190–232.
- [DP06] R. Dittmar und B.-M. Pfeiffer. „Modellbasierte prädiktive Regelung in der industriellen Praxis“. In: *at - Automatisierungstechnik* 54.12 (2006), S. 590–601.
- [DWL18] Y. Ding, L. Wang, Y. Li u. a. „Model predictive control and its application in agriculture: A review“. In: *Computers and Electronics in Agriculture* 151 (2018), S. 104–117.
- [EBV13] A. Ebadat, G. Bottegal, D. Varagnolo u. a. „Estimation of building occupancy levels through environmental signals deconvolution“. In: *Proceedings of the 5th ACM Workshop on Embedded Systems For Energy-Efficient Buildings* (2013), S. 1–8.
- [EBV15] A. Ebadat, G. Bottegal, D. Varagnolo u. a. „Blind identification strategies for room occupancy estimation“. In: *Proceedings of the European Control Conference* (2015), S. 1315–1320.
- [EKR04] H.-F. Eckey, R. Kosfeld und C. Dreger. *Ökonometrie: Grundlagen, Methoden, Beispiele*. Gabler Verlag, 2004.
- [Eykh63] P. Eykhoff. „Some fundamental aspects of process-parameter estimation“. In: *IEEE Transactions on Automatic Control* 8.4 (1963), S. 347–357.
- [Foe11] O. Föllinger. *Laplace-, Fourier- und z-Transformation*. VDE-Verlag, 2011.
- [FP08] M. Farina und L. Piroddi. „Some Convergence Properties of Multi-Step Prediction Error Identification Criteria“. In: *Proceedings of the 47th IEEE Conference on Decision and Control* (2008), S. 756–761.
- [FP11] M. Farina und L. Piroddi. „Simulation error minimization identification based on multi-stage prediction“. In: *International Journal of Adaptive Control and Signal Processing* (2011).
- [FS10] L. Ferkl und J. Široký. „Ceiling radiant cooling: Comparison of ARMAX and subspace identification modelling methods“. In: *Building and Environment* 45.1 (2010), S. 205–212.
- [FWM17] M. Fiorentini, J. Wall, Z. Ma u. a. „Hybrid model predictive control of a residential HVAC system with on-site thermal energy generation and storage“. In: *Applied Energy* 187 (2017), S. 465–479.

- [FWW17] D. Fischer, T. Wolf, J. Wapler u. a. „Model-based flexibility assessment of a residential heat pump pool“. In: *Energy* 118 (2017), S. 853–864.
- [GB10] F. Giri und E.-W. Bai, Hrsg. *Block-oriented Nonlinear System Identification*. Springer, 2010.
- [GDJ18] D. Geysen, O. De Somer, C. Johansson u. a. „Operational thermal load forecasting in district heating networks using machine learning and expert advice“. In: *Energy and Buildings* 162 (2018), S. 144–153.
- [GEL86] G. C. Goodwin, R. J. Evans, R. L. Leal u. a. „Sinusoidal disturbance rejection with application to helicopter flight data estimation“. In: *IEEE Transactions on Acoustics, Speech, and Signal Processing* 34.3 (1986), S. 479–484.
- [Geo19] *GeoHybrid*. Geo-En Energy Technologies GmbH. Juli 2019. URL: https://www.geo-en.de/images/pdf/pdf/GeoHybrid_2019.pdf (besucht am 07.05.2020).
- [GGS13] M. Gwender, D. Gyalistras, C. Sagerschnig u. a. *Final Report: Use of Weather And Occupancy Forecasts For Optimal Building Climate Control –Part II: Demonstration (OptiControl-II)*. Techn. Ber. Automatic Control Laboratory, ETH Zurich, Switzerland, 2013.
- [GHP80] G. Gardner, A. C. Harvey und G. D. A. Phillips. „An Algorithm for Exact Maximum Likelihood Estimation of Autoregressive-Moving Average Models by Means of Kalman Filtering“. In: *Journal of the Royal Statistical Society* 29.3 (1980), S. 311–322.
- [GMW82] P. E. Gill, W. Murray und M. H. Wright. *Practical Optimization*. 11. Aufl. Academic Press, 1982.
- [GP00] N. E. Goodzeit und M. Q. Phan. „System Identification in the Presence of Completely Unknown Periodic Disturbances“. In: *Journal of Guidance, Control and Dynamics* 23.2 (2000), S. 251–259.
- [GPS02] R. B. Gopaluni, R. S. Patwardhan und S. L. Shah. „Bias Distribution in MPC Relevant Identification“. In: *Proceedings of 15th IFAC World Congress, Barcelona* (2002), S. 2196–2201.
- [GPS05] I. Goethals, K. Pelckmans, J. A. K. Suykens u. a. „Subspace Identification of Hammerstein Systems Using Least Squares Support Vector Machines“. In: *IEEE Transactions on Automatic Control* 50.10 (2005), S. 1509–1519.
- [Gro14] L. Gröll. „Methodik zur Integration von Vorwissen in die Modellbildung“. Habilitation. Karlsruher Institut für Technologie, 2014.
- [Gue86] M. Günther. *Zeitdiskrete Steuerungssysteme*. VEB Verlag Technik, Berlin, 1986.
- [Har15] E. N. Hartley. „A Tutorial on Model Predictive Control for Spacecraft Rendezvous“. In: *Proceedings of the European Control Conference*. 2015, S. 1355–1361.

- [HBB12] R. Halvgaard, P. Bacher, P. Bengt u. a. „Model predictive control for a smart solar tank based on weather and consumption forecasts“. In: *Energy Procedia* 30 (2012), S. 270–278.
- [HC18] B. Hörner und M. Casties, Hrsg. *Handbuch der Klimatechnik, Band 2: Anwendungen*. 7. Aufl. 2018.
- [HCT12] D. Hrovat, S. Di Cairano, H. E. Tseng u. a. „The Development of Model Predictive Control in Automotive Industry: A Survey“. In: *Proceedings of the IEEE International Conference on Control Applications* (2012), S. 295–302.
- [HGP12] I. Hazyuk, C. Ghiaus und D. Penhouet. „Optimal temperature control of intermittently heated buildings using Model Predictive Control: Part II - Control algorithm“. In: *Building and Environment* 51 (2012), S. 388–394.
- [HK86] I. W. Hunter und M. J. Korenberg. „The Identification of Nonlinear Biological Systems: LNL Cascade Models“. In: *Biological Cybernetics* 55 (1986), S. 125–134.
- [HKS16] T. Hilliard, M. Kavgic und L. Swan. „Model predictive control for commercial buildings: trends and opportunities“. In: *Advances in Building Energy Research* 10.2 (2016), S. 172–190.
- [HLW08] A. Hagenblad, L. Ljung und A. Wills. „Maximum likelihood identification of Wiener models“. In: *Automatica* 44.11 (2008), S. 2697–2705.
- [HMT03] B. Huang, A. Malhotra und E. C. Tamayo. „Model predictive control relevant identification and validation“. In: *Chemical Engineering Science* 58.11 (2003), S. 2389–2401.
- [HO81] V. Haggan und T. Ozaki. „Modelling nonlinear random vibrations using an amplitude-dependent autoregressive time series model“. In: *Biometrika* 68.1 (1981), S. 189–196.
- [HSJ21] C. Huang, S. Seidel, X. Jia, F. Paschke und J. Bräunig. „Energy Optimal Control of a Multivalent Building Energy System using Machine Learning“. In: *Proceedings of the 10th International Conference on Smart Cities and Green ICT Systems (SMARTGREENS)* (2021), S. 57–66.
- [HVB16] R. Halvgaard, L. Vandenberghe, N. Kjølstad Poulsen u. a. „Distributed Model Predictive Control for Smart Energy Systems“. In: *IEEE Transactions on Smart Grid* 7 (2016), S. 1675–1682.
- [HW14] R. Hoffmann und M. Wolff. *Intelligente Signalverarbeitung*. Springer Vieweg, 2014.
- [HW99] B. Huang und Z. Wang. „The Role of data prefiltering for integrated identification and model predictive control“. In: *Proceedings of the 14th IFAC World Congress* (1999), S. 6751–6756.
- [IM11] R. Isermann und M. Münchhof. *Identification of Dynamic Systems*. Springer-Verlag Berlin Heidelberg, 2011.

- [Irl16] J. P. Irl. „Analyse und Entwurf von Beobachtern mit unbekanntem Signalen und Parametern“. Diss. Karlsruher Institut für Technologie (KIT), Fakultät für Maschinenbau, 2016.
- [Ise06] R. Isermann. *Fault-Diagnosis Systems: An Introduction from Fault Detection to Fault Tolerance*. Springer, 2006.
- [Ise92a] R. Isermann. *Identifikation dynamischer Systeme 1*. 2. Aufl. Springer-Verlag Berlin Heidelberg, 1992.
- [Ise92b] R. Isermann. *Identifikation dynamischer Systeme 2*. 2. Aufl. Springer-Verlag Berlin Heidelberg, 1992.
- [Jak79] A. J. Jakeman. „Multivariable Instrumental Variable Estimators: The Choice between Alternatives“. In: *IFAC Proceedings Volumes* 12.8 (1979), S. 459–467.
- [Jaz70] A. H. Jazwinski. *Stochastic Processes and Filtering Theory*. Academic Press, 1970.
- [JBH19] F. Jorissen, W. Boydens und L. Helsens. „TACO, an automated toolchain for modelpredictive control of building systems: implementation and verification“. In: *Journal of Building Performance Simulation* 12.2 (2019), S. 180–192.
- [Jin12] H. Jin. „Parameter Estimation Based Models of Water Source Heat Pumps“. Diss. Oklahoma State University, 2012.
- [Jor18] F. Jorissen. „Toolchain for Optimal Control and Design of Energy Systems in Buildings“. Diss. KU Leuven – Faculty of Engineering Science, 2018.
- [JPS21] F. Jorissen, D. Picard, K. Six u. a. „Detailed White-Box Non-Linear Model Predictive Control for Scalable Building HVAC Control“. In: *Proceedings of the 14th International Modelica Conference* (2021).
- [JW68] G. M. Jenkins und D. G. Watts. *Spectral Analysis and Its Applications*. Holden-Day, San Francisco, 1968.
- [Kai80] T. Kailath. *Linear Systems*. Prentice Hall, 1980.
- [Kal60a] R. E. Kálmán. „A New Approach to Linear Filtering and Prediction Problems“. In: *Transactions of the ASME–Journal of Basic Engineering* 82 (1960), S. 35–45.
- [Kal63] R. E. Kálmán. „Lyapunov Functions for the Problem of Lur’e in Automatic Control“. In: *Proceedings of the National Academy of Sciences of the United States of America* 49.2 (1963), S. 201–205.
- [KB11] A. Kelman und F. Borrelli. „Bilinear Model Predictive Control of a HVAC System Using Sequential Quadratic Programming“. In: *IFAC Proceedings Volumes* 44.1 (2011), S. 9869–9874.
- [KB61] R. E. Kálmán und R. S. Bucy. „New Results in Linear Filtering and Prediction Theory“. In: *Journal of Basic Engineering* (1961), S. 95–108.

- [Kel99] C. T. Kelley. *Iterative Methods for Optimization*. Society for Industrial and Applied Mathematics Philadelphia, 1999.
- [KH15] M. Kavgić und T. Hilliard. „Opportunities for implementation of MPC in commercial buildings“. In: *Energy Procedia* 78 (2015), S. 2148–2153.
- [KN06] J. P. Kreiss und G. Neuhaus. *Einführung in die Zeitreihenanalyse*. Springer-Verlag Berlin Heidelberg, 2006.
- [Knu94] T. Knudsen. „A New Method for Estimating ARMAX Models“. In: *IFAC Proceedings Volumes* 27 (8 1994), S. 895–901.
- [Knu96] T. Knudsen. „The initialization problem in parameter estimating for general SISO models“. In: *IFAC Proceedings Volumes* 29 (1 1996), S. 4646–4651.
- [KSR08] M. Komareji, J. Stoustrup, H. Rasmussen u. a. „Optimal Model-Based Control in HVAC Systems“. In: *Proceedings of the American Control Conference* (2008), S. 1443–1448.
- [Kub20] S. Kuboth. „Modellprädiktive Regelung von Wärmepumpensystemen“. Diss. Universität Bayreuth, 2020.
- [KZ13] R. Kwadzogah und M. Zhou. „Model Predictive Control for HVAC Systems - A Review“. In: *Proceedings of the IEEE International Conference on Automation Science and Engineering (CASE)* (2013).
- [LAB12a] M. Y. Lamoudi, M. Alamir und P. Báguer. „Model predictive control for energy management in buildings Part 1: Zone Model Predictive Control“. In: *IFAC Proceedings Volumes* 45.17 (2012), S. 21–26.
- [LAB12b] M. Y. Lamoudi, M. Alamir und P. Báguer. „Model predictive control for energy management in buildings, Part 2: Distributed Model Predictive Control“. In: *IFAC Proceedings Volumes* 45.17 (2012), S. 226–231.
- [Lar83] W. E. Larimore. „System Identification, Reduced-Order Filtering and Modeling via Canonical Variate Analysis“. In: *Proceedings of the American Control Conference* (1983).
- [LB79] G. M. Ljung und G. E. P. Box. „The Likelihood Function of Stationary Autoregressive-Moving Average Models“. In: *Biometrika* 66.2 (1979), S. 265–270.
- [LC98] E. L. Lehmann und G. Casella. *Theory of Point Estimation*. Springer, 1998.
- [LE17] J. Linder und M. Enqvist. „Identification of Systems with Unknown Inputs Using Indirect Input Measurements“. In: *International Journal of Control* 90.4 (2017), S. 729–745.
- [LGG13] B. Lehmann, D. Gyalistras, M. Gwender u. a. „Intermediate complexity model for Model Predictive Control of Integrated Room Automation“. In: *Energy and Buildings* 58 (2013), S. 250–262.
- [Lin17] J. Linder. „Indirect System Identification for Unknown Input Problems With Applications to Ships“. Diss. Linköping University, 2017.

- [Lju12] L. Ljung. *System Identification: Theory for the User*. 2. Prentice Hall, 2012.
- [Lju76] L. Ljung. „On the Consistency of Prediction Error Identification Methods“. In: *Mathematics in Science and Engineering* 126 (1976), S. 121–164.
- [LLM98] I. D. Landau, R. Lozano und M. M’Saad. *Adaptive Control*. Springer-Verlag London, 1998.
- [LMT12] Y. Lin, T. Middelkoop und P. Barooah. „Identification of control-oriented thermal models of rooms in multi-room buildings“. In: *IEEE Transactions on Control Systems Technology* (2012).
- [Lud95] G. Ludyk. *Theoretische Regelungstechnik 1 und 2*. Springer Verlag, 1995.
- [Lun16a] J. Lunze. *Regelungstechnik 1: Systemtheoretische Grundlagen, Analyse und Entwurf einschleifiger Regelungen*. Springer Vieweg, 2016.
- [Lun16b] J. Lunze. *Regelungstechnik 2: Mehrgrößensysteme, Digitale Regelung*. Springer Vieweg, 2016.
- [LW14] H. Lutz und W. Wendt. *Taschenbuch der Regelungstechnik*. 10. Aufl. Verlag Europa-Lehrmittel, 2014.
- [LW19] H. Lutz und W. Wendt. *Taschenbuch der Regelungstechnik*. 11. Aufl. Verlag Europa-Lehrmittel, 2019.
- [LZ06] I. D. Landau und G. Zito. *Digital Control Systems*. Springer-Verlag London, 2006.
- [Mac00] J. M. Maciejowski. *Predictive control with constraints*. Prentice Hall, 2000.
- [MB12] Y. Ma und F. Borrelli. „Fast Stochastic Predictive Control for Building Temperature Regulation“. In: *Proceedings of the 2012 American Control Conference* (2012), S. 3075–3080.
- [MBD10] P.-D. Morosan, R. Bourdais, D. Dumur u. a. „Building temperature regulation using a distributed model predictive control“. In: *Energy and Buildings* 42 (9 2010), S. 1445–1452.
- [MBH09] Y. Ma, F. Borrelli, B. Hency u. a. „Model Predictive Control of Thermal Energy Storage in Building Cooling Systems“. In: *Proceedings of the Joint 48th IEEE Conference on Decision and Control and 28th Chinese Control Conference* (2009), S. 392–397.
- [MBH12] Y. Ma, F. Borrelli, B. Hency u. a. „Model Predictive Control for the Operation of Building Cooling Systems“. In: *IEEE Transactions on Control Systems Technology* 20.3 (2012), S. 796–803.
- [MCL10] G. Mustafaraj, J. Chen und G. Lowry. „Development of room temperature and relative humidity linear parametric models for an open office using BMS data“. In: *Energy and Buildings* 42.3 (2010), S. 348–356.
- [MCL11] G. Mustafaraj, J. Chen und G. Lowry. „Prediction of room temperature and relative humidity by autoregressive linear and nonlinear neural network models for an open office“. In: *Energy and Buildings* 42 (2011), S. 348–356.

- [MD16] A. Mirakhorli und B. Dong. „Occupancy behavior based model predictive control for building indoor climate - A critical review“. In: *Energy and Buildings* 129 (2016), S. 499–513.
- [Mes11] T. Mestekemper. „Energy Demand Forecasts and Dynamic Water Temperature Management“. Diss. Universität Bielefeld, Fakultät für Wirtschaftswissenschaften, 2011.
- [MH95] H. Madsen und J. Holst. „Estimation of continuous-time models for the heat dynamics of a building“. In: *Energy and Buildings* 22 (1 1995), S. 67–79.
- [MSV13] A. Van Mulders, J. Schoukens und L. Vanbeylen. „Identification of systems with localised nonlinearity: From state-space to block-structured models“. In: *Automatica* 49 (2013), S. 1392–1396.
- [MZ04] A. Mechaqrane und M. Zouak. „A comparison of linear and neural network ARX models applied to a prediction of the indoor temperature of a building“. In: *Neural Computing & Applications* 13 (2004), S. 32–37.
- [NA05] K. S. Narendra und A. M. Annaswamy. *Stable Adaptive Systems*. Dover Publications, Inc., 2005.
- [Nel01] O. Nelles. *Nonlinear System Identification*. Springer, 2001.
- [Nic12] A. Nicolai. *Physikalische Grundlagen des thermischen Raummodells THE-RAKLES*. Techn. Ber. TU Dresden, Institut für Bauklimatik, 2012.
- [NMI19] T. Naderi, D. Materassi, G. Innocenti u. a. „Revisiting Kalman and Aizerman Conjectures via a Graphical Interpretation“. In: *IEEE Transactions on Automatic Control* 64.2 (2019), S. 670–682.
- [NP90] K. S. Narendra und K. Parthasarathy. „Identification and control of dynamical systems using neural networks“. In: *IEEE Transactions on Neural Networks* 1.1 (1990).
- [OM93] P. Van Overschee und B. De Moor. „N4SID: Numerical Algorithms for State Space Subspace System Identification“. In: *IFAC Proceedings Volumes* 26.2 (1993), S. 55–58.
- [OM94] P. Van Overschee und B. De Moor. „Subspace algorithms for the identification of combined deterministic-stochastic systems“. In: *Automatica* 30.1 (1994), S. 75–93.
- [OM95] P. Van Overschee und B. De Moor. „A unifying theorem for three subspace system identification algorithms“. In: *Automatica* 31.12 (1995), S. 1853–1864.
- [OM96] P. Van Overschee und B. De Moor. *Subspace Identification for Linear Systems*. Kluwer Academic Publishers, 1996.
- [OPJ10] F. Oldewurtel, A. Parisio, C. N. Jones u. a. „Energy Efficient Building Climate Control using Stochastic Model Predictive Control and Weather Predictions“. In: *Proceedings of the American Control Conference* (2010), S. 5100–5105.

- [OPJ12] F. Oldewurtel, A. Parisio, C. N. Jones u. a. „Use of model predictive control and weather forecasts for energy efficient building climate control“. In: *Energy and Buildings* 45 (2012), S. 15–27.
- [ORD14] H. Ohlsson, L. Ratliff, R. Dong u. a. „Blind Identification via Lifting“. In: *IFAC Proceedings Volumes* 47.3 (2014), S. 10367–10372.
- [Oza81] T. Ozaki. „Nonlinear Threshold Autoregressive Models for Nonlinear Random Vibrations“. In: *Journal of Applied Probability* 18 (1981), S. 443–451.
- [Pas20] F. Paschke. „PEM-Identification of a block-oriented nonlinear stochastic model with application to room temperature modeling“. In: *Proceedings of the 28th Mediterranean Conference on Control and Automation* (2020), S. 351–356.
- [PCV12] S. Privara, J. Cigler, Z. Váňa u. a. „Incorporation of system steady state properties into subspace identification algorithm“. In: *International Journal of Modelling Identification and Control* 16.2 (2012), S. 159–167.
- [PCZ13] S. Privara, J. Cigler, V. Zdeněk u. a. „Building modeling as a crucial part for building predictive control“. In: *Energy and Buildings* 56 (2013), S. 8–22.
- [PFH16] F. Paschke, M. Franke und J. Haufe. „Modellprädiktive Einzelraumregelung auf Basis empirischer Modelle“. In: *Proceedings of the Central European Symposium on Building Physics 2016 - Bausim 2016* (2016), S. 405–411.
- [PFH17] F. Paschke, M. Franke und J. Haufe. „Modellprädiktive Einzelraumregelung auf Basis empirischer Modelle“. In: *Gebäudetechnik in Wissenschaft und Praxis* 138.6 (2017), S. 468–476.
- [PKJ11] F. L. Puente, U. Kiencke und H. Jäkel. *Signale und Systeme*. 5. Aufl. Oldenbourg Verlag, 2011.
- [PRG14] A. Parisio, E. Rikos und L. Glielmo. „A Model Predictive Control Approach to Microgrid Operation Optimization“. In: *IEEE Transactions on Control Systems Technology* 22.5 (2014), S. 1813–1827.
- [Pri13] S. Privara. „Building Modeling and Identification For Predictive Control“. Diss. CTU Prague, Department of Control Engineering, 2013.
- [PSF11] S. Privara, J. Široký, L. Ferkl u. a. „Model predictive control of a building heating system: The first experience“. In: *Energy and Buildings* 43 (2-3 2011), S. 564–572.
- [PWK17] A. Parisio, C. Wiezorek, T. Kyntäjä u. a. „Cooperative MPC-Based Energy Management for Networked Microgrids“. In: *IEEE Transactions on Smart Grid* 8.6 (2017), S. 3066–3074.
- [PZ18] F. Paschke und T. Zaiczek. „Identification of dynamic models for the short-term temperature prediction in a single room“. In: *IFAC-PapersOnLine* 51.2 (2018), S. 79–84.

- [PZR19] F. Paschke, T. Zaiczek und K. Röbenack. „Identification of room temperature models using k -step PEM for Hammerstein systems“. In: *Proceedings of the 23rd International Conference on System Theory, Control and Computing* (2019), S. 320–325.
- [Qin06] S. J. Qin. „An overview of subspace identification“. In: *Computers and Chemical Engineering* 30 (2006), S. 1502–1513.
- [Qua13] V. Quasching. *Regenerative Energiesysteme*. 8. Aufl. Hanser Verlag München, 2013.
- [Rak80] H. Rake. „Step Response And Frequency Response Methods“. In: *Automatica* 16 (1980), S. 519–526.
- [RBD90] D. T. Reindl, W. A. Beckman und J. A. Duffie. „Diffuse Fraction Correlations“. In: *Solar Energy* 45.1 (1990), S. 1–7.
- [Rei06] K. Reinschke. *Lineare Regelungs- und Steuerungstheorie*. Springer-Verlag, 2006.
- [RH11] J. Rehrl und M. Horn. „Temperature Control for HVAC Systems based on Exact Linearization and Model Predictive Control“. In: *Proceedings of the IEEE International Conference on Control Applications* (2011), S. 1119–1124.
- [Row70] I. H. Rowe. „A bootstrap method for the statistical estimation of model parameters“. In: *International Journal of Control* 12.5 (1970), S. 721–738.
- [RRT78] J. Richalet, A. Rault, J. L. Testud u. a. „Model predictive heuristic control: Applications to industrial processes“. In: *Automatica* 14 (5 1978), S. 413–428.
- [RSS11] H. Recknagel, E. Sprenger, E.-R. Schramek u. a., Hrsg. *Taschenbuch für Heizung und Klimatechnik*. 75. Aufl. Oldenbourg Industrieverlag München, 2011.
- [Rug81] W. J. Rugh. *Nonlinear System Theory - The Volterra/Wiener Approach*. The Johns Hopkins University Press, 1981.
- [RZV05] R. Raich, G. Zhou und M. Viberg. „Subspace based approaches for Wiener system identification“. In: *IEEE Transactions on Automatic Control* 50.10 (2005), S. 1629–1634.
- [SBB12] I. Staffell, D. Brett, N. Brandon u. a. „A review of domestic heat pumps“. In: *Energy & Environmental Science* 5.11 (2012), S. 9291–9306.
- [Sch00] B. De Schutter. „Minimal state-space realization in linear system theory: an overview“. In: *Journal of Computational and Applied Mathematics* 121.1-2 (2000), S. 331–354.
- [Sch80] M. Schetzen. *The Volterra and Wiener theories of nonlinear systems*. Wiley, 1980.

- [SFC18] G. Serale, M. Fiorentini, A. Capozzoli u. a. „Model Predictive Control (MPC) for Enhancing Building and HVAC System Energy Efficiency: Problem Formulation, Applications and Opportunities“. In: *Energies* 11 (3 2018), S. 631–665.
- [SGG86] C. E. De Souza, M. R. Gevers und G. C. Goodwin. „Riccati equations in optimal filtering of nonstabilizable systems having singular state transition matrices“. In: *IEEE Transactions on Automatic Control* 31.9 (1986), S. 831–838.
- [SGM08] J. Schoukens, L. Gommé, W. Van Moer u. a. „Identification of a Block-Structured Nonlinear Feedback System, Applied to a Microwave Crystal Detector“. In: *IEEE Transactions on Instrumentation and Measurement* 57 (8 2008), S. 1734–1740.
- [SGM16] D. Sturzenegger, D. Gyalistras, M. Morari u. a. „Model Predictive Climate Control of a Swiss Office Building: Implementation, Results, and Cost–Benefit Analysis“. In: *IEEE Transactions on Control Systems Technology* 24.1 (2016), S. 1–12.
- [SGS14] D. Sturzenegger, D. Gyalistras, V. Semeraro u. a. „BRCM Matlab Toolbox: Model Generation for Model Predictive Building Control.“ In: *Proceedings of the American Control Conference* (2014).
- [Sin80] Y. P. Singh und S. Subramanian. „Frequency-response identification of structure of nonlinear systems“. In: *IEE Proceedings D - Control Theory and Applications* 127.3 (1980), S. 77–82.
- [SKM14] D. Sturzenegger, D. Keusch, L. Muffato u. a. „Frequency-Domain Identification of a Ventilated Room for Model Based Control“. In: *IFAC Proceedings Volumes* 47.3 (2014), S. 593–5.
- [SLD15] T. B. Schön, F. Lindsten, J. Dahlin u. a. „Sequential Monte Carlo Methods for System Identification“. In: *IFAC-PapersOnLine* 48.28 (2015), S. 775–786.
- [SM05] P. Stoica und R. Moses. *Spectral Analysis of Signals*. Prentice Hall, 2005.
- [SMS92] D. S. Shook, C. Mohtadi und S. L. Shah. „A Control-Relevant Identification Strategy for GPC“. In: *IEEE Transactions on Automatic Control* 37.7 (1992), S. 975–980.
- [Soe02] T. Söderström. *Discrete-time Stochastic Systems*. Springer-Verlag London, 2002.
- [Soe92] R. Soeterboek. *Predictive control: A unified approach*. Prentice Hall, 1992.
- [SPP19] M. F. Shakib, A. Y. Pogromsky, A. Pavlov u. a. „Fast Identification of Continuous-Time Lur’e-type Systems with Stability Certification“. In: *IFAC PapersOnLine* 52.16 (2019), S. 227–232.

- [SRM09] P. Lopes dos Santos, J. Ramos und J. L. Martins de Carvalho. „Identification of a Benchmark Wiener-Hammerstein System by Bilinear and Hammerstein-Bilinear Models“. In: *IFAC Proceedings Volumes* 42.10 (2009), S. 832–837.
- [SS02] T. Söderström und P. Stoica. „Instrumental variable methods for system identification“. In: *Circuits Systems Signal Processing* 21.1 (2002), S. 1–9.
- [SS16] J. Sachs und O. Sawodny. „A Two-Stage Model Predictive Control Strategy for Economic Diesel-PV-Battery Island Microgrid Operation in Rural Areas“. In: *IEEE Transactions on Sustainable Energy* 7.3 (2016), S. 903–913.
- [SS83a] P. Stoica und T. Söderström. „Optimal instrumental-variable methods for identification of multivariable linear systems“. In: *Automatica* 19 (4 1983), S. 425–429.
- [SS83b] T. Söderström und P. Stoica. *Instrumental Variable Methods for System Identification*. Springer-Verlag, 1983.
- [SS84] T. Söderström und P. Stoica. „Uniqueness of estimated k -step prediction models of ARMA processes“. In: *Systems & Control Letters* 4.6 (1984), S. 325–331.
- [SS89] T. Söderström und P. Stoica. *System Identification*. Prentice Hall International, 1989.
- [SSA85] P. Stoica, T. Söderström, A. Ahlén u. a. „On the convergence of pseudo-linear regression algorithms“. In: *International Journal of Control* 41 (1985), S. 1429–1444.
- [ST17] M. Schoukens und K. Tiels. „Identification of Block-Oriented Nonlinear Systems starting from Linear Approximations: A Survey“. In: *Automatica* 85 (2017), S. 272–292.
- [Ste04] M. Steinmair. „Identifikation und Linearisierung nichtlinearer Leistungsverstärker mit Voltterrareihen“. Magisterarb. Technische Universität Wien: Fakultät für Elektrotechnik und Informationstechnik, 2004.
- [Stu14] D. Sturzenegger. „Model predictive building climate control - Steps towards practice“. Diss. ETH Zürich, 2014.
- [SWN11] T. Schön, A. Wills und B. Ninness. „System identification of nonlinear state-space models“. In: *Automatica* 47.1 (2011), S. 39–49.
- [TH05] P. Trnka und V. Havlena. „Subspace identification as multi-step predictions optimization“. In: *Proceedings of the IASTED International Conference on Modelling* (2005).
- [Toh08] E. Tohme. „Initialization of Output Error Identification Algorithms“. Diss. Universität Poitiers, 2008.
- [Tow94] N. C. E. Townsend. „Implementation of Control Policies for ARMAX Models“. In: *ULAM Quarterly* 2.4 (1994).

- [Unb11] H. Unbehauen. *Regelungstechnik III*. 7. Aufl. Vieweg+Teubner, 2011.
- [Van96] G. Vandersteen. „Non-parametric identification of the linear dynamic parts of non-linear systems containing one static non-linearity“. In: *Proceedings of 35th IEEE Conference on Decision and Control* (1996), S. 1099–1102.
- [VDI19] P. Stephan, S. Kabelac, M. Kind u. a. *VDI-Wärmeatlas*. 12. Aufl. 2019.
- [VDL12] C. Verhelst, D. Degrauwe, F. Logist u. a. „Multi-objective optimal control of an air-to-water heat pump for residential heating“. In: *Building Simulation* 5 (2012), S. 281–291.
- [Vie11] *Planungshandbuch Wärmepumpen*. Viessmann GmbH. Sep. 2011. URL: https://www.viessmann.de/content/dam/vi-brands/DE/PDF/Planungshandbuch/ph-waermepumpen.pdf/_jcr_content/renditions/original.media_file.download_attachment.file/ph-waermepumpen.pdf (besucht am 30.06.2021).
- [Vie19] *Energiesysteme für jede Anwendung*. Viessmann GmbH. Juli 2019. URL: https://www.viessmann.de/content/dam/vi-brands/DE/PDF/Programm/energiesysteme-fuer-jede-anwendung.pdf/_jcr_content/renditions/original.media_file.download_attachment.file/energiesysteme-fuer-jede-anwendung.pdf (besucht am 07.05.2020).
- [VS99] G. Vandersteen und J. Schoukens. „Measurement and identification of nonlinear systems consisting of linear dynamic blocks and one static nonlinearity“. In: *IEEE Transactions on Automatic Control* 44.6 (1999), S. 1266–1271.
- [VV07] M. Verhaegen und V. Verdult. *Filtering and System Identification: A Least Squares Approach*. Cambridge University Press, 2007.
- [Wah91] B. Wahlberg. „System Identification Using Laguerre Models“. In: *IEEE Transactions on Automatic Control* 36.5 (1991), S. 551–562.
- [Wah94] B. Wahlberg. „System Identification Using Kautz Models“. In: *IEEE Transactions on Automatic Control* 39.6 (1994), S. 1276–1282.
- [WB04] H. L. Wei und S. A. Billings. „A Unified Wavelet-Based Modelling Framework for Nonlinear System Identification the WANARX Model Structure“. In: *International Journal of Control* 77.4 (2004), S. 351–366.
- [Wet21] M. Wetter. *Modelica Buildings Library*. 19. März 2021. URL: <https://simulationresearch.lbl.gov/modelica/>.
- [Wig93] T. Wigren. „Recursive Prediction Error Identification Using the Nonlinear Wiener Model“. In: *Automatica* 29.4 (1993), S. 1011–1025.
- [Wim05] R. W. Wimmer. „Regelung einer Wärmepumpenanlage mit Model Predictive Control“. Diss. ETH Zürich, 2005.
- [WIO98] T. Wada, M. Ikeda, Y. Ohta u. a. „Parametric Absolute Stability of Lur’e Systems“. In: *IEEE Transactions on Automatic Control* 43.11 (1998), S. 1649–1653.

- [Wol38] H. O. A. Wold. „A Study in the Analysis of Stationary Time Series“. Diss. Univ. Stockholm, 1938.
- [WR86] M. A. Wincek und G. C. Reinsel. „An Exact Maximum Likelihood Estimation Procedure for Regression-ARMA Time Series Models with Possibly Nonconsecutive Data“. In: *Journal of the Royal Statistical Society* 48.3 (1986), S. 303–313.
- [WSL13] A. Wills, T. B. Schön, L. Ljung u. a. „Identification of Hammerstein-Wiener Models“. In: *Automatica* 49.1 (2013), S. 70–81.
- [WV96] D. Westwick und M. Verhaegen. „Identifying MIMO Wiener systems using subspace model identification methods“. In: *Signal Processing* 52.2 (1996), S. 235–258.
- [WZN11] M. Wetter, W. Zuo und T. S. Nouidui. „Modeling of Heat Transfer in Rooms in the Modelica Buildings Library“. In: *Proceedings of the 12th Conference of International Building Performance Simulation Association*. 2011, S. 1096–1103.
- [YP06] S. Yuan und R. Perez. „Multiple-zone ventilation and temperature control of a single-duct VAV system using model predictive strategy“. In: *Energy and Buildings* 38 (2006), S. 1248–1261.
- [Yul27] G. U. Yule. „On a Method of Investigating Periodicities in Disturbed Series, with Special Reference to Wolfer’s Sunspot Numbers“. In: *Philosophical Transactions of the Royal Society of London* 226 (1927), S. 267–298.
- [ZAN14] T. Zakula, P. R. Armstrong und L. Norford. „Modeling environment for model predictive control of buildings“. In: *Energy and Buildings* 85 (2014), S. 549–559.
- [ZB93] Q. M. Zhu und S. A. Billings. „Parameter Estimation for Stochastic Nonlinear Rational Models“. In: *International Journal of Control* 57.2 (1993), S. 309–33.
- [ZL05] G. Zito und I. D. Landau. „NARMAX model identification of a Variable Geometry Turbocharged diesel engine“. In: *Proceedings of the American Control Conference* (2005), S. 1021–1026.
- [ZPV11] E. Žáčková, S. Privara und Z. Váňa. „Model predictive control relevant identification using partial least squares for building modeling“. In: *Proceedings of the Australian Control Conference* (2011).
- [ZPY16] J. Zhao, K. Pho Lam, E. B. Ydstie u. a. „Occupant-oriented mixed-mode EnergyPlus predictive control simulation“. In: *Energy and Buildings* 117 (2016), S. 362–371.
- [ZSS18] E. Zhang, M. Schoukens und J. Schoukens. „Structure detection of Wiener-Hammerstein systems with process noise“. In: *IEEE Transactions on Instrumentation and Measurement* 66.3 (2017), S. 569–576.

- [ZZP14] J. Zhao, Y. Zhu und R. Patwardhan. „Identification of k -step-ahead prediction error model and MPC control“. In: *Journal of Process Control* 24 (2014), S. 48–56.

Abbildungsverzeichnis

2.1	Schematische Systemdarstellung	5
2.2	Lineares stochastisches Modell nach Gleichung (2.7)	9
2.3	Lineares stochastisches Modell im geschlossenen Regelkreis	17
2.4	Hammerstein-Wiener-Modell bzw. Wiener-Hammerstein-Modell	24
2.5	Lur'e und LFR-Modell	25
3.1	Klassifikation von Identifikationsverfahren für LTI-Systeme	30
3.2	Höhenlinien des quad. Kostenfunktional für einen ARX-Prozess	46
3.3	Amplitudenfrequenzgang für ein ARX-Modell	47
3.4	Höhenlinien des quad. Gütefunktional für ein System erster Ordnung mit Ausgangsquantisierung	48
3.5	Bilineares Beispielsystem	53
3.6	Simulation des bilinearen Beispielsystems	55
4.1	Funktionsebenen eines Gebäudeenergiesystems	58
4.2	Zuluft-Abluft-Kaskadenregelung	59
4.3	Vorlauftemperaturregelung und Heizkurvendiagramm	60
4.4	Beispielhafter Anlagenaufbau eines Mehrzonensystems	65
5.1	Schematische Darstellung des Konferenzraumes	75
5.2	RMMSE unterschiedlicher Hammerstein-Modelle (Konferenzraum)	78
5.3	Bode-Diagramm des Hammersteinmodells (Konferenzraum)	79
5.4	Box-Whisker-Diagramm und Autokovarianzfolge (Konferenzraum)	80
5.5	RMMSE unterschiedlich parametrierter nichtlinearer Modelle mit Ausgangsrückkopplung (Konferenzraum)	82
5.6	Box-Whisker-Diagramm und Autokovarianzfolge (Konferenzraum)	82
5.7	Prädiktion der Raumtemperatur (Konferenzraum)	83
5.8	Vergleich ausgangsrückgekoppelter Modelle mit und ohne saisonalem Rauschmodell (Büro)	85
5.9	Box-Whisker-Diagramm und Autokovarianzfolge (Turnhalle)	86
5.10	Turnhalle mit Deckenheizsystem	87
5.11	Box-Whisker-Diagramm und Autokovarianzfolge (Turnhalle)	88
5.12	Prädiktion der Temperatur einer Turnhalle	89
5.13	Schematische Darstellung der Hörsäle mit Lüftung	90
5.14	Beispielhaft aufgezeichnete Daten eines Hörsaals.	91
5.15	Schematische Darstellung eines Gas-Brennwertkessels	94
5.16	Heizleistung des Kessels in Abhängigkeit der Rücklauftemperatur	96
5.17	Prädiktion des Kesselmodells	97

5.18	Schematische Darstellung einer Wärmepumpe	98
5.19	Kennlinienfelder der Wärmepumpe	100
5.20	Prädiktion des Wärmepumpenmodells	101
6.1	Schematische Darstellung des Digitalregelkreises	104
6.2	Modelica Einzelraummodell mit Fußbodenheizung	106
6.3	Bode-Diagramm des Raummodells	107
6.4	Box-Whisker-Diagramm und Autokovarianzfolge des Raummodells mit Fußbodenheizung	108
6.5	Prädiktion der Raumtemperatur (Raum mit Fußbodenheizung)	109
6.6	Komfortabhängige Kosten (6.13) und Raumtemperaturgrenzwerte . . .	113
6.7	Vergleich einer konventionellen mit einer prädiktiven Regelstrategie für einen Raum mit Fußbodenheizung	114
6.8	Jahresvergleich einer konventionellen mit einer prädiktiven Regelstrategie für einen Raum mit Fußbodenheizung	115
6.9	Darstellung des Hörsaalgebäudes mit zentraler Lüftungsanlage	116
6.10	Grundriss und Materialparameter des Hörsaalgebäudes	117
6.11	Modelica Model des Hörsaalgebäudes	117
6.12	Box-Whisker-Diagramm und Autokovarianzfolge der Hörsaalmodelle . .	120
6.13	Prädiktion der Hörsaaltemperaturen	121
6.14	Referenzregelstrategie für das Hörsaalgebäude	123
6.15	Kosten- und Jahresenergievergleich für das Hörsaalgebäude	125
6.16	Verlauf der Jahresenergiekosten für das Hörsaalgebäude	125
6.17	Vergleich der konventionellen mit der MPC-Regelung (Hörsaalgebäude)	126
6.18	Optimierte Stell- bzw. Regelgrößenfolge (Hörsaalgebäude)	127
A.1	Dendrogramm spezieller stochastischer Prozesse	138
B.1	Vergleich der MkQ- und der Backforecasting-Methode	147
C.1	Lösungen der Heizkörpergleichungen	152
C.2	Dichte feuchter Luft	153
D.1	Vergleich System- und Modellausgang (Beispiel 1)	158
D.2	Vergleich Simulation System und Modell (Beispiel 3)	162

Tabellenverzeichnis

1	Typische SISO-Modellstrukturen als Spezialfälle von (2.36)	19
2	Übersicht blockorientierter Modellstrukturen	25
3	Kovarianzmatrizen für das ARX-Modell	46
4	Parameter der Modelle (3.51)	54
5	Klassifikation der publizierten MPC-Ansätze für Gebäude	64
6	Bedeutung der verwendeten Symbolik	73
7	Parameterschätzwerte des Büromodells	86
8	Geschätzte Parameter des Modells (Turnhalle)	88
9	Parameterschätzwerte des Kesselmodells	96
10	Parameterschätzwerte des Wärmepumpenmodells	101
11	Wandaufbau und -ausrichtung sowie Randbedingungen des Modelica Einzelraummodells.	105
12	Parameterschätzwerte des Raummodells mit Fußbodenheizung	107
13	Polynommatrizen des Modells (D.1)	155
14	Attribute der Klasse <code>idModels.NsfPolyModel</code>	156
15	Wichtige Optionen der Methode <code>identify()</code>	157
16	Wichtige Methoden der Klasse <code>idModels.NsfPolyModel</code>	157